UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE ODONTOLOGÍA



TESIS DOCTORAL

Odontología basada en la evidencia (OBE): de la evidencia a la recomendación

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Álvaro Limones Burgos

DIRECTORES

Juan Antonio Martínez Vázquez de Parga
María Concepción Martín Álvaro

© Álvaro Limones Burgos, 2024

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE ODONTOLOGÍA

TESIS DOCTORAL

ODONTOLOGÍA BASADA EN LA EVIDENCIA (OBE): DE LA EVIDENCIA A LA RECOMENDACIÓN

MEMORIA PARA OPTAR AL TÍTULO DE DOCTOR
PRESENTADA POR

Álvaro Limones Burgos

DIRECTORES

Juan Antonio Martínez Vázquez de Parga

María Concepción Martín Álvaro

A mis padres, Diego y Marisa,

A mis hermanos, Diego y Arturo.

AGRADECIMIENTOS

Al Profesor Doctor Juan Antonio Martínez Vázquez de Parga, por quien siento un agradecimiento que va más allá de lo que las palabras pueden expresar, por ser un regalo en mi vida, tanto personal como profesional.

A la Profesora Doctora María Concepción Martín Álvaro a quien admiro y respeto profundamente, convirtiéndose en mi referente investigador. Por su sensibilidad, amabilidad y alegría, por acompañarme en este proyecto tan excepcional superado con mucho humor y arduo trabajo.

A la Profesora Doctora Alicia Celemín Viñuela, por su generosidad, su cariño y atención constantes.

A todos los coautores de las publicaciones, con quienes ha sido un placer trabajar en equipo.

A todas las personas que de forma anónima aportaron sus esfuerzos altruistamente con palabras de ánimo.

Finalmente, de manera muy especial a D. Diego Espinoza por su amistad con mi recuerdo in memoriam.

ÍNDICE

1. RESUMEN 1

ABSTRACT 3

2. INTRODUCCIÓN 5

2.1. Contexto Histórico – *La transmisión del saber* **5**

2.2. Odontología Basada en la Evidencia (OBE) **8**

2.2.1. Fundamentos de la OBE: de la evidencia a la recomendación 9

2.3. Colaboración Cochrane **12**

2.4. Revisiones sistemáticas **14**

2.4.1. Definición 14

2.4.2. Tipos de Revisiones sistemáticas: 14

2.4.2.1. Revisiones sistemáticas de efectividad de las intervenciones 14

2.4.2.1.1. Evidencia sobre efectos adversos 15

2.4.2.1.2. Evidencia sobre las perspectivas de los pacientes 15

2.4.2.1.3. Evidencia económica 15

2.4.2.2. Revisiones sistemáticas de precisión de pruebas diagnósticas 16

2.4.2.3. Revisiones sistemáticas de prevalencia e incidencia 16

2.4.2.4. Revisiones sistemáticas de asociaciones etiológicas o factores de riesgo 17

2.4.2.5. Revisiones sistemáticas de estudios pronósticos y modelos predictivos 18

2.4.2.6. Revisiones sistemáticas de instrumentos y variables de medición 18

2.4.3. Herramientas de evaluación de una revisión sistemática: 18

2.4.3.1. Herramienta de evaluación metodológica: AMSTAR-II 18

2.4.3.2. Herramienta de evaluación del riesgo de sesgo: ROBIS 21

2.4.3.3. Herramienta de evaluación de la certeza de la evidencia: GRADE 25

2.4.3.3.1. Dominio 1 (*de disminución*) – Evaluación del riesgo de sesgo de los estudios primarios 27

2.4.3.3.2. Dominio 2 *(de disminución)* – Evaluación de la inconsistencia....38

2.4.3.3.2.1. Estadísticos para evaluar la heterogeneidad 40

2.4.3.3.2.2. Gráficos para evaluar la heterogeneidad 41

2.4.3.3.2.2.1. Gráficos de L'Abbe 41

2.4.3.3.2.2.2. Gráfico radial o Diagrama de Galbraith 42

2.4.3.3.2.3. Manejo de la heterogeneidad 42

2.4.3.3.3. Dominio 3 *(de disminución)* – Evaluación del carácter indirecto de la evidencia....44

2.4.3.3.4. Dominio 4 *(de disminución)* – Evaluación de la imprecisión 46

2.4.3.3.5. Dominio 5 *(de disminución)* – Evaluación del sesgo de publicación....49

2.4.3.3.5.1. Método visual.... 50

2.4.3.4.5.1.1. *Funnel plot o* diagrama de embudo 50

2.4.3.3.5.2. Métodos estadísticos 50

2.4.3.3.5.2.1. Método de Begg 51

2.4.3.3.5.2.2. Método de Egger 51

2.4.3.3.5.2.3. Método de Harbord.... 51

2.4.3.3.5.3. Métodos combinados de detección estadística y visual.... 51

2.4.3.3.5.3.1. Método de podar y completar (*Trim & fill*) 51

2.4.3.3.6. Dominio 6 *(de incremento)* – Evaluación de la existencia de "gran magnitud del efecto" 52

2.4.3.3.7. Dominio 7 *(de incremento)* – Evaluación del Gradiente Dosis-Respuesta (GDR).....53

2.4.3.3.8. Dominio 8 *(de incremento)* – Evaluación de factores de confusión plausibles que aumentan la confianza en el efecto estimado 54

2.4.3.4. Herramienta para guiar el reporte de revisiones sistemáticas: PRISMA.... 54

2.4.4. Métodos de presentación de la certeza de la evidencia: "Tablas resumen de hallazgos o de la evidencia".... 54

2.5. Otros métodos de síntesis de la evidencia 56

2.5.1. *Living systematic reviews* o Revisiones sistemáticas a tiempo real 56

2.5.2. *Overview / Meta-review / Umbrella review o* Revisión de revisiones sistemáticas 57

2.5.3. *Scoping reviews* o Revisiones de alcance.... 58

2.5.4. *Mapping reviews* o Revisiones de mapeo de la evidencia.... 58

© Álvaro Limones Burgos, 2024

2.5.5. *State-of-the-art reviews* o Revisiones del estado del arte 59

2.5.6. *Rapid reviews* o Revisiones rápidas 59

2.5.7. *Narrative reviews* o Revisiones narrativas o de la literatura 59

2.6. Metaanálisis 60

2.6.1 Definición 60

2.6.1.1. *Forest plot* o diagrama de Bosque o de efectos 60

2.6.1.2. Uso de medias relativas o absolutas 61

2.6.1.3. Modelo de efectos fijos o aleatorios 62

2.6.2. Metaanálisis de variables dicotómicas: 63

2.6.2.1. Método de Inverso de la varianza 63

2.6.2.2. Método de Mantel-Haenszel 63

2.6.2.3. Métodos para eventos raros o cero eventos 63

2.6.2.3.1. Factor de corrección 64

2.6.2.3.2. Método de Peto 64

2.6.3. Metaanálisis de variables continuas: 65

2.6.3.1. Método de Inverso de la Varianza 65

2.6.3.2. Metaanálisis de variables de cambio 65

2.6.4. Metaanálisis de combinación de variables dicotómicas y continuas 66

2.6.5. Metaanálisis de variables tiempo-hasta-el-evento: 67

2.6.6. Análisis de sensibilidad 67

2.7. Otros métodos metaanalíticos: 69

2.7.1. Metaanálisis acumulado 69

2.7.2. Metaanálisis en red *o Network Meta-analysis* 70

2.7.3. Metaanálisis Multivariante 71

2.7.4. Metaregresión 71

2.7.5. Gráfico de Albatross 71

2.8. Guías de práctica clínica y formulación de recomendaciones – Enfoque GRADE 73

2.9.1. Definición 73

2.9.2. Recomendaciones fuertes 74

2.9.3. Recomendaciones débiles 75

2.9.4. ¿Cuándo no hacer recomendaciones? 75

2.9.5. Factores que influyen en la dirección y la fuerza de la recomendación 75

2.9.6. Redacción de recomendaciones 78

2.9.7. Marcos de la Evidencia a la Decisión (EtD) 79

2.9. Iniciativa COMET para la estandarización de variables en ensayos clínicos 82

3. JUSTIFICACION 84

4. OBJETIVOS 87

4.1. Objetivo primario de carácter general en Odontología 87

4.2. Objetivo secundario de carácter específico en el campo de la Odontología Conservadora y la Prótesis Bucofacial 87

5. HIPÓTESIS 89

5.1. Hipótesis nula y alternativa primaria de carácter general en Odontología 89

5.2. Hipótesis nula y alternativa secundaria de carácter específico en el campo de la Odontología Conservadora y Prótesis Bucofacial 89

6. MATERIAL Y MÉTODO 91

6.1. Componente Teórica – Revisión del estado del arte: 91

6.1.1. Una revisión del estado del arte: 91

6.1.1.1. Criterios de inclusión 91

6.1.1.2. Criterio de exclusión 91

6.1.1.3. Estrategias de búsqueda 91

6.1.1.4. Extracción de datos 94

6.1.1.5. Método de síntesis 94

6.2. Componente Práctica – Revisiones sistemáticas publicadas: 95

6.2.1. Cuatro revisiones sistemáticas originales en diversas disciplinas odontológicas: 95

6.2.2. Una revisión sistemática de calidad, variables e instrumentos de medición en el campo de la Odontología Conservadora y Prótesis Bucofacial: 97

© Álvaro Limones Burgos, 2024

6.3. Recursos: 97

7. RESULTADOS 99

7.1. Componente Teórica – Revisión del estado del arte 99

7.1.1. Limitaciones intrínsecas al propio proceso de revisión, por la calidad metodológica 102

7.1.2. Limitaciones extrínsecas al proceso de revisión, por la evidencia disponible 104

7.2. Componente Práctica – Revisiones sistemáticas publicadas 106

7.2.1. PRIMERA REVISIÓN SISTEMÁTICA 106

7.2.1.1. AMSTAR-II: Resultados de la evaluación metodológica 106

7.2.1.2. ROBIS: Resultados de la evaluación del riesgo de sesgo 107

7.2.1.3. GRADE: Resultados de la evaluación de la certeza de la evidencia 107

7.2.1.4. PRISMA: Comprobación del cumplimiento de la guía de reporte 107

7.2.1.5. Resumen de la evaluación – limitaciones identificadas 108

7.2.1.6. Tabla resumen de hallazgos o de la evidencia 109

7.2.2. SEGUNDA REVISIÓN SISTEMÁTICA 110

7.2.2.1. AMSTAR-II: Resultados de la evaluación metodológica 110

7.2.2.2. ROBIS: Resultados de la evaluación del riesgo de sesgo 110

7.2.2.3. GRADE: Resultados de la evaluación de la certeza de la evidencia 111

7.2.2.4. PRISMA: Comprobación del cumplimiento de la guía de reporte 111

7.2.2.5. Resumen de la evaluación – limitaciones identificadas 111

7.2.2.6. Tabla resumen de hallazgos o de la evidencia 112

7.2.3. TERCERA REVISIÓN SISTEMÁTICA 114

7.2.3.1. AMSTAR-II: Resultados de la herramienta de evaluación metodológica 114

7.2.3.2. ROBIS: Resultados de la herramienta de evaluación del riesgo de sesgo 114

7.2.3.3. GRADE: Resultados de la evaluación de la certeza de la evidencia 114

7.2.3.4. PRISMA: Comprobación del cumplimiento de la guía de reporte 115

7.2.3.5. Resumen de la evaluación – limitaciones identificadas 115

7.2.3.6. Tabla resumen de hallazgos o de la evidencia 116

7.2.4. CUARTA REVISIÓN SISTEMÁTICA 118

7.2.4.1. AMSTAR-II: Resultados de la herramienta de evaluación metodológica 118

7.2.4.2. ROBIS: Resultados de la herramienta de evaluación del riesgo de sesgo 118

7.2.4.3. GRADE: Resultados de la evaluación de la certeza de la evidencia 118

7.2.4.4. PRISMA: Comprobación del cumplimiento de la guía de reporte 119

7.2.4.5. Resumen de la evaluación – limitaciones identificadas 119

7.2.4.6. Tabla resumen de hallazgos o de la evidencia 120

7.2.5. QUINTA REVISIÓN SISTEMÁTICA 122

7.2.5.1. AMSTAR II: Resultados de la herramienta de evaluación metodológica 122

7.2.5.2. ROBIS: Resultados de la herramienta de evaluación del riesgo de sesgo 122

7.2.5.3. PRISMA: Comprobación del cumplimiento de la guía de reporte 122

7.2.5.4. ROB 2.0 + CONSORT: Comprobación del riesgo de sesgo y la calidad del reporte de los ensayos clínicos publicados sobre prótesis fija dentosoportada 123

7.2.5.5. Resumen de la evaluación – limitaciones identificadas 126

7.2.5.6. Presentación de las características de los estudios primarios 126

7.2.6. TABLA RESUMEN DE LAS LIMITACIONES ENCONTRADAS 129

8. DISCUSIÓN 131

8.1. Componente Teórica – Revisión del estado del arte 131

8.1.1. Odontología Basada en la Evidencia: de la evidencia a la recomendación 131

8.1.2. Ventajas de las revisiones sistemáticas 132

8.1.3. Soluciones a las limitaciones de las revisiones sistemáticas 133

8.1.4. Importancia de la selección de marcos de relevancia clínica y el impacto del P-valor en la investigación odontológica 135

8.1.5. Necesidad de actualización de las revisiones sistemáticas 136

8.1.6. Evolución de las revisiones sistemáticas, un nuevo enfoque de la pirámide de la evidencia científica .. 136

8.1.7. Inteligencia artificial aplicada a la confección de las revisiones sistemáticas 138

8.2. Componente Práctica – Revisiones sistemáticas publicadas y evaluadas 139

8.2.1. Limitaciones globales identificadas que comprometen la formulación de recomendaciones fuertes en guías de práctica clínica 139

8.2.2. Situación actual de la investigación clínica publicada en próstesis fija dentosoporatada 144

© Álvaro Limones Burgos, 2024

9. LIMITACIONES 149

10. CONCLUSIONES 151

11. ANEXOS 153

ANEXO I: Manuscrito de la primera revisión sistemática 153

ANEXO 1.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II)..... 165

ANEXO 1.2: Resultados de la evaluación del riesgo de sesgo (ROBIS). 167

ANEXO 1.3: Resultados de la evaluación de la certeza de la evidencia científica (GRADE). 168

ANEXO 1.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA). .. 170

ANEXO II: Manuscrito de la segunda revisión sistemática 173

ANEXO 2.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II)..... 197

ANEXO 2.2: Resultados de la evaluación del riesgo de sesgo (ROBIS). 199

ANEXO 2.3 Resultados de la evaluación de la certeza de la evidencia científica (GRADE). 200

ANEXO 2.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA). .. 203

ANEXO III: Manuscrito de la tercera revisión sistemática 206

ANEXO 3.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II)..... 230

ANEXO 3.2: Resultados de la evaluación del riesgo de sesgo (ROBIS). 233

ANEXO 3.3: Resultados de la evaluación de la certeza de la evidencia científica (GRADE). 234

ANEXO 3.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA). .. 238

ANEXO IV: Manuscrito de la cuarta revisión sistemática 241

ANEXO 4.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II)..... 501

ANEXO 4.2: Resultados de la evaluación del riesgo de sesgo (ROBIS) 503

ANEXO 4.3: Resultados de la evaluación de la certeza de la evidencia científica (GRADE). 504

ANEXO 4.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA). .. 507

ANEXO V: Manuscrito de la quinta revisión sistemática 510

ANEXO 5.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II)..... 534

ANEXO 5.2: Resultados de la evaluación del riesgo de sesgo (ROBIS). 536

ANEXO 5.3: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA). .. 537

12. REFERENCIAS 540

GLOSARIO

ADA	*American Dental Association*
AGREE	*Appraisal of Guidelines for REsearch & Evaluation*
AMSTAR	*A MeaSurement Tool to Assess systematic Reviews*
AWMF	*Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften*
AXIS	*Appraisal tool for Cross-Sectional Studies*
BoP	*Bleeding on Probing*
CAL	*Clinical Attachment Loss*
CDA	*California Dental Association*
CHEERS	*Consolidated Health Economic Evaluation Reporting Standards*
CIE	*Commission Internationale de l'Eclairage*
COMET	*Core Outcome Measures in Effectiveness Trials*
CONSORT	*Consolidated Standards of Reporting Trials*
COS	*Core Outcome Set*
COSMIN	*COnsensus-based Standards for the selection of health Measurement INstruments*
COSMIN	*COnsensus based Standards for the selection of health Measurement Instruments*
CPG	*Clinical Practice Guideline*
DTA	*Diagnostic Test Accuracy*
ECA	Ensayo Controlado y Aleatorizado
EtD	*Evidence to Decision* (de la Evidencia a la Decisón)
FDI	Federación Dental Internacional
FDP	*Fixed Dental Prosthesis*
GDR	Gradiente Dosis-Respuesta
GPC	Guías de Práctica Clínica
GRADE	*Grading of Recommendations, Assessment, Development and Evaluation*
HR	*Hazard Ratio*
IC	Intervalo de Confianza
ITI	*International Team of Implantology*
ITT	*Intention-To-Treat*
MBE	Medicina Basada en la Evidencia

MOOSE	*Meta-analysis Of Observational Studies in Epidemiology*
MRONJ	*Medication-Related Osteonecrosis of the Jaws*
OBE	Odontología Basada en la Evidencia
OHIP	*Oral Health Impact Profile*
OIDP	*Oral Impacts on Daily Performances*
OIS	*Optimal Information Size*
OMS	Organización Mundial de la Salud
OR	*Odds Ratio*
OSS	*Oral Satisfaction Assessment*
PBE	Práctica basada en la Evidencia
PCR	*Plaque Control Record*
PEO	*Population, Exposure, Outcome*
PES / WES	*Pink Esthetic Score / White Esthetic Score*
PFD	Prótesis Fija Dentosoportada
PICO	*Population, Intervention, Comparator, Outcomes, Study, Time*
PP	*Per Protocol*
PPD	*Probing Pocket Depth*
PRISMA	*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*
QoL	*Quality of Life*
RCT	*Randomized Controlled Trial*
RoB	*Risk of Bias*
ROBIS	*Risk of Bias in Systematic Reviews*
ROC	*Receiver Operating Characteristic*
RR	*Risk Ratio*
RS	Revisión Sistemática
SoFt	*Summary of Findings Table*
SEPES	Sociedad Española de Prótesis Estomatológica y Estética
SR	*Systematic Review*
STARR	*SelecTing Approaches for Rapid Review*
USPHS	*U.S. Public Health Service*
VAS	*Visual Analogue Scale*

© Álvaro Limones Burgos, 2024

1. RESUMEN

Antecedentes. Estamos viviendo un aumento significativo y en alza de la producción científica, que representa un desafío constante para los profesionales de la Odontología en su búsqueda por mantenerse actualizados en beneficio del paciente. Esto resalta la importancia de la síntesis y el filtrado de la evidencia mediante Revisiones Sistemáticas (RS), y la formulación de recomendaciones mediante Guías de Práctica Clínica (GPC) para trasladar la ciencia al quehacer diario de los clínicos.

Objetivos. El objetivo primario de la presente disertación de Tesis Doctoral fue analizar las limitaciones de las RS en Odontología y su repercusión en el desarrollo de recomendaciones fuertes a través de GPC. El objetivo secundario fue analizar la investigación clínica publicada en Prótesis Fija Dentosoportada (PFD).

Materiales y Métodos. El análisis de las limitaciones se llevó a cabo desde un enfoque dual: teórico y práctico. La concepción teórica consistió en una revisión del estado del arte para identificar las limitaciones de las RS de manera conceptual; y la práctica, en la realización de 4 RS originales en diferentes especialidades odontológicas (Prótesis, Medicina oral, Implantología y Ortodoncia), para explorar y confirmar estas limitaciones en contextos reales. Una 5ª RS evaluó la cantidad, calidad, consistencia metodológica y representatividad clínica de los Ensayos clínicos Controlados y Aleatorizados (ECA) publicados en los últimos 10 años en el campo de la PFD.

Resultados y Discusión. La concepción teórica identificó y clasificó las limitaciones de las RS en: intrínsecas (relacionadas con la propia metodología de las RS) y extrínsecas (vinculadas a los estudios incluidos). Evaluó las fortalezas, soluciones a las limitaciones, el estado actual y las perspectivas futuras de este diseño de estudio. La concepción práctica mostró que, en un contexto real, la presencia de limitaciones extrínsecas en una de nuestras RS comprometió el establecimiento de recomendaciones fuertes a través de una GPC nivel S3. Esto también se observó en el resto de las RS consideradas en esta disertación. El sustrato para desarrollar recomendaciones fuertes, presentado en Tablas Resumen de Hallazgos, se vio comprometido principalmente por una baja o muy baja certeza de la evidencia, atribuida a una escasez de ECA y participantes, no alcanzando, a menudo, el tamaño óptimo de información. Las inconsistencias en las intervenciones

© Álvaro Limones Burgos, 2024

y las variables de evaluación fueron el segundo motivo más común para disminuir la certeza de la evidencia, aunque se resolvió moderadamente mediante análisis de subgrupos. También se observó el carácter indirecto de la evidencia, debido a que se disponía de resultados a corto plazo para fundamentar decisiones clínicas que cobran relevancia en el largo plazo.
En la evaluación de la investigación clínica en PFD, fue notoria la diversidad de variables e instrumentos de medición utilizados, la no consideración de efectos adversos y algunas preocupaciones metodológicas, no significativas, que pueden mejorarse.

Conclusiones. El enfoque dual de esta disertación demostró cómo la escasez de ECA, la presencia de inconsistencias entre estudios, el carácter indirecto de la evidencia, el riesgo de sesgo y la no consideración de variables críticas como efectos adversos, análisis económicos, o satisfacción del paciente, comprometen el desarrollo de GPC; e identificó fortalezas y limitaciones de las RS, sugiriendo soluciones, y enfatizando la importancia de seleccionar marcos de relevancia clínica, actualizar periódicamente las revisiones, y explorar las nuevas aplicaciones de inteligencia artificial para su confección. La evaluación de la investigación clínica publicada en PFD subrayó la necesidad de estandarizar un conjunto mínimo de variables relevantes en este campo y de potenciar la calidad y claridad metodológica en la presentación de las publicaciones.

© Álvaro Limones Burgos, 2024

ABSTRACT

Background. We are experiencing a significant and rising increase in scientific production, which poses a constant challenge for dental professionals in their quest to stay updated for the benefit of the patient. This further highlights the importance of synthesizing and filtering evidence through Systematic Reviews (SRs), and the formulation of recommendations through Clinical Practice Guidelines (CPGs) to translate science into the daily practice of clinicians.

Objectives. The primary objective of this Doctoral thesis dissertation was to analyze the limitations of SRs in Dentistry and their impact on the development of strong recommendations through CPGs. The secondary objective was to analyze clinical research in tooth-supported Fixed Dental Prosthesis (FDP).

Materials and Methods. The analysis of limitations was carried out from a dual approach: theoretical and practical. The theoretical conception involved a state-of-the-art review to identify the conceptual limitations of SRs; and the practical, in the conduct of 4 original SRs in different dental specialties (prosthesis, oral medicine, implantology, and orthodontics), to explore and confirm these limitations in real contexts. A 5th SR assessed the quantity, quality, methodological consistency, and clinical representativeness of Randomized Controlled Trials (RCTs) published in the last 10 years in the field of tooth-supported FDP.

Results and Discussion. The theoretical conception identified and classified the limitations of SRs into intrinsic (related to the methodology of SRs itself) and extrinsic (linked to the included studies). It evaluated the advantages, the current status, and future perspectives of this study design. The practical conception showed that, in a real context, the presence of extrinsic limitations in one of our SRs compromised the establishment of strong recommendations through a level S3 CPG. This was also observed in the rest of the SRs considered in this dissertation. The substrate for developing strong recommendations, presented in Summary of Findings Tables, was mainly compromised by low or very low certainty of evidence, attributed to a shortage of RCTs and

© Álvaro Limones Burgos, 2024

participants, often not reaching the optimal information size. Inconsistencies in interventions and evaluation variables were the second most common reason for decreasing the certainty of evidence, although it was moderately resolved through subgroup analysis. The indirect nature of the evidence was also noted, due to the availability of short-term results to support clinical decisions that are relevant in the long term.

In the evaluation of clinical research in tooth-supported FDP, the diversity of outcome measures and outcome instruments used, the non-consideration of adverse effects, and some, non-significant, methodological concerns that can be improved were notable.

Conclusions. The dual approach of this dissertation demonstrated how the scarcity of RCTs, the presence of inconsistencies among studies, the indirect nature of the evidence, the risk of bias, and the failure to consider critical variables such as adverse effects, economic analyses, or patient satisfaction, compromise the development of CPGs; and identified strengths and limitations of systematic reviews, suggesting solutions, and emphasizing the importance of periodically updating reviews, selecting frameworks of clinical relevance, and exploring new applications of artificial intelligence in their development. The evaluation of clinical research published in PDFs highlighted the need to standardize a minimal set of relevant variables in this field and to enhance the quality and methodological clarity in the presentation of publications.

© Álvaro Limones Burgos, 2024

© Álvaro Limones Burgos, 2024

2. INTRODUCCIÓN

2.1. Contexto Histórico – *La transmisión del saber*

Desde los albores de la prehistoria, **la primera transmisión del conocimiento humano se realizaba de manera oral y práctica**. Los hombres primitivos, cazadores y recolectores, poseían conocimiento de diferentes técnicas de caza, pesca, fabricación de armas, utensilios y vestimentas; así como métodos para conservar alimentos mediante el secado y el ahumado (Vallejo, 2023).

Posteriormente, las **primitivas pinturas rupestres,** como las halladas en las cuevas de Altamira (Cantabria, Santander) y Lascaux (Dordoña, Francia), constituyeron un intento temprano por documentar la experiencia humana, sus actividades diarias y creencias, a través de imágenes.

Hacia el año 3000 a.C., emergieron en las grandes civilizaciones de Egipto y Mesopotamia dos **sistemas de escritura** revolucionarios: los jeroglíficos egipcios y la escritura cuneiforme mesopotámica. Estas formas de escritura permitieron una documentación más detallada y permanente de datos, acontecimientos históricos, leyes, y avances técnicos y científicos. Sin embargo, fue con los griegos y la aparición de influyentes filósofos como Sócrates, Platón y Aristóteles, además de destacados científicos como Arquímedes y Euclides, que se sentaron las bases de diversas disciplinas científicas. La civilización griega contribuyó significativamente a la difusión del conocimiento mediante la creación de **los primeros libros**, inicialmente en forma de colecciones de rollos de papiro. A medida que la demanda de materiales de escritura aumentaba, en Pérgamo se innovó utilizando pieles de animales tratadas para escribir, dando origen al pergamino, y su agrupación, a los códices. El papel nace en China hacia el año 200 a.C., pero fue introducido en Europa por los árabes, estableciéndose los primeros talleres de papel en Córdoba en 1036. El primer molino de papel en Europa se fundó en Játiva en 1144. En España, Alfonso X El Sabio adoptó el papel para la documentación oficial, superando las limitaciones del pergamino, cuyo costo y escasez restringían su uso (Vallejo, 2023).

Las **bibliotecas** han sido históricamente bastiones para la protección y el almacenamiento del conocimiento. Un caso destacado es la Biblioteca de Alejandría, cuyo impacto en la conservación del conocimiento y la ciencia fue trascendental, representando un esfuerzo

significativo por preservar la sabiduría de la humanidad. Siguiendo los pasos de la civilización helénica, los romanos expandieron sus conocimientos en áreas como la filosofía, destacando figuras como Séneca; la ingeniería con Apolodoro de Damasco como referente; la medicina de la mano de Galeno; y las estrategias militares ilustradas por Julio César y sus legiones. En esa época, florecieron las bibliotecas públicas, como la fundada por el emperador Trajano en el Foro de Trajano, aunque también era habitual que los nobles romanos mantuvieran bibliotecas personales llenas de rollos de papiros. Globalmente, civilizaciones como la bizantina, la musulmana y las de Extremo Oriente siguieron avanzando, intercambiando y conservando el conocimiento a través de registros y bibliotecas en lugares emblemáticos como Estambul, Córdoba, Damasco, Japón y China.

Tras la caída del imperio romano de Occidente ante las invasiones bárbaras, se inicia un período en el cual la cultura y la conservación del conocimiento en diversas áreas se concentran principalmente en **los monasterios**. Estos adoptaron prácticas para la preservación cultural, incluyendo talleres de copia de libros y códices, y el intercambio o venta de estos a otras comunidades religiosas o laicas, tales como otros monasterios, obispados e incluso cortes reales (Yeves, 2019). Los musulmanes jugaron un papel crucial al reintroducir textos de filósofos griegos previamente perdidos y avanzados conceptos matemáticos, como el cero, originario de India; propiciando significativos progresos en matemáticas y astronomía. La creación de la Escuela de Traductores de Toledo por Alfonso X "El Sabio" fue fundamental en este sentido. Además, las cruzadas y los intercambios con Bizancio fomentaron el desarrollo científico y técnico, impulsando avances en agricultura, ganadería, arquitectura, y técnicas en metalurgia, química y la introducción de la pólvora, inventada por los chinos.

En el siglo XIII, el franciscano inglés Roger Bacon, filósofo, precursor de la ciencia moderna y teólogo, fue **uno de los primeros pensadores del moderno método científico**. Además, durante este siglo y el siguiente, las primeras universidades emergen de las escuelas episcopales en lugares como Bolonia (Italia), París (Francia), y Salamanca (España), promoviendo la enseñanza y el debate de ideas en un entorno de libertad académica. Estas instituciones no se limitaban a la teología, extendiendo su saber al derecho, economía, medicina, farmacia y otros campos del conocimiento. Sin embargo, es en el Renacimiento, con su redescubrimiento de la cultura clásica, cuando pensadores, artistas y arquitectos como Leonardo Da Vinci fomentan una

© Álvaro Limones Burgos, 2024

profunda exploración de la naturaleza y sus leyes, marcando el comienzo de una era de transformación (Asimov, 2014).

El siglo XV presenció un avance crucial en la historia y difusión del conocimiento con la **invención de la imprenta** por Johannes Gutenberg en 1440, facilitando la reproducción de textos a bajo costo. Este invento coincidió con un aumento significativo en la alfabetización, gracias a las escuelas eclesiásticas y urbanas. La imprenta permitió replicar rápidamente textos existentes a un precio accesible y en grandes cantidades. También, el descubrimiento de América por Cristóbal Colón en 1492 abrió un nuevo continente lleno de conocimientos, culturas y oportunidades, impulsando esfuerzos técnicos, económicos y científicos para integrar esta riqueza de información. En 1536, la primera imprenta se estableció en México, seguida por la fundación de universidades en Lima y México en 1551, y la Universidad de Harvard en 1636 en las colonias inglesas (Montaner, 1887).

La aplicación práctica de conocimientos fundamentados en la experiencia empírica y la observación sistemática, tal como se describe en los libros y tratados de Gilbert, Torricelli, Pascal o Boyle, entre otros, señala **el comienzo del desarrollo del método científico**. Este enfoque se formalizó en el siglo XVIII con Isaac Newton, quien propuso una metodología científica rigurosa en sus estudios sobre la descomposición de la luz en diferentes colores y formulación de las leyes de la gravitación universal. Este período también vio el nacimiento del cálculo infinitesimal, una herramienta matemática esencial para la ciencia, desarrollada independientemente por Gottfried Leibniz (Asimov, 2014).

En los siglos XVIII y XIX, la difusión del conocimiento científico trascendió los límites de libros y universidades para incluir sociedades científicas, museos, ateneos, conferencias, etc.

En el siglo XX, las **publicaciones científicas** empezaron a convertirse en el principal canal para la difusión y el intercambio de conocimiento científico. Fue entonces cuando Eugene Garfield abordó la necesidad de organizar la información científica en su obra *Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas*", siendo el precursor de los índices de impacto y el *Journal Citation Reports* (JCR), una herramienta que rige, a día de hoy, el prestigio de las revistas científicas (Garfield, 2006).

En el siglo XXI, la digitalización, **Internet**, el desarrollo de sistemas computacionales capaces de gestionar mayores volúmenes de datos, las mejoras en las herramientas de búsqueda, la creación de registros internacionales y plataformas *online* de intercambio de datos, entre otros

© Álvaro Limones Burgos, 2024

avances, han impulsado de manera significativa la producción científica, creando una cantidad abrumadora de información para los profesionales de la salud. A menudo, los clínicos pueden cometer el error de confiar en estudios publicados en revistas de alto impacto o con muchas citaciones, como si estos fueran indicadores sólidos y fiables para fundamentar sus prácticas clínicas con seguridad. Sin embargo, un alto número de citas no es una razón plausible de calidad; los artículos pueden ser citados por ser controvertidos, provocadores o incluso erróneos. Además, la prevalencia de **artículos con manipulación de datos y metodologías cuestionables** subraya la importancia de un escrutinio riguroso y métodos de síntesis para transmitir el saber científico, a la recomendación clínica (Andrade, 2021; Fire & Guestrin, 2019).

2.2. Odontología Basada en la Evidencia (OBE)

En el pasado, se solía dar por sentado que los tratamientos médicos se basaban en la mejor evidencia científica disponible, pero en realidad, esto no siempre era el caso. Fue la conceptualización de la Medicina Basada en la Evidencia (MBE) en la década de 1990, la que cambió esta situación para mejor (Evidence-Based Medicine Working Group, 1992). Este concepto fue acuñado por el Dr. David Sackett, el Dr. Gordon Guyatt y su equipo de trabajo de la Universidad McMaster (Hamilton, Ontario, Canadá) en el artículo clásico titulado: "*Evidence-Based Medicine: What It Is and What It Isn't"* (Sackett et al., 1996). En dicho artículo, se asentó **la importancia de combinar el juicio clínico de los profesionales de la salud, con las preferencias de los pacientes y la evidencia científica proveniente de investigaciones rigurosas** para tomar decisiones médicas informadas, posicionando a la MBE como la piedra angular de la práctica médica moderna.

El mismo enfoque se orientó hacia la Odontología (Richards & Lawrence, 1995), en un afán por cimentar la toma de decisiones clínicas también en este campo. La *American Dental Association* (ADA) definió la Odontología Basada en la Evidencia (OBE) como un enfoque de atención bucodental que requiere una prudente integración de: evaluaciones sistemáticas de variables de interés clínico (variables críticas), el historial médico y bucodental del paciente, la idoneidad clínica del odontólogo, y las necesidades y preferencias del paciente (Figura 1) (O'Loughlin, 2020).

© Álvaro Limones Burgos, 2024

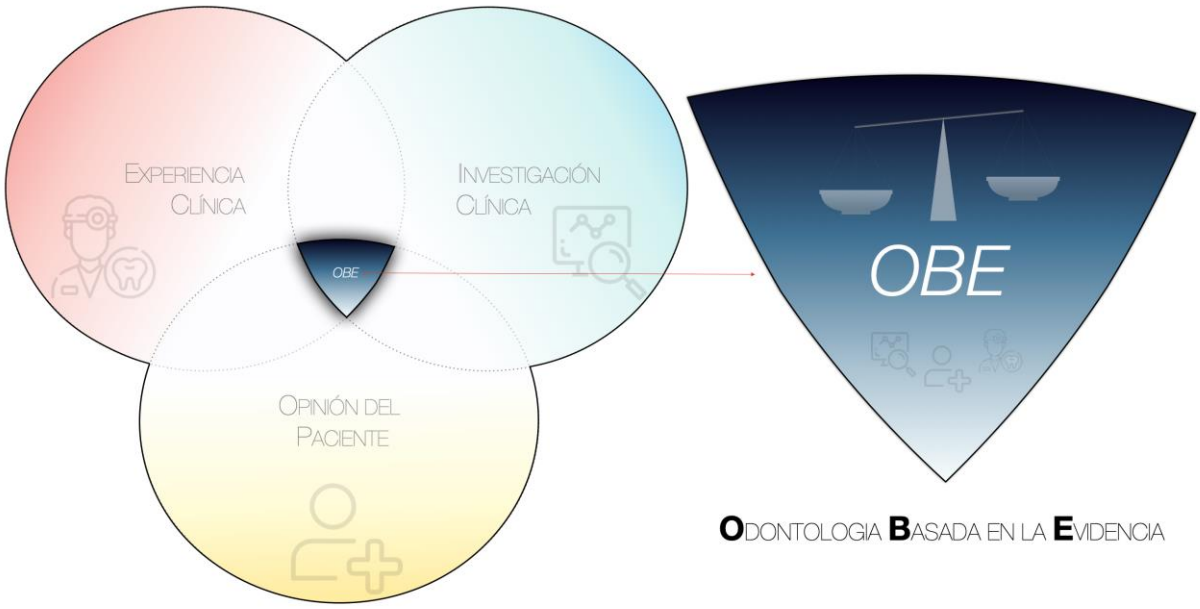


Figura 1. Infografía de la OBE.

En la actualidad, se espera que un odontólogo tenga un amplio conocimiento de todas las opciones terapéuticas disponibles, lo que incluye, en algunos casos, la decisión de no intervenir. Existe, por tanto, una tremenda necesidad de acercar la investigación a la práctica clínica dental y optimizar la información disponible de alta calidad para los clínicos y pacientes, ya que a menudo se ven abrumados por la extensa y exponencialmente creciente cantidad de literatura existente.

Sin lugar a dudas, la OBE demuestra su trascendencia en este punto al implementar Guías de Práctica Clínica (GPC) fundamentadas y amigables, en beneficio de nuestros pacientes y de la excelencia en la atención odontológica.

2.2.1. Fundamentos de la OBE: de la evidencia a la recomendación

El proceso de transformación de la evidencia a la recomendación se basa en una serie de fundamentos. Comienza con la realización de una Revisión Sistemática (RS) y avanza hacia el desarrollo de GPC, las cuales se distribuyen a través de sociedades científicas, llegando finalmente a los entornos de atención odontológica (Figura 2).

© Álvaro Limones Burgos, 2024

Todo comienza con la formulación de una pregunta de investigación clínica bien definida, utilizando el formato PICO (por sus siglas en inglés, *Population, Intervention, Comparator, Outcomes*), lo que implica identificar la Población de interés, la Intervención de estudio, el Comparador (si es relevante) y las Variables críticas (Outcome, en inglés), importantes o no críticas para la toma de decisiones.

Una vez que la pregunta PICO está bien estructurada, se procede a realizar una búsqueda exhaustiva y sistemática de la evidencia científica pertinente y disponible a través de búsquedas electrónicas. Los resultados de los estudios seleccionados se sintetizan para identificar patrones, tendencias y conclusiones generales. Esto incluye calcular estimaciones combinadas del efecto, evaluar su consistencia, riesgo de sesgo y publicación, y determinar el nivel de confianza en la evidencia final mediante herramientas específicas, que se explicarán a lo largo de esta disertación (i.e. la herramienta GRADE, por sus siglas en inglés, *Grading of Recommendations, Assessment, Development and Evaluation*). Es aquí donde termina la función de una RS, con la presentación de "Tabla resumen de Hallazgos o de Evidencia" (*Summary of Findings Table*, SoFt, por sus siglas en inglés), que contiene las estimaciones del efecto y el nivel de confianza que disponemos en esas estimaciones, lo que se conoce como certeza de la evidencia.

Posteriormente, basándose en los hallazgos de las RS presentados en "Tabla resumen de Hallazgos o de Evidencia", se formulan recomendaciones clínicas dentro de GPC elaboradas por consenso de un panel de expertos. Estas recomendaciones se sustentan en la evaluación del balance neto entre los efectos beneficiosos y perjudiciales, las preferencias de los pacientes y las consideraciones económicas, entre otros, para establecer una dirección (a favor o en contra) y un grado (fuerte o débil) de recomendación. Las GPC se revisan y actualizan regularmente a medida que surgen nuevas investigaciones, lo que garantiza que las recomendaciones clínicas estén respaldadas por la información más actualizada.

En la fase final, las sociedades científicas desempeñan un papel crucial al difundir estas GPC, permitiendo así que los profesionales de la salud dental puedan beneficiarse de ellas y fundamentar su ejercicio diario en la OBE. Esto contribuye a la mejora de la seguridad de los pacientes, mantiene a los clínicos actualizados en su práctica profesional y promueve la excelencia en su desempeño.

© Álvaro Limones Burgos, 2024

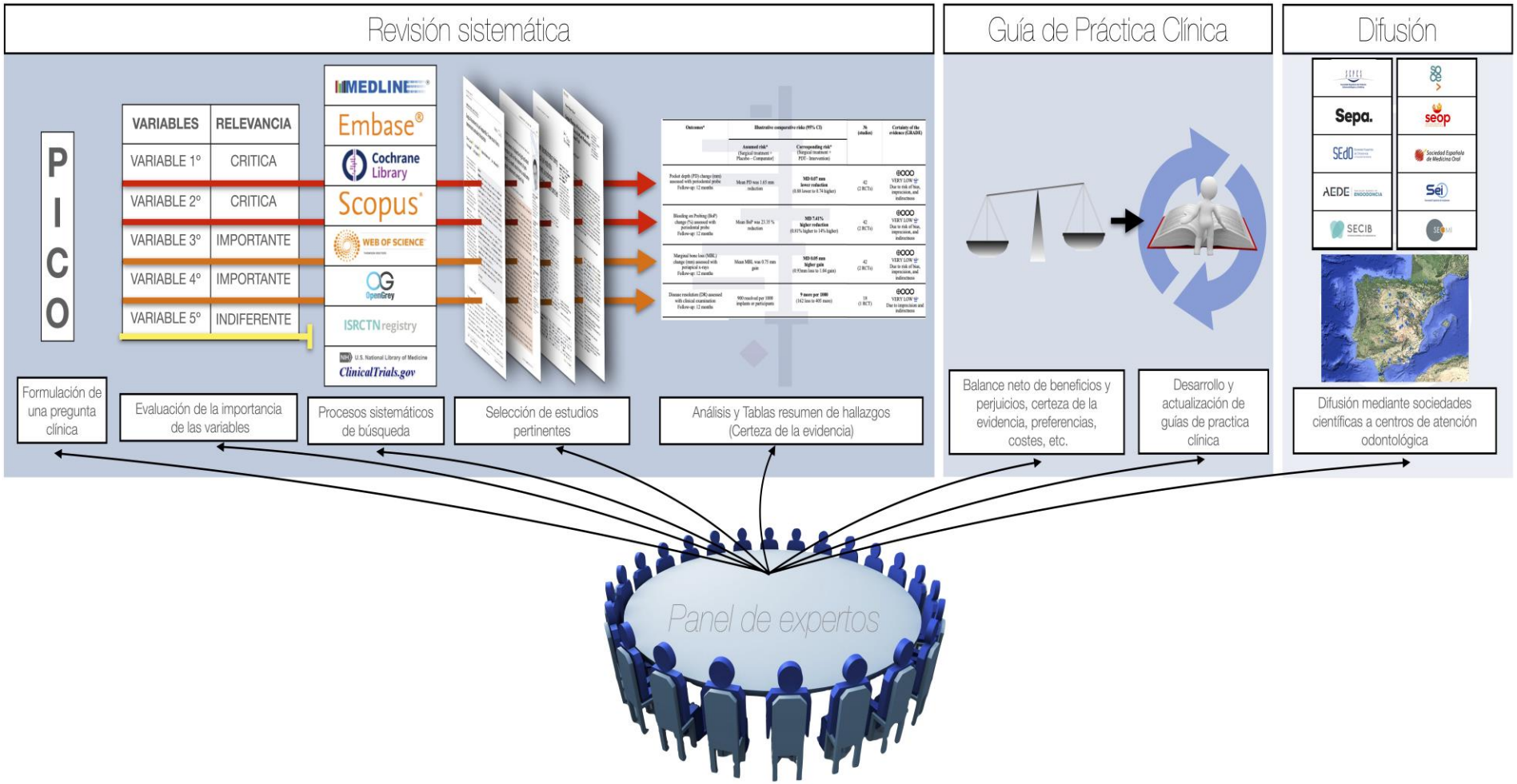


Figura 2. Visión esquemática de la transición de la síntesis de la evidencia a la formulación de recomendaciones.

© Álvaro Limones Burgos, 2024

2.3. Colaboración Cochrane

La Colaboración Cochrane es la organización de mayor relevancia internacional dedicada a la preparación y el mantenimiento de RS de alta calidad sobre intervenciones sanitarias y precisión de pruebas diagnósticas, entre otras, con el objetivo de proporcionar evidencia confiable para la toma de decisiones en salud. Esta red internacional, abarca todas las áreas de la salud y todos los aspectos de la atención sanitaria, desde la prevención y detección de enfermedades, hasta el tratamiento y la rehabilitación.

Tomó su nombre en honor al considerado padre de la MBE, el Dr. Archie Cochrane, un destacado investigador médico británico que resaltó la importancia de fundamentar la atención médica en evidencia científica sólida (Archibald L. Cochrane, 1973). En 1993, el Dr. Iain Chalmers estableció oficialmente la Colaboración Cochrane en Oxford (Reino Unido) con el propósito de materializar la visión de Archie Cochrane.

En los últimos 30 años (Figura 3), la Colaboración Cochrane ha experimentado un notable crecimiento, convirtiéndose en una organización multinacional que involucra a más de 130 países de todo el mundo y cuenta con más de 11.000 miembros voluntarios que no reciben una compensación directa.



Figura 3. Treinta años de la Colaboración Cochrane.

Es una organización sin ánimo de lucro y con independencia financiera, que no acepta financiación de intereses comerciales ni tiene conflictos de interés, lo que asegura la imparcialidad

© Álvaro Limones Burgos, 2024

y la integridad de las revisiones que produce. Esto ha llevado a que las revisiones Cochrane sean ampliamente reconocidas y utilizadas en GPC. Un indicador importante de su impacto es que el 75% de las recientes directrices de la Organización Mundial de la Salud (OMS) han incorporado revisiones Cochrane como parte esencial de su fundamento (Fig. 4).

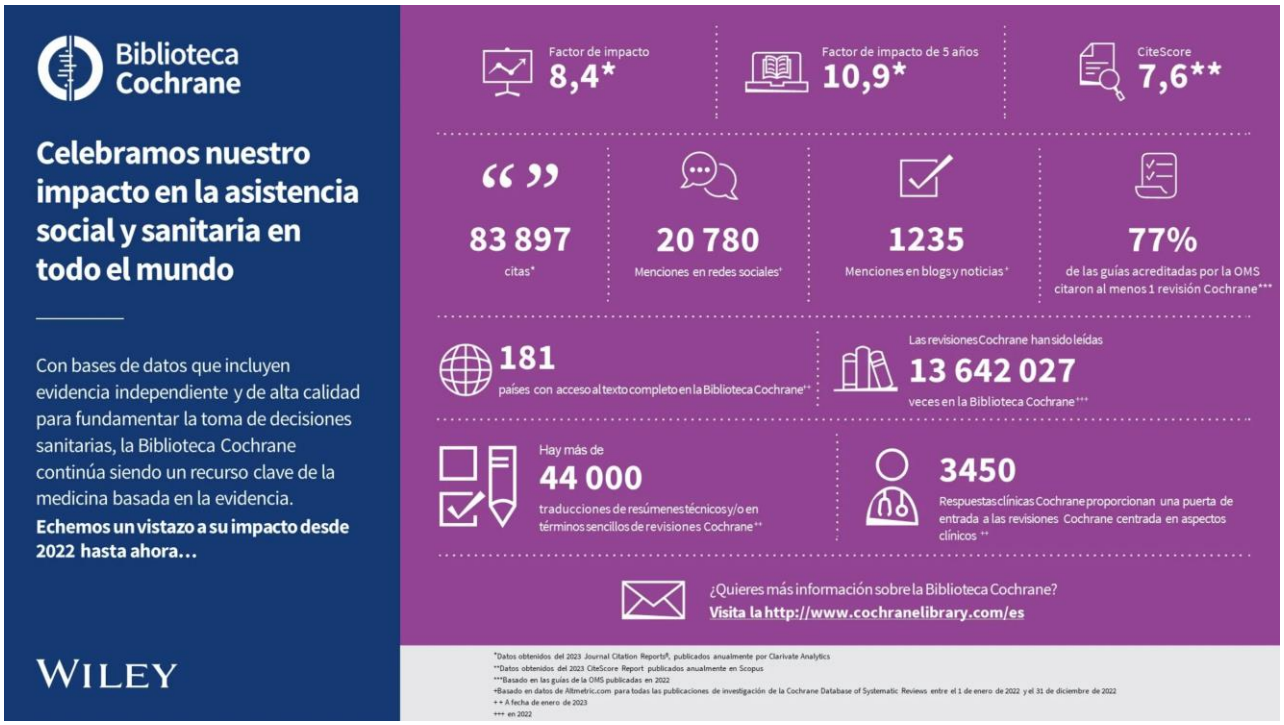


Figura 4. Métricas más recientes de la revista *Cochrane Database of Systematic Reviews.*

© Álvaro Limones Burgos, 2024

2.4. Revisiones sistemáticas

2.4.1. Definición

La Colaboración Cochrane, institución de reconocido prestigio, y de la que forma parte nuestra Facultad de Odontología de la Universidad Complutense de Madrid, define la Revisión Sistemática (RS) como un diseño de estudio científico que busca responder a una pregunta de investigación clínica mediante la identificación, selección, evaluación y el análisis descriptivo y/o estadístico (metaanálisis) de todos los estudios relevantes con métodos sistemáticos y explícitos (Higgins et al., 2022), culminando con la presentación de una "Tabla Resumen de Hallazgos o de Evidencia" (Carrasco-Labra et al., 2016), que incluye un epítome con información sobre la magnitud estimada del efecto y el grado de confianza (certeza de la evidencia) que tenemos en que esa estimación refleje el efecto real.

El manual Cochrane ofrece una guía para su elaboración (Higgins et al., 2022) y la declaración PRISMA (de sus siglas en inglés: *Preferred Reporting Items for Systematic Reviews and Meta-Analyses)* una guía para su reporte en publicaciones científicas (Page et al., 2021).

2.4.2. Tipos de Revisiones sistemáticas:

Las RS pueden clasificarse de acuerdo a sus objetivos en:

2.4.2.1. Revisiones sistemáticas de efectividad de las intervenciones

Las RS de efectividad de las intervenciones tienen como objetivo evaluar la efectividad y/o seguridad de una o más intervenciones destinadas a la prevención, tratamiento y/o rehabilitación de una enfermedad o afección específica. Este tipo de revisión generalmente se enfoca en la búsqueda y recopilación de ensayos clínicos aleatorizados (ECA), aunque en ocasiones también puede incluir estudios observacionales cuando los ECA no proporcionan suficientes datos sobre efectos adversos u otros aspectos relevantes. Además de examinar aspectos clínicos, estas revisiones también abordan la evidencia sobre efectos adversos, la evidencia económica a través

© Álvaro Limones Burgos, 2024

de relaciones coste-efectividad, coste-beneficio, coste-utilidad o coste-minimización (Gomersall et al., 2014); y la evidencia sobre la opinión y preferencias de los pacientes y clínicos (McArthur et al., 2020) de cada intervención. A continuación, se describen estos aspectos más detalladamente:

2.4.2.1.1. Evidencia sobre efectos adversos

La evidencia relacionada con los efectos adversos de una intervención desempeña un papel crucial en el desarrollo de GPC, como veremos más adelante. La detección y evaluación de efectos adversos pueden requerir diversos enfoques metodológicos. Algunos son poco comunes o se manifiestan considerablemente después de la intervención, mientras que otros son más frecuentes y se presentan poco tiempo después de esta. Los ensayos clínicos de pequeño tamaño con seguimiento a corto plazo pueden ser adecuados para identificar efectos adversos comunes y de aparición inmediata. No obstante, para evaluar efectos adversos raros o de larga duración, es ocasiones es necesario incluir estudios observacionales, como estudios de cohortes o estudios de casos y controles.

2.4.2.1.2. Evidencia sobre las perspectivas de los pacientes

Es de suma importancia para el desarrollo de GPC tener en cuenta también las perspectivas de los pacientes sobre las intervenciones que se estén evaluando, a menudo recopiladas a través de cuestionarios de autoevaluación. Esto incluiría cualquier variable que refleje el estado de salud tal como lo informa el propio paciente, sin la necesidad de la interpretación de un profesional u operador.

Aunque es posible llevar a cabo una RS con un enfoque exclusivo en la evaluación de las perspectivas y preferencias de los pacientes, por lo general, esta evaluación se considera como objetivo secundario dentro del contexto de una RS de efectividad de las intervenciones.

2.4.2.1.3. Evidencia económica

También juega un papel fundamental en el desarrollo de GPC la evaluación de la evidencia económica, que se enfoca en analizar cómo se asignan los recursos, generalmente limitados, a las intervenciones que se están evaluando. Estos recursos no se limitan únicamente a aspectos monetarios, sino que también incluyen otros recursos como el

© Álvaro Limones Burgos, 2024

tiempo, el personal, la energía u otros elementos utilizados para llevar a cabo una intervención específica. Aunque se puede realizar una RS con un enfoque exclusivo en la evidencia económica, normalmente es evaluada como un objetivo secundario dentro de una RS de efectividad de las intervenciones, ya que la evidencia económica no debe separarse de la evidencia de la efectividad de las intervenciones, para evitar interpretaciones sesgadas. Las guías de reporte CHEERS (por sus siglas en inglés, *Consolidated Health Economic Evaluation Reporting Standards*) son útiles para la presentación de evaluaciones económicas (Husereau et al., 2022).

2.4.2.2. Revisiones sistemáticas de precisión de pruebas diagnósticas

Las RS de precisión de pruebas diagnósticas tienen como objetivo evaluar la evidencia disponible sobre la precisión de una o más pruebas diagnósticas para identificar la presencia o ausencia de una enfermedad o afección específica en una población determinada. Estas revisiones incluyen estudios de precisión diagnóstica que evalúan parámetros como la sensibilidad, la especificidad, los valores predictivos y las curvas ROC (por sus siglas en inglés, *Receiver Operating Characteristic*) (Higgins et al., 2022). La Colaboración Cochrane ha publicado un manual para guiar a los autores en la elaboración de RS de precisión de pruebas diagnósticas (Deeks et al., 2023), y la declaración PRISMA ha desarrollado una extensión específica para guiar a los autores en el reporte de este tipo de revisiones (Salameh et al., 2020) conocida como PRISMA-DTA (por sus siglas en inglés: *Diagnostic Test Accuracy)*.

2.4.2.3. Revisiones sistemáticas de prevalencia e incidencia

Las RS de prevalencia e incidencia sintetizan datos sobre la proporción de individuos afectados por determinadas enfermedades en poblaciones específicas, utilizando frecuentemente diseños de estudio transversal o prospectivo, respectivamente (Borges Migliavaca et al., 2020). Este enfoque ha ganado impulso en la última década, evidenciando su importancia en la mejora de la toma de decisiones para profesionales de la salud y responsables políticos. Al ofrecer una visión clara de la distribución y tendencias temporales de las enfermedades, estas revisiones apoyan la asignación eficaz de recursos y el desarrollo de políticas de salud pública basadas en evidencia,

© Álvaro Limones Burgos, 2024

marcando una herramienta indispensable para la gestión de la salud poblacional. Sin embargo, no cuenta todavía con una guía de presentación de informes aceptada y publicada.

2.4.2.4. Revisiones sistemáticas de asociaciones etiológicas o factores de riesgo

Las RS de estudios pronósticos tienen como objetivo identificar y resumir la evidencia sobre las asociaciones entre factores de exposición/riesgo y resultados en salud. Estos análisis generalmente involucran uno o más tipos de asociaciones y se basan en estudios observacionales, como estudios de cohortes, estudios de casos y controles, o estudios transversales, ya que estos estudios son más adecuados para investigar asociaciones epidemiológicas que los ECA. La iniciativa MOOSE (por sus siglas en inglés: *Meta-analysis Of Observational Studies in Epidemiology*) proporciona pautas para la conducción y presentación de RS y metaanálisis de estudios observacionales (Stroup et al., 2000). Generalmente, en este tipo de revisiones, no se recomienda la síntesis cuantitativa de resultados (metaanálisis), ya que los estudios incluidos generalmente presentan heterogeneidad en el diseño y las poblaciones estudiadas, lo que limita la capacidad para combinar resultados. En estos casos, los gráficos de Albatross pueden ser una herramienta útil para mostrar una visión general de la distribución del efecto de las intervenciones (Figura 5) (Harrison et al., 2017).

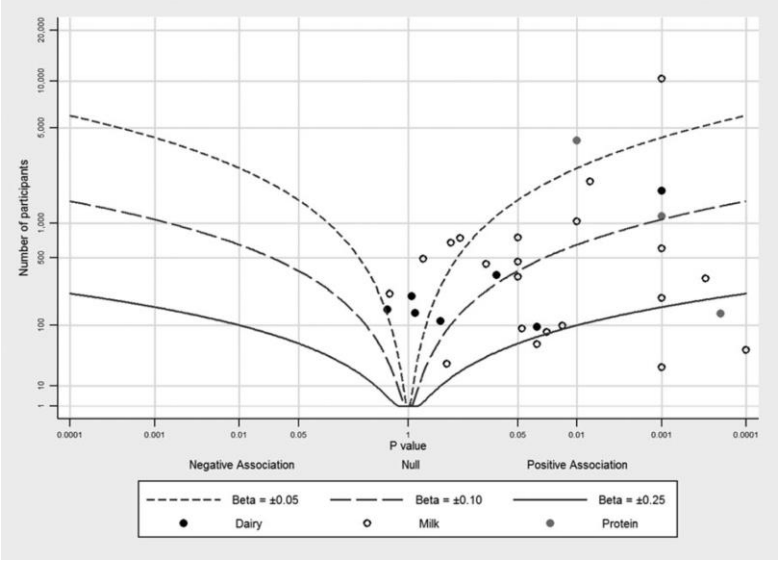


Figura 5. Gráfico de Albatross representado un efecto favorable del factor de exposición (Harrison et al., 2017)

© Álvaro Limones Burgos, 2024

2.4.2.5. Revisiones sistemáticas de estudios pronósticos y modelos predictivos

Una RS de estudios pronósticos evalúa críticamente y recopila datos de estudios que incluyen biomarcadores, factores y modelos pronósticos multivariables, para estimar la probabilidad o riesgo de que una enfermedad o condición esté presente (modelos de diagnóstico) o de que ocurra un evento específico en el futuro (modelos pronósticos) (Damen et al., 2023). Esto mejora significativamente la toma de decisiones clínicas y la planificación de tratamientos.

2.4.2.6. Revisiones sistemáticas de instrumentos y variables de medición

Las RS de instrumentos y variables de medición tienen como objetivo identificar de manera sistemática el número y la diversidad de variables e instrumentos de medición utilizados para evaluar la efectividad de las intervenciones. Una selección adecuada es fundamental para asegurar la representatividad clínica y la homogeneidad entre los estudios. Con este propósito, la Iniciativa COMET (*Core Outcome Measures in Effectiveness Trials*) se dedica a promover la estandarización de un conjunto mínimo de variables, conocido como *Core Outcome Sets* (COS), que deben medirse y reportarse en todos los estudios en una área de salud específica. Esta iniciativa busca mejorar la consistencia de la investigación y la calidad. Actualmente, se está elaborando una extensión PRISMA para guiar a los autores en el reporte de RS de instrumentos y variables de medición, conocida como PRISMA-COSMIN, de sus siglas en inglés: *COnsensus based Standards for the selection of health Measurement Instruments* (Elsman et al., 2022).

2.4.3. Herramientas de evaluación de una revisión sistemática:

En la literatura científica existen diversos métodos de evaluación de las RS, abordados desde distintas perspectivas:

2.4.3.1. Herramienta de evaluación metodológica: AMSTAR-II

La herramienta AMSTAR (por sus siglas en inglés: *A MeaSurement Tool to Assess systematic Reviews*) fue publicada en 2007 originalmente para evaluar la calidad

© Álvaro Limones Burgos, 2024

metodológica de RS de ensayos clínicos controlados y aleatorizados (Shea et al., 2007). Sin embargo, la inclusión de estudios no aleatorizados en RS se ha hecho cada vez más común y plantea nuevos sesgos que requieren evaluaciones diferentes. Por este motivo, una nueva actualización de esta herramienta fue publicada en 2017, bajo el nombre AMSTAR-II (Shea et al., 2017) con una versión intuitiva en línea a través de la página web https://amstar.ca . La herramienta consta de 16 preguntas que evalúan diferentes aspectos de una RS (Tabla 1).

Tabla 1. Preguntas de evaluación metodológica de la herramienta AMSTAR-II.

Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]
Pregunta 1	-	Pregunta de investigación y criterios de inclusión: ¿Se establecieron preguntas de investigación y criterios de inclusión que incluyeron los componentes de PICO (Población, Intervención, Grupo comparador, Resultado)?
Pregunta 2	**CRÍTICA**	**Protocolo registrado antes del inicio de la revisión:** ¿Se registró o publicó el protocolo de la revisión antes de empezar su realización y se han justificado las desviaciones significativas del protocolo?
Pregunta 3	-	Diseño de estudios seleccionados: ¿Se explica por los autores cual es el diseño de estudio que se ha seleccionado para confeccionar la revisión?
Pregunta 4	**CRÍTICA**	**Búsqueda exhaustiva de la literatura:** ¿Los autores de la revisión utilizaron una estrategia de búsqueda exhaustiva?
Pregunta 5	-	Duplicación en la selección de estudios: ¿Los autores de la revisión realizaron la selección de estudios por duplicado?
Pregunta 6		Duplicación en la extracción de datos: ¿Los autores de la revisión realizaron la extracción de datos por duplicado?
Pregunta 7	**CRÍTICA**	**Justificación para excluir estudios individuales:** ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?
Pregunta 8	-	Descripción de estudios incluidos: ¿Los autores de la revisión describieron los estudios incluidos con suficiente detalle?
Pregunta 9	**CRÍTICA**	**Evaluación del riesgo de sesgo de los estudios primarios incluidos:** ¿Utilizaron los autores de la revisión una técnica satisfactoria para evaluar el

© Álvaro Limones Burgos, 2024

Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]
		riesgo de sesgo (RoB) en los estudios individuales que se incluyeron en la revisión?
Pregunta 10	-	Fuentes de financiación: ¿Los autores de la revisión informaron sobre las fuentes de financiación de los estudios incluidos en la revisión?
Pregunta 11	**CRÍTICA**	**Adecuados métodos analíticos:** Si se realizó un metaanálisis, ¿los autores de la revisión utilizaron métodos apropiados para la combinación estadística de resultados?
Pregunta 12	-	Impacto del RoB en la metaanálisis: Si se realizó un metaanálisis, ¿los autores de la revisión evaluaron el impacto potencial de RoB de los estudios individuales sobre los resultados del metaanálisis u otra síntesis de evidencia?
Pregunta 13	**CRÍTICA**	**Impacto del RoB en los resultados y discusión:** ¿Los revisores tuvieron en cuenta el RoB de los estudios primarios al interpretar/discutir los resultados de la revisión?
Pregunta 14	-	Evaluación de la Heterogeneidad: ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?
Pregunta 15	**CRÍTICA**	**Evaluación de la presencia e impacto del sesgo de publicación:** Si se realizó una síntesis cuantitativa, ¿se llevó a cabo una investigación adecuada del sesgo de publicación (sesgo de estudios pequeños) y se discutió su probable impacto en los resultados de la revisión?
Pregunta 16	-	Conflictos de interés: ¿Se informaron las posibles fuentes de conflicto de interés, incluyendo cualquier financiamiento recibido para llevar a cabo la revisión?

La herramienta AMSTAR-II no está diseñada para generar una puntuación general, ya que una alta puntuación puede ocultar debilidades críticas en aspectos específicos de la revisión. Su evaluación general se fundamenta en revisar si cumple o no cumple cada dominio presentado en la guía, con especial énfasis en el impacto que puedan tener los **dominios críticos** en la validez de la revisión, y que son fundamentales para asegurar una alta calidad metodológica. Así, la confianza general en los resultados de una RS se evalúa en cuatro niveles: alta confianza, confianza moderada, baja confianza y baja confianza crítica de acuerdo con los criterios presentados en la Tabla 2.

© Álvaro Limones Burgos, 2024

Tabla 2. Calificación de la confianza general en los resultados de la revisión mediante la herramienta AMSTAR-II.

Confianza	Valoración	Interpretación
Alta Confianza	Se aplica cuando la revisión sistemática no tiene ninguna debilidad crítica y máximo una debilidad no crítica.	En este caso, se considera que la revisión proporciona un resumen preciso y completo de los resultados de los estudios disponibles que abordan la pregunta de interés.
Confianza Moderada	Se otorga cuando la revisión sistemática no tiene ninguna debilidad crítica pero más de una debilidad no crítica. Aun que múltiples debilidades no críticas pueden disminuir la confianza en la revisión y puede ser apropiado bajar la evaluación general de confianza moderada a baja.	A pesar de estas debilidades, la revisión puede ofrecer un resumen preciso de los resultados de los estudios disponibles incluidos en la revisión.
Baja Confianza	Esta categoría se asigna cuando la revisión tiene al menos una debilidad crítica, independientemente de si también tiene debilidades no críticas.	En este caso, la revisión puede no proporcionar un resumen preciso y completo de los estudios disponibles que abordan la pregunta de interés.
Baja Confianza Crítica	Se aplica cuando la revisión tiene más de una debilidad crítica, ya sea con o sin debilidades no críticas.	En esta situación, la revisión no debe considerarse confiable para proporcionar un resumen preciso y completo de los estudios disponibles.

2.4.3.2. Herramienta de evaluación del riesgo de sesgo: ROBIS

Desarrollada en 2016, la herramienta ROBIS (*Risk of Bias in Systematic Reviews*, por sus siglas en inglés) se centra en analizar el riesgo de sesgo en las RS, mientras que la herramienta AMSTAR-II se enfoca en evaluar la calidad metodológica. Aunque ambos conceptos están estrechamente relacionados, no pueden tratarse de manera equivalente (Whiting et al., 2016). Consta de tres fases: en primer lugar, evaluar la idoneidad de la pregunta enfocada (opcional); en segundo lugar, identificar problemas en el proceso de revisión; y, por último, juzgar el riesgo de sesgo global.

© Álvaro Limones Burgos, 2024

La primera fase es opcional y su propósito es evaluar si la pregunta de investigación expuesta por los autores en su publicación se adecúa al formato PICO, o sus análogos (i.e. PEO – *Population, Exposure, Outcome*), y a los objetivos de la revisión. Si una o más de las categorías PICO no coinciden entre los evaluadores y los autores de la revisión, se califica como "No". Por ejemplo, si la pregunta se refiere a adultos, pero la RS se limita a participantes mayores de 60 años (no estaría teniendo en cuenta los adultos comprendidos entre 18 y 59 años).

La segunda fase se organiza en cuatro dominios, que incluyen preguntas para identificar sesgos, a las cuales se responde con "sí (S)", "probablemente sí (PS)", "probablemente no (PN)", "no (N)" o "sin información (NI)". Un resumen de la herramienta se presenta en la Tabla 3.

Finalmente, la tercera fase considera el riesgo de sesgo global de la RS a partir de tres preguntas claves:

1. La primera pregunta se enfoca en identificar la presencia de algún sesgo durante la segunda fase entre los dominios del uno al cuatro. El riesgo de sesgo se clasifica en varios niveles (Inexistente/bajo/alto/incierto) de acuerdo con las respuestas proporcionadas (S/PS/N/PN/I) en los dominios. La inexistencia de riesgo de sesgo se establece si todas las respuestas a las preguntas de los diferentes dominios indican que no existe riesgo de sesgo y se han respondido con un “Sí”. Por otro lado, se considera riesgo de sesgo bajo cuando las respuestas son principalmente "sí" o "probablemente sí" en todas las preguntas de los diferentes dominios. En contraste, se habla de un riesgo de sesgo alto cuando las respuestas son "no" o "probablemente no" en alguna de las preguntas de los diferentes dominios. La categoría de riesgo de sesgo incierto se utiliza solo en casos en los que no se disponga de datos suficientes para emitir un juicio definitivo sobre el riesgo de sesgo.
1. La segunda pregunta verifica si los estudios incluidos son adecuados para responder a la pregunta de investigación en términos de su aplicabilidad clínica y validez externa.

© Álvaro Limones Burgos, 2024

2. La tercera pregunta comprueba si al evaluar múltiples variables se proporciona una descripción equilibrada de todas ellas, evitando mostrar preferencia por las que son estadísticamente significativas. Esto es importante para no sobreestimar interpretaciones y la introducción de sesgos.

© Álvaro Limones Burgos, 2024

1ª FASE	2ª FASE				3ª FASE
Identificación de la pregunta de investigación y sus componentes	**Criterios de elegibilidad (Dominio 1)**	**Identificación y selección de estudios (Dominio 2)**	**Recopilación de datos y evaluación de estudios (Dominio 3)**	**Síntesis y conclusiones (Dominio 4)**	**Riesgo de sesgo global**
¿La pregunta abordada en la revisión está correctamente definida de acuerdo a sus objetivos? S/PS/PN/N/NI* Revisiones de intervención: *P: Población* *I: Intervención* *C: Comparador* *O: Variables* Revisiones de pruebas de precisión diagnóstica: *P: Población* *I: Test diagnóstico* *C: Estándar de referencia* *O: Enfermedad objetivo* Revisiones de asociaciones etiológicas o factores de riesgo: *P: Población* *I: Exposición* *C: Comparador* *O: Variables* Revisiones de estudios pronóstico: *P: Población* *I: Factor pronóstico o modelo predictivo* *O: Desenlace a predecir* Revisiones de variables e instrumentos de medición: *P: Población/Campo de conocimiento* *O: Variables e instrumentos de medición*	1.1 ¿Se ajustó la revisión a los objetivos y criterios de elegibilidad predefinidos? S/PS/PN/N/NI* 1.2 ¿Los criterios de elegibilidad eran apropiados para la pregunta de la revisión? S/PS/PN/N/NI* 1.3 ¿Eran inequívocos los criterios de elegibilidad? S/PS/PN/N/NI* 1.4 ¿Eran apropiadas las restricciones de los criterios de elegibilidad basadas en las características del estudio (por ejemplo, fecha, tamaño de la muestra, calidad del estudio, resultados medidos, etc.)? S/PS/PN/N/NI* 1.5 ¿Era adecuada alguna restricción en los criterios de elegibilidad basada en las fuentes de fuentes de información (por ejemplo, estado o formato de publicación, idioma, disponibilidad de datos)? S/PS/PN/N/NI*	2.1 ¿Incluyó la búsqueda una gama apropiada de bases de datos/fuentes electrónicas para informes publicados y no publicados? S/PS/PN/N/NI* 2.2 ¿Se utilizaron métodos adicionales a la búsqueda en bases de datos para identificar informes relevantes? S/PS/PN/N/NI* 2.3 ¿Los términos y la estructura de la estrategia de búsqueda permitieron recuperar el mayor número posible de estudios elegibles? S/PS/PN/N/NI* 2.4 ¿Fueron apropiadas las restricciones basadas en la fecha, el formato de publicación o el idioma? S/PS/PN/N/NI* 2.5 ¿Se realizaron esfuerzos para minimizar los errores en la selección de los estudios? S/PS/PN/N/NI*	3.1 ¿Se hicieron esfuerzos para minimizar los errores en la recogida de datos? S/PS/PN/N/NI* 3.2 ¿Se dispuso de suficientes características de los estudios para que tanto los autores de la revisión como los lectores pudieran interpretar los resultados? S/PS/PN/N/NI* 3.3 ¿Se recogieron todos los resultados relevantes de los estudios para su uso en la síntesis? S/PS/PN/N/NI* 3.4 ¿Se evaluó formalmente el riesgo de sesgo (o la calidad metodológica) utilizando criterios apropiados? S/PS/PN/N/NI* 3.5 ¿Se hicieron esfuerzos para minimizar los errores en la evaluación del riesgo de sesgo? S/PS/PN/N/NI*	4.1 ¿Incluyó la síntesis todos los estudios que debía? S/PS/PN/N/NI* 4.2 ¿Se comunicaron todos los análisis predefinidos o se explicaron las desviaciones? S/PS/PN/N/NI* 4.3 ¿Fue adecuada la síntesis dada la naturaleza y la similitud de las preguntas de investigación, los diseños de los estudios y los resultados de los estudios incluidos? S/PS/PN/N/NI* 4.4 ¿La variación entre estudios (heterogeneidad) fue mínima o se abordó en la síntesis? S/PS/PN/N/NI* 4.5 ¿Fueron sólidos los hallazgos, por ejemplo, como se demostró mediante el gráfico en embudo o los análisis de sensibilidad? S/PS/PN/N/NI* 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? S/PS/PN/N/NI*	A. ¿Se han abordado en la interpretación de los resultados todos los problemas señalados en los dominios del uno al cuatro? S/PS/PN/N/NI* B. ¿Se consideró adecuadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? S/PS/PN/N/NI* C. ¿Evitaron los revisores destacar los resultados en función de su significación estadística? S/PS/PN/N/NI*
¿Coincide la pregunta abordada por la revisión con la pregunta objetivo SI/INCIERTO/NO	**Preocupación por la especificación de los criterios de elegibilidad del estudio** ALTA/INCIERTA/BAJA	**Preocupación por los métodos utilizados para identificar y/o seleccionar los estudios** ALTA/INCIERTA/BAJA	**Preocupación por los métodos de recogida de datos y evaluación de los estudios** ALTA/INCIERTA/BAJA	**Preocupaciones sobre la síntesis y los resultados** ALTA/INCIERTA/BAJA	**Riesgo de sesgo global considerado** ALTO/INCIERTO/ BAJO/INEXISTENTE

Tabla 3. Resumen de la herramienta ROBIS (adaptada de la publicación Whiting et al., 2016).
S: Sí; PS: Probablemente sí; PN: Probablemente no; N: No; NI: No hay información

© Álvaro Limones Burgos, 2024

2.4.3.3. Herramienta de evaluación de la certeza de la evidencia: GRADE

El Grupo de Trabajo GRADE (*Grading of Recommendations, Assessment, Development and Evaluation,* por sus siglas en inglés) propone un método con el mismo nombre para calificar la certeza de la evidencia disponible. La certeza de la evidencia se definiría como el grado de confianza que tenemos en que la magnitud del efecto estimado represente al efecto real de las intervenciones que se evalúan. La certeza de la evidencia se evalúa por separado en cada variable. Otros sinónimos usados en la literatura científica son: calidad de la evidencia, fuerza de la evidencia y confianza en el estimativo. El enfoque GRADE categoriza la certeza de la evidencia en cuatro grados (Tabla 4).

Tabla 4. Grados de la certeza de la evidencia científica según el enfoque GRADE.

CERTEZA	DEFINICIÓN
⊕⊕⊕⊕ ALTA	Estamos muy seguros de que la estimación del efecto se aproxima al efecto real.
⊕⊕⊕◯ MODERADA	Tenemos una confianza moderada en la estimación del efecto. Es probable que la estimación del efecto se aproxime al efecto real, pero existe la posibilidad de que sea sustancialmente diferente.
⊕⊕◯◯ BAJA	Nuestra confianza en la estimación del efecto es limitada. El efecto estimado puede ser sustancialmente diferente del efecto real.
⊕◯◯◯ MUY BAJA	Tenemos muy poca confianza en la estimación del efecto. Es probable que el efecto estimado sea sustancialmente diferente del efecto real.

La categorización de la certeza de la evidencia comienza con la indicación del diseño de los estudios incluidos y, a continuación, aborda cinco dominios para degradar (Tabla 5) y tres consideraciones para incrementar (Tabla 6) el grado de certeza de la evidencia.

© Álvaro Limones Burgos, 2024

Tabla 5. Dominios que disminución la certeza de la evidencia.

Factor	**Disminución de la certeza de la evidencia**
1. Riesgo de sesgo *(ver punto 2.4.3.4.1)*	↓ 1 o 2 niveles
2. Inconsistencia en los resultados *(ver punto 2.4.3.4.2)*	↓ 1 o 2 niveles
3. Carácter indirecto de la evidencia *(ver punto 2.4.2.4.3)*	↓ 1 o 2 niveles
4. Imprecisión *(ver punto 2.4.3.4.4)*	↓ 1, o 2 niveles
5. Sesgo de publicación *(ver punto 2.4.3.4.5)*	↓ 1 o 2 niveles

No obstante, **la consideración de los factores que disminuyen la calidad de la evidencia debe preceder a la de los factores para aumentarla**. La decisión para aumentar la calidad de la evidencia solo debe realizarse **solo cuando no hay limitaciones serias** en cualquiera de los 5 dominios de disminución de la certeza de la evidencia.

Tabla 6. Dominios que incrementan la certeza de la evidencia.

Factor	**Incremento de la certeza de la evidencia**
1. Gran magnitud del efecto *(ver punto 2.4.3.4.6)*	↑ 1 o 2 niveles
2. Gradiente dosis - respuesta *(ver punto 2.4.3.4.7)*	↑ 1 nivel
3. Efecto de los potenciales factores de confusión residual *(ver punto 2.4.3.4.8)*	↑ 1 nivel

Es raro que se justifique el incremento en la calidad de la evidencia en estudios de intervención, siendo más aplicable a investigaciones observacionales (tales como estudios de cohortes, de casos y controles etc.) o de intervenciones no aleatorizadas. Aunque en teoría se podría considerar elevar la calidad de los resultados obtenidos de ECA, Guyatt et al. refieren no identificar ejemplos claros en el que esto sea apropiado, por lo que no es una actuación que se deba plantear de forma rutinaria (Guyatt, Oxman, Sultan, et al., 2011).

© Álvaro Limones Burgos, 2024

2.4.3.3.1. Dominio 1 (*de disminución*) – Evaluación del riesgo de sesgo de los estudios primarios

Un sesgo se define como una desviación de la realidad en los resultados obtenidos de un estudio (Grimes & Schulz, 2002), produciéndose una subestimación o sobreestimación del efecto evaluado. Para evaluarlos, la comunidad científica ha estipulado utilizar herramientas validadas de evaluación del riesgo de sesgo que dependerán del diseño de cada estudio (Tabla 6). Estas herramientas se enfocan en los sesgos más significativos, aunque a menudo incorporan secciones abiertas que permiten una consideración específica de sesgos particulares durante la:

a. **Selección de la muestra o muestreo,** durante la que pueden ocurrir un:
   - Sesgo de selección. Este sesgo ocurre cuando los participantes o grupos del estudio difieren significativamente de la población de interés, lo que puede llevar a errores al extrapolar los resultados a esa población.
   - Sesgo de disponibilidad. Este sesgo se produce cuando se utiliza información o una población que es más fácilmente accesible en lugar de la que es verdaderamente representativa, lo que puede distorsionar los resultados.
   - Sesgo voluntario. Cuando los participantes que se ofrecen como voluntarios para el estudio poseen características intrínsecas diferentes a las de la población general de interés.
   - Sesgo de adherencia. Este sesgo se manifiesta cuando los participantes que siguen estrictamente un protocolo de estudio o intervención difieren de aquellos que no lo siguen, y esa diferencia afecta al resultado de interés.
   - Sesgo de asignación. Se refiere a las diferencias en la forma en que se asignan los participantes a los grupos de comparación en un ensayo clínico.
   - Sesgo por falta de cegamiento. La falta de enmascaramiento de una intervención o tratamiento control en un ensayo clínico puede influir en los resultados.

© Álvaro Limones Burgos, 2024

- Sesgo de prevalencia-incidencia (Neyman). Se da por la exclusión de personas con enfermedades graves o leves, lo que puede llevar a errores en la estimación de la relación entre una exposición y un resultado.
- Sesgo racial. Este sesgo se origina en prejuicios, ya sean explícitos o implícitos, basados en la raza o etnia, que afectan la planificación, métodos, resultados, interpretación, difusión y aplicación de la investigación en salud.
- Sesgo de referencia. Cuando la derivación de pacientes de un nivel de atención (primaria, hospitales, centros especializados) a otro provoca una concentración de casos más graves, complejos o con peores resultados puede alterar la magnitud del efecto evaluado.
- Sesgo centrípeto. Se produce cuando la reputación de médicos e instituciones atrae a pacientes con trastornos específicos, lo que puede llevar a interpretaciones exageradas.
- Sesgo de tiempo de inicio. Ocurre cuando no se puede determinar una hora de inicio común para una exposición o enfermedad.
- Sesgo cronológico. Cuando los participantes asignados a una intervención difieren en exposiciones o riesgos de los reclutados más tarde.
- Sesgo incorrecto del tamaño de la muestra. Este sesgo surge cuando se utiliza un tamaño de muestra incorrecto en un estudio, lo que puede llevar a hallazgos aleatorios con tamaños de muestra pequeños o hallazgos estadísticamente significativos pero clínicamente irrelevantes con tamaños de muestra grandes.

b. Ejecución de una intervención, prueba diagnóstica, modelo predictivo o exposición; durante la que pueden ocurrir un:

- Sesgo del operador o terapeuta. Este sesgo se produce cuando los tratamientos no son ciegos, y las creencias del operador pueden influir en los resultados.
- Sesgo por contaminación. Cuando miembros del grupo control reciben la maniobra o exposición experimental sin intención, reduciendo las diferencias entre los grupos experimental y control.

© Álvaro Limones Burgos, 2024

- Sesgo por el efecto *Hawthorne*. Este sesgo se manifiesta cuando las personas modifican su comportamiento al saber que están siendo observadas.
- Sesgo de cumplimiento. Cuando los participantes que cumplen con una intervención difieren de los que no cumplen, lo que puede afectar al resultado de interés.
- Sesgo por desenmascaramiento. Se produce cuando ocurre una exposición inocente de la intervención, por ejemplo, cuando tienen lugar signos o síntomas relacionados con uno de los grupos.
- Sesgo de acceso al diagnóstico. Por diferencias económicas o temporales en el acceso a pruebas diagnósticas.
- Sesgo por sospecha diagnóstica. Cuando el conocimiento de que un paciente ha estado expuesto a una causa potencial influye en la intensidad del efecto y el proceso diagnóstico.
- Sesgo por falso control. Este sesgo surge cuando los pacientes asignados a una intervención experimental fallecen o enferman y son reasignados al grupo de control, lo que puede hacer que la intervención experimental parezca erróneamente superior.

c. Medición de las intervenciones, pruebas diagnósticas, modelos predictivos o exposiciones; durante la que pueden ocurrir un:

- Sesgo de desgaste, abandonos o pérdida de seguimiento. Ocurre cuando los participantes se pierden o abandonan el estudio de manera desigual entre las intervenciones, lo que puede afectar la validez de los resultados.
- Sesgo por factores de confusión. Se produce cuando una variable no medida o controlada influye en la relación aparente entre la variable independiente y la variable dependiente, llevando a conclusiones incorrectas.
- Sesgo de detección. Este sesgo ocurre cuando el observador influye en los resultados de un ensayo, posiblemente debido a expectativas preconcebidas.

© Álvaro Limones Burgos, 2024

- Sesgo de instrumento. Se refiere a desviaciones de los valores correctos debido a defectos de calibración o mantenimiento de los instrumentos de medición.
- Sesgo de información. Surge de diferencias en la recopilación, recuerdo, registro o manejo de información utilizada en un estudio.
- Sesgo de medida insensible. Este sesgo se produce cuando el método de medición no es lo suficientemente preciso para detectar cambios o diferencias relevantes en el resultado de interés.
- Sesgo de no respuesta. Ocurre debido a las diferencias entre los participantes que responden y los que no responden.
- Sesgo de obsequiosidad. Cuando los sujetos modifican sus respuestas en cuestionarios para satisfacer las expectativas del investigador.
- Sesgo de seguimiento por la causa subyacente. Se da cuando se hacen preguntas sobre exposiciones específicas con mayor frecuencia en el grupo de exposición en comparación con el grupo control, lo que puede aumentar la probabilidad de encontrar una asociación debido a un mayor escrutinio.

d. Análisis de datos, durante el que pueden ocurrir un:

- Sesgo de dragado de datos o *p-hacking*. Este sesgo surge cuando se presentan resultados de pruebas estadísticas no planificadas como si fueran parte de un análisis previamente especificado.
- Sesgo por limpieza de datos. Se refiere a la exclusión de casos aislados o resultados atípicos sin justificación estadística adecuada.
- Sesgo por excavación de los datos. Ocurre cuando se exploran datos en busca de múltiples asociaciones sin una hipótesis previa, lo que aumenta la probabilidad de encontrar resultados significativos por azar.

e. Interpretación de análisis, durante el que pueden ocurrir un:

- Sesgo de patrocinio de la industria. La tendencia a que los resultados y métodos de un estudio respalden los intereses de la organización financiadora.

© Álvaro Limones Burgos, 2024

- Sesgo novedoso. La percepción errónea de que una intervención es mejor solo por ser nueva.
- Sesgo de giro. La interpretación distorsionada, ya sea intencionada o no, de los resultados de la investigación, lo que puede llevar a conclusiones engañosas.
- Sesgo de magnitud. Cuando las diferencias mínimas clínicamente relevantes entre autores afectan significativamente las interpretaciones.

f. Redacción del manuscrito, durante el que pueden ocurrir un:

- Sesgo de confirmación o uso de bibliografía tendenciosa. La búsqueda y uso selectivo de información para respaldar ideas, creencias o hipótesis preconcebidas.
- Sesgos de la retórica. El uso de argumentos persuasivos que no se basan en la razón o la evidencia.
- Sesgo lingüístico. Publicar los resultados en un idioma distinto al reconocido internacionalmente.
- Sesgo de referencia unilateral. Cuando los autores se limitan a citar obras que apoyan su posición.
- Sesgo en la presentación de informes de resultados. La divulgación selectiva de resultados preespecificados en ensayos clínicos publicados.
- Sesgo por bibliografía incorrecta. No todas las publicaciones tienen la misma fiabilidad.

g. Publicación del estudio, durante el que pueden ocurrir un:

- Sesgo por novedad. Cuando un tema de investigación está de moda, los revisores pueden ser menos críticos en su evaluación y los editores pueden estar más inclinados a publicar los resultados.
- Sesgo de publicación. Cuando la probabilidad de publicar un estudio se ve influenciada por sus hallazgos.

© Álvaro Limones Burgos, 2024

- Sesgo de publicación por resultados positivos. La tendencia a presentar, aceptar y publicar resultados positivos en lugar de resultados no significativos o negativos.

En definitiva, estudios con un alto riesgo de sesgo pueden distorsionar significativamente los resultados, lo que justifica la degradación de la evidencia científica (herramienta GRADE). Pero, ¿cuáles son los fundamentos de esta degradación y cómo se aplican en la práctica? La degradación de la evidencia científica debido al riesgo de sesgo se justifica, de acuerdo con el enfoque GRADE por los siguientes fundamentos:

1. No se debe promediar la evaluación individual del riesgo de sesgo de los estudios primarios para proporcionar una valoración global del riesgo de sesgo. En lugar de ello, se puede optar por una evaluación cualitativa, la cual implica valorar la contribución de cada estudio en el metaanálisis junto con su nivel de riesgo de sesgo. O bien, se puede llevar a cabo una evaluación cuantitativa mediante la realización de una combinación secuencial de estudios en el metaanálisis de acuerdo a su calidad (metaanálisis acumulado), o mediante la comparación de estimaciones con y sin los estudios que presenten limitaciones significativas. Estas estrategias permiten una mejor comprensión de cómo el riesgo de sesgo puede influir en la estimación del efecto (sesgándolo), y así emitir juicios razonados que degraden justificadamente la certeza de la evidencia en uno o dos niveles.
2. Es importante estar seguro de que existe un riesgo sustancial de sesgo antes de degradar la evidencia y debe ir acompañado de su justificación. No se debe disminuir la calidad de la evidencia de manera automática o sin fundamentos sólidos. GRADE junto con la Colaboración Cochrane proponen una tabla explicativa para ayudar a los revisores en sus interpretaciones (Tabla 7) (Schünemann et al., 2013).

Con frecuencia, es evidente la utilización de herramientas de evaluación de la calidad metodológica para analizar el riesgo de sesgo y viceversa, aunque sean conceptualmente diferentes. A pesar de que abordan aspectos comunes, es importante destacar que un estudio puede ser riguroso desde un punto de vista metodológico y, al

© Álvaro Limones Burgos, 2024

mismo tiempo, mostrar sesgos significativos. Varios estudios han comparado herramientas de evaluación del riesgo de sesgo con herramientas de evaluación de la calidad metodológica, como ROBIS con AMSTAR-II o ROBINS-1 con NOS, concluyendo que ambas son efectivas y complementarias para detectar limitaciones en el diseño y la ejecución de los estudios (Perry et al., 2021; Zhang et al., 2021) (Tabla 8).

También es crucial no confundir ninguna de estas herramientas con las guías de reporte que, únicamente, ayudan a los autores en la redacción de sus publicaciones (Tabla 9).

Tabla 7. Guía para evaluar el riesgo de sesgo propuesta por la Colaboración Cochrane y la calidad de la evidencia correspondiente según GRADE.

Riesgo de sesgo	Presencia de riesgo de sesgo en los estudio primarios	Interpretación	Consideraciones	Enfoque GRADE para evaluar las limitaciones de los estudios
Bajo	La mayoría de la información es de estudios con bajo riesgo de sesgo	Poco probable que el sesgo plausible afecte seriamente los resultados	Sin limitaciones aparentes	Sin limitaciones serias, no disminuir
No claro	La mayoría de la información es de estudios con bajo riesgo de sesgo o riesgo no claro	Sesgo plausible que genera cierta duda acerca de los resultados	Es poco probable que las potenciales limitaciones disminuyan la confianza en las estimaciones del efecto.	Sin limitaciones serias, no disminuir
No claro	La mayoría de la información es de estudios con bajo riesgo de sesgo o riesgo no claro	Sesgo plausible que genera cierta duda acerca de los resultados	Es probable que las potenciales limitaciones disminuyan la confianza en el efecto estimativo.	Limitaciones serias, disminuir un nivel
Alto	La proporción de información de los estudios con alto riesgo de sesgo es suficiente como para afectar la interpretación de los resultados	Sesgo plausible que debilita seriamente la confianza en los resultados	Limitación crucial para un criterio, o algunas limitaciones en múltiples criterios, suficientes para disminuir la confianza en las estimaciones del efecto.	Limitaciones serias, disminuir un nivel

© Álvaro Limones Burgos, 2024

Tabla 8. Herramientas de evaluación del riesgo de sesgo o de la calidad metodológica para cada diseño de estudios.

Diseño de estudio	Riesgo de Sesgo	Calidad Metodológica
Guías de Práctica Clínica (GPC)	/	AGREE GRS (Brouwers et al., 2012)
Revisión de Revisiones Sistemáticas o *Overviews*	/	JBI *Critical Appraisal Checklist* (Aromataris et al., 2015)
Revisiones Sistemáticas (RS)	ROBIS (Whiting et al., 2016)	AMSTAR-II *Checklist* (Shea et al., 2007; Shea et al., 2017) COSMIN RoB *Checklist* (Mokkink et al., 2010) (solo RS de variables e instrumentos de medición) CHARMS (Moons et al., 2014) (solo RS de modelos predictivos) CASP *Checklist* (https://casp-uk.net/casp-tools-checklists/) SING *Checklist* (https://www.sign.ac.uk/what-we-do/methodology/checklists/)
Ensayos clínicos aleatorizados y controlados (ECA)	ROB 2.0 (Sterne et al., 2019)	JBI *Critical Appraisal Checklist* (Barker et al., 2023) CASP Checklist (https://casp-uk.net/casp-tools-checklists/) SING Checklist (https://www.sign.ac.uk/what-we-do/methodology/checklists/)
Protocolos de ensayos clínicos aleatorizados y controlados	/	/
Estudios clínicos no aleatorizados (Cohortes, Casos-Controles, Trasnversales, etc.)	ROBINS-I (Sterne et al., 2016)	*Newcastle-Ottawa Scale* (NOS) (Wells et al., 2000) JBI *Critical Appraisal Checklist* (https://jbi.global/critical-appraisal-tools) CASP *Checklist* (https://casp-uk.net/casp-tools-checklists/) SING *Checklist* (https://www.sign.ac.uk/what-we-do/methodology/checklists/) GRACE (Dreyer et al., 2016) Q-Coh (Jarde et al., 2013) (solo Estudios de Cohortes) AXIS (Downes et al., 2016) (solo Estudios Transversales)
Estudios de factores pronósticos / modelos predictivos	QUIPS (Hayden et al., 2013) PROBAST (Wolff et al., 2019)	/

Estudios de precisión diagnóstica	/	QUADAS-2 (Whiting et al., 2003; Whiting et al., 2011) JBI *Critical Appraisal Checklist* (Campbell et al., 2020) CASP *Checklist* (https://casp-uk.net/casp-tools-checklists/) SING *Checklist* (https://www.sign.ac.uk/what-we-do/methodology/checklists/)
Estudio preclínicos en experimentación animal	SYRCLE RoB (Hooijmans et al., 2014)	/
Estudios In vitro	QUIN (Sheth et al., 2022)	/
Evaluaciones económicas	/	JBI *Critical Appraisal Checklist* (https://jbi.global/critical-appraisal-tools) CASP *Checklist* (https://casp-uk.net/casp-tools-checklists/) SING *Checklist* (https://www.sign.ac.uk/what-we-do/methodology/checklists/)
Series de casos	/	JBI *Critical Appraisal Checklist* (Munn et al., 2020)
Casos clínicos	/	JBI *Critical Appraisal Checklist* (https://jbi.global/critical-appraisal-tools)

© Álvaro Limones Burgos, 2024

Tabla 9. Guías de reporte para cada diseño de estudio.

Diseño de estudio	Guías de Reporte
Guías de Práctica Clínica (GPC)	AGREE (Brouwers et al., 2016) RIGHT*[1] (Chen et al., 2017) *[1] Extensión RIGHT-Ad@pt: para actualizaciones de CPC (Song et al., 2019)
Revisión de Revisiones Sistemáticas o *Overviews*	PRIOR (Gates et al., 2022)
Revisiones Sistemáticas (RS)	PRISMA*[2] (Page et al., 2021) [antiguo QUOROM (Moher et al., 1999)] *[2] Extensión PRISMA-*Abstracts*: para Abstracts de RS (Beller et al., 2013) *[2] Extensión PRISMA-P: para Protocolos de RS (Moher et al., 2015) *[2] Extensión PRISMA-NMA: para Metaanálisis en red (Hutton et al., 2015) *[2] Extensión PRISMA-DTA: para RS de estudios de precisión diagnóstica (Salameh et al., 2020) *[2] Extensión PRISMA-DTA *for Abstracts*: para Abstracts de RS de estudios de precisión diagnóstica (Cohen et al., 2021) *[2] Extensión PRISMA-ScR: para *Scoping reviews* (Tricco et al., 2018) *[2] Extensión PRISMA-RR (en desarrollo): para *Rapid reviews* (Stevens et al., 2018) *[2] Extensión PRISMA-COSMIN (en desarrollo): para RS de instrumentos y variables de medición (Elsman et al., 2022) *[2] Extensión PRISMA-S: para búsquedas de la literatura (Rethlefsen et al., 2021) MECIR MOOSE para RS de estudios observacionales (Stroup et al., 2000) SYRCLE para RS de estudios preclínicos (De Vries et al., 2015)
Ensayos clínicos aleatorizados y controlados (ECA)	CONSORT*[3] (Schulz et al., 2010) *[3] Extensión para ECA con múltiples intervenciones (Juszczak et al., 2019) *[3] Extensión para ECA de intervenciones cruzadas (Dwan et al., 2019)

	*[3] Extensión CONSORT *Non-inferiority*: para ECA de ensayos de no inferioridad o equivalencia (Piaggio et al., 2012) *[3] Extensión CONSORT-AI: para intervenciones que involucran inteligencia artificial (Liu, Rivera, et al., 2020) *[3] Extensión CONSORT-*Abstracts*: para Abstracts (Hopewell et al., 2008) *[3] Extensión para ECA en ortodoncia (Pandis et al., 2015)
Protocolos de ensayos clínicos aleatorizados y controlados	SPIRIT*[4] (Chan et al., 2013) *[4] Extensión SPIRIT-AI: para protocolos de ECA que involucran inteligencia artificial (Cruz Rivera et al., 2020) *[4] Extensión SPIRIT-PRO: para protocolos de ECA que informar de variables relacionadas con la opinión de los pacientes (Calvert et al., 2018)
Estudios clínicos no aleatorizados (Cohortes, Casos-Controles, Trasnversales, etc.)	STROBE (von Elm et al., 2007)
Estudios de factores pronósticos / modelos predictivos	TRIPOD (Collins et al., 2015)
Estudios de precisión diagnóstica	STARD (Bossuyt et al., 2015)
Estudio preclínicos en experimentación animal	ARRIVE (Percie du Sert et al., 2020)
Estudios In vitro	/
Evaluaciones económicas	CHEERS (Husereau et al., 2022)
Series de casos	PROCESS (Agha et al., 2020) (Solo para estudios de intervenciones quirúrgicas)
Casos clínicos	CARE (Gagnier et al., 2014)

© Álvaro Limones Burgos, 2024

2.4.3.3.2. Dominio 2 *(de disminución)* – Evaluación de la inconsistencia

El término inconsistencia se refiere a la existencia de heterogeneidad inexplicada en la estimación combinada del efecto en el metaanálisis. Los autores de RS tiene que estar preparados para encontrar algún tipo de heterogeneidad y preparar hipótesis a priori de lo que puede explicar esa heterogeneidad para poder explorarla. Si no logran identificar una explicación plausible, la calidad de la evidencia debe ser reducida. La decisión de reducirla en uno o dos niveles dependerá del grado de inconsistencia en los resultados.

En el año 2011, GRADE publicó directrices para reducir el nivel de certeza de la evidencia debido a la presencia de Inconsistencia (heterogeneidad) (Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Glasziou, et al., 2011). Recientemente, en 2023, se ha realizado una actualización de estas directrices (Guyatt et al., 2023). La evaluación de la inconsistencia responde a los siguientes criterios:

1. **Semejanza de las estimaciones del efecto.** Puede identificarse visualmente cuando todas las estimaciones individuales de los estudios se localicen dentro de los umbrales marcados por el investigador de: no relevancia clínica, efecto leve, efecto moderado o efecto grave/severo (Figura 6).

   En 2017, GRADE establece la necesidad de marcar un "marco de nulidad" y "marco de decisión clínica o de diferencia mínima clínicamente relevante" (Hultcrantz et al., 2017), de tal forma que aunque los estudios puedan diferir en sus resultados, si todas las estimaciones son consistentes dentro de un umbral elegido, degradar por inconsistencia no sería apropiado. Si por el contrario las estimaciones caen a un lado y al otro de uno de los umbrales establecidos, degradar por inconsistencia, sí, sería en este caso apropiado.

© Álvaro Limones Burgos, 2024

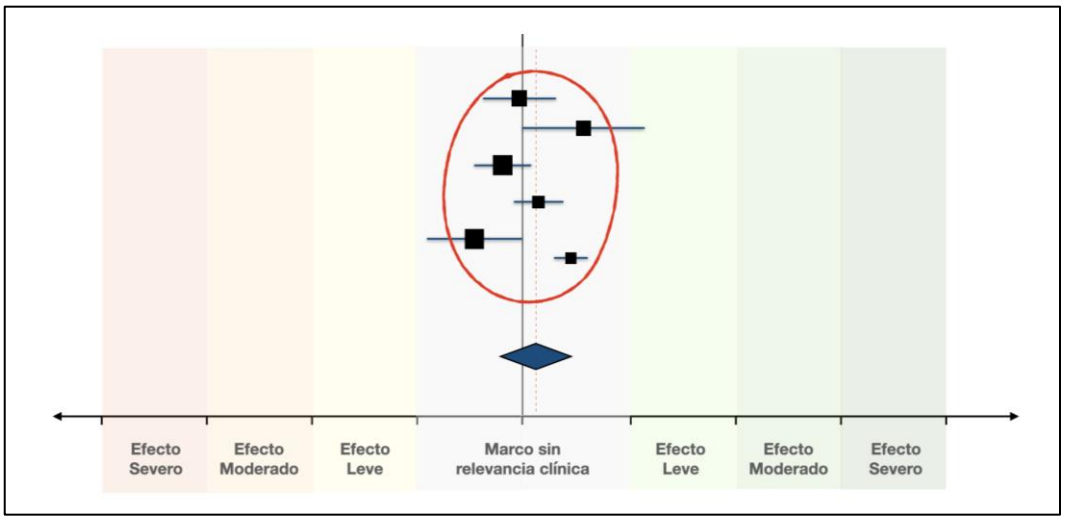


Figura 6. Detección visual de la ausencia de heterogeneidad en el *Forest plot* de acuerdo con los umbrales marcados por el investigador.

2. **Superposición de los intervalos de confianza** de forma contundente, mínima o inexistente (Figura 7); y amplitud de los intervalos de confianza. Es importante contextualizar el $I^{2,}$ ya que menores amplitudes pueden sobreestimar I^2 y no considerarse como la existencia de una heterogeneidad real.

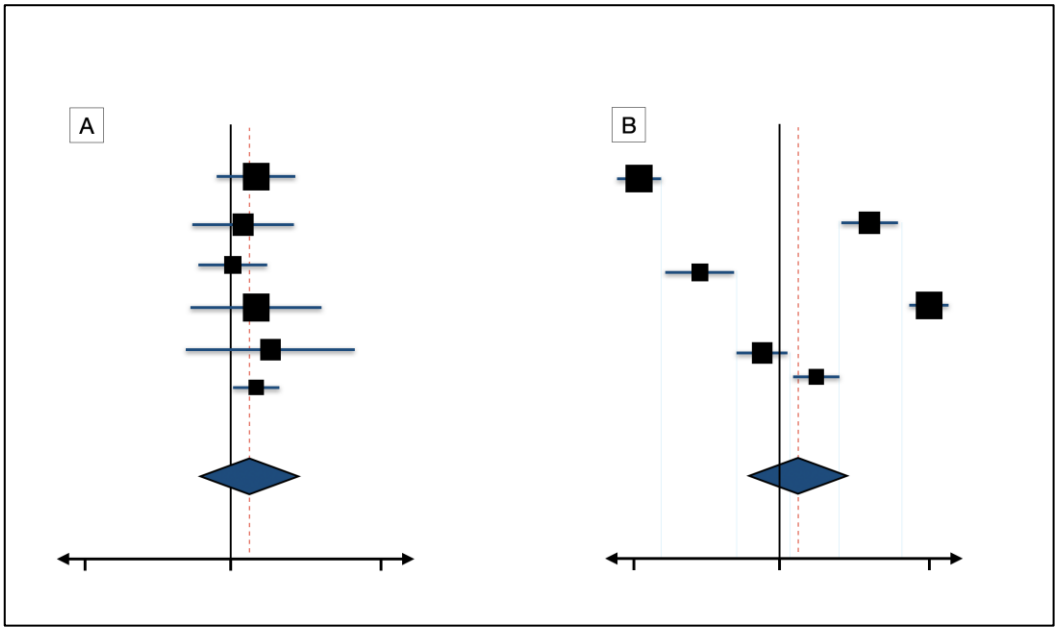


Figura 7. Detección visual de la heterogeneidad en el *Forest plot* de acuerdo con la superposición de los intervalos de confianza. A) Superposición contundente; B) Superposición inexistente.

© Álvaro Limones Burgos, 2024

3. **Estadísticos como el Test de Heterogeneidad o Prueba Q de Cochrane, I^2 y Tau^2**. Esta heterogeneidad estadística se manifiesta cuando los diversos estudios arrojan estimaciones del efecto muy diferentes entre sí, e incluso contrarias. La existencia de una amplia heterogeneidad estadística requiere la búsqueda de una explicación. Esto puede deberse a dos situaciones muy diferentes de acuerdo con el enfoque GRADE:
    a) bien al diseño del estudio fundamentado en el formato PICO, donde puede existir variabilidad entre: los participantes/población, las características de las intervenciones/comparadores (heterogeneidad clínica), o entre los tipos de estudios incluidos (diseño), los instrumentos, las variables o los tiempos de evaluación (heterogeneidad metodológica).
    b) o bien a los resultados obtenidos.

4. **Métodos visuales como los gráficos de L'Abbe y Galbraith**.

2.4.3.3.2.1. Estadísticos para evaluar la heterogeneidad

Los estadísticos con los que más frecuentemente se evalúa la heterogeneidad son:

A. **Prueba Q de Cochrane*** que usa el estadístico Chi^2, χ^2 o chi cuadrado. Evalúa si las diferencias observadas en los resultados son consistentes o no con el azar. Una $p < 0.10$ nos sugiere que los resultados difieren entre sí más de lo que cabría esperar por el mero azar
B. **Estadístico I^2 de Higgins***, que cuantifica la proporción de variación en las estimaciones del efecto a través de la fórmula: $I^2 = \left(\frac{Q-df}{Q}\right) x\ 100\%$, en la que Q es la estadística de χ^2 y df los grados de liberdad. Aunque no existe una regla fija, se puede utilizar la siguiente guía:
    I^2 <40% puede considerarse bajo y no ser importante.
    I^2 entre 30% y 60% puede representar una heterogeneidad moderada.

© Álvaro Limones Burgos, 2024

I^2 entre 50% y 90% puede representar una heterogeneidad sustancial.

I^2 entre 75% y 100% representa una heterogeneidad considerable.

*Ambos estadísticos son sensibles a tamaños muéstrales pequeños, a la combinación de pocos estudios o a identificar diferencias estadísticas clínicamente irrelevantes si hay muchos estudios. Los tamaños de muestra pequeños pueden generar diferencias sustanciales incluso con valores bajos de I^2 debido al azar, mientras que tamaños de muestra grandes pueden mostrar diferencias relativas pequeñas y, sin embargo, resultar en un valor alto de I^2. Por lo tanto, también es útil examinar el estadístico τ^2 (tau cuadrado), que mide la variabilidad sin depender del tamaño de la muestra.

C. **Estadístico τ^2** . Mide la variabilidad sin depender del tamaño de la muestra.

2.4.3.3.2.2. Gráficos para evaluar la heterogeneidad

2.4.3.3.2.2.1. Gráficos de L'Abbe

Se trata de un procedimiento visual para valorar la heterogeneidad aplicable a variables dicotómicas (ocurrencia o no de un suceso clínico) (L'Abbé et al., 1987). La diagonal en el gráfico representa la igualdad entre el grupo test y control, lo que sugiere un efecto nulo. Cada circulo en el gráfico representa un estudio, donde el tamaño del circulo es proporcional al tamaño del estudio y su dispersión indica heterogeneidad (Figura 8). Este gráfico es intuitivo y permite identificar fácilmente los estudios responsables de la heterogeneidad.

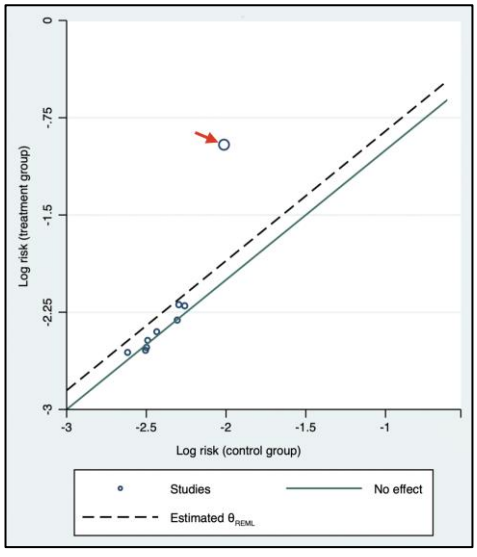


Figura 8. Diagrama de L'Abbe. La presencia de heterogeneidad se observa por el estudio marcado con la flecha roja.

© Álvaro Limones Burgos, 2024

2.4.3.3.2.2.2. Gráfico radial o Diagrama de Galbraith

Se trata de un procedimiento visual para valorar la heterogeneidad aplicable a variables dicotómicas y continuas (Galbraith, 1988). "Representa la relación entre la precisión de cada estudio (representada por el inverso del error estándar) y el efecto estandarizado, mediante una la línea de regresión ajustada y su intervalo de confianza. Los estudios que caen fuera de este intervalo son los que más contribuyen a la heterogeneidad (Figura 9)". Esto permite detectar puntos anómalos o valores atípicos. Además, la posición de los estudios en el eje de abscisas permite identificar visualmente aquellos que tienen un mayor peso en el metaanálisis.

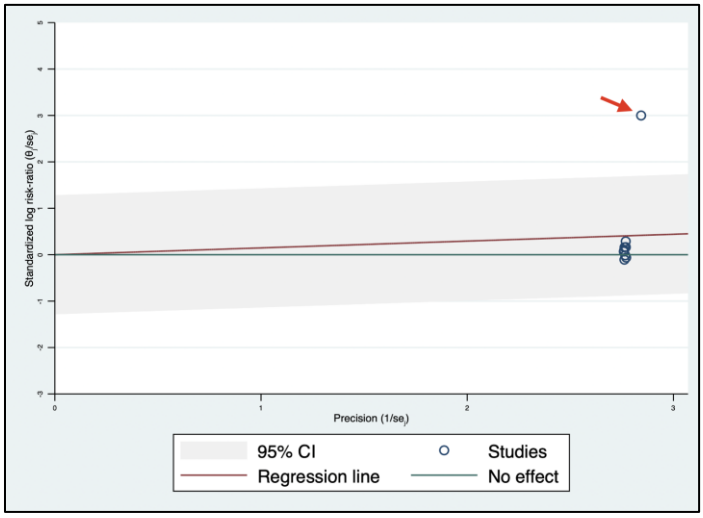


Figura 9. Diagrama de Galbraith. La presencia de heterogeneidad se observa por el estudio marcado con la flecha roja.

2.4.3.3.2.3. Manejo de la heterogeneidad

Algunos autores proponen los marcos: **P<0.1 para la Prueba Q y $I^2 = 30\%$** para determinar si se debe buscar una explicación.

La heterogeneidad, si es correctamente identificada su causa, puede resolverse con un análisis por subgrupos sin tener que disminuir la certeza de la evidencia científica. Si por el contrario no pueden identificar una explicación plausible, la calidad de la evidencia se debe disminuir de acuerdo con el enfoque GRADE (Schünemann et al., 2013). No obstante, los análisis

© Álvaro Limones Burgos, 2024

de subgrupos implican ofrecer recomendaciones especificas en el posterior desarrollo de recomendaciones a través de las GPC.

Para explorar si realmente tiene sentido un análisis de subgrupos por que existe un factor modificador sobre el que subanalizar, se puede aplicar la herramienta ICEMAN de credibilidad de existencia de un modificador de la estimación del efecto que justifique un análisis por subgrupos mediante la siguiente Tabla extraida del artículo de Guyatt et al. (2023):

1: Is the analysis of effect modification based on comparison within rather than between trials?			
[ ] Completely between	[ ] Mostly between or unclear	[ ] Mostly within	[ ] Completely within
Subgroup analysis or meta-regression comparing overall effects of each individual trial. This is typical for aggregate data meta-analysis.	*Subgroup analysis or meta-regression with most information coming from overall effects, but some trials providing within-trial subgroup information*	*Most trials providing within-trial subgroup information; or individual participant data analysis that combines within and between trial information*	*All trials providing within-trial subgroup information or individual participant data; and the analysis separates within from between trial information, e.g., meta-analysis of interactions*
2: For within-trial comparisons, is the effect modification similar from trial to trial? [ ] Not applicable: no or one within-RCT comparison			
[ ] Definitely not similar	[ ] Probably not similar or unclear	[ ] Mostly similar	[ ] Definitely similar
Effect modification reported for two or more trials and clearly different directions	*Effect modification not reported for individual trials or too imprecise to tell*	*Effect modification reported for two or more trials, mostly similar in direction, but considerable differences in magnitude*	*Effect modification reported for two or more trials, similar in direction, only some differences in magnitude*
3: For between-trial comparisons, is the number of trials large? [ ] Not applicable: no between RCT comparison			
[ ] Very small	[ ] Rather small or unclear	[ ] Rather large	[ ] Large
1 or 2 or in smallest subgroup; 5 or less in continuous meta-regression	*3-4 in smallest subgroup; 6-10 in continuous meta-regression*	*5-9 in smallest subgroup; 11 to 15 in continuous meta-regression*	*10 or more in smallest subgroup; more than 15 in continuous meta-regression*
4: Was the direction of effect modification correctly hypothesized a priori?			
[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Clearly post-hoc or results inconsistent with hypothesized direction or biologically very implausible	*Vague hypothesis or hypothesized direction unclear*	*No prior protocol available but unequivocal statement of a priori hypothesis with correct direction of effect modification*	*Prior protocol available and includes correct specification of direction of effect modification, e.g., based on a biologic rationale*
5: Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification? (consider irrespective of number of effect modifiers)			
[ ] Chance a very likely explanation	[ ] Chance a likely explanation or unclear	[ ] Chance may not explain	[ ] Chance an unlikely explanation
Interaction or meta-regression p-value >0.05	*Interaction or meta-regression p-value ≤0.05 and >0.01, or no test of interaction reported and not computable*	*Interaction or meta-regression p-value ≤0.01 and >0.005*	*Interaction or meta-regression p-value ≤0.005*
6: Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis?			
[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Explicitly exploratory analysis or large number of effect modifiers tested (e.g., greater than 10) and multiplicity not considered in analysis	*No mention of number or 4-10 effect modifiers tested and number not considered in analysis*	*No protocol available but unequivocal statement of 3 or fewer effect modifiers tested*	*Protocol available and 3 or fewer effect modifiers tested or number considered in analysis*
7: Did the authors use a random effects model? [ ] Not applicable			
[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Fixed (or common) effect or fixed effects model explicitly stated	*Probably fixed effect(s) model*	*Probably random (or mixed) effects*	*Random (or mixed) effects explicitly stated*
8: If the effect modifier is a continuous variable, were arbitrary cut points avoided? [ ] not applicable: not continuous			
[ ] Definitely no	[ ] Probably no or unclear	[ ] Probably yes	[ ] Definitely yes
Analysis based on exploratory cut point(s), e.g., picking cut point associated with highest interaction p-value	*Analysis based on cut point(s) of unclear origin*	*Analysis based on pre-specified cut point(s), e.g., suggested by prior RCT*	*Analysis based on the full continuum, e.g., assuming a linear or logarithmic relationship*
9 Optional: Are there any additional considerations that may increase or decrease credibility? (manual section 3.9) [ ] not applicable			
10: How would you rate the overall credibility of the proposed effect modification?			
		X	
Very low credibility	**Low credibility**	**Moderate credibility**	**High credibility**

La puntuación de la credibilidad según la herramienta ICEMAN, sigue las siguientes indicaciones:

- Si dos o más respuestas ***definitivamente*** disminuyen la credibilidad, se considerará como credibilidad baja – *low* o muy baja – *very low*.

© Álvaro Limones Burgos, 2024

- Si una respuesta ***definitivamente*** disminuye la credibilidad, o dos respuestas ***probablemente*** reducen la credibilidad, se considerará como moderada – *moderate* o baja – *low*.
- Si no hay ninguna respuesta que ***definitivamente*** o menos de dos que ***probablemente*** disminuyan la credibilidad del análisis por subgrupos, entonces se considerará credibilidad alta – *high*.

2.4.3.3.3. Dominio 3 *(de disminución)* – Evaluación del carácter indirecto de la evidencia

Es crucial evaluar el carácter indirecto de la evidencia o validez externa, ya que la credibilidad del estimativo puede disminuir si existen discrepancias entre la población estudiada, las intervenciones aplicadas y/o las variables evaluadas, y el entorno clínico sobre el que se pretende extrapolar conclusiones. La evidencia directa proviene de la comparación directa de las intervenciones de interés, en la población que es de interés y en aquellas variables que son importantes para los pacientes (Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Falck-Ytter, et al., 2011). De este modo, la evidencia se considera indirecta de las siguientes 4 formas posibles:

1. Discrepancias en la **Población**. Cuando la evidencia proviene de estudios en poblaciones que no reflejan exactamente a la población de interés (i.e. diferentes edades, países, o estudios en animales o in vitro), se deben extrapolar conclusiones con precaución. Por ejemplo, los estudios en animales pueden ofrecer visiones preliminares, apoyar hipótesis o advertir de toxicidades, pero la extrapolación de sus resultados a humanos puede no ser directa debido a diferencias biológicas y fisiológicas.

2. Discrepancias en las **Intervenciones**. Los autores deberán considerar reducir la evidencia si las intervenciones no son similares entre estudios, o si no se puede implementar el mismo rigor o sofisticación técnica en los entornos clínicos diarios. Esta discrepancia puede llevar a cuestionar si los resultados observados se replicarían bajo condiciones menos controladas o con diferentes enfoques técnicos en la práctica clínica diaria de los clínicos.

© Álvaro Limones Burgos, 2024

3. Discrepancias en las **Variables de evaluación**. Al comenzar una RS o el desarrollo de una GPC, es primordial identificar las variables de resultado que tienen un significado directo y práctico para los pacientes (variables críticas). La consideración de variables subrogadas, aunque no son directamente significativas para los pacientes, se examinan bajo la suposición de que pueden indicar cambios en los resultados clínicamente importantes. Además, el tiempo de seguimiento en un estudio es un factor crítico. Si hay una discrepancia significativa entre la duración del seguimiento en el estudio y el periodo que es clínicamente relevante para la pregunta de investigación o la decisión clínica en cuestión, la aplicabilidad de los resultados puede verse comprometida. En tales casos, GRADE aconseja considerar la degradación de la evidencia, lo que puede resultar en reducir la certeza de la evidencia uno o dos niveles.

4. **Comparaciones indirectas**. Las comparaciones indirectas juegan un papel significativo en la síntesis de evidencia científica, especialmente cuando no existen estudios directos que comparen intervenciones de interés (Hao et al., 2023). Sin embargo, estas comparaciones indirectas llevan inherente un nivel de incertidumbre significativamente mayor debido a la suposición de similitud entre los estudios. Se sostiene la suposición de que las variaciones en el diseño del estudio, la población de pacientes, las intervenciones, las formas en que se miden los resultados, y la calidad metodológica, no afectan de manera significativa a los efectos observados entre las intervenciones comparadas indirectamente mediante metaanálisis en red. Si las comparaciones indirectas, por su naturaleza, comprometieran la solidez de las conclusiones se debería disminuir la certeza de la evidencia.

© Álvaro Limones Burgos, 2024

2.4.3.3.4. Dominio 4 *(de disminución)* – Evaluación de la imprecisión

Los resultados tienden a ser imprecisos cuando los estudios incluyen relativamente pocos pacientes y pocos eventos, lo que se refleja en intervalos de confianza (IC) amplios en el estimativo del efecto. En tales casos, uno puede considerar que la calidad de la evidencia es menor debido a incertidumbre en los resultados.

La imprecisión se evalúa de acuerdo con los siguientes criterios (Schünemann et al., 2022):

1. **Amplitud y cruce** del intervalo de confianza (del 95%) con los "umbrales de decisión o relevancia clínica del efecto" (Figura 10), que son: marco sin relevancia clínica, marco de efecto leve, marco de efecto moderado o marco de efecto grave (marcados por el investigador). Los IC demasiado anchos podrían implicar una reducción de la certeza de la evidencia por imprecisión uno o dos, o incluso tres niveles, si sobrepasan uno, dos o tres umbrales de relevancia clínica.

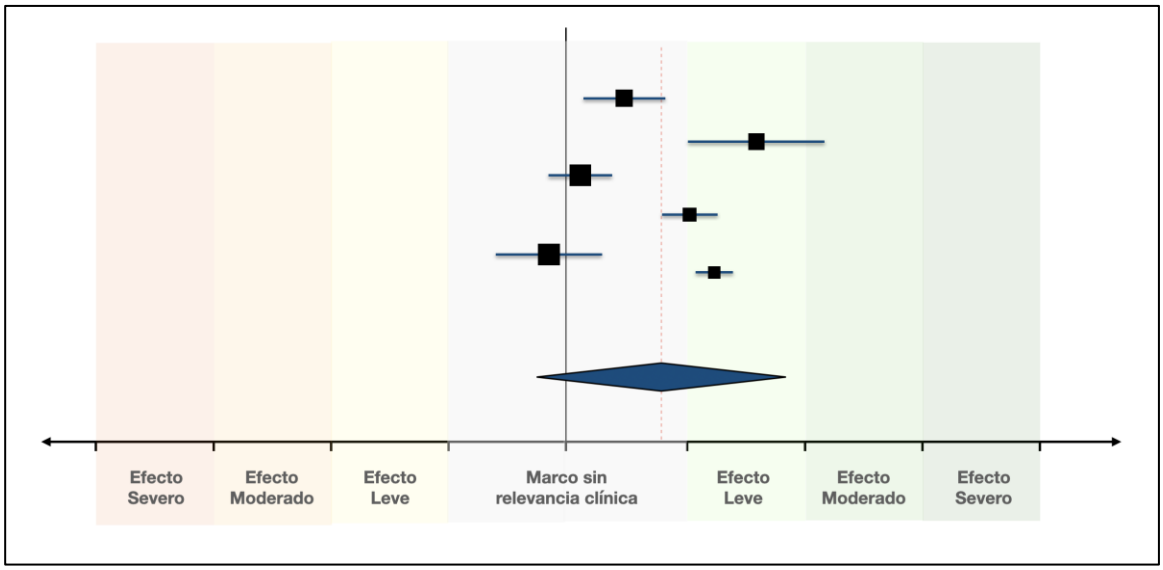


Figura 10. Detección visual de la imprecisión en el *Forest plot* por la amplitud del efecto estimado (representado con el diamante azul) que cruza el marco sin relevancia clínica y el del efecto leve al mismo tiempo.

© Álvaro Limones Burgos, 2024

2. Sobrepasar:
   a) el "**Tamaño Óptimo de Información**" (OIS, por sus siglas en inglés, *Optimal Information Size*). Comúnmente conocido como el "cálculo del tamaño muestral" para un ensayo. Necesario para hacer inferencias con suficiente poder estadístico. Se calcula a partir de estudios previos o de delta (Δ) determinado por el investigador y definido por Pogue y Yusuf como "el efecto mínimo de tratamiento que es considerado como apreciable o biológicamente plausible" (Pogue & Yusuf, 1997).

   b) el "**Tamaño Revisable de Información**" (RIS, por sus siglas en inglés, *Review Information Size*), calcula el tamaño muestral necesario para hacer inferencias en cada "umbral de relevancia clínica del efecto" (severo, moderado y leve) definidos por el investigador (Figura 11), del siguiente modo:

      i. Si el tamaño muestral agrupado en el metaanálisis no supera al tamaño muestral calculado para el considerado umbral-de-efecto-severo, se reducirá por imprecisión un nivel la certeza de la evidencia científica.
      ii. Si el tamaño muestral agrupado en el metaanálisis no supera al tamaño muestral calculado para el considerado umbral-de-efecto-moderado, se reducirá por imprecisión dos niveles la certeza de la evidencia científica.
      iii. Si el tamaño muestral agrupado en el metaanálisis no supera al tamaño muestral calculado para el considerado umbral-de-efecto-leve, se reducirá por imprecisión tres niveles la certeza de la evidencia científica.

© Álvaro Limones Burgos, 2024

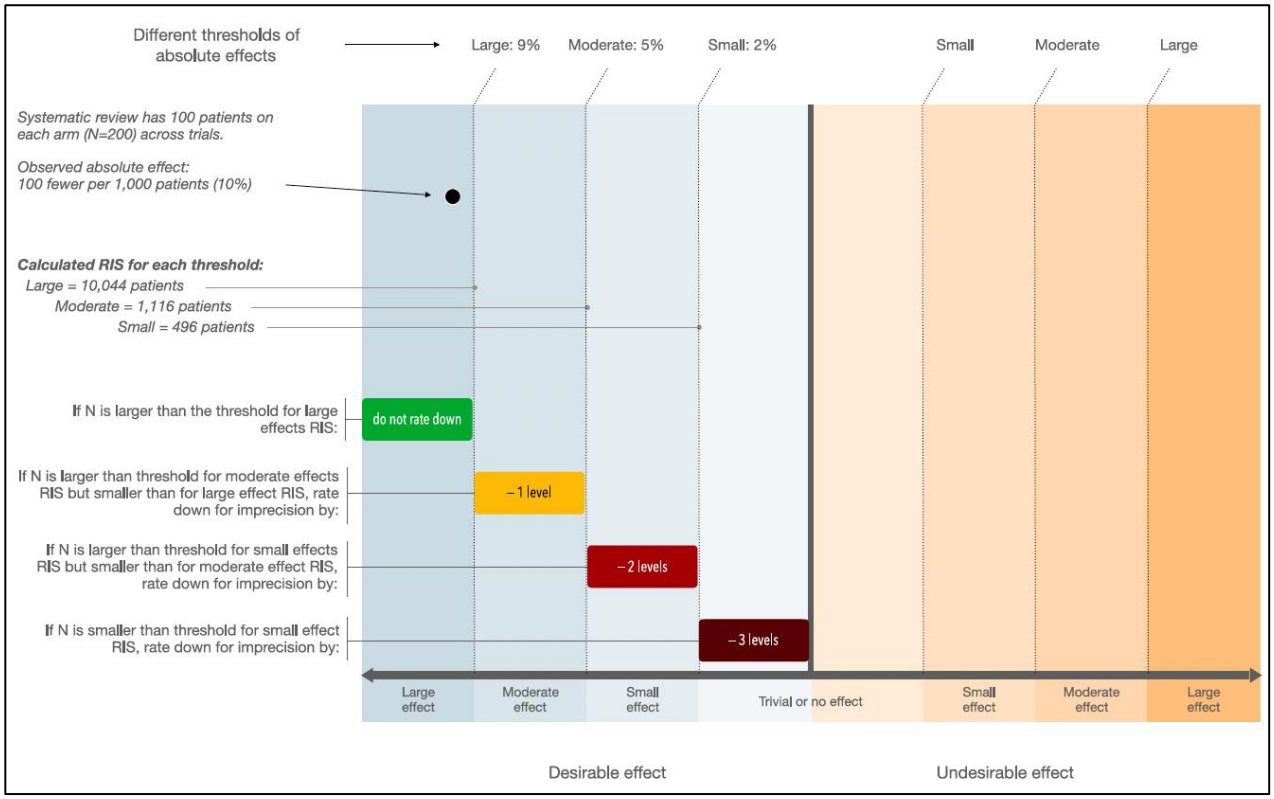


Figura 11. Imagen explicativa del *Review Información Size*, extraída del artículo de Schünemann et al. (2022)

La recomendación de GRADE en 2011 reflejó como referencia, que para variables continuas, una muestra inferior a 400 participantes, o para variables dicotómicas, una muestra inferior a 2000 participantes, para degradar la certeza de la evidencia por imprecisión (Guyatt, Oxman, Kunz, Brozek, et al., 2011). Sin embargo, la última actualización de 2022 recomienda utilizar el OIS o el RIS (Schünemann et al., 2022).

© Álvaro Limones Burgos, 2024

2.4.3.3.5. Dominio 5 *(de disminución)* – Evaluación del sesgo de publicación

Cuando se produce una sub- o sobrestimación de los hallazgos obtenidos en una RS con metaanálisis por una publicación selectiva de estudios, dependiendo de la naturaleza y/o dirección de los hallazgos. La evidencia empírica sugiere que los estudios con resultados positivos o significativos, que muestran efectos grandes o en la dirección esperada, tienen mayor probabilidad de ser publicados, mientras que aquellos con resultados no significativos, negativos o que contradicen las expectativas planteadas a menudo permanecen inéditos o publicados en revistas locales no accesibles a la comunidad internacional (Guyatt, Oxman, Montori, et al., 2011). El sesgo de publicación se debe diferenciar del sesgo por selección (de resultados positivos). Enfoque que se aborda dentro de la evaluación del riesgo de sesgo de los estudios primarios. Otros contribuidores al sesgo de publicación son:

1. **Sesgo lingüístico**. Investigadores de países donde el inglés no es el idioma principal a menudo optan por divulgar sus hallazgos negativos o menos impactantes en revistas locales en revistas de menor impacto, utilizando su lengua materna . Esta práctica limita la visibilidad y accesibilidad de estos estudios a la comunidad científica global, contribuyendo al sesgo de publicación.

2. **Literatura gris**. Los resultados negativos frecuentemente se encuentran en lo que se denomina literatura gris, que incluye Tesis Doctorales, capítulos de libros, y resúmenes de congresos, entre otros. Aunque estos trabajos son valiosos para comprender el panorama completo de la investigación en un área determinada, su inclusión en RS y metaanálisis es menos común debido a la dificultad para localizarlos y acceder a ellos, si no se realizan exhaustivas estrategias de búsqueda.

3. **Patrocinios de la industria – conflictos de interés**. Cuando son pocos estudios y todos financiados, o cuando los resultados de los estudios financiados obtienen resultados negativos que las compañías financiadoras prefieren no divulgar.

© Álvaro Limones Burgos, 2024

4. **Novedad del tema – sesgo de retraso**. La RS realizadas demasiado apresuradas antes de que existan suficientes estudios, caen en un fenómeno que se conoce como "*lag bias*" (Hopewell et al., 2009).

Todo lo anterior, conduce a una representación desbalanceada y sesgada de la evidencia disponible, ya que los estudios no publicados, que pueden mostrar efectos muy pequeños o incluso nulos, son sistemáticamente omitidos de las revisiones. Inflando artificialmente la eficacia percibida de una intervención o tratamiento. Esto puede llevar a la adopción de prácticas clínicas basadas en una base de evidencia incompleta o erróneamente optimista. Se ha documentado que el sesgo de publicación es más frecuente en los metaanálisis sobre efectos adversos que en los de efectividad de las intervenciones (Kicinski et al., 2015). Se realiza de forma visual y estadística o combinada.

2.4.3.3.5.1. Método visual

Se realiza a través del diagrama de embudo:

2.4.3.4.5.1.1. *Funnel plot o* diagrama de embudo

Visualmente el sesgo de publicación se estudia mediante el Diagrama de Embudo o *Funnel Plot*, valorando si existe una distribución asimétrica de los estudios en el gráfico, a pesar de que algunos autores consideran que es un método que infunda errores de interpretación (Terrin et al., 2005).

2.4.3.3.5.2. Métodos estadísticos

Se han desarrollado método estadísticos para evaluar cuantitativamente el sesgo de publicación mediante una prueba estadística de asimetría. "Como regla de oro, la gráfica de embudo y las pruebas estadísticas de asimetría se deben usar para detectar sesgo de publicación si existen al menos **10 estudios incluidos en el metaanálisis** (algunos autores dicen que al menos 5 estudios)".

© Álvaro Limones Burgos, 2024

2.4.3.3.5.2.1. Método de Begg

La prueba de Begg evalúa la existencia de una asociación entre las estimaciones de los efectos de los estudios y sus respectivas varianzas (Begg & Mazumdar, 1994). La presencia de correlación entre estos dos elementos sugiere la posibilidad de un sesgo de publicación. Sin embargo, un aspecto crítico de esta prueba es su limitada potencia estadística, lo que la hace poco confiable para analizar conjuntos de datos pequeños o cuando el número de estudios primarios involucrados es reducido.

2.4.3.3.5.2.2. Método de Egger

La prueba de Egger proporciona un enfoque más específico que la prueba de Begg para detectar sesgos de publicación. Implica el trazado de una recta de regresión entre la precisión de los estudios (como variable independiente) y el efecto estandarizado (como variable dependiente) (Egger et al., 1997). Para aumentar la precisión de esta prueba, la regresión se pondera por el inverso de la varianza de cada estudio, añadiendo complejidad estadística al análisis. Tradicionalmente, se considera que un valor de p menor a 0,1 en la regresión sugiere sesgo de publicación, en lugar del umbral más común de menos de 0,05.

2.4.3.3.5.2.3. Método de Harbord

Harbord et al. (2006) presentaron una mejora en las propiedades estadísticas del test de Egger, aunque su uso es menos frecuente que el del método original. Este ajuste busca refinar la capacidad del test de Egger para detectar menos falsos positivos del sesgos de publicación, ofreciendo una alternativa que podría mejorar la precisión en la evaluación de la asimetría en metaanálisis.

2.4.3.3.5.3. Métodos combinados de detección estadística y visual

2.4.3.3.5.3.1. Método de podar y completar (*Trim & fill*)

El método *Trim and Fill* (o “Podar y Completar” en español, suele retener su denominación original anglosajona debido a la complejidad de su traducción directa.

© Álvaro Limones Burgos, 2024

Comienza con el cálculo de una media ponderada utilizando todos los estudios disponibles. A continuación, emplea un algoritmo sencillo para identificar y mantener únicamente aquellos estudios que contribuyen a la simetría del gráfico, eliminando los estudios asimétricos. Tras esta eliminación, se calcula nuevamente una media ponderada.

El siguiente paso consiste en estimar los estudios faltantes en el lado asimétrico del gráfico, cuya inclusión restablecería la simetría. Estos estudios estimados, sumados a los ya identificados en la parte asimétrica, permiten alcanzar una nueva media ponderada que sería comparable a la calculada inicialmente con los estudios simétricos. Este proceso ayuda a corregir el sesgo de publicación y proporciona una estimación del efecto más precisa del efecto global (ver ejemplo en el ANEXO 3.4).

2.4.3.3.6. Dominio 6 *(de incremento)* – Evaluación de la existencia de "gran magnitud del efecto"

Cuando un estudio*[1] presenta estimaciones de gran magnitud, entonces es razonablemente deducible que el efecto pueda ser real (indicios de causalidad) (Guyatt, Oxman, Sultan, et al., 2011). La estimación del efecto debe ser precisa: sin IC demasiado amplios y sin superponer "umbrales de decisión clínica". Existe mayor probabilidad de aumentar la calidad de la evidencia por magnitud de efecto grande o muy grande si:

- El efecto es rápido.
- El efecto es consistente a lo largo de los sujetos.
- La gran magnitud del efecto está soportada por evidencia indirecta.

GRADE propone la siguiente Tabla 10 para definir qué es un efecto grande o muy grande y el número de niveles a incrementar:

[1] Generalmente aplicable a estudios observacionales (i.e. estudios de cohortes, de casos y controles, etc.) o de intervenciones no aleatorizadas.

© Álvaro Limones Burgos, 2024

Tabla 10. Criterios GRADE de consideración de la existencia de "Gran Magnitud del efecto.

Magnitud del efecto	Definición	Incremento de la certeza de la evidencia
Grande	**RR>2 o <0.5** (Basado en evidencia directa, sin posibles factores de confusión)	↑ 1 nivel
Muy grande	**RR>5 o >0.2** (Basado en evidencia directa, sin riesgo de sesgo, sin imprecisión, sin posibles factores de confusión)	↑ 2 niveles

2.4.3.3.7. Dominio 7 *(de incremento)* – Evaluación del Gradiente Dosis-Respuesta (GDR)

La existencia de un gradiente dosis-respuesta ha sido tradicionalmente considerada como un indicador clave para validar relaciones causa-efecto (Guyatt, Oxman, Sultan, et al., 2011). La observación de este gradiente puede reforzar nuestra seguridad en las conclusiones extraídas de estudios observacionales, y por tanto, la certeza de la evidencia. La última actualización GRADE propone evaluar la credibilidad de que exista un gradiente dosis-respuesta mediante 5 preguntas (Murad et al., 2023):

- ¿**Es adecuado el enfoque analítico** para detectar el GDR?
- ¿**Es un factor de confusión** la causa del GDR?
- ¿Hay serias preocupaciones sobre el **sesgo ecológico**?
- ¿Es **consistente** el GDR a lo largo de los estudios?
- ¿**Existe evidencia indirecta** que apoye el GDR?

Se aconseja añadir la variables subrogadas que proporcionan evidencia indirecta (si existe) que justifique el GDR en n las "Tablas resumen de Hallazgos".

© Álvaro Limones Burgos, 2024

2.4.3.3.8. Dominio 8 *(de incremento)* – Evaluación de factores de confusión plausibles que aumentan la confianza en el efecto estimado

Este fenómeno destaca que, en ciertos contextos, los factores de confusión no ajustados pueden incrementar, en lugar de disminuir, nuestra confianza en los resultados de una intervención. Cuando una intervención se enfrenta a condiciones adversas específicas de su grupo, marcadas por un factor de confusión no ajustado que, en teoría, debería mermar su efectividad, y aún así muestra resultados favorables a la intervención, nuestra certeza en la autenticidad y fortaleza de ese efecto se fortalece. Este contexto indica que si una intervención logra beneficios a pesar de las adversidades inherentes a su grupo de estudio, ello constituye una señal potente de su verdadera eficacia (Guyatt, Oxman, Sultan, et al., 2011).

2.4.3.4. Herramienta para guiar el reporte de revisiones sistemáticas: PRISMA

La declaración PRISMA (de sus siglas en inglés: *Preferred Reporting Items for Systematic Reviews and Meta-Analyses)* ofrece una guía para garantizar un reporte adecuado de las RS en las publicaciones científicas (Page et al., 2021) a través de una lista de comprobación o *Checklist*. Cuenta con varias extensiones de acuerdo con el tipo de RS, descritas en la Tabla 9.

2.4.4. Métodos de presentación de la certeza de la evidencia: "Tablas resumen de hallazgos o de la evidencia"

El sistema GRADE propone presentar los resultados de la evaluación de la certeza de la evidencia en lo que se conoce como "Tablas de evidencia o resumen de hallazgos", o en inglés *Summary of Findings (SoFs) tables,* siguiendo el formato propuesto por Carrasco-Labra y cols. 2016 (Carrasco-Labra et al., 2016), y la última actualización de Morgan y

© Álvaro Limones Burgos, 2024

cols. 2023 (Morgan et al., 2023), que incluye el número de pacientes con necesidad de tratar (NNT) (Tabla 11). Las tablas SoFs suelen incluir las variables principales consideradas críticas o importantes; para de forma transparente, visual y concisa proporcionar la información clave sobre la que fundamentar recomendaciones mediante las GPC. El formato estándar de una Tabla SoF incluye:

- número de estudios, diseño y número de participantes.
- lista de variables más relevantes para evaluar las intervenciones.
- los estimadores del efecto observado en términos de efecto relativo y efecto absoluto con un 95% de IC y el número de pacientes con necesidad de tratar (NNT).
- la certeza de la evidencia evaluada de acuerdo con el sistema GRADE. Se puede incluir en la propia tabla columnas adicionales con los dominios que ponderan la certeza de la evidencia (limitaciones del diseño de los estudios, inconsistencias, imprecisión, carácter indirecto y/o sesgo de publicación); o bien incluir esta información a modo de superíndices con su respectiva descripción en la leyenda a pie de tabla.

Tabla 11. Ejemplo de "Tabla de evidencia o Tabla resumen de hallazgos" extraída de la publicación de Morgan et al. (2023).

Home visiting programs compared to usual community care for patients with heart failure recently discharge from hospital

Patient or population: Patients with heart failure recently discharged from hospital
Setting: Outpatient
Intervention: Home visiting programs
Comparison: Usual community care

Outcome № of participants (studies)	Relative effect (95% CI)	Anticipated absolute effects (95% CI)			Number needed to treat (95%CI)	Quality	What happens
		Without Home visiting	With Home visiting	Difference			
Mortality follow up: range 3 months to 6 months № of participants: 1693 (8 RCTs)	**RR 0.77** (0.60 to 0.99)	138 per 1,000	**107 per 1,000** (83 to 137)	**31 fewer** (55 fewer to 1 fewer)	**31** (19 to 1000)	⊕⊕⊕◯ MODERATE [1]	Home visiting programs probably reduce mortality slightly
All-cause readmission follow up: range 3 months to 6 months № of participants: 1661 (9 RCTs)	**RR 0.75** (0.68 to 0.86)	410 per 1,000	**307 per 1,000** (278 to 352)	**102 fewer** (131 fewer to 57 fewer)	**10** (8 to 18)	⊕⊕⊕⊕ HIGH	Home visiting programs reduce all cause readmission
Quality of Life (QoL) assessed with: Minnesota Living with Heart Failure Questionnaire. (MLWHFQ); MID: 5 follow up: 6 months № of participants: 336 (2 RCTs)	-	-	-	SMD **0.04 SD lower** (0.35 lower to 0.26 higher)	-	⊕◯◯◯ VERY LOW [2,3,4]	We are uncertain whether home visiting programs increase QoL as the quality of the evidence has been assessed as very low
Resource use follow up: mean 12 months № of participants: 239 (1 RCT)	Mean cost savings of $4,845 per patient in the intervention group; 95% CI of true difference in mean total cost $8,976–$1,301; P=0.002. Total and mean costs (reimbursements) per patient were lower in the intervention group than in the control group. Evidence based on one trial of 239 participants. [5]				-	⊕⊕◯◯ LOW [6,7,8]	We are uncertain whether home visiting programs decrease cost as the quality of the evidence has been assessed as very low

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk ratio; **SD**: Standard deviation; **SMD:** Standardized mean difference

2.5. Otros métodos de síntesis de la evidencia

En las últimas tres décadas, han surgido y evolucionado diferentes métodos de síntesis de la evidencia (Figura 12).

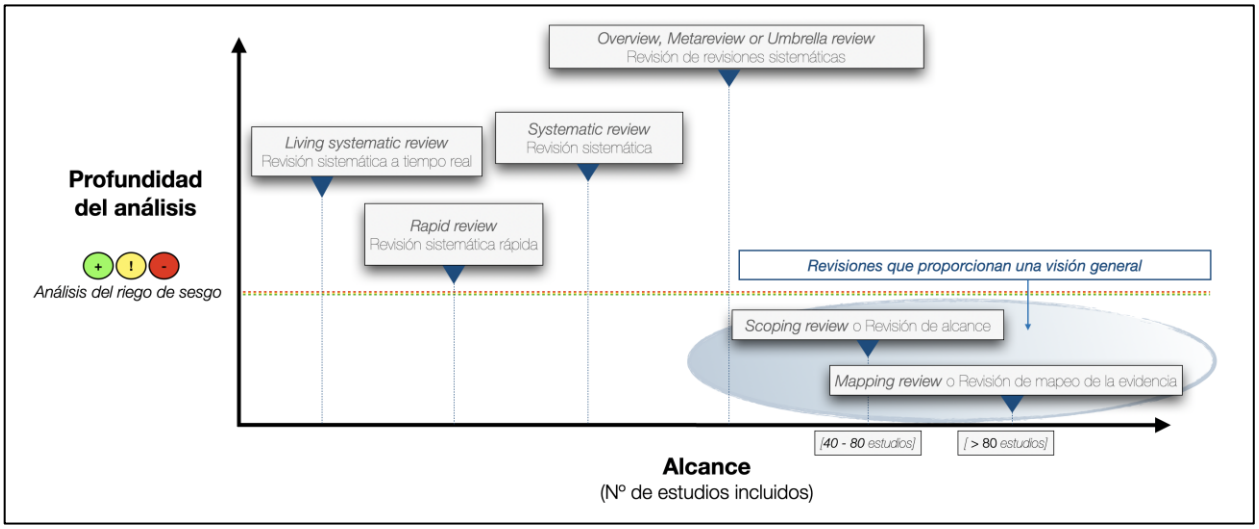


Figura 12. Representación gráfica de los tipos de revisiones de acuerdo con la profundidad del análisis y el alcance de la revisión. Modificación ampliada de la representación gráfica propuesta por Saran and White (2018).

A continuación, se describen los más frecuentemente utilizados.

2.5.1. *Living systematic reviews* o Revisiones sistemáticas a tiempo real

Una RS a tiempo real se define como una RS que se actualiza continuamente, incorporando nuevos estudios a medida que están disponibles. No todas las RS son susceptibles de este enfoque, sino que deben reunir los siguientes criterios:

a) Que sea una revisión prioritaria o de suficiente importancia.

b) Que la certeza de la evidencia existente sea baja o muy baja.

c) Que sea un campo en el que la evidencia está emergiendo con rapidez, lo cual podría cambiar el curso de decisiones políticas o prácticas clínicas.

© Álvaro Limones Burgos, 2024

Embarcarse en este tipo de revisiones no supone una "cadena perpetua", sino que conviene poner fin a la actualización cuando dejen de cumplirse alguno de los criterios descritos anteriormente. Sus fundamentos son compartidos con las RS convencionales, pero estas deben incluir de forma explícita, transparente y predefinida con qué frecuencia se actualizarán las búsquedas, los métodos utilizados para mantener la revisión constantemente actualizada y cómo se publicarán los resultados (Akl et al., 2017; Elliott et al., 2017; Simmonds et al., 2017; Thomas et al., 2017). Para lograrlo, se emplean búsquedas electrónicas automatizadas con sistemas de alerta para localizar nuevas publicaciones (por ejemplo, en MEDLINE/Pubmed, y otras bases de datos) y software basado en *crowdsourcing* y *machine-learning* para un mejor cribado de las publicaciones potenciales (i.e. *Evidence Pipeline, Cochrane Crowd, Screen4ME*), así como plataformas en línea para compartir datos en tiempo real tras la publicación de la revisión (i.e. en *Open Science Framework*).

2.5.2. *Overview / Meta-review / Umbrella review o* Revisión de revisiones sistemáticas

Las revisiones de RS, conocidas en inglés como *Overviews, Meta-reviews o Umbrella reviews*, son un método de síntesis que se utiliza cuando ya existen múltiples RS sobre temas similares o relacionados. Su objetivo principal es reunir, analizar y sintetizar la sobrecarga de información utilizando RS como unidad de análisis (en lugar de los estudios primarios), mediante métodos sistemáticos, reproducibles y explícitos (Higgins et al., 2022).

La herramienta AMSTAR (Shea et al., 2007), en su versión actualizada AMSTAR II (Shea et al., 2017), se utiliza para evaluar la calidad metodológica y la herramienta ROBIS para evaluar el riesgo de sesgo de las RS incluidas (Whiting et al., 2016). Además, los autores de *Overviews* deben extraer e informar también sobre el análisis del riesgo de sesgo de los estudios primarios contenidos dentro de cada RS incluida. Si esta evaluación no se hubiera realizado en la RS original, los autores de *Overviews* deben llevar a cabo este análisis utilizando la herramienta correspondiente para cada estudio.

En cuanto al análisis de la magnitud del efecto, existen dos formas principales de abordarlo, bien presentando los datos exactamente como se muestran en las RS incluidas; o bien volviendo a analizar los datos con un enfoque diferente, como, por ejemplo, analizar

© Álvaro Limones Burgos, 2024

subgrupos específicos o considerar variables o tiempos de seguimiento diferentes a partir de los estudios primarios.

2.5.3. *Scoping reviews* o Revisiones de alcance

Las revisiones de alcance son un método de síntesis que fue concebido inicialmente por Arksey y O'Malley en 2005 y perfeccionadas por Levac et al. y el *Joanna Briggs Institute*, cuyo principal objetivo es contextualizar un área temática de investigación (Arksey & O'Malley, 2005; Levac et al., 2010; Peters et al., 2015; Peters et al., 2020). Analizan descriptivamente conceptos, definiciones clave y términos, identifican investigaciones prioritarias y posibles lagunas de conocimiento, y proporcionan orientación para futuras investigaciones primarias o RS específicas. Aunque siguen un proceso estructurado similar a las RS convencionales, abordan temas más amplios sin necesariamente realizar un análisis exhaustivo del riesgo de sesgo de los estudios incluidos. Por lo general, incluyen alrededor de 40-80 estudios (Campbell et al., 2023). Existe una extensión PRISMA propia para este tipo de estudio, conocida como PRIMA-ScR, para guiar a los autores en el reporte de la publicación (Tricco et al., 2018).

2.5.4. *Mapping reviews* o Revisiones de mapeo de la evidencia

Las revisiones de mapeo son un método de síntesis que surge en 2003 para manejar grandes volúmenes de literatura científica, superando a otros métodos de síntesis de evidencia (Katz et al., 2003). Generalmente, incluyen más de 80 estudios (Campbell et al., 2023). Su objetivo principal es identificar y presentar la ubicación de la evidencia utilizando formatos visuales, a modo de "mapas". Esto se hace sin extraer de forma exhaustiva los datos o realizar una evaluación del riesgo de sesgo de los estudios incluidos, dado el gran volumen de literatura involucrado. En estas revisiones, se puede incluir y analizar cualquier diseño de estudio centrado en una pregunta de investigación común. En 2018, la Colaboración Campbell publicó un manual para llevar a cabo este tipo de revisiones (Saran & White, 2018).

© Álvaro Limones Burgos, 2024

2.5.5. *State-of-the-art reviews* o Revisiones del estado del arte

Las revisiones del Estado del Arte son un método de síntesis que ofrece una visión contemporánea y contextualizada en el tiempo para la comprensión de un campo de conocimiento. Normalmente, abordan el estado actual, la evolución histórica, y las direcciones futuras, proporcionando una síntesis de actualidad (Barry et al., 2022).

2.5.6. *Rapid reviews* o Revisiones rápidas

Las revisiones rápidas son un método de síntesis que simplifica u omite ciertos pasos del proceso de realización de una RS tradicional, con el objetivo de agilizar la producción de información para los tomadores de decisiones en situaciones que requieren respuestas rápidas o para optimizar el uso de recursos (Hamel et al., 2021). La elaboración de revisiones rápidas ha experimentado un marcado aumento en los últimos años, con la pandemia de la COVID-19 como el ejemplo más destacado, durante la cual se llevaron a cabo numerosas revisiones rápidas.

La colaboración Cochrane (Garritty et al., 2021), la Universidad de McMaster (Dobbins, 2017), o la herramienta STARR (por sus siglas en inglés, *SelecTing Approaches for Rapid Review*) ofrecen recomendaciones prácticas sobre cómo realizar revisiones rápidas (Pandor et al., 2019). Está elaborándose actualmente una extensión PRISMA para el reporte de revisiones rápidas, conocida como PRISMA-RR (Stevens et al., 2018).

2.5.7. *Narrative reviews* o Revisiones narrativas o de la literatura

Las revisiones narrativas o de la literatura son un método de síntesis basadas en la revisión de los estudios y trabajos previamente publicados sobre un tema específico, pero a diferencia de una RS, no sigue un protocolo formal y estructurado con métodos explícitos de búsqueda, selección y análisis de datos. En ellas, el autor utiliza un enfoque más descriptivo y cualitativo para resumir, analizar y discutir la literatura relevante, a menudo proporcionando una narrativa coherente y contextualizada sobre el tema en cuestión.

© Álvaro Limones Burgos, 2024

2.6. Metaanálisis

2.6.1 Definición

El metaanálisis es una herramienta de combinación estadística de resultados de dos o más estudios primarios y homogéneos, localizados mediante una RS. Las posibles ventajas de los metaanálisis incluyen:

a. Mejorar la precisión.
b. Aumentar la validez externa de la conclusiones.
c. Permitir responder a preguntas no planteadas por estudios individuales, planteando nuevas hipótesis.
d. Permitir resolver controversias que surgen de estudios aparentemente contradictorios.

El metaanálisis cuenta con dos etapas:

1º Cálculo del efecto en cada estudio bajo un mismo estadístico.
2º Estimación combinada y ponderada del efecto.

2.6.1.1. *Forest plot* o diagrama de Bosque o de efectos

El metaanálisis se visualiza a través de los conocidos como Diagramas de Bosque o *Forest plot*. En este gráfico, se ubica la magnitud del efecto y IC del 95% en el eje horizontal (abscisas), mientras que el eje vertical enumera los nombres de los estudios, acompañados habitualmente por su año de publicación (ordenadas). La magnitud del efecto se muestra con una figura geométrica, generalmente un cuadrado para los estudios individuales, cuyo tamaño refleja el peso del estudio en el análisis global. La estimación global, basada en la ponderación de todos los estudios, se señala usualmente con un rombo o diamante, cuya amplitud representa el rango del IC del 95% (Figura 13).

Este diagrama muestra las medidas de efecto de cada estudio a ambos lados de la línea de no efecto, que se representa con el 0 para las medidas absolutas *Mean difference* (MD) o *Risk Difference* (RD), y con el 1 para las medidas relativas *Odds ratio* (OR), *Risk Ratio* (RR) y *Hazard Ratio* (HR). Los IC ofrecen información sobre la precisión de cada

© Álvaro Limones Burgos, 2024

estudio e identifican aquellos con resultados estadísticamente significativos cuando su IC no cruza la línea de no efecto.

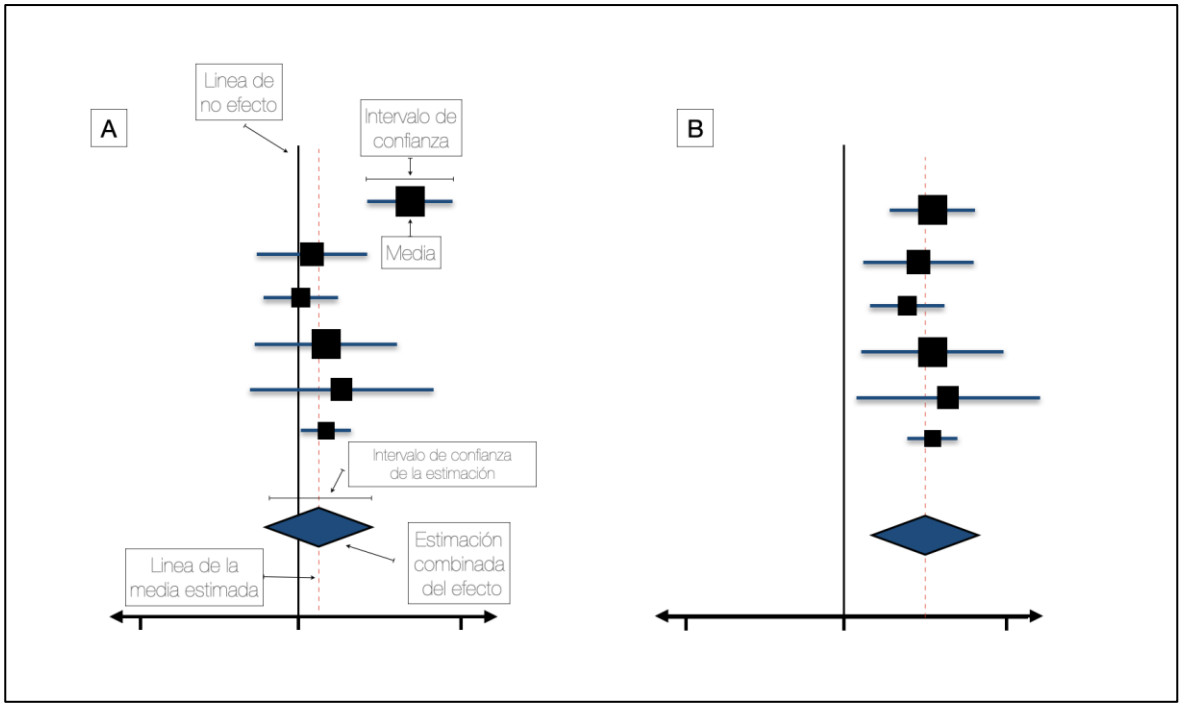


Figura 13. Diagrama de bosque o *Forest plot*. A: Representación visual de una estimación combinada del efecto no estadísticamente significativa. B. Representación visual de una estimación combinada del efecto estadísticamente significativa, el diamante no cruza la línea de no efecto.

2.6.1.2. Uso de medias relativas o absolutas

La evidencia empírica sugiere que las medidas del efecto relativas (i.e. RR o OR) tienden a ser, en promedio, más consistentes que las medidas del efecto absolutas (i.e. RD o NNT) (Deeks, 2002). A pesar de esto, las medidas absolutas suelen ser más fáciles de interpretar. Frecuentemente, es prudente emplear una estadística para el metaanálisis y, posteriormente, expresar los resultados utilizando una segunda estadística más interpretable. Por ejemplo, el metaanálisis puede realizarse de manera más efectiva utilizando medidas del efecto relativas, y luego, los resultados reinterpretarlos usando medidas del efecto absolutas. Esta práctica es uno de los fundamentos para las “Tablas resumen de hallazgos".

© Álvaro Limones Burgos, 2024

2.6.1.3. Modelo de efectos fijos o aleatorios

El modelo de efectos aleatorios parte de la suposición de que los estudios, a pesar de presentar diferentes magnitudes del efecto, están relacionados (DerSimonian & Laird, 1986). Este modelo permite el análisis a pesar de la existencia de heterogeneidad no explicada por factores conocidos.

El modelo de efectos fijos parte de la suposición de que en todos los estudios el verdadero efecto de la intervención (tanto en magnitud, como en dirección) es el mismo (es decir, fijo en todos los estudios) y que las diferencias entre ellos se deben únicamente al azar.

El Manual Cochrane destaca que la elección entre un modelo de efectos fijos y uno de efectos aleatorios ha sido un tema de debate controvertido, sin ofrecer una recomendación universal. Subraya que la decisión no debe basarse únicamente en pruebas estadísticas de heterogeneidad (Higgins et al., 2022). Adoptar un enfoque pragmático, que considere ambos modelos, es recomendado. El método de efectos aleatorios y el método de efecto fijo darán resultados idénticos cuando no exista heterogeneidad. Cuando hay heterogeneidad, el IC de la estimación global ponderada por un modelo de efectos aleatorios es más amplio que con un modelo de efectos fijos. Esto sucederá siempre que la estadística de I^2 sea mayor que cero, incluso si la heterogeneidad no es detectada por la Prueba Q de Cochrane.

Un posible fundamento para presentar únicamente uno u otro es reconocer si se dispone de argumentos teóricos sólidos, además de estadísticos, que soporten una uniformidad de los efectos a través de los estudios, pudiendo preferir en este caso un modelo de efectos fijos. En cambio, si es poco probable que los efectos de una intervención sean consistentes en diferentes contextos y poblaciones, el modelo de efectos aleatorios sería más apropiado, proporcionando una estimación más generalizable.

Los intervalos de predicción en el modelo de efectos aleatorios son útiles para mostrar el rango de variabilidad entre los estudios cuando exista heterogeneidad. Los intervalos de predicción han demostrado ser una forma popular de expresar la cantidad de heterogeneidad en un metaanálisis (Riley et al., 2011). Sin embargo, se basan fuertemente en la suposición de una distribución normal de los efectos a través de los estudios, y pueden ser muy problemáticos cuando el número de estudios es pequeño, en cuyo caso pueden parecer extremadamente anchos o extremadamente estrechos. Sin embargo, es

© Álvaro Limones Burgos, 2024

recomendable su uso cuando el número de estudios es razonable (i.e. más de diez) y no hay una asimetría clara del diagrama de embudo.

Si además de heterogeneidad, hay evidencias de sesgo de publicación, se debería considerar la posibilidad de omitir la presentación de resultados en forma de metaanálisis o de realizar un análisis de sensibilidad.

2.6.2. Metaanálisis de variables dicotómicas:

Hay cuatro métodos comúnmente utilizados para el metaanálisis de variables dicotómicas: tres métodos de efectos fijos (Mantel-Haenszel, Peto e inverso de la varianza) y un método de efectos aleatorios (varianza inversa de DerSimonian y Laird). El método de Peto solo puede combinar OR, mientras que los otros tres métodos pueden combinar RR, OR y RD.

2.6.2.1. Método de Inverso de la varianza

Es el método más común, se llama así porque el peso dado a cada estudio se establece por el inverso de la varianza de la estimación del efecto (es decir, sobre el cuadrado de su error estándar) (DerSimonian & Laird, 1986). De este modo, los estudios más grandes, que tienen errores estándar más pequeños, tienen más peso que los estudios más pequeños, que tienen errores estándar más grandes. Esta elección de pesos minimiza la imprecisión (incertidumbre) de la estimación del efecto agrupado.

2.6.2.2. Método de Mantel-Haenszel

Presenta mejores propiedades estadísticas para metaanálisis de pocos eventos o tamaños muéstrales reducidos (Mantel & Haenszel, 1959).

2.6.2.3. Métodos para eventos raros o cero eventos

Problemas computacionales pueden ocurrir cuando no se observan eventos en uno o ambos grupos de un estudio primario. Los métodos meta-analíticos de Inverso de la Varianza calculan la estimación del efecto y su error estándar para cada estudio. Para los estudios en los que no se observaron eventos en uno o ambos brazos, estos cálculos a menudo implican la división por cero, lo que resulta en un error computacional. Para solucionarlo se aplica el:

© Álvaro Limones Burgos, 2024

2.6.2.3.1. Factor de corrección

La mayoría de los programas de metaanálisis aplican automáticamente una corrección fija (usualmente de 0.5) a todas las celdas de tablas 2x2 en situaciones donde uno de los grupos registra cero eventos, para prevenir problemas de cálculo. Aunque esta corrección evita errores computacionales, a menudo introduce un sesgo en las estimaciones del estudio, especialmente cuando hay desigualdad en los tamaños de los grupos del estudio. Este sesgo es más pronunciado en estudios no aleatorizados que en ensayos controlados aleatorios. Sweeting et al. (2004) han investigado correcciones alternativas, incluyendo ajustes proporcionales al tamaño de los grupos de estudio, para mitigar este efecto (Sweeting et al., 2004).

El método de Mantel-Haenszel solo requiere correcciones para celdas con cero eventos si ocurre en todos los estudios incluidos, haciendo que la necesidad de corrección sea menos frecuente. Si bien, en muchos *softwares* se aplican las mismas reglas de corrección tanto para los métodos Mantel-Haenszel como para los de Inverso de la Varianza, con lo cual puede no ser lo más adecuado.

2.6.2.3.2. Método de Peto

No requiere factor de corrección por cero eventos en alguno de los grupos de intervención. Además, a tasas de eventos inferiores al 1%, se ha observado que el método Peto es menos sesgado y más poderoso que otros métodos en metaanálisis de eventos raros (Bradburn et al., 2007), siempre que no hubiera un desequilibrio sustancial en el tamaño muestral de los grupos de comparación y que los efectos del tratamiento no fueran excepcionalmente grandes.

© Álvaro Limones Burgos, 2024

2.6.3. Metaanálisis de variables continuas:

Las dos principales medidas de efecto utilizadas en metaanálisis de variables continuas son la diferencia de medias (MD) y la diferencia de medias estandarizada (SMD).

En el caso de la MD, se utilizan las desviaciones estándar (SD) y los tamaños muestrales de los estudios para determinar el peso asignado a cada estudio. A los estudios con SD pequeña se les da un peso relativamente mayor, mientras que a los estudios con SD más grande se les dan pesos relativamente más pequeños.

Por otro lado, el enfoque de SMD utiliza la SD para estandarizar las diferencias medias a una escala común, afectando también el cálculo de los pesos de los estudios. Así, estudios con SD pequeñas resultan en estimaciones más altas de SMD y viceversa.

Una suposición importante para datos continuos es que cada brazo de intervención en cada estudio sigue una distribución normal. Aunque esta suposición puede no cumplirse siempre, su importancia disminuye en estudios de gran tamaño.

La comprobación se calcula restando la media observada menos el valor más bajo posible (o el valor más alto posible menos la media observada), y dividir esto por la desviación estándar (SD) (Altman & Bland, 1996). Una proporción menor a 2 sugiere asimetría. Si la proporción es menor a 1, hay una fuerte evidencia de una distribución sesgada. La transformación de los datos originales a escala logarítmica puede reducir sustancialmente la asimetría.

2.6.3.1. Método de Inverso de la Varianza

Es el método más comúnmente utilizado. Descrito en el apartado 2.6.2.1.

2.6.3.2. Metaanálisis de variables de cambio

Un metaanálisis con variables de cambio (restando medición final menos inicial) será más eficiente y poderoso que uno con variables finales post-intervención, ya que elimina la componente de variabilidad intra-sujeto.

En la práctica, es común que un autor de RS se tope con una mezcla de estudios que incluyen variables de cambio y variables de medición final post-intervención. Sin embargo, no hay ninguna razón estadística por la que los estudios con variables de cambio no deban combinarse con estudios con variables finales post-intervención cuando se utiliza

© Álvaro Limones Burgos, 2024

el método MD (no estandarizado) ya que la diferencia de los valores medios post-intervención será, en promedio, la misma que la diferencia de las variables de cambio. Si el uso de variables de cambio aumenta la precisión, los estudios que las usen tendrán pesos más altos en el metaanálisis ya que tendrán SD más pequeños (Higgins et al., 2022).

Sin embargo, el cálculo de variables de cambio implica la medición del resultado dos veces y, en la práctica, puede ser menos eficiente cuando los resultados son poco consistentes o difíciles de medir con precisión, donde el error de medición puede ser mayor que la verdadera variabilidad intra-sujeto.

2.6.4. Metaanálisis de combinación de variables dicotómicas y continuas

En ocasiones, los autores de RS se enfrentan al desafío de que la misma evaluación del efecto se ha realizado con una variable dicotómica y una variable continua. Por ejemplo, los estudios de satisfacción de retención post-ortodoncia, que pueden ser removibles o fijos pueden evaluarse de dos formas: mediante categorías como satisfecho/si o insatisfecho/no (variables dicotómicas) o utilizando una escala visual analógica [entre 1-100 mm] (variable continua). Este tipo de información suele ser más fácil de entender, y más útil, cuando se dicotomiza. Sin embargo, decidir sobre un punto de corte puede ser arbitrario, y la información se pierde cuando los datos continuos se transforman en datos dicotómicos.

Existen métodos estadísticos disponibles que permiten expresar las OR como SMD, y viceversa, lo que permite combinar datos dicotómicos y continuos (Anzures-Cabrera et al., 2011); de acuerdo con la siguiente fórmula:

$$SMD = \frac{\sqrt{3}}{\pi} \ln OR$$

Esto se basa en la suposición de que las mediciones continuas en cada grupo de intervención siguen una distribución logística (una distribución simétrica similar a la normal, pero con mayor cantidad de datos en las colas), y que la variabilidad de los resultados es equivalente tanto en los participantes experimentales como en los comparadores.

Una vez que se han calculado los SMD (o OR) y sus errores estándar para todos los estudios en el metaanálisis, se pueden combinar utilizando el método genérico de Inverso de la Varianza.

© Álvaro Limones Burgos, 2024

2.6.5. Metaanálisis de variables tiempo-hasta-el-evento:

Las variables de "tiempo hasta el evento" o *time-to-event outcome*s se refieren al tiempo transcurrido hasta que ocurre un evento de interés, y son conocidas generalmente como datos de supervivencia. Estos datos incluyen pares de observaciones para cada individuo: primero, un período durante el cual no se observa ningún evento, y segundo, un indicador que señala si al final de ese período se produjo un evento o simplemente finalizó la observación. Los participantes cuyos períodos de tiempo no terminan en un evento se consideran "censurados". Su tiempo libre de eventos aporta información valiosa y se incluye en el análisis. Las variables de "tiempo hasta el evento" se pueden analizar como datos dicotómicos, si se estudia el estado de todos los pacientes en un momento fijo del tiempo.

Las variables de "tiempo hasta el evento" pueden analizarse como datos dicotómicos si se evalúa el estado de todos los pacientes en un momento específico. No es adecuado utilizar métodos para variables continuas en el análisis de datos de tiempo hasta el evento, ya que los tiempos relevantes solo se conocen para aquellos participantes que han experimentado el evento. Excluir a los participantes censurados introduce casi con seguridad un sesgo.

La forma más adecuada de resumir los datos de "tiempo hasta el evento" es utilizar métodos de análisis de supervivencia y usar la medida del efecto relativa Hazard Ratio (HR).

2.6.6. Análisis de sensibilidad

Un desafío importante en los metaanálisis y RS es la inclusión frecuente de estudios con resultados atípicos, lo que puede resultar en conclusiones erróneas y afectar negativamente la solidez de la toma de decisiones. Los análisis de sensibilidad deben realizarse para examinar si los hallazgos son sólidos para fundamentar decisiones potencialmente influyentes.

El enfoque tradicional para analizar la sensibilidad fue descrito por Viechtbauer y Cheung en 2010, mediante el método de "dejar un estudio fuera" o en inglés, *the leave-one-study-out procedure* para evaluar cómo cada estudio individual afecta el resultado general (Viechtbauer & Cheung, 2010). No obstante, este método ha sido objeto de críticas recientes, ya que, al evaluar la influencia de un estudio potencialmente atípico, la presencia de otros estudios atípicos en el

© Álvaro Limones Burgos, 2024

análisis puede distorsionar significativamente la evaluación, es decir, solo funciona bien cuando un único estudio presenta resultados atípicos en el metaanálisis, pero no cuando existen múltiples estudios con resultados atípicos. Para solucionar esto, Meng et al. (2024) proponen un modelo interactivo para detectar valores atípicos potenciales arrojando menor sesgo.

© Álvaro Limones Burgos, 2024

2.7. Otros métodos metaanalíticos:

2.7.1. Metaanálisis acumulado

Usados principalmente en las *Living Systematic Reviews* o Revisiones sistemáticas en tiempo real que se actualiza continuamente, incorporando nuevos estudios a medida que están disponibles. Requiere la selección de una variable específica para ordenar los estudios según un criterio predefinido, como por ejemplo la fecha de publicación. Tras establecer este orden, los estudios se incorporan secuencialmente, uno tras otro, permitiendo la recalculación y estimación del parámetro resumen en cada etapa.

Los resultados de este proceso se visualizan a través de gráficos que representan el valor de la medida del efecto (RR, OR etc.) después de la adición de cada nuevo estudio. Este enfoque gráfico facilita la interpretación de cómo evoluciona la estimación del efecto con la inclusión de más datos y cómo cada estudio individual contribuye al conocimiento general sobre el tema en cuestión.

La organización de los estudios según la fecha de publicación o la fecha en que se recolectaron los datos es el enfoque más comúnmente utilizado en los metaanálisis acumulados. Este método permite evaluar la influencia de cada estudio en la base de evidencia disponible en un momento dado, proporcionando una perspectiva dinámica sobre el desarrollo del conocimiento en el área de interés.

Además de la temporalidad, los metaanálisis acumulados pueden organizarse en torno a otras variables cuantitativas, como la calidad general de cada estudio o la frecuencia del efecto en el grupo de referencia. Este enfoque facilita la exploración de la heterogeneidad entre los estudios, permitiendo a los investigadores comprender mejor cómo varía el resultado del metaanálisis en función de características específicas de los estudios incluidos.

© Álvaro Limones Burgos, 2024

2.7.2. Metaanálisis en red *o Network Meta-analysis*

Se trata de un método estadístico que permite la comparación simultánea de múltiples intervenciones, integrando evidencia directa e indirecta bajo supuestos de transitividad fundamentados en comparadores comunes y metodologías homogéneas y consistentes entre los estudios incluidos que conforman la red (Catalá-López & Tobías, 2013).

Dicha red se presenta en Diagramas de Red o *Network diagrams,* que contiene nodos, para representar cada intervención (cuyo tamaño varía según el número de ensayos incluidos), y líneas coloreadas que unen estos nodos para indicar la evidencia directa (trazo único) o indirecta (puntuada) de diferentes grosores de acuerdo con el número de estudios incluidos, y el riesgo de sesgo asociado (Rojo – alto riesgo de sesgo, Verde – bajo riesgo de sesgo, Amarillo – algunas preocupaciones) (Figura 14). En una red conectada, el número de comparaciones que deben estimarse en un metaanálisis en red es igual al número de intervenciones menos una.

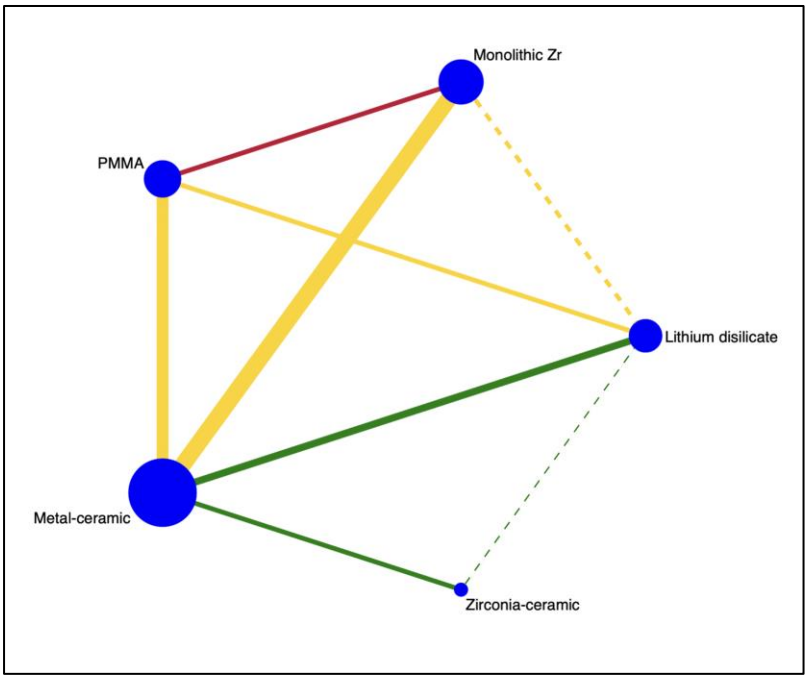


Figura 14. Ejemplo de Diagrama en Red o *Network diagram.*

Su principal ventaja es que proporciona información de comparaciones que nunca han sido evaluadas en ECAs individuales, y permite establecer un ranking de las intervenciones comparadas.

© Álvaro Limones Burgos, 2024

2.7.3. Metaanálisis Multivariante

El metaanálisis multivariado realiza el análisis combinado de varias variables teniendo en cuenta sus intercorrelaciones aumentando la eficiencia analítica y produciendo conclusiones que pueden diferir de aquellas obtenidas a través de metaanálisis univariantes (Hattle et al., 2022).

2.7.4. Metaregresión

La metarregresión es otro método estadístico utilizado en metaanálisis para explorar cómo diferentes factores o características de los estudios incluidos pueden afectar la estimación combinada del efecto (Tipton et al., 2019). A menudo se les llama "modificadores potenciales del efecto" o covariables. No debe considerarse cuando hay menos de diez estudios en un metaanálisis y permite examinar tanto factores categóricos como continuos.

El coeficiente de regresión obtenido de un análisis de metarregresión describirá cómo la estimación combinada del efecto cambia con un aumento unitario en una variable explicativa. La significancia estadística del coeficiente de regresión es una prueba que indica si hay una relación lineal entre la estimación combinada del efecto y la variable explicativa.

2.7.5. Gráfico de Albatross

Este gráfico muestra la distribución de la magnitud del efecto de los estudios cuando no es posible realizar un meta-análisis (Harrison et al., 2017), ofreciendo una idea general de hacia dónde se inclina una determinada comparación (Figura 15). El gráfico Albatros se obtiene a partir del número de participantes, de los valores P y las direcciones de efecto. Los métodos estadísticos que actualmente soporta son: proporciones; MD; SMD; RR y OR.

© Álvaro Limones Burgos, 2024

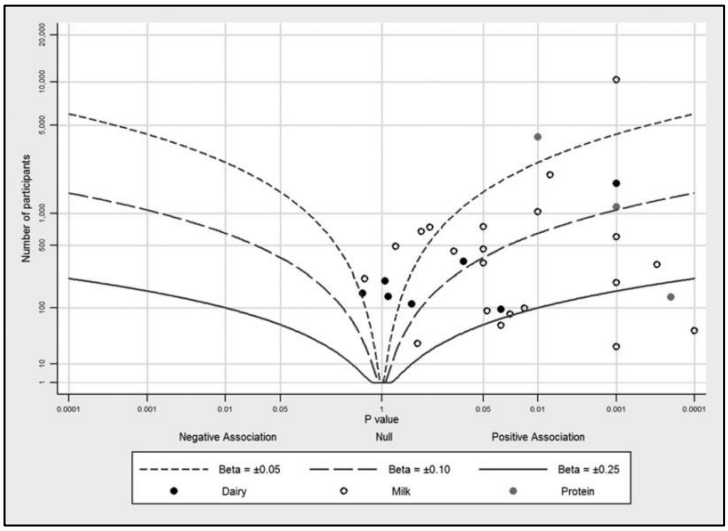


Figura 15. Gráfico de Albatross representado un efecto favorable de la intervención obtenido del artículo de Harrison et al. (2017).

© Álvaro Limones Burgos, 2024

2.8. Guías de práctica clínica y formulación de recomendaciones – Enfoque GRADE

2.9.1. Definición

Las guías de práctica clínica (GPC) son documentos sistematizados que contienen recomendaciones clínicas, fundamentadas en la evidencia científica más actual identificada a través de RS, para **optimizar y mejorar el cuidado del paciente (**Brouwers et al., 2016**)**. La Asociación científico-médica alemana o *Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften* (AWMF), ha establecido un sistema de clasificación que organiza las GPC en cuatro niveles, detallados en la Tabla 10:

- Nivel S1: se basan en agrupaciones de opiniones individuales de expertos y no se desarrollan mediante un proceso sistemático.
- Nivel S2: se divide en:
    a. Nivel S2e: fundamentadas en la evidencia científica recopilada a través de RS.
    b. Nivel S2k: originadas a partir de reuniones de consenso mediante métodos estructurados y miembros expertos de comités o sociedades representativas de del tema de estudio.
- Nivel S3: representa el escalón más elevado, integran la síntesis de la evidencia científica mediante RS con reuniones del consenso por las entidades más relevantes.

Tabla 10. Clasificación de las guías de práctica clínica.

Clasificación		Tipo	
S1		Guías basadas en opiniones individuales de expertos	Nivel de relevancia
S2	**S2k**	Guías basadas en reuniones de consenso	
	S2e	Guías basadas en revisiones sistemáticas	
S3		Guías basadas en revisiones sistemáticas consensuadas en reuniones de consenso	

© Álvaro Limones Burgos, 2024

El enfoque GRADE proporciona un método estructurado tanto para el desarrollo de guías de práctica clínica como para la evaluación de la certeza de la evidencia científica en RS, tal como ya hemos visto en secciones anteriores. No obstante, establece una nítida distinción entre ambos procesos. Según su enfoque, las RS tienen la responsabilidad principal de proporcionar un resumen exhaustivo de la evidencia disponible, pero, en general, **no deben incluir recomendaciones clínicas directas**. Por lo tanto, la función de las RS finaliza después de calificar la certeza de la evidencia y de presentar los hallazgos de manera clara en una Tabla Resumen de Hallazgos.

Los pasos siguientes son responsabilidad de las GPC nivel S3, que utilizan las RS como base para su elaboración. En ocasiones, en lugar de llevar a cabo revisiones específicas con ese propósito, se pueden emplear RS previas de alta calidad. El objetivo de las GPC son proporcionar recomendaciones clínicas que guíen a los profesionales de la salud en la toma de decisiones sobre cómo tratar a sus pacientes de manera más eficiente y segura. Para lograr esto, se evalúa cuidadosamente el balance entre los efectos positivos y negativos de cada intervención en estudio.

Los desarrolladores de las guías toman dos aspectos para formular recomendaciones: la dirección (a favor o en contra de la intervención) y la fuerza (fuerte o débil) de la recomendación. La fuerza de una recomendación refleja la confianza que el panel de expertos tiene en que los beneficios de una intervención superan claramente los posibles efectos adversos, o viceversa.

Cuando el panel de expertos no tiene claro el balance entre beneficios y riesgos, o no dispone de información suficiente, se tiende a emitir recomendaciones débiles, o incluso pueden optar por no hacer ninguna recomendación (Schünemann et al., 2013).

2.9.2. Recomendaciones fuertes

Una recomendación es considerada fuerte cuando el panel de expertos de la guía clínica tiene **alta confianza** en que los beneficios esperados de una intervención superan de manera significativa los efectos negativos (recomendación fuerte a favor de la intervención) o, por el contrario, cuando los efectos negativos son claramente superiores a los beneficios esperados (recomendación fuerte en contra de la intervención). Esto implica que la mayoría o incluso todos los individuos se beneficiarían notablemente al seguir el curso de acción recomendado (Schünemann et al., 2013).

© Álvaro Limones Burgos, 2024

2.9.3. Recomendaciones débiles

Una recomendación débil se caracteriza por presentar **una incertidumbre apreciable,** a pesar de que el efecto deseable probablemente supere los efectos indeseables (recomendación débil a favor de una intervención), o los efectos indeseables probablemente superen los efectos deseables (recomendación débil en contra de una intervención). En tales casos, es necesario considerar más cuidadosamente otros aspectos como: las circunstancias individuales, las preferencias y la opinión de los pacientes (Schünemann et al., 2013).

2.9.4. ¿Cuándo no hacer recomendaciones?

Tres son las razones por las cuales los panelistas de guías pueden mostrar cierta reticencia a emitir una recomendación a favor o en contra, e incluso no recomendarla para fines de investigación:

1. Cuando la confianza en la **estimación del efecto es tan baja** que el panel considera que una recomendación **sería altamente especulativa**. En otras palabras, a mayor certeza de la evidencia, mayor probabilidad de emitir una recomendación fuerte.
2. Cuando el **equilibrio entre los beneficios y los riesgos es muy pequeño o igualado**. En estas situaciones, el panel enfrenta dificultades considerables para determinar la dirección de la recomendación (a medida que el balance neto entre las consecuencias deseables e indeseables aumenta, aumenta la probabilidad de justificar una recomendación fuerte).
3. Cuando las **intervenciones comparadas conducen a desenlaces diferentes y poco deseables**, es probable que cada paciente responda de formas muy diversas y no existan preferencias comunes.

2.9.5. Factores que influyen en la dirección y la fuerza de la recomendación

- **Balance neto** entre variables deseables y no deseables teniendo en cuenta lo siguiente:
    - Prioridad de las variables críticas, generalmente, las primarias sobre las secundarias.
    - La magnitud del efecto de las variables deseables y no deseables. Cuanto mayor sean las diferencias entre las variables deseables y no deseables, mayor

© Álvaro Limones Burgos, 2024

probabilidad de justificar una recomendación fuerte. Cuanto más pequeño sea el beneficio neto y menos apreciable sea el beneficio o perjuicio, mayor probabilidad de garantiza una recomendación débil.

- Que la mayoría de las variables apunten en la misma dirección, hacia efectos deseables o hacia efectos indeseables (i.e. grandes magnitudes del efecto en direcciones opuestas implicarán una recomendación débil).

- **Nivel de certeza de la evidencia**. Cuanto mayor sea la certeza de la evidencia, mayor es la probabilidad de una recomendación fuerte. Típicamente, una recomendación fuerte se asocia a alta certeza o al menos moderada en las estimaciones del efecto de las variables principales. Si se tiene alta certeza en la variable primaria pero baja certeza en otras variables consideradas críticas, se garantiza por lo general una recomendación débil; incluso cuando aparentemente exista un balance neto claramente a favor o en contra. Rara vez una baja certeza en las variables principales se asocia a recomendaciones fuertes. En general, el sistema GRADE desalienta a los panelistas de las guías a hacer recomendaciones fuertes cuando la certeza en las variables críticas sea baja o muy baja; aunque ha identificado cinco situaciones paradigmáticas en las cuales se garantizan recomendaciones fuertes o muy fuertes a pesar de la baja o muy baja calidad de la evidencia (Tabla 11).

Tabla 11. Situaciones paradigmáticas en las cuales se garantiza una recomendación fuerte a pesar de la baja o muy baja confianza en el estimativo del efecto con ejemplos. La Tabla esta extraída del manual del enfoque GRADE (Schünemann et al., 2013).

Situación paradigmática		Ejemplos
1	Cuando evidencia de baja calidad sugiere beneficio en una situación que amenace la vida (la evidencia acerca de los riesgos puede ser baja o alta)	1. Plasma fresco congelado o vitamina K en un paciente recibiendo warfarina que tiene un INR elevado y tiene sangrado intracraneal. Solo evidencia de baja calidad soporta el beneficio de limitar el grado de sangrado. 2. Anfotericina B vs. itraconazol en blastomicosis diseminada que amenaza la vida. Evidencia de alta calidad sugiere que la anfotericina B es más tóxica que el itraconazol, y evidencia de baja calidad sugiere que ésta reduce la mortalidad en este contexto.
2	Cuando evidencia de baja calidad sugiere beneficio y evidencia de alta calidad sugiere riesgo o costos muy altos	Tamizaje de cáncer con TAC/RMN de cabeza a pies. Evidencia de baja calidad de beneficios de detección temprana, pero evidencia de alta calidad de posibles riesgos y/o evidencia de costos altos ∑ (recomendación fuerte en contra de esta estrategia)

© Álvaro Limones Burgos, 2024

3	Cuando evidencia de baja calidad sugiere equivalencia de dos alternativas, pero evidencia de alta calidad de menor riesgo para una de las alternativas	Erradicación de Helicobacter pylori en pacientes con linfoma gástrico MALT en fases tempranas con H. pilory positivo. Evidencia de baja calidad sugiere que la erradicación inicial de H. pillory resulta en tasas similares de respuesta completa en comparación con las alternativas de radioterapia o gastrectomía; evidencia de alta calidad sugiere menos daño/morbilidad.
4	Cuando evidencia de alta calidad sugiere equivalencia de dos alternativas y evidencia de baja calidad sugiere riesgo en una de las alternativas	Hipertensión en mujeres planeando concebir y en embarazo. Recomendaciones fuertes para labetalol y nifedipino y recomendaciones fuertes en contra de inhibidores de la enzima convertidora de angiotensina (IECA) y antagonistas de receptores de angiotensina (ARA), todos los agentes tienen evidencia de alta calidad para desenlaces equivalentes beneficiosos, con evidencia de baja calidad de efectos adversos mayores con IECAs y ARAs
5	Cuando evidencia de alta calidad sugiere beneficio moderado y evidencia de baja o muy baja calidad sugiere posibilidad de riesgo catastrófico.	Testosterona en hombres con cáncer de próstata o en riesgo de tenerlo. Evidencia de alta calidad para beneficios moderados del tratamiento con testosterona en hombres con deficiencia de andrógenos sintomática en mejorar la densidad mineral ósea y la fuerza muscular. Evidencia de baja calidad para riesgos en pacientes con cáncer de próstata o en riesgo de tenerlo.

INR – razón normalizada internacional; TAC – tomografía axial computarizada; RMN– Resonancia magnética nuclear; MALT – Tejido linfoide asociado a mucosa.

- **Preferencias de los pacientes y su variabilidad.** Cuanto mayor sea la variabilidad en las preferencias de los pacientes, mayor probabilidad de una recomendación débil. La obtención de esta información se lleva a cabo mediante RS de efectividad de las intervenciones que exploran la evidencia relacionada con la calidad de vida, los valores y las preferencias informados por los pacientes. También se puede obtener a través de consultas individuales a pacientes o grupos de pacientes. En última instancia, si no se dispone de estas fuentes, se puede basar en revisión no sistemáticas de la literatura que aborden la experiencia derivada de las interacciones con los pacientes.

- **Costes y recursos**. Cuanto mayor sean los costes de una intervención (más recursos consumidos), menor probabilidad de una recomendación fuerte. Aunque los costes y recursos se pueden considerar como una variable potencialmente importante, como mortalidad, morbilidad y/o calidad de vida, los panelistas pueden o no considerar el uso de los costes y recursos para determinar la dirección y la fuerza de las recomendaciones. Las razones para no tener en cuenta los costes y recursos podrían ser: la falta de datos confiables, la posibilidad de que la intervención no aporte beneficios significativos en relación al esfuerzo necesario para calcular el uso de recursos, la preponderancia del efecto deseado sobre cualquier efecto indeseable, que las consideraciones de los recursos no influyan en la decisión final o que se haya decidido delegar estas consideraciones a otras

© Álvaro Limones Burgos, 2024

instancias de decisión. No obstante, los panelistas deben detallar su elección de no tener en cuenta los costes y recursos, y proporcionar las razones explícitas detrás de esta decisión.

- **Perspectiva/Entorno**. Se debe informar de la audiencia a la que está destinada: un hospital individual, un país, o una audiencia internacional.
- **Equidad**. Se debe considerar si las intervenciones son accesibles para todo el mundo, sin atender a cuestiones económicas, demográficas, etc.
- **Aceptabilidad**. Se refiere a cómo de aceptable es una opción para las partes clave involucradas, como pacientes, profesionales de la salud, políticos, y otros interesados.
- **Factibilidad**. Se refiere a cómo de sencillo es realizar una opción terapéutica o diagnóstica en términos prácticos. Si una opción presenta demasiadas barreras o dificultades para ser implementada, es menos probable que se recomiende.

2.9.6. Redacción de recomendaciones

La redacción de una recomendación debe proporcionar al clínico la mayor cantidad posible de indicadores para su comprensión e interpretación.

Las recomendaciones deben siempre responder a la pregunta clínica original. Por lo tanto, es esencial especificar los pacientes o la población (definidos por la enfermedad y otros factores identificativos) a los que se dirige la recomendación, así como la intervención recomendada, de manera tan específica y detallada como sea posible. A menos que sea evidente, también se debe indicar el comparador. En algunos casos, la recomendación puede hacer referencia a un escenario específico (p.ej. atención primaria o terciaria etc.).

Las recomendaciones formuladas en voz pasiva pueden ser poco claras, por lo que GRADE sugiere que los desarrolladores de guías utilicen la voz activa.

Para las recomendaciones fuertes, el grupo de trabajo GRADE ha propuesto utilizar la terminología: “nosotros recomendamos…” o “los clínicos deberían…”, “los clínicos no deberían…” o “Haga…”, “No haga…”.

Para recomendaciones débiles, el grupo de trabajo GRADE sugiere emplear frases menos definitivas, como “nosotros sugerimos” o “los clínicos podrían…” o “Nosotros recomendamos condicionalmente…” o “Nosotros hacemos una recomendación calificada que…”

© Álvaro Limones Burgos, 2024

2.9.7. Marcos de la Evidencia a la Decisión (EtD)

El grupo de trabajo GRADE ha creado los "marcos de la evidencia a la decisión" (EtD, por sus siglas en inglés, *Evidence to Decision*). Estos marcos EtD tienen como finalidad asistir a los panelistas expertos y desarrolladores de GPC

en la presentación estructurada, transparente y contextualizada de las decisiones, juicios y la evidencia que las sustenta para formular recomendaciones clínicas. Los marcos EtD se organizan en una estructura uniforme que incluye tres secciones: 1. Formulación de una pregunta de investigación, 2. Juicios basados en la evidencia, 3. Conclusiones (Alonso-Coello et al., 2018). En la Tabla 12, se detallan minuciosamente los aspectos a considerar en el marco EtD.

Tabla 12. Marco EtD.

Dominio	Criterio	Juicios específicos
Sección 1: Presentación explícita de una pregunta de investigación		
Sección 2: Juicios basados en la evidencia		
Grado de prioridad del problema	¿El problema, constituye una prioridad?	¿Las consecuencias del problema son serias (graves o importantes en cuanto a los beneficios o el ahorro potencial)? ¿Es urgente? ¿Es una prioridad reconocida? (como por ejemplo basada en una decisión política o de política sanitaria)?
Beneficios	¿Cuál es la magnitud de los efectos deseados que se prevén?	Juicios para cada resultado para el cual hay un efecto deseado
Daños	¿Cuál es la magnitud de los efectos no deseados que se prevén?	Juicios para cada resultado para el cual hay un efecto no deseado
Certeza (calidad o confianza) de la evidencia	¿Cuál es la certeza global de la evidencia de los efectos?	Ver la guía GRADE sobre los juicios detallados de la certeza de la evidencia o confianza en las estimaciones de los efectos
Importancia del resultado	¿Hay incertidumbre, o variabilidad, importante en el valor que dan las personas a los resultados principales?	¿Hay incertidumbre importante en el valor que dan a cada una de las variables principales? ¿Hay variabilidad importante en el valor que dan a cada una de las variables principales? Justifica qué punto influyen las siguientes consideraciones en el equilibrio entre los efectos deseados y no deseados: ¿Cuánto menos valoran las personas los resultados futuros en comparación con los actuales (las tasas de descuento)? La actitud de las personas ante los efectos no deseados, ¿cuál es su aversión hacia el riesgo?

© Álvaro Limones Burgos, 2024

		La actitud de las personas ante los efectos deseados, ¿hasta qué punto asumen riesgos?
Uso de recursos	¿Cuál es la certeza de la evidencia de recursos necesarios?	¿Se han identificado todos las variables importantes relacionadas con los recursos necesarios que diferencian las opciones consideradas? ¿Cuál es la certeza de la evidencia de las variables importantes relacionadas con los recursos necesarios que diferencian las opciones consideradas?
Balance	¿Valen los beneficios netos el coste incremental?	¿Es la razón coste-efectividad sensible a los análisis unilaterales? ¿Es la razón coste-efectividad sensible a los análisis multifactoriales? ¿Es fiable la evaluación económica en que se basa la estimación del coste-efectividad aplicable al marco de interés? ¿Es la evaluación económica en que se basa la estimación del coste-efectividad aplicable a las intervenciones consideradas?
Equidad	¿Cuál sería el impacto sobre la equidad en salud?	Hay grupos o ámbitos que puedan estar en desventaja con relación al problema o las intervenciones (opciones) consideradas? ¿Existen razones plausibles para anticipar diferencias en la efectividad relativa de la intervención (opción)? ¿Hay condiciones iniciales diferentes entre los grupos o ámbitos que influyan sobre la efectividad absoluta de la intervención o la comparación del problema para grupos o ámbitos en desventaja? ¿Deben tenerse en cuenta consideraciones importantes a la hora de implementar la intervención (opción) para garantizar que disminuyen las desigualdades, si es posible, o que no aumentan?
Aceptabilidad	¿Es aceptable la intervención/opción para los grupos de interés?	¿Hay grupos de interés principales que no aceptarían la distribución de los beneficios, daños y costes? ¿Hay grupos de interés principales que no aceptarían los costos o los efectos no deseados a corto plazo para beneficiarse de unos efectos deseados (beneficios) en el futuro? ¿Hay grupos de interés principales que no estarían de acuerdo con la importancia (el valor) otorgada a los efectos deseados o no deseados (por cómo puede afectarles personalmente o por sus percepciones de la importancia relativa de los efectos para otros)? ¿Tendría la intervención efectos negativos sobre la autonomía de las personas? ¿Hay grupos de interés principales que no aprobarían la intervención moralmente por razones distintas de sus efectos sobre la autonomía de las personas (con respecto a principios éticos como no maleficencia, beneficencia o justicia)?
Factibilidad	¿Es factible la implementación de la intervención?	¿Es sostenible la opción o intervención? ¿Hay obstáculos importantes que probablemente limiten la viabilidad de implementar la intervención (opción) o que requieran consideraciones al implementarla? ¿Es sostenible la cobertura sanitaria de la intervención? ¿Es factible garantizar su uso apropiado para las indicaciones apropiadas? ¿Es el uso inapropiado (indicaciones no aprobadas) un problema importante? ¿Es el acceso a la intervención un problema importante? ¿Hay restricciones legales o burocráticas importantes que puedan dificultar o impedir la cobertura sanitaria de la intervención? ¿Es sostenible la opción o intervención?

© Álvaro Limones Burgos, 2024

Sección 3: Conclusiones

Recomendación
Certeza de la evidencia
Grado de consenso
Justificación
Consideraciones de subgrupos
Consideraciones para la implementación
Prioridades de estudio

© Álvaro Limones Burgos, 2024

2.9. Iniciativa COMET para la estandarización de variables en ensayos clínicos

La Iniciativa COMET (*Core Outcome Measures in Effectiveness Trials*) reúne a diversas partes interesadas en el desarrollo y la aplicación de conjuntos de variables esenciales (*Core Outcome Sets*, COS) que deben evaluarse y reportarse en todos los ensayos clínicos sobre una condición específica. El objetivo es facilitar la consistencia entre estudios y, en última instancia, fortalecer la certeza de la evidencia científica para así mejorar la formulación de recomendaciones clínicas (Williamson et al., 2017). Se trata de la selección de un mínimo de variables estandarizadas, relevantes e importantes para evaluar el efecto de las intervenciones.

El uso de COS promueve ensayos de mayor calidad, facilita la comparación y combinación de resultados, reduce la heterogeneidad entre estudios, el desperdicio de recursos y el riesgo de sesgo por notificación selectiva de resultados que puede llevar al uso de intervenciones ineficaces o incluso dañinas.

En los ensayos clínicos, generalmente se incluyen múltiples variables de interés. La variable principal se elige por su importancia terapéutica y suele ser determinante para el cálculo del tamaño de la muestra; aunque a veces, se proponen varias variables primarias críticas y de igual importancia terapéutica.

Los resultados secundarios evalúan efectos beneficiosos o dañinos adicionales o ayudan a explicar más detalles de la intervención. No obstante, es crucial considerar siempre los efectos nocivos, independientemente de si se etiquetan como resultados primarios o secundarios. Además, suelen valorarse otros aspectos como la aceptabilidad y rentabilidad de las intervenciones. Algunos resultados clínicos son combinaciones de varios elementos, conocidos como resultados compuestos.

El desarrollo de COS se estructura en el siguiente esquema:

1. **Identificación de la condición a estudiar**: ¿Qué se va a medir?
2. **Definición de los métodos de medición**: ¿Cómo se va a medir? Variables e instrumentos de medición.
3. **Determinación de una métrica específica**: ¿En qué unidades? Escalas y unidades.

© Álvaro Limones Burgos, 2024

4. **Establecimiento del método de agregación** (es decir, como se presentan los datos) **y el marco de relevancia clínica.**

Incluir un conjunto de variables esenciales (COS) en el diseño de ECA no implica que las variables evaluadas deban restringirse únicamente a las especificadas en dicho conjunto.

© Álvaro Limones Burgos, 2024

3. JUSTIFICACION

A lo largo de las últimas cuatro décadas, el concepto de Odontología Basada en la Evidencia (OBE) ha emergido y consolidado su presencia en el panorama científico dental global.

Los profesionales de la Odontología se ven generalmente abrumados con vastas cantidades de información científica sobre salud oral, pudiendo pasar por alto criterios importantes y siendo poco probable que cuenten con el tiempo y los recursos necesarios para identificar, evaluar y discernir la evidencia científica de alta calidad que fundamente la toma de decisiones clínicas diarias. El desafío constante que representa mantenerse actualizado, resalta aún más la importancia de la síntesis y la formulación de recomendaciones clínicas en beneficio de los pacientes **(**Clarkson & Worthington, 2019**)**.

El notable aumento de las revistas de pago en el ámbito de la Odontología, que dependen en parte o en su totalidad de las contribuciones monetarias de los autores para su financiación, con procesos de revisión extremadamente rápidos que permiten la publicación en pocos días (Eliades et al., 2024), así como el uso de métricas para evaluar a investigadores, grupos de investigación y universidades, están incrementando exponencialmente la producción científica, lo que subraya aún más la importancia de llevar a cabo una evaluación rigurosa, un escrutinio y una síntesis de las publicaciones sobre las que fundamentar la práctica clínica con seguridad.

Varias investigaciones ya han señalado que los patrones de publicación están evolucionando en una dirección menos deseable, debido a lo que se conoce como la "Ley de *Goodhart*", que sostiene que "cuando una medida se convierte en un objetivo, deja de ser una buena medida". Este fenómeno, que no es ajeno a la Odontología, fomenta comportamientos no deseados y éticamente cuestionables, como atribuciones de autorías ficticias, autocitas forzadas, manipulación de datos, métricas o análisis, publicación de un mismo estudio por fascículos, recurrencia a revistas de pago con procesos de publicación extremadamente rápidos, o revistas fuera del campo odontológico, e incluso el plagio por parte de revisores de revistas científicas **(**Fire & Guestrin, 2019**)**.

Entonces, es natural que muchos clínicos se pregunten: ¿cómo puedo identificar la producción científica de alta calidad que respalde mi práctica clínica en la actualidad? La respuesta

© Álvaro Limones Burgos, 2024

reside en la OBE, avalada por reputados grupos de investigación o prestigiosas instituciones como la Colaboración Cochrane, el *International Team of Implantology* (ITI), o la *European Federation of Periodontology* (EFP) entre otras, que actúan como un filtro a través del cual se examina la ciencia y se vislumbran conclusiones comprensibles tanto para los entornos clínicos cotidianos como para los pacientes, y permiten también la toma de decisiones políticas en materia sanitaria.

No obstante, a pesar de la aplicación rigurosa de metodologías sistemáticas y explícitas en los procesos de síntesis de la evidencia, es posible que surjan limitaciones extrínsecas al proceso de revisión, relacionadas con los estudios incluidos como: la existencia de heterogeneidad, riesgo de sesgo/publicación, insuficiente número de participantes, etc. Es decir, no son sólo los métodos que usamos de síntesis, sino también el sustrato del que disponemos. Estas limitaciones podrían dificultar la formulación de recomendaciones fuertes tanto a favor como en contra de una intervención específica.

Por esta razón, la presente disertación de Tesis Doctoral tendrá como unidad de análisis la RS, abordando minuciosamente las limitaciones que comprometen el desarrollo de recomendaciones fuertes en GPC, en un proceso de transición de la evidencia a la práctica clínica. Se llevará a cabo desde un enfoque dual: teórico y práctico, con el apoyo de los siguientes artículos publicados en revistas científicas indexadas en *Journal Citation Reports* (JCR) entre los años 2020 a 2023 en las categorías de "*DENTISTRY, ORAL SURGERY & MEDICINE" y "MEDICINE, GENERAL & INTERNAL"*.

1er Artículo: Limones, A., Molinero-Mourelle, P., Azevedo, L., Romeo-Rubio, M., Correia, A., & Gómez-Polo, M. (2020). Zirconia-ceramic versus metal-ceramic posterior multiunit tooth-supported fixed dental prostheses: A systematic review and meta-analysis of randomized controlled trials. *J Am Dent Assoc*, 151(4), 230-238.e237.

2do Artículo: Limones, A., Saez-Alcaide, L. M., Diaz-Parreno, S. A., Helm, A., Bornstein, M. M., & Molinero-Mourelle, P. (2020). Medication-related osteonecrosis of the jaws (MRONJ) in cancer patients treated with denosumab VS. zoledronic acid: A systematic review and meta-analysis. *Med Oral Patol Oral Cir Bucal*, *25*(3), e326-e336.

© Álvaro Limones Burgos, 2024

3er Artículo: Wilensky A, Shapira L, Limones A, Martin C. (2023). The efficacy of implant surface decontamination using chemicals during surgical treatment of peri-implantitis: A systematic review and meta-analysis. *Journal of Clinical Periodontology*, 50 (26), 336-58.

4er Artículo: Martin, C., Littlewood, S. J., Millett, D. T., Doubleday, B., Bearn, D., Worthington, H. V., & Limones, A. (2023). Retention procedures for stabilising tooth position after treatment with orthodontic braces. *Cochrane Database Syst Rev*, *5*(5), Cd002283.

5to Artículo: Limones, A., Celemín-Viñuela, A., Romeo-Rubio, M., Castillo-Oyagüe, R., Gómez-Polo, M., & Martínez Vázquez de Parga, J. A. (2022). Outcome measurements and quality of randomized controlled clinical trials of tooth-supported fixed dental prostheses: A systematic review and qualitative analysis. *J Prosthet Dent*. Published online ahead of print, 2022 Sep.

© Álvaro Limones Burgos, 2024

4. OBJETIVOS

Los objetivos de la presente Tesis Doctoral fueron fundamentalmente dos:

4.1. Objetivo primario de carácter general en Odontología

Analizar críticamente las limitaciones de las revisiones sistemáticas en Odontología, a través de un enfoque dual: teórico y práctico, a partir de una revisión del estado del arte, y 4 revisiones sistemáticas originales en contextos reales de diferentes áreas temáticas odontológicas (prótesis, medicina oral, implantología y ortodoncia); y su repercusión en el desarrollo de recomendaciones fuertes a través de guías de práctica clínica (*1º, 2º, 3º, 4º artículos*).

4.2. Objetivo secundario de carácter específico en el campo de la Odontología Conservadora y la Prótesis Bucofacial

Analizar críticamente la investigación clínica publicada en prótesis fija dentosoportada mediante la evaluación de la calidad, consistencia metodológica y representatividad clínica de los ensayos clínicos controlados y aleatorizados publicados en los últimos 10 años en el área temática, mediante una RS de instrumentos y variables de medición (*5º artículo*).

© Álvaro Limones Burgos, 2024

© Álvaro Limones Burgos, 2024

5. HIPÓTESIS

Las hipótesis de la presente Tesis Doctoral fueron:

5.1. Hipótesis nula y alternativa primaria de carácter general en Odontología

La hipótesis nula primaria de esta Tesis Doctoral fue:

HØ$_1$: Las revisiones sistemáticas no generan un sustrato adecuado para la consecución de recomendaciones fuertes a través de guías de práctica clínica, cuando presentan limitaciones en su metodología o en los estudios primarios incluidos

(*1º, 2º, 3º, 4º artículo).*

La hipótesis alternativa primaria de esta Tesis Doctoral fue:

HA$_1$: Las revisiones sistemáticas generan un sustrato adecuado para la consecución de recomendaciones fuertes a través de guías de práctica clínica, cuando carecen de limitaciones en su metodología o en los estudios primarios incluidos

(*1º, 2º, 3º, 4º artículo).*

5.2. Hipótesis nula y alternativa secundaria de carácter específico en el campo de la Odontología Conservadora y Prótesis Bucofacial

La hipótesis nula secundaria en el campo de la Odontología Conservadora y Prótesis Bucofacial fue:

HØ$_2$: Los ensayos clínicos en prótesis fija dentosoportada no son de alta calidad metodológica, consistentes en cuanto a sus variables e instrumentos de medición mediante una metodología estandarizada, o representativos de la práctica clínica diaria.

(5º artículo)

La hipótesis alternativa secundaria en el campo de la Odontología Conservadora y Prótesis Bucofacial fue:

HA$_2$: Los ensayos clínicos en prótesis fija dentosoportada son de alta calidad metodológica, consistentes en cuanto a sus variables e instrumentos de medición mediante una metodología estandarizada, o representativos de la práctica clínica diaria.

(5º artículo)

© Álvaro Limones Burgos, 2024

© Álvaro Limones Burgos, 2024

6. MATERIAL Y MÉTODO

La disertación de la presente Tesis Doctoral se constituyó a partir de una:

6.1. Componente Teórica – Revisión del estado del arte:

Fundamentada en:

6.1.1. Una revisión del estado del arte:

Con el propósito de actualizar y analizar de manera conceptual las limitaciones de las RS que obstaculizan la consecución de recomendaciones fuertes a través de GPC. La pregunta de investigación establecida fue: ¿cuáles son y cómo afectan las limitaciones de las RS al desarrollo de recomendaciones fuertes a través de GPC en Odontología?

6.1.1.1. Criterios de inclusión

Artículos de consenso contemporáneos o revisiones sistemáticas que aborden metodologías actuales, limitaciones y líneas futuras de la Práctica Basada en la Evidencia (PBE).

6.1.1.2. Criterio de exclusión

Otros diseños de estudio a los expuestos en los criterios de inclusión como ensayos clínicos, estudios observacionales, series de casos o casos clínicos.

6.1.1.3. Estrategias de búsqueda

Se llevaron a cabo cinco búsquedas electrónicas en tres bases de datos diferentes: MEDLINE/Pubmed, Embase y SCOPUS para abordar la pregunta de investigación de esta revisión del estado del arte. Las estrategias específicas utilizadas se detallan en la Tabla 13. No se incluyó la base de datos *Cochrane Library*, dado que su enfoque en revisiones sistemáticas de salud no coincide con los objetivos y criterios de inclusión de nuestra revisión.

© Álvaro Limones Burgos, 2024

La búsqueda fue completada revisando las referencias de los artículos incluidos y utilizando sistemas de alerta en todas las bases de datos para identificar nuevas referencias hasta el 30 de Abril de 2024.

Tabla 13. Estrategia de búsqueda de la revisión del estado del arte. Los términos de búsqueda se ha preparado de acuerdo con la sintaxis especifica de cada base de datos.

Primera búsqueda:	**Objetivo: localizar literatura relacionada con Prácticas Basadas en la Evidencia**
Bases de datos:	**Términos de búsqueda:**
MEDLINE/Pubmed:	("**Evidence-Based Dentistry**"[Title] OR "**Evidence-Based Practice**"[Title] OR "**Evidence-Based Medicine**"[Title] OR "**Evidence Gaps**"[Title] OR "**Practice Guidelines**"[Title] OR "**GRADE approach**"[Title]) AND (Limitation*[Title] OR Advantage*[Title] OR misunderstandings[Title] OR "**Artificial intelligence**"[Title]) *NOTA: En negrita se marcan los términos MeSH.*
Embase:	('evidence-based dentistry':ti OR 'evidence-based practice':ti OR 'evidence-based medicine':ti OR 'evidence gaps':ti OR 'practice guidelines':ti OR 'grade approach':ti) AND (limitation*:ti OR advantage*:ti OR misunderstandings:ti OR 'artificial intelligence':ti)
SCOPUS	TITLE ("Evidence-Based Dentistry" OR "Evidence-Based Practice" OR "Evidence-Based Medicine" OR "Evidence Gaps" OR "Practice Guidelines" OR "GRADE approach") AND TITLE ("Limitation*" OR "Advantage*" OR "misunderstandings" OR "Artificial intelligence")
Segunda búsqueda	**Objetivo: localizar literatura relacionada con los dominios modificadores de la certeza de la evidencia.**
Bases de datos:	**Términos de búsqueda:**
MEDLINE/Pubmed:	("**GRADE approach**"[Title] OR GRADE[Title]) AND (Certainty[Title] OR Imprecision[Title] OR Inconsistency[Title] OR "Risk of bias"[Title] OR Indirectness[Title] OR "**Publication bias**"[Title] OR "Dose-response gradient"[Title] OR "Large magnitude of an effect"[Title] OR "Effect of plausible residual confounding"[Title]) *NOTA: En negrita se marcan los términos MeSH.*
Embase:	('grade approach':ti OR grade:ti) AND (certainty:ti OR imprecision:ti OR inconsistency:ti OR 'risk of bias':ti OR indirectness:ti OR 'publication bias':ti OR 'dose-response gradient':ti OR 'large magnitude of an effect':ti OR 'effect of plausible residual confounding':ti)
SCOPUS	TITLE ("GRADE approach" OR grade) AND TITLE (certainty OR imprecision OR inconsistency OR "Risk of bias" OR indirectness OR "Publication bias" OR "Dose-response gradient" OR "Large magnitude of an effect" OR "Effect of plausible residual confounding"

© Álvaro Limones Burgos, 2024

Tercera búsqueda:	**Objetivo: localizar literatura relacionada con limitaciones, ventajas, nuevos enfoques y uso de la inteligencia artificial en la confección de revisiones sistemáticas.**
Bases de datos:	**Términos de búsqueda:**
MEDLINE/Pubmed:	((“**Systematic review***”[Title]) AND (Problem*[Title] OR Limitation*[Title] OR Advantage*[Title] OR "Artificial intelligence"[Title])) AND ("Journal of Clinical Epidemiology"[Journal] OR "Systemtic reviews"[Journal] OR "Statistical Methods in Medical Research"[Journal] OR "Clinical Trials"[Journal] OR "Evidence-Based Dentistry"[Journal]) *NOTA: En negrita se marcan los términos MeSH.*
Embase:	(“systematic review*”:ti) AND (problem*:ti OR limitation*:ti OR advantage*:ti OR 'artificial intelligence':ti) AND ('journal of clinical epidemiology':jt OR 'systemtic reviews':jt OR 'statistical methods in medical research':jt OR 'clinical trials':jt OR 'evidence-based dentistry':jt)
SCOPUS	TITLE ("Systematic review*") AND TITLE (problem* OR limitation* OR advantage* OR "Artificial intelligence") AND SRCTITLE ("Journal of Clinical Epidemiology" OR "Systemtic reviews" OR "Statistical Methods in Medical Research" OR "Clinical Trials" OR "Evidence-Based Dentistry")
Cuarta búsqueda	**Objetivo: localizar herramientas de evaluación y guías para la confección de revisiones sistemáticas.**
Bases de datos:	**Términos de búsqueda:**
MEDLINE/Pubmed:	(PRISMA[Title] OR AMSTAR[Title] OR ROBIS[Title]) AND ("Journal of Clinical Epidemiology"[Journal] OR "Systemtic reviews"[Journal] OR "Statistical Methods in Medical Research"[Journal] OR "Clinical Trials"[Journal] OR "Evidence-Based Dentistry"[Journal])
Embase:	(prisma:ti OR amstar:ti OR robis:ti) AND ('journal of clinical epidemiology':ti OR 'systematic reviews':jt OR 'statistical methods in medical research':jt OR 'clinical trials':jt OR 'evidence-based dentistry':jt)
SCOPUS	TITLE (PRISMA OR AMSTAR OR ROBIS) AND SRCTITLE ("Journal of Clinical Epidemiology" OR "Systemtic reviews" OR "Statistical Methods in Medical Research" OR "Clinical Trials" OR "Evidence-Based Dentistry")
Quinta búsqueda	**Objetivo: localizar limitaciones, conceptos y desafíos para el desarrollo de guías de práctica clínica.**
Bases de datos:	**Términos de búsqueda:**
MEDLINE/Pubmed:	"Practice guideline*"[Title] AND (limitation*[Title] OR Concept*[Title] OR Challenge*[Title])
Embase:	'practice guidelines':ti AND (limitation*:ti OR concept*:ti OR challenge*:ti)
SCOPUS	TITLE ("Practice guideline*") AND TITLE (limitation* OR concept* OR challenge*)

© Álvaro Limones Burgos, 2024

6.1.1.4. Extracción de datos

La extracción de datos fue llevada a cabo por dos revisores (A.L.B. & M.C.M.A.). Se extrajeron conclusiones descriptivas de cada artículo incluido.

6.1.1.5. Método de síntesis

Se empleó un método de síntesis narrativa para identificar áreas de consenso, tendencias emergentes y discrepancia entre los artículos revisados, ofreciendo una comprensión profunda de la evolución y el estado actual de la OBE.

© Álvaro Limones Burgos, 2024

6.2. Componente Práctica – Revisiones sistemáticas publicadas:

Fundamentada en:

6.2.1. Cuatro revisiones sistemáticas originales en diversas disciplinas odontológicas:

Con el propósito de analizar de manera práctica**,** de principio a fin y en contextos reales, las limitaciones de las RS en diversas disciplinas odontológicas que obstaculizan la consecución de recomendaciones fuertes a través de GPC.

Para el análisis crítico de las limitaciones de las RS en contextos reales, se utilizaron las siguientes herramientas de evaluación y guías de reporte, sobre las RS elaboradas y publicadas:

- **AMSTAR-II**: herramienta de evaluación metodológica (Shea et al., 2017).
- **ROBIS**: herramienta de evaluación del riesgo de sesgo (Whiting et al., 2016).
- **PRISMA**: comprobación de la guía de reporte (Moher et al., 2015; Page et al., 2021).
- **GRADE**: herramienta de evaluación de la certeza de la evidencia (Hultcrantz et al., 2017).

Posteriormente, se contrastaron las limitaciones identificadas en la componente práctica con los factores que influyen en la fuerza de la recomendación en GPC.

Cada RS que conforma la presente disertación se centró en una disciplina odontológica particular (medicina oral, prostodoncia, implantología y ortodoncia). La metodología empleada en cada una de ellas se encuentra detallada en los artículos publicados (Limones et al., 2022; Limones, Molinero-Mourelle, et al., 2020; Limones, Saez-Alcaide, et al., 2020; Martin et al., 2023; Wilensky et al., 2023) y adjuntos en los ANEXOS del I al IV. El objetivo de cada RS consistió en responder a las siguientes preguntas clínicas de investigación, de acuerdo con el formato PICO[S][T] (por sus siglas en inglés, *Population, Intervention, Comparator, Outcomes, Study, Time*):

PRIMERA REVISIÓN SÍSTEMÁTICA en MEDICINA ORAL

En pacientes oncológicos bajo terapia antiresortiva [*Población*], ¿presenta el denosumab [*Intervención*] comparado con el ácido zoledrónico [*Comparador*] una incidencia o resolución diferentes de osteonecrosis de los maxilares relacionada con la

© Álvaro Limones Burgos, 2024

medicación *[Variables – Outcomes]* en ensayos clínicos controlados aleatorizados *[Estudio]*?

SEGUNDO REVISIÓN SÍSTEMÁTICA en PROSTODONCIA

En pacientes parcialmente desdentados rehabilitados con prótesis fija dentosoportada *[Población]*, ¿hay diferencias entre el material de restauración circonio-cerámica *[Intervención]* comparado con el material de restauración metal-cerámica *[Comparador]* al evaluarse las tasas de éxito, supervivencia, complicaciones biológicas y mecánicas *[Variables – Outcomes]* en ensayos clínicos controlados aleatorizados *[Estudio]*?

TERCER REVISIÓN SÍSTEMÁTICA en IMPLANTOLOGÍA

En pacientes adultos con periimplantitis *[Población]*, ¿el tratamiento quirúrgico junto con procedimientos de descontaminación química de la superficie de los implantes *[Intervención]* en comparación con el tratamiento quirúrgico solo o junto con placebo *[Comparador]*, presenta diferencias al evaluar parámetros como la pérdida ósea marginal, la profundidad de sondaje o el sangrado/supuración al sondaje *[Variables – Outcomes]*, en ensayos clínicos controlados aleatorizados (ECA) *[Estudio]* y no aleatroizados con al menos 6 meses de seguimiento *[Tiempo]*?

CUARTA REVISIÓN SÍSTEMÁTICA en ORTODONCIA

En pacientes tratados con ortodoncia mediante aparatología fija *[Población]*, ¿hay diferencias entre los sistemas de retención removibles *[Intervención]* comparados con los sistemas de retención fija *[Comparador]* al evaluarse parámetros de estabilidad, ¿satisfacción del paciente, salud oral y efectos adversos *[Variables – Outcomes]* en ensayos clínicos controlados aleatorizados *[Estudio]*?

© Álvaro Limones Burgos, 2024

6.2.2. Una revisión sistemática de calidad, variables e instrumentos de medición en el campo de la Odontología Conservadora y Prótesis Bucofacial:

Con el propósito de analizar críticamente la calidad y la consistencia de los ensayos clínicos controlados y aleatorizados publicados en la última década e informar de las limitaciones presentes en los estudios primarios. La metodología empleada se encuentra detallada en el artículo publicado (Limones et al., 2022) y adjunto en el ANEXO V. El objetivo de esta RS consistió en responder a la siguiente pregunta de investigación:

QUINTA REVISIÓN SISTEMÁTICA

Los ensayos clínicos controlados y aleatorizados sobre prótesis fija dentosoportada *[Población]*, ¿siguen las guías actuales, son de alta calidad metodológica, representativos de la práctica clínica y consistentes en sus instrumentos y variables de medición *[Variables – Outcomes]*?

6.3. Recursos:

Para la consecución de la presente Tesis Doctoral se emplearon los siguientes recursos:

- Programa de edición de texto: Microsoft Word (v.2021, Microsoft Corporation, Estados Unidos).
- Hojas de cálculo: Microsoft Excel (v.2021, Microsoft Corporation, Estados Unidos).
- Programas de edición: Keynote (v.11.1, Apple Inc., Estados Unidos).¡
- Bases de datos electrónicas: Embase (Elsevier, Países Bajos), Medline (Biblioteca Nacional de Medicina de EE. UU.), Cochrane Library (Cochrane Collaboration), WOS (Clarivate Analytics) y SCOPUS (Elsevier, Países Bajos).
- Plataformas de registro internacional: PROSPERO (York, Reino Unido), US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov), World

© Álvaro Limones Burgos, 2024

Health Organization International Clinical Trials Registry Platform (WHO-ICTRP), Open Science Framework (Estados Unidos).

- Programas de revisión por pares: Covidence (Veritas Health Innovation, Australia), DistillerSR (Evidence Partners, Canadá).
- Softwares estadísticos: Comprehensive Meta-Analysis (v.3, Biostat, Estados Unidos), STATA (v. 17, StataCorp LLC, Estados Unidos), RevMan 5 (v.5.4, Cochrane Collaboration), Mix (BiostatXL).
- Gestores bibliográficos: EndNote (v.21, Clarivate Analytics).
- Biblioteca, hemeroteca y mediateca: Facultad de Odontología de la Universidad Complutense de Madrid.

© Álvaro Limones Burgos, 2024

7. RESULTADOS

7.1. Componente Teórica – Revisión del estado del arte

Se incluyeron un total de 39 artículos en la componente teórica, cuyas características se presentan en la Tabla 14.

Tabla 14. Características de los artículos incluidos.

Limitaciones intrínsecas de las RS				
Autor (Año)	**Diseño de estudio**	**Análisis descriptivo**		
Uttley et al. (2023)	RS a tiempo real	Analiza de forma continuada en el tiempo las limitaciones intrínsecas de las RS		
Limitaciones extrínsecas de las RS – enfoque GRADE				
Autor (Año)	**Diseño de estudio**	**Temática & Dominio**	**Guía**	**Análisis descriptivo**
Guyatt, Oxman, Akl, et al. (2011)	Consenso de expertos	Introducción	1	Introducción al enfoque GRADE y Tablas Resumen de Hallazgos.
Guyatt, Oxman, Kunz, Atkins, et al. (2011)	Consenso de expertos	Pregunta de investigación Selección de variables críticas	2	Enfoque de la pregunta de investigación y selección de variables críticas.
Schünemann, Mustafa, et al. (2019)	Consenso de expertos	Selección de variables críticas	22	Consideración de variables críticas sobre las perspectivas de los pacientes.
Balshem et al. (2011)	Consenso de expertos	Evaluación de la evidencia	3	Evaluación de la certeza de la evidencia.
(Guyatt, Oxman, Sultan, et al., 2013)	Consenso de expertos	Evaluación de la evidencia	11	Valoración general de la confianza en las estimaciones de efecto.
Zeng et al. (2021)	Consenso de expertos	Evaluación de la evidencia	32	Objetivos para la calificación de la certeza de la evidencia.

© Álvaro Limones Burgos, 2024

Brunetti et al. (2013)	Consenso de expertos	Evaluación de la evidencia	10	Evaluación de la certeza de la evidencia económica y recursos.
Xie et al. (2023)	Consenso de expertos	Evaluación de la evidencia	23	Evaluación de la certeza de la evidencia sobre el coste-efectividad de las intervenciones.
Foroutan et al. (2020)	Consenso de expertos	Evaluación de la evidencia	28	Evaluación de la certeza de la evidencia en RS de factores pronósticos.
Goldkuhle et al. (2021)	Consenso de expertos	Evaluación de la evidencia	29	Evaluación de la certeza de la evidencia de variables de tiempo hasta el evento (limitaciones).
Yang et al. (2021)	Consenso de expertos	Evaluación de la evidencia	31	Evaluación de la certeza de la evidencia en RS de precisión de pruebas diagnósticas.
Zeng et al. (2021)	Consenso de expertos	Evaluación de la evidencia	32	Objetivos para calificación de la certeza de la evidencia.
Guyatt, Oxman, Vist, et al. (2011)	Consenso de expertos	Riesgo de sesgo	4	Evaluación del riesgo de sesgo.
Guyatt et al. (2017)	Consenso de expertos	Riesgo de sesgo	17	Actualización en la evaluación del riesgo de sesgo por datos perdidos.
Schünemann, Cuello, et al. (2019)	Consenso de expertos	Riesgo de sesgo	18	Actualización en la evaluación del riesgo de sesgo de ensayos no aleatorizados.
Zhang, Alonso-Coello, et al. (2019)	Consenso de expertos	Riesgo de sesgo Carácter indirecto de la evidencia	19	Actualización en la evaluación del riesgo de sesgo y el carácter indirecto de la evidencia.
(Schünemann et al., 2020a, 2020b)	Consenso de expertos	Riesgo de sesgo Carácter indirecto de la evidencia	21	Riesgo de sesgo y carácter indirecto en RS de precisión de pruebas diagnósticas.
Guyatt, Oxman, Montori, et al. (2011)	Consenso de expertos	Sesgo de publicación	5	Evaluación del sesgo de publicación.
Guyatt, Oxman, Kunz, Brozek, et al. (2011)	Consenso de expertos	Imprecisión	6	Evaluación de la imprecisión.
Zhang, Coello, et al. (2019)	Consenso de expertos	Imprecisión Inconsistencia	20	Actualización en la evaluación de la inconsistencia, imprecisión y otros dominios.
Brignardello-Petersen et al. (2021)	Consenso de expertos	Imprecisión	33	Evaluación de la imprecisión en metaanálisis en red.
Zeng et al. (2022)	Consenso de expertos	Imprecisión	34	Actualización sobre la evaluación de la imprecisión utilizando un enfoque mínimamente contextualizado.
Schünemann et al. (2022)	Consenso de expertos	Imprecisión	35	Actualización sobre la evaluación de la imprecisión contextualizada y toma de decisiones.

© Álvaro Limones Burgos, 2024

Mustafa et al. (2024)	Consenso de expertos	Imprecisión	37	Evaluación de la imprecisión en RS de precisión de pruebas diagnósticas.
Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Glasziou, et al. (2011)	Consenso de expertos	Inconsistencia	7	Evaluación de las inconsistencias.
Guyatt et al. (2023)	Consenso de expertos	Inconsistencia	36	Actualización del enfoque GRADE para evaluar la inconsistencia.
Cuello-Garcia et al. (2022)	Consenso de expertos	Inconsistencia	24	Integración de estudios aleatorizados y no aleatorizados de intervenciones.
Guyatt, Oxman, Kunz, Woodcock, Brozek, Helfand, Alonso-Coello, Falck-Ytter, et al. (2011)	Consenso de expertos	Carácter indirecto de la evidencia	8	Evaluación del carácter indirecto de la evidencia.
Guyatt, Oxman, Sultan, et al. (2011)	Consenso de expertos	Motivos de aumento de la evidencia	9	Razones de aumento de la certeza de la evidencia.
Murad et al. (2023)	Consenso de expertos	Gradiente dosis-respuesta	38	Guía actualizada para aumentar la certeza de la evidencia debido a un gradiente de dosis-respuesta.
Santesso et al. (2020)	Consenso de expertos	Presentación de hallazgos	26	Declaraciones informativas para comunicar los hallazgos de RS de intervenciones.
Guyatt, Oxman, Santesso, et al. (2013)	Consenso de expertos	Presentación de hallazgos	12	Tablas Resumen de Hallazgos para variables binarias.
Guyatt, Thorlund, et al. (2013)	Consenso de expertos	Presentación de hallazgos	13	Tablas Resumen de Hallazgos para variables continuas.
Skoetz et al. (2020)	Consenso de expertos	Presentación de hallazgos	27	Tablas Resumen de Hallazgos para variables de tiempo hasta el evento
Implicaciones en las guías de práctica clínica - enfoque GRADE				
Autor (Año)	**Diseño de estudio**	**Temática & Dominio**	**Guía**	**Análisis descriptivo**
J. Andrews et al. (2013)	Consenso de expertos	/	14	De la evidencia a la recomendación – presentación de recomendaciones.
J. C. Andrews et al. (2013)	Consenso de expertos	/	15	Determinantes de la dirección y la fuerza de las recomendaciones.
Schünemann et al. (2016)	Consenso de expertos	/	16	Marcos de la EtD.
Brozek et al. (2021)	Consenso de expertos	/	30	Visión general de la toma de decisiones en salud.

© Álvaro Limones Burgos, 2024

De acuerdo con la literatura revisada, clasificamos las limitaciones de las RS en:

7.1.1. Limitaciones intrínsecas al propio proceso de revisión, por la calidad metodológica

Los errores metodológicos son solventables y se evalúan mediante las herramientas AMSTAR-II y ROBINS. Se subclasifican en los siguientes dominios de acuerdo con Uttley et al. (2023):

- **Exhaustividad**

  Debe incluir todos los estudios pertinentes. Los problemas surgen cuando se omiten o ignoran estudios relevantes, lo que puede limitar la validez de la RS. Esto puede ocurrir por los siguientes motivos:

  - Criterios de inclusión demasiado estrictos que afectan a la validez externa.
  - Errores u omisiones en la estrategia de búsqueda.
  - Literatura gris excluida.
  - Búsquedas bibliográficas insuficientes.
  - Falta de búsquedas complementarias más allá de las bases de datos.
  - Restricciones lingüísticas.
  - Búsquedas obsoletas.
  - Errores en la inclusión de estudios u omisión de estudios relevantes.
  - Dependencia de los ensayos controlados aleatorizados para los datos de daños/seguridad.
  - Escasa consideración del sesgo de publicación.
  - No se actualiza regularmente.
  - Revisión redundante/solapada/duplicada; conduce al desperdicio de la investigación.
  - Inoportuna (lleva demasiado tiempo) o consume muchos recursos.

- **Rigurosidad**

  Se deben utilizar métodos adecuados para garantizar solidez metodológica. Errores en el enfoque de la revisión o la falta de experiencia pueden poner en peligro la validez interna de la RS. Esto puede ocurrir por los siguientes motivos:

© Álvaro Limones Burgos, 2024

- Intervención no definida correctamente.
- Errores en la extracción de datos.
- No evaluar los conflictos de intereses o la financiación de los estudios primarios.
- Evaluación limitada o ausencia de evaluación del riesgo de sesgo.
- Errores en los cálculos de las estimaciones de los efectos o en la síntesis de los datos.
- Análisis inadecuado de la heterogeneidad.
- Interpretación incorrecta o error de diferencia estadística del metaanálisis.
- Ejecución deficiente de la síntesis narrativa.
- Riesgo de sesgo no incorporado en las conclusiones de la revisión.
- Número reducido de ensayos en los metaanálisis.
- Análisis de sensibilidad o de subgrupos no planificados o injustificados.
- Falta de orientación o coherencia en la revisión.
- Baja calidad metodológica (AMSTAR II).

- **Transparencia**

  Se debe informar de los métodos con suficiente detalle para permitir la reproducción. Si la información es deficiente o los métodos son inaccesibles, entonces las RS no es reproducible. Esto puede ocurrir por los siguientes motivos:

  - Falta de pre-especificación en los criterios de elegibilidad.
  - Métodos no descritos para permitir la replicación.
  - Ausencia de protocolo registrado o publicado.
  - Desviaciones no documentadas o injustificadas del protocolo de revisión.
  - Falta la declaración de conflicto de intereses o la divulgación de los autores de la revisión.
  - No se informa sobre la financiación o el patrocinador de la RS.
  - No se proporciona la estrategia de búsqueda.
  - No se informa con suficiente detalle las características individuales de los estudios.

© Álvaro Limones Burgos, 2024

- No se proporcionan las razones para excluir estudios potencialmente elegibles.
- Notificación selectiva de efectos adversos.
- Poco manejables / difíciles de leer.
- Informes de baja calidad (PRISMA).

- **Objetividad**

  Debe realizarse de forma imparcial, por equipos con experiencia adecuada y no por personas con intereses predefinidos. Esto puede ocurrir por los siguientes motivos:

  - No tener en cuenta la equidad, los diferentes grupos socioeconómicos o las poblaciones desfavorecidas.
  - Falta de definición de resultados clínicamente significativos.
  - Falta de perspectiva de expertos clínicos/partes interesadas/pacientes.
  - Pregunta de revisión no justificada/irrelevante.
  - Búsquedas bibliográficas no validadas por especialistas en información.
  - Un único revisor/falta de doble comprobación.
  - Falta de experiencia estadística en el manejo de datos cuantitativos.
  - Interpretación hilada o subjetiva de los resultados.
  - Interpretados sin tener en cuenta la certeza de la evidencia científica.
  - Conflictos de intereses económicos de los revisores.
  - Conflictos de intereses no financieros de los revisores.
  - Autoría invitada/regalo.
  - Alto riesgo de sesgo.

7.1.2. Limitaciones extrínsecas al proceso de revisión, por la evidencia disponible

No son solventables en el momento de la confección la revisión y son directamente proporcionales a la cantidad y calidad de los estudios incluidos. Las limitaciones

© Álvaro Limones Burgos, 2024

extrínsecas que degradan la evidencia científica se pueden identificar mediante el sistema GRADE valorando:

- **Dominio 1: Imprecisión**
- **Dominio 2: Carácter indirecto, del inglés *Indirectness***
- **Dominio 3: Inconsistencia**
- **Dominio 4: Riesgo de sesgo**
- **Dominio 5: Riesgo de publicación**

Estos dominios han sido explicados en la sección de Introducción de la presente Tesis Doctoral (Apartado 2.4.3.3.)

© Álvaro Limones Burgos, 2024

7.2. Componente Práctica – Revisiones sistemáticas publicadas

Se realizaron cinco RS originales en distintas especialidades odontológicas, con el objetivo de identificar de manera práctica y en contextos reales las limitaciones intrínsecas y extrínsecas:

7.2.1. PRIMERA REVISIÓN SISTEMÁTICA

(Limones, Saez-Alcaide, et al., 2020)

El manuscrito publicado se encuentra en el ANEXO I. Su objetivo primario fue responder a la siguiente pregunta de investigación: *En pacientes oncológicos bajo terapia antiresortiva, ¿presenta el denosumab comparado con el ácido zoledrónico una incidencia diferente de osteonecrosis de los maxilares relacionada con la medicación según se informa en ensayos clínicos controlados y aleatorizados?* Para analizar críticamente las limitaciones de esta RS, se utilizaron las siguientes herramientas, cuyos resultados se presentan a continuación:

7.2.1.1. AMSTAR-II: Resultados de la evaluación metodológica

De acuerdo con la herramienta AMSTAR-II, esta RS concluyó ser de BAJA confianza. Esto se debió al no haberse publicado un protocolo previo al comienzo del estudio en una plataforma de registro internacional accesible para contrastar posibles desviaciones del protocolo (aspecto crítico) (Tabla del ANEXO 1.1). Otra limitación, en este caso, parcial, fue la presentación de las justificaciones de los estudios excluidos dentro del diagrama de flujo acompañada de una descripción en la sección de resultados, en lugar de proporcionar un documento explicito como material suplementario.

© Álvaro Limones Burgos, 2024

7.2.1.2. ROBIS: Resultados de la evaluación del riesgo de sesgo

De acuerdo con la herramienta ROBIS, esta RS se clasificó como de ALTO riesgo de sesgo. Esto se debió a la omisión del registro previo del protocolo en un registro internacional de acceso público. Además, los autores no evitaron destacar los resultados en función de su relevancia estadística, lo que pudo generar sobreestimaciones en las conclusiones. La evaluación detallada del riesgo de sesgo se encuentra en la Tabla del ANEXO 1.2.

7.2.1.3. GRADE: Resultados de la evaluación de la certeza de la evidencia

Se consideró que la certeza de la evidencia para la variable principal (incidencia de osteonecrosis de los maxilares o MRONJ por sus siglas en ingles, *Medication-Related Osteonecrosis of the Jaws*) fue MODERADA. El computo final de la evaluación de los dominios relacionados con el aumento y disminución de la certeza de la evidencia condujo a la decisión de disminuir la certeza en dos niveles debido al riesgo de sesgo e imprecisión. Sin embargo, se incrementó un nivel de certeza basándose en la presencia de un gradiente dosis-respuesta. En definitiva, la certeza de la evidencia se redujo un nivel.

Se consideró que la certeza de la evidencia para la variable secundaria (resolución de la MRONJ) fue MUY BAJA. El tamaño muestral agrupado fue extremadamente bajo en esta variable, lo que obligó a reducir la evidencia 3 niveles.

Las justificaciones detalladas para cada dominio se presentan en la Tabla del ANEXO 1.3 .

7.2.1.4. PRISMA: Comprobación del cumplimiento de la guía de reporte

La comprobación del cumplimiento de la guía de reporte se evaluó a partir de la lista de comprobación propuesta por la última versión de la declaración PRISMA (Moher et al., 2015; Page et al., 2021) y su extensión para resúmenes (*abstracts)* (Beller et al., 2013) en un diseño de tabla combinado y mostrado en la Tabla del ANEXO 1.4 Se detectó un incumplimiento de 63% de los ítems de la extensión para resúmenes y un 21% de los ítems de la lista de comprobación original actualizada, debido a que esta publicación se basó en la declaración PRISMA de 2015 (Moher et al., 2015) y no en la actualizada de 2021 (Page et al., 2021), publicada después de nuestra revisión. A pesar de ello, la comprobación del cumplimiento de la guía de reporte identificó varias limitaciones:

1. Las normas de las revistas científicas cuentan con limitaciones de palabras en los resúmenes, que impiden informar de aspectos requeridos en la extensión PRISMA para

© Álvaro Limones Burgos, 2024

resúmenes; no habiendo informado en el resumen de los criterios de inclusión y exclusión, las bases de datos consultadas, los métodos para evaluar el riesgo de sesgo, los métodos utilizados para presentar y sintetizar los resultados, así como los aspectos que redujeron la certeza de la evidencia o la financiación del estudio.

2. A pesar de no encontrarse heterogeneidad y no ser necesario realizar análisis de la sensibilidad, análisis de subgrupos o metarregresiones, esto no fue mencionado de forma explicita en el manuscrito.
3. No se incluyó una reflexión sobre las limitaciones que pudo tener el proceso de revisión utilizado (i.e. solo se incluyeron estudios en inglés, francés y español).
4. No se informó de un registro previo del protocolo.
5. No se puso a disposición del público las plantillas de extracción de datos o cualquier otro material utilizado en la revisión.

7.2.1.5. Resumen de la evaluación – limitaciones identificadas

A la vista del análisis de los resultados obtenidos tras la aplicación de las herramientas mencionadas anteriormente, y siguiendo la clasificación de las limitaciones presentada en el apartado 7.1 de esta Tesis Doctoral (resultados correspondientes a la componente teórica), dividimos las limitaciones encontradas en intrínsecas y extrínsecas.

Dentro de las limitaciones <u>intrínsecas</u> propias de esta RS destacan:

1. No registrar del protocolo del estudio en un registro internacional de acceso público (signo de rigurosidad).
2. No reflejar la no necesidad de análisis de sensibilidad, análisis de subgrupos o metarregresiones (signo de rigurosidad).
3. No poner a disposición del público las plantillas de extracción de datos o cualquier otro material utilizado en la revisión (signo de transparencia).
4. No discutir las limitaciones que pudo tener la no inclusión de estudios en idiomas distintos al inglés, francés o español (signo de exhaustividad)
5. No definir qué es un resultado clínicamente significativo (signo de objetividad).

Dentro de las limitaciones <u>extrínsecas</u> ajenas a la metodología de esta RS destacan:

© Álvaro Limones Burgos, 2024

1. todos los estudios primarios incluidos estuvieron financiados por la empresa distribuidora y fabricante de la intervención estudiada. Esto representa un riesgo potencial de sesgo y enfatiza la necesidad de manejar con cautela las relaciones contractuales entre investigadores y empresas financiadoras.
2. No se superó el tamaño óptimo de información.

7.2.1.6. Tabla resumen de hallazgos o de la evidencia

Se presentaron los hallazgos obtenidos en esta revisión en la siguiente Tabla resumen de la evidencia siguiendo el formato propuesto por Carrasco-Labra et al. (2016), y la última actualización de Morgan et al. (2023) no seguida en la publicación Limones, Saez-Alcaide, et al. (2020):

Denosumab comparado con Acido Zoledrónico en la incidencia de MRONJ

Población/Participantes: Pacientes oncológicos
Intervención: Denosumab
Comparación: Ácido Zoledrónico
Seguimiento: 2 años

Variables CRÍTICAS Nº de participantes (Estudios)	Efecto relativo (IC 95%)	Efecto absoluto previsto			Número de pacientes con necesidad de tratar (IC 95%)	Certeza de la evidencia	Interpretación
		Ácido Zoledrónico	**Denosumab**	**Diferencia (IC 95%)**			
INCIDENCIA DE OSTEONECROSIS DE LOS MAXILARES INDUCIDA POR MEDICAMENTOS **Nº de participantes**: 11976 (6 ECAs)	**RR 1.448** (1.114 a 1.881)	**13 por 1000**	**18 por 1000**	**4 mas** (1 a 9 más)	**200** (106 a 1530)	⊕⊕⊕◯ MODERADA (Ver ANEXO 1.3)	El denosumab **SI demostró una incidencia superior,** comparado con el ácido zoledrónico.
RESOLUCIÓN DE OSTEONECROSIS DE LOS MAXILARES INDUCIDA POR MEDICAMENTOS **Nº de participantes**: 127 (3 ECAs)	**OR 1.730** (0.801 a 3.739)	**254 por 1000**	**342 por 1000**	**117 mas** (40 a 306 más)	No aplicable	⊕⊕⊕◯ MODERADA (Ver ANEXO 1.3)	La MRONJ producida por denosumab **NO demostró un resolución diferente,** comparado con ácido zoledrónico.

Todas las justificaciones detalladas de la degradación y aumento de la certeza de la evidencia de acuerdo con el enfoque GRADE, se presentaron en el ANEXO 1.3.

© Álvaro Limones Burgos, 2024

7.2.2. SEGUNDA REVISIÓN SISTEMÁTICA

(Limones, Molinero-Mourelle, et al., 2020)

El manuscrito publicado se encuentra en el ANEXO II. Su objetivo primario fue responder a la siguiente pregunta de investigación: *En pacientes parcialmente desdentados rehabilitados con prótesis fija dentosoportada, ¿hay diferencias entre el material de restauración circonio-cerámica comparado con metal-cerámica al evaluarse las tasas de supervivencia, complicaciones mecánicas, biológicas y estéticas según se informa en ensayos clínicos controlados aleatorizados?* Para analizar críticamente las limitaciones de esta RS, se utilizaron las siguientes herramientas, cuyos resultados se presentan a continuación:

7.2.2.1. AMSTAR-II: Resultados de la evaluación metodológica

De acuerdo con la herramienta AMSTAR-II, se considera que es una revisión de MODERADA confianza por la metodología empleada. A pesar de que los autores de la revisión no informaron sobre las fuentes de financiación de los estudios incluidos (aspecto no crítico), se concluye que la revisión puede proporcionar un resumen preciso de los resultados de los estudios disponibles que abordan la pregunta de interés, ya que no se identificaron aspectos críticos que limitasen la confianza en los hallazgos obtenidos en la RS. Otra limitación parcial fue la presentación de las justificaciones de los excluidos dentro del diagrama de flujo acompañada de una descripción en la sección de resultados, en lugar de proporcionar un documento explicito como material suplementario. La evaluación de la calidad metodológica de la publicación de Limones, Molinero-Mourelle, et al. (2020) se presenta en la Tabla del ANEXO 2.1.

7.2.2.2. ROBIS: Resultados de la evaluación del riesgo de sesgo

De acuerdo con la herramienta ROBIS, esta RS se clasificó como de BAJO riesgo de sesgo, no destacándose preocupaciones importantes en ninguna de las fases de evaluación. Dada la consistencia de los hallazgos y el reducido número de estudios incluidos en el metaanálisis (sin resultados atípicos) no fue relevante realizar análisis de la sensibilidad. La heterogeneidad existente en algunas de las variables críticas secundarias fue explorada mediante análisis de

© Álvaro Limones Burgos, 2024

subgrupos. La evaluación detallada del riesgo de sesgo se encuentra en la Tabla del ANEXO 2.2.

7.2.2.3. GRADE: Resultados de la evaluación de la certeza de la evidencia

El computo final de la evaluación de los dominios relacionados con el aumento y disminución de la certeza de la evidencia condujo a la decisión de disminuir la certeza por presencia de carácter indirecto de la evidencia, imprecisión e inconsistencia. Esto se debió a que todos los estudios incluidos informaban de resultados a un tiempo de seguimiento a medio plazo (de 3 a 5 años) y no a largo plazo (más de 10 años) que es el tiempo de seguimiento de relevancia clínica que marca la toma de decisiones en prótesis fija dentosoportada; por la disparidad de instrumentos de evaluación de complicaciones mecánicas (sistema USPHS (por las siglas en inglés *United States Public Health Service*) o CDA (por sus siglas en inglés *California Dental Association*); por no superar el tamaño optimo de información o por el reducido tamaño muestral de algunas de las variables. Más información en el ANEXO 2.3.

7.2.2.4. PRISMA: Comprobación del cumplimiento de la guía de reporte

La comprobación del cumplimiento de la guía de reporte se evaluó a partir de la lista de comprobación propuesta por la última versión de la declaración PRISMA (Moher et al., 2015; Page et al., 2021) y su extensión para resúmenes (*abstracts*) (Beller et al., 2013) en un diseño de tabla combinado y mostrado en la Tabla del ANEXO 2.4. Se detectó un incumplimiento de 36% de los ítems de la extensión para *abstracts* y un 17% de los ítems de la lista de comprobación original actualizada. La comprobación del cumplimiento de la guía de reporte identificó varias limitaciones:

1. No se evaluó en la publicación la certeza de la evidencia científica mediante el enfoque GRADE.
2. No se realizaron análisis de sensibilidad.

7.2.2.5. Resumen de la evaluación – limitaciones identificadas

A la vista del análisis de los resultados obtenidos tras la aplicación de las herramientas mencionadas anteriormente, y siguiendo la clasificación de las limitaciones presentada en el

© Álvaro Limones Burgos, 2024

apartado 7.1 de esta Tesis Doctoral (resultados correspondientes a la componente teórica), dividimos las limitaciones encontradas en intrínsecas y extrínsecas.

Dentro de las limitaciones intrínsecas propias de esta RS destacan:

1. No se realizaron análisis de sensibilidad (signo de rigurosidad).
2. No se evaluó la certeza de la evidencia científica mediante en enfoque GRADE (signo de rigurosidad).
3. No se presentó un listado explicito de las referencias de los estudios excluidos (signo de transparencia).
4. No se definió qué es un resultado clínicamente significativo para cada variable crítica (signo de objetividad).

Dentro de las limitaciones extrínsecas ajenas a la metodología de esta RS destacan:

1. Se encontró heterogeneidad entre las variables e instrumentos de medición de complicaciones mecánicas/técnicas y biológicas.
2. No se superó el tamaño optimo de información en algunas de las variables críticas.
3. No se informó de todas de variables críticas (i.e. sangrado al sondaje).
4. No se registró el protocolo de los estudios primarios con carácter prospectivo.

7.2.2.6. Tabla resumen de hallazgos o de la evidencia

Se presentaron los hallazgos obtenidos en esta revisión en la siguiente Tabla resumen de la evidencia siguiendo el formato propuesto por Carrasco-Labra et al. (2016), y la última actualización de Morgan et al. (2023) no presentada en la publicación Limones, Molinero-Mourelle, et al. (2020):

Circonio-cerámica comparado con metal-cerámica en prótesis fija de múltiples unidades en sector posterior

Población/Participantes: Edéntulos parciales, clase III de Kennedy en sector posterior
Intervención: Prótesis fija de múltiples unidades confeccionada en circonio-cerámica
Comparación: Prótesis fija de múltiples unidades confeccionada en metal-cerámica
Seguimiento: 4.07 años

© Álvaro Limones Burgos, 2024

Variables CRÍTICAS Nº de participantes (Estudios)	Efecto relativo (IC 95%)	Efecto absoluto previsto			Número de pacientes con necesidad de tratar (IC 95%)	Certeza de la evidencia	Interpretación
		Circonio-cerámica	**Metal-cerámica**	**Diferencia (IC 95%)**			
TASA DE SUPERVIVENCIA **Nº de participantes**: 330 (5 ECAs)	-	**79 cada 1000**	**29 cada 1000**	**RD 0.017** (-0.02 a 0.05)	**20** (10 a 308)	⊕⊕⊕◯ MODERADA (Ver ANEXO 2.3)	El circonio-cerámica **NO demostró una supervivencia diferente,** comparado con meta-cerámica.
DESLAMINADO CERÁMICO (Chipping) **Nº de participantes**: 330 (5 ECAs)	-	**590 cada 1000**	**294 cada 1000**	**RD 0.223** (0.03 a 0.416)	**7** (5 a 16)	⊕◯◯◯ MUY BAJA (Ver ANEXO 2.3)	El circonio-cerámica **SI demostró mayor deslaminado cerámico,** comparado con meta-cerámica
PÉRDIDA DE RETENCIÓN **Nº de participantes**: 149 (3 ECAs)	-	**38 cada 1000**	**1 cada 1000**	**RD 0.023** (-0.031 a 0.078)	**54** (17 a 46)	⊕⊕◯◯ BAJA (Ver ANEXO 2.3)	El circonio-cerámica **NO demostró una pérdida de retención diferente,** comparado con meta-cerámica.
PROFUNDIDAD DE SONDAJE **Nº de participantes**: 103 (2 ECAs)	-	**Media 0**	**Media 0**	**MD 0.00** (-0.39 a 0.39)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 2.3)	El circonio-cerámica **NO demostró una profundidad de sondaje diferente,** comparado con meta-cerámica.
SANGRADO AL SONDAJE **Nº de participantes**: 69 (1 ECAs)	-	**Media 32.8**	**Media 29.8**	**MD 3.00** (-9.24 a 15.24)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 2.3)	El circonio-cerámica **NO demostró un sangrado al sondaje diferente,** comparado con meta-cerámica.
CARIES SECUNDARIA **Nº de participantes**: 0 (0 ECAs)	-	-	-	-	-	-	Sin datos.
PÉRDIDA DE VITALIDAD **Nº de participantes**: 183 (4 ECAs)	-	**31 cada 1000**	**12 cada 1000**	**RD 0.20** (-0.03 a 0.07)	**28** (13 a 249)	⊕⊕⊕◯ MODERADA (Ver ANEXO 2.3)	El circonio-cerámica **NO demostró una pérdida de vitalidad diferente,** comparado con meta-cerámica.

Todas las justificaciones detalladas de la degradación de la certeza de la evidencia de acuerdo con el enfoque GRADE, se presentaron en el ANEXO 2.3.

© Álvaro Limones Burgos, 2024

7.2.3. TERCERA REVISIÓN SISTEMÁTICA

(Wilensky et al., 2023)

El manuscrito publicado se encuentra en el ANEXO III. Su objetivo primario fue responder a la siguiente pregunta de investigación: *En pacientes adultos con periimplantitis, ¿cuál es la eficacia del tratamiento quirúrgico de la periimplantitis con coadyuvantes de descontaminación química de las superficies de los implantes en comparación con el tratamiento quirúrgico solo o con placebo, en términos de reducción de la profundidad del sondaje y el sangrado al sondaje según se informa en ensayos clínicos controlados aleatorizados y no aleatorizados (RCTs) con un seguimiento de al menos 6 meses?* Para analizar críticamente las limitaciones de esta RS, se utilizaron las siguientes herramientas, cuyos resultados se presentan a continuación:

7.2.3.1. AMSTAR-II: Resultados de la herramienta de evaluación metodológica

De acuerdo con la herramienta AMSTAR-II, se considera que es una revisión de ALTA confianza por la metodología empleada. Se concluye que la revisión puede proporcionar un resumen preciso y completo de los resultados de los estudios disponibles que abordan la pregunta de interés. La única limitación parcial encontrada fue la limitación de idioma a estudios publicados en inglés por la limitación de tiempo que tenía esta revisión por ser por encargo para el desarrollo de una GPC. La evaluación de la calidad metodológica de la publicación (Wilensky et al., 2023) se presenta en la Tabla del ANEXO 3.1.

7.2.3.2. ROBIS: Resultados de la herramienta de evaluación del riesgo de sesgo

De acuerdo con la herramienta ROBIS, esta RS se clasificó como de INEXISTENTE riesgo de sesgo, no destacándose preocupaciones importantes en ninguna de las fases de evaluación. La evaluación detallada del riesgo de sesgo se encuentra en la Tabla del ANEXO 3.2.

7.2.3.3. GRADE: Resultados de la evaluación de la certeza de la evidencia

Se consideró que la certeza de la evidencia para la mayoría de las variables de esta RS fue MUY BAJA. El computo final de la evaluación de los dominios relacionados con el aumento y

© Álvaro Limones Burgos, 2024

disminución de la certeza de la evidencia condujo a la decisión de disminuir la certeza en tres niveles debido al riesgo de sesgo, imprecisión e inconsistencia. Más información en la Tabla del ANEXO 3.3.

7.2.3.4. PRISMA: Comprobación del cumplimiento de la guía de reporte

La comprobación del cumplimiento de la guía de reporte se evaluó a partir de la lista de comprobación propuesta por la última versión de la declaración PRISMA (Moher et al., 2015; Page et al., 2021) y su extensión para resúmenes (*abstracts)* (Beller et al., 2013) en un diseño de tabla combinado y mostrado en la Tabla del ANEXO 3.4. Se detectó un incumplimiento de 27% de los ítems de la extensión para resúmenes y un 100% de cumplimiento de los ítems de la lista de comprobación original actualizada. La comprobación del cumplimiento de la guía de reporte no identificó ninguna limitación.

7.2.3.5. Resumen de la evaluación – limitaciones identificadas

A la vista del análisis de los resultados obtenidos tras la aplicación de las herramientas mencionadas anteriormente, y siguiendo la clasificación de las limitaciones presentada en el apartado 7.1 de esta Tesis Doctoral (resultados correspondientes a la componente teórica), dividimos las limitaciones encontradas en intrínsecas y extrínsecas.

Dentro de las limitaciones intrínsecas propias de esta RS solo se identificó la restricción de idioma a estudios escritos en inglés por la limitación de tiempo por tratarse de una revisión por encargo. La no existencia de limitaciones pudo deberse a:

- Experiencia de los investigadores.
- Revisión por encargo de un consenso de expertos, quienes desarrollaron el protocolo para la elaboración de la GPC.

Dentro de las limitaciones extrínsecas ajenas a la metodología de esta RS destacan:

1. Riesgos de sesgo en los estudios primarios por falta de cegamiento, reporte selectivo de resultados o presencia de conflictos de interés
2. Un tamaño muestral insuficiente por la complejidad de ejecutar las intervenciones.
3. Presencia de heterogeneidad no explicada.

© Álvaro Limones Burgos, 2024

7.2.3.6. Tabla resumen de hallazgos o de la evidencia

Se presentaron los hallazgos obtenidos en esta revisión en la siguiente Tabla resumen de la evidencia siguiendo el formato propuesto por Carrasco-Labra et al. (2016), y la última actualización de Morgan et al. (2023):

Coadyuvantes químicos vs Placebo en el tratamiento quirúrgico de la periimplantitis

Población/Participantes: Pacientes con implantes diagnosticados con periimplantitis

Intervención: uso de coadyuvantes químicos + tratamiento quirúrgico de la periimplantitis

Comparación: placebo + tratamiento quirúrgico de la periimplantitis

Tiempo de evaluación: 6 meses (excepto en variables fracaso y resolución)

Unidad de evaluación: a nivel de implante

Variables CRÍTICAS Nº de participantes (Estudios)	**Efecto relativo (IC 95%)**	**Efecto absoluto previsto**			**Número de pacientes con necesidad de tratar (IC 95%)**	**Certeza de la evidencia**	**Interpretación**
		Placebo	**Coadjuvantes químicos**	**Diferencia comparativa (IC 95%)**			
PRODUNDIDAD DE SONDAJE [mm] **Nº de participantes**: 376 (6 ECAs)	-	**Media 2.01** de reducción	**Media 2.21** de reducción	**MD 0.27** (-0.29 a 0.82)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 3.3)	Los coadyuvantes químicos **NO demostraron tener efecto** sobre la profundidad de sondaje
SANGRADO AL SONDAJE [%] **Nº de participantes**: 198 (5 ECAs)	-	**Media 25%** de reducción	**Media 36%** de reducción	**MD 7.3%** de menor sangrado al sondaje (1.56 a 13)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 3.3)	Los coadyuvantes químicos **SI demostró una reducción del sangrado al sondaje**, pero con una relevancia clínica cuestionable (<10%).
PERDIDA ÓSEA MARGINAL **Nº de participantes**: 180 (4 ECAs)	-	**Media -0.03** de aumento	**Media 0.40** de reducción	**MD 0.48** (-0.31 a 0.99)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 3.3)	Los coadyuvantes químicos **NO demostraron tener efecto** sobre la pérdida ósea marginal
SUPERVIVIENCIA DEL IMPLANTE **Seguimiento**: 6 + 12 meses **Nº de participantes**: 127 (3 ECAs)	-	**4 por 10000**	**1 por 10000**	**RD -0.03** (-0.09 a 0.03)	**23** (12 a 253)	⊕◯◯◯ MUY BAJA (Ver ANEXO 3.3)	Los coadyuvantes químicos **NO demostraron tener efecto** sobre la supervivencia de los implantes.

© Álvaro Limones Burgos, 2024

RESOLUCIÓN DE LA PATOLOGÍA **Seguimiento**: 12 meses **Nº de participantes**: 274 (3 ECAs)	**RR 0.948** (0.72 a 1.27)	**4 por 1000**	**4 por 1000**	**13 casos menos de resolución** (de 82 menos a 59 más)	No aplicable	⨁◯◯◯ MUY BAJA (Ver ANEXO 3.3)	Los coadyuvantes químicos **NO demostraron tener efecto** sobre la resolución de la periimplantitis.

Todas las justificaciones detalladas de la degradación de la certeza de la evidencia de acuerdo con el enfoque GRADE, se presentaron en el ANEXO 3.3.

© Álvaro Limones Burgos, 2024

7.2.4. CUARTA REVISIÓN SISTEMÁTICA

(Martin et al., 2023)

El manuscrito publicado se encuentra en el ANEXO IV. Su objetivo primario fue responder a la siguiente pregunta de investigación: *En pacientes tratados con ortodoncia mediante aparatología fija, ¿hay diferencias entre los sistemas de retención removibles comparados con los sistemas de retención fija al evaluarse parámetros de estabilidad, satisfacción del paciente, salud oral y efectos adversos según se informa en ensayos clínicos controlados aleatorizados?* Para analizar críticamente las limitaciones de esta RS, se utilizaron las siguientes herramientas, cuyos resultados se presentan a continuación:

7.2.4.1. AMSTAR-II: Resultados de la herramienta de evaluación metodológica

De acuerdo con la herramienta AMSTAR-II, se considera que es una revisión de ALTA confianza por la metodología empleada. Se concluye que la revisión puede proporcionar un resumen preciso y completo de los resultados de los estudios disponibles que abordan la pregunta de interés. La evaluación de la calidad metodológica de la publicación (Martin et al., 2023) se presenta en la Tabla del ANEXO 4.1.

7.2.4.2. ROBIS: Resultados de la herramienta de evaluación del riesgo de sesgo

De acuerdo con la herramienta ROBIS, esta RS se clasificó como de INEXISTENTE riesgo de sesgo, no destacándose preocupaciones importantes en ninguna de las fases de evaluación. La heterogeneidad existente en algunas de las variables secundarias fue explorada mediante análisis de subgrupos. La evaluación detallada del riesgo de sesgo se encuentra en la Tabla del ANEXO 4.2.

7.2.4.3. GRADE: Resultados de la evaluación de la certeza de la evidencia

Se consideró que la certeza de la evidencia para la mayoría de las variables de esta RS fue MUY BAJA. El computo final de la evaluación de los dominios relacionados con el aumento y disminución de la certeza de la evidencia condujo a la decisión de disminuir la certeza en tres niveles debido al riesgo de sesgo, e imprecisión (Tabla del ANEXO 4.3).

© Álvaro Limones Burgos, 2024

7.2.4.4. PRISMA: Comprobación del cumplimiento de la guía de reporte

La comprobación del cumplimiento de la guía de reporte se evaluó a partir de la lista de comprobación propuesta por la última versión de la declaración PRISMA (Moher et al., 2015; Page et al., 2021) y su extensión para resúmenes (*abstracts*) (Beller et al., 2013) en un diseño de tabla combinado y mostrado en la Tabla del ANEXO 4.4. Se detectó un incumplimiento de 18% de los ítems de la extensión para resúmenes y un 100% de cumplimiento de los ítems de la lista de comprobación original actualizada. La comprobación del cumplimiento de la guía de reporte no identificó ninguna limitación.

7.2.4.5. Resumen de la evaluación – limitaciones identificadas

A la vista del análisis de los resultados obtenidos tras la aplicación de las herramientas mencionadas anteriormente, y siguiendo la clasificación de las limitaciones presentada en el apartado 7.1 de esta Tesis Doctoral (resultados correspondientes a la componente teórica), dividimos las limitaciones encontradas en intrínsecas y extrínsecas.

No se detectaron limitaciones intrínsecas propias de esta RS, debido a:

- Experiencia del equipo investigador.
- Revisión actualizada por cuarta vez consecutiva.
- Apoyo de una institución de reconocido prestigio (Cochrane)

Dentro de las limitaciones extrínsecas ajenas a la metodología de esta RS destacan:

1. Riesgos de sesgo en los estudios primarios por falta de cegamiento en la asignación/evaluación o reporte selectivo de resultados.
2. Un tamaño muestral insuficiente por falta de consistencia entre las intervenciones de los estudios, características, tiempos de uso de retención y tiempos de seguimiento.

© Álvaro Limones Burgos, 2024

7.2.4.6. Tabla resumen de hallazgos o de la evidencia

Se presentaron los hallazgos obtenidos en esta revisión en la siguiente Tabla resumen de la evidencia siguiendo el formato propuesto por Carrasco-Labra et al. (2016), y la última actualización de Morgan et al. (2023):

Retención removible vs retención fija tras del tratamiento con aparatología fija

Población/Participantes: Pacientes sometidos a tratamiento ortodoncico con aparatología fija

Intervención: Retenedor removible plástico transparente (uso a tiempo parcial)

Comparación: Retenedor fijo

Tiempo de evaluación: 12 meses

Arcada: Inferior

Variables CRÍTICAS Nº de participantes (Estudios)	**Efecto relativo (IC 95%)**	**Efecto absoluto previsto**			**Número de pacientes con necesidad de tratar (IC 95%)**	**Certeza de la evidencia**	**Interpretación**
		Retenedor fijo	**Retenedor removable**	**Diferencia comparativa (IC 95%)**			
ÍNDICE DE IRREGULARIDAD [mm] **Nº de participantes**: 56 (1 ECAs)	-	**Media 0.77mm** de recidiva	**Media 1.69mm** de recidiva	**MD 0.92mm** más de recidiva de apiñamiento (0.23 a 1.61)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 4.3)	**Ligeramente mayor recidiva con el retenedor removible**, pero la diferencia media fue menor de 1 mm, por lo que no es clínicamente significativo.
SUPERVIVENCIA DEL RETENEDOR **Nº de participantes**: 57 (1 ECAs)	**RR 0.44** (0.22 a 0.98)	**500 cada 1000**	**222 cada 1000**	**280 menos por cada 1000** (400 a 100 menos)	**7** (3 a 26)	⊕◯◯◯ MUY BAJA (Ver ANEXO 4.3)	**Mayor supervivencia del retenedor removible** (uso a tiempo parcial).
SATISFACCIÓN DEL RETENEDOR (Incomodidad) **Nº de participantes**: 57 (1 ECAs)	**RR 12.22** (1.69 a 88.52)	**33 cada 1000**	**407 cada 1000**	**374 más cada 1000** (23 a 1000 más)	**4** (2 a 10)	⊕◯◯◯ MUY BAJA (Ver ANEXO 4.3)	**Más incómodo el retenedor removible** que el retenedor fijo.
ÍNDICE GINGIVAL **Nº de participantes**: 59 (1 ECAs)	-	**Media 0.92**	**Media 0.58**	**MD 0.34 menor** (0.66 a 0.02 menor)	-	⊕◯◯◯ MUY BAJA (Ver ANEXO 4.3)	**Mejor salud periodontal con el retenedor removible** que el retenedor fijo.

Todas las justificaciones detalladas de la degradación de la certeza de la evidencia de acuerdo con el enfoque GRADE, se presentaron en el ANEXO 4.3.

© Álvaro Limones Burgos, 2024

7.2.5. QUINTA REVISIÓN SISTEMÁTICA

(Limones et al., 2022)

El manuscrito publicado se encuentra en el ANEXO V. Su objetivo primario fue responder a la siguiente pregunta de investigación: *los ensayos clínicos controlados y aleatorizados sobre prótesis fija dentosoportada, ¿siguen las guías actuales, son de alta calidad metodológica y hay homogeneidad en sus instrumentos y variables de medición*? Para analizar críticamente las limitaciones intrínsecas de esta RS, se utilizaron las siguientes herramientas: AMSTAR-II, ROBIS, PRISMA. Para analizar críticamente las limitaciones extrínsecas de esta RS, se utilizaron las siguientes herramientas: ROB 2.0 y CONSORT, en lugar de la herramienta GRADE por tratarse de una RS de instrumentos y variables de medición. Los resultados se presentan a continuación:

7.2.5.1. AMSTAR II: Resultados de la herramienta de evaluación metodológica

De acuerdo con la herramienta AMSTAR-II, se considera que es una revisión de ALTA confianza por la metodología empleada. Se concluye que la revisión puede proporcionar un resumen preciso y completo de los resultados de los estudios disponibles que abordan la pregunta de interés (Tabla del ANEXO 5.1).

7.2.5.2. ROBIS: Resultados de la herramienta de evaluación del riesgo de sesgo

De acuerdo con la herramienta ROBIS, esta RS se clasificó como de INEXISTENTE riesgo de sesgo, no destacándose preocupaciones importantes en ninguna de las fases de evaluación. La evaluación detallada del riesgo de sesgo se encuentra en la Tabla del ANEXO 5.2.

7.2.5.3. PRISMA: Comprobación del cumplimiento de la guía de reporte

La comprobación del cumplimiento de la guía de reporte se evaluó a partir de la lista de comprobación propuesta por la última versión de la declaración PRISMA (Moher et al., 2015; Page et al., 2021) y su extensión para resúmenes (*abstracts)* (Beller et al., 2013) en un diseño de tabla combinado y mostrado en la Tabla del ANEXO 5.3. Se detectó un incumplimiento de

© Álvaro Limones Burgos, 2024

18% de los ítems de la extensión para resúmenes y un 100% de cumplimiento de los ítems de la lista de comprobación original actualizada. La comprobación del cumplimiento de la guía de reporte no identificó ninguna limitación.

7.2.5.4. ROB 2.0 + CONSORT: Comprobación del riesgo de sesgo y la calidad del reporte de los ensayos clínicos publicados sobre prótesis fija dentosoportada

De forma particular en esta RS se evaluó el riesgo de sesgo y la calidad del reporte en los estudios primarios para responder al objetivo secundario que consintió en analizar críticamente la investigación clínica publicada en prótesis fija dentosoportada mediante la evaluación de la calidad, consistencia metodológica y representatividad clínica de los ensayos clínicos controlados y aleatorizados publicados en los últimos 10 años en el área temática, mediante una RS de instrumentos y variables de medición. El riesgo de sesgo analizado en los estudios primarios y el cumplimiento de las guías de reporte para ECA se presentan en la Tabla 15. Un total del 20% de los estudios tuvieron un bajo RoB, el 79% presentó algunas preocupaciones, y el 1% tuvo un alto RoB. Un total del 72% de los estudios no mencionó el uso de la declaración de reporte CONSORT, el 69% no describió un cálculo del tamaño de la muestra, y el 50% no describió el resultado primario de la investigación.

Tablas 15. Característica de los ensayos incluidos, análisis del riesgo de sesgo con la herramienta RoB 2.0 y cumplimiento CONSORT. Los superíndices hacen referencia a la publicación Limones et al. (2022).

Estudio	Año	Diseño	Temática	Revista	Pais	RoB 2.0	Puntuación CONSORT
Agustín-Panadero et al [16]	2021	ECA	Técnicas de preparación dental	J Prosthet Dent	España	SC	-
Ahrberg et al [17]	2015	ECA	PF unitaria	Clin Oral Investig	Alemania	SC	-
Akin et al [18]	2014	ECA	PF unitaria	J Prosthodont	Turquía	SC	-
Benic et al[30–32]	2016	ECA	PF unitaria	J Prosthet Dent	Suiza	SC	-
Bömickle et al [33]	2019	ECA	PF múltiples unidades	J Prosthodont Res	Alemania	SC	21
Bömickle et al [34]	2021	ECA	PF múltiples unidades	J Prosthodont	Alemania	SC	-
Botelho et al [35]	2016	ECA	Diseños protésicos	J Dent	Alemania	SC	-
Botelho et al [36]	2020	ECA	Diseños protésicos	J Dent	Alemania	L	-

© Álvaro Limones Burgos, 2024

Breemer et al [37,38]	2019	ECA	Inlays	Oper Dent	Paises Bajos	L	23
Cagidiaco et al [15]	2019a	ECA	PF unitaria	BMC Oral Health	Italia	L	22
Cagidiaco et al [39]	2019b	ECA	Técnicas de preparación dental	Int J Periodontics Restorative Dent	Italia	SC	20
Cheng et al [40]	2021	ECA	Flujos de trabajo	J Prosthet Dent	China	SC	-
Christensen et al [41]	2010	ECA	PF múltiples unidades	JADA	EEUU	SC	-
Crins et al [43]	2021	ECA	Onlays	Dent Mater	Paises Bajos	SC	-
El-Ma'aita et al [43]	2022	ECA	Endocoronas	J Prosthodont	Jordan	L	23
Eltoukhy et al [44]	2021	ECA	Inlay FDPs	J Dent	Egipto	SC	27
Esquival-Upshaw et al [45]	2013	ECA	PF unitaria	J Prosthodont	EEUU	SC	-
Etman et al [46]	2010	ECA	PF unitaria	J Prosthet Dent	Reino Unido	SC	-
Ferrari et a l [47]	2019	ECA	PF unitaria	J Dent	Italia	L	25
Gjelvold et al [48]	2016	ECA	Flujos de trabajo	J Prosthodont	Suecia	SC	-
Gotfredsen et al [49]	2021	ECA	PF unitaria	J Adhes Dent	Dinamarca	SC	-
Gresnigt et al [50]	2012	ECA	Carillas	J Dent	Paises Bajos	SC	-
Gresnigt et al [52,52]	2013	ECA	Carillas	J Dent & J Adhes Dent	Paises Bajos	L	25
Grohmann et al [53]	2015	ECA	PF múltiples unidades	Quintessence Int	Suiza	SC	-
Haddadi et al [54]	2018	ECA	Flujos de trabajo	Int J Prosthodont	Dinamarca	SC	20
Halawani et al [55]	2021	ECA	PF unitaria	Int J Prosthodont	Egipto	SC	-
Hammoudi et al [56]	2022	ECA	PF múltiples unidades	J Prosthet Dent	Suecia	SC	-
Hu et al [57]	2010	ECA	PF múltiples unidades	Int J Prosthodont	China	SC	-
Huettig et al [58]	2016	ECA	PF múltiples unidades	Clin Oral Investig	Alemania	SC	21
Izgi et al [59]	2011	ECA	PF unitaria	J Adhes Dent	Turquía	H	-
Juloski et al [60]	2014	ECA	PF unitaria	J Dent Res	Italia	L	26
Jurema et al [61]	2020	ECA	PF unitaria	J Esthet Restor Dent	Brasil	L	28
Klur et al [62]	2019	ECA	PF unitaria & múltiple	Clin Oral Investig	Alemania	SC	-
Kozmacs et al [63]	2017	ECA	PF unitaria	J Prosthet Dent	Alemania	SC	-
Makarouna et al [19]	2011	ECA	PF múltiples unidades	Int J Prosthodont	Alemania	SC	-
Malmstrom et al [64]	2015	ECA	PF múltiples unidades	J Oral Rehabil	EEUU	SC	-
Marchionatti et al [65]	2017	ECA	Carillas	J Prosthet Dent	Brasil	SC	24
McKenna et al [66]	2014	ECA	Prótesis fija de múltiples unidades y removible	Community Dent Oral Epidemiol	Irlanda	SC	-
Metiner et al [20]	2021	ECA	Flujos de trabajo	Int. J. Comput. Dent	Turquía	SC	17

© Álvaro Limones Burgos, 2024

Monaco et al [67]	2017	ECA	PF unitaria	J Dent	Italia	L	25
Naenni et al [68]	2015	ECA	PF múltiples unidades	J Dent	Suiza	SC	-
Nicolaisen et al [69,70]	2016	ECA	PF múltiples unidades	Int J Prosthodont	Dinamarca	SC	-
Paniz et al [71]	2017	ECA	Técnicas de preparación dental	Int J Prosthodont	Italia	SC	-
Pelaez et al [72,73]	2012	ECA	PF múltiples unidades	J Prosthet Dent & Int J Prosthodont	España	SC	-
Peumans et al [74]	2013	ECA	Inlays/Onlays	Clin Oral Investig	Bélgica	SC	-
Pousette-Lundgren et al[75,76]	2015	ECA	PF unitaria	J Dent	Suecia	L	-
Reitemeier et al [77,78]	2013	ECA	PF unitaria & múltiple	J Prosthet Dent	Alemania	SC	-
Rinke et al [79]	2022	ECA	Inlays/Onlays	Clin Oral Investig	Alemania	SC	-
Sailer et al [80–82]	2009	ECA	PF múltiples unidades	Int J Prosthodont & J Dent	Suiza	L	-
Sailer et al [83–85]	2019	ECA	Flujos de trabajo	J Prosthet Dent	Suiza	SC	-
Seidel et al [86]	2020	ECA	PF unitaria	J Dent	Alemania	SC	-
Selz et al [87]	2014	ECA	PF unitaria	Clin Oral Investig	Alemania	SC	-
Seydler et al [88]	2015	ECA	PF unitaria	J Prosthet Dent	Alemania	SC	-
Shetty et al [89]	2017	ECA	PF unitaria	Int J Prosthodont	India	SC	-
Smitter et al[90]	2011	ECA	PF unitaria	Quintessence Int	Alemania	SC	-
Soares-Rusu et al [91]	2021	ECA	Carillas	Oper Dent	Brasil	SC	19
Sterzenbach et al [92]	2012	ECA	PF unitaria	J Endod	Alemania	L	28
Suarez et al [93]	2019	ECA	PF múltiples unidades	J Prosthodont	España	SC	-
Taschner et al [94]	2012	ECA	PF unitaria	Dent Mater	Alemania	SC	23
Vanoorbeek et al [21]	2010	ECA	PF unitaria	Int J Prosthodont	Bélgica	SC	-
Vigolo et al [95]	2012	ECA	PF unitaria	J Prosthodont	Italia	SC	-
Wolfart et al [96–99]	2014	ECA	Diseños protésicos	Clin Oral Investig, J Dent, J Prosthodont Res & J Oral Rehabil	Alemania	SC	-
Zenthöfer et al [100]	2015	ECA	Diseños protésicos	J Prosthet Dent	Alemania	SC	-
Zhang et al [101]	2018	ECA	PF unitaria	J Prosthet Dent	China	L	-

© Álvaro Limones Burgos, 2024

7.2.5.5. Resumen de la evaluación – limitaciones identificadas

A la vista del análisis de los resultados obtenidos tras la aplicación de las herramientas mencionadas anteriormente, y siguiendo la clasificación de las limitaciones presentada en el apartado 7.1 de esta Tesis Doctoral (resultados correspondientes a la componente teórica), dividimos las limitaciones encontradas en intrínsecas y extrínsecas.

No se detectaron limitaciones intrínsecas propias de esta RS, debido a:

- Experiencia del equipo investigador.

Dentro de las limitaciones extrínsecas ajenas a la metodología de esta RS destacan:

1. Heterogeneidad de índices de evaluación de complicaciones, mecánicas, biológicas y estéticas.
2. Moderado cumplimiento de las guías CONSORT.
3. Riesgos de sesgo en los estudios primarios por falta de cegamiento en la asignación o evaluación y reporte selectivo de resultados.

7.2.5.6. Presentación de las características de los estudios primarios

Las características de los 64 ECA identificados y publicados en 79 artículos científicos durante los últimos 10 años en el campo de la prótesis fija, fueron presentados en la Tabla 16.

Tablas 16. Característica de los ensayos incluidos.

Características de diseño	**N**	**(%)**
Número de ECA	64	/
Número de publicaciones	79	/
Diseño del ECA		
Boca dividida	12	18.7%
Cruzado	5	7.8%
Paralelo	45	70.3%
Número de centros donde tiene lugar el ECA		
Unicéntrico	56	87.5%
Multicéntrico	8	12.5%
Entorno donde transcurre el estudio		
Práctica privada	11	17.2%

© Álvaro Limones Burgos, 2024

Servicio de salud pública	6	9.4%
Entorno universitario	47	73.4%
Aprobación del comité de ética	59	92.2%
Registrado en una plataforma internacional accesible	22	34.4%
Adherencia a la declaración CONSORT 2010	18	28.1%
Mención explícita del resultado primario	32	50.0%
Cálculo del tamaño de la muestra	20	31.2%
Generación de la secuencia aleatoria		
No mencionado	24	37.5%
Generación de secuencia no aleatoria	0	0.0%
Lanzamiento de moneda, tirar dados o repartir cartas barajadas	10	15.6%
Lista de aleatorización generada por computadora	30	46.9%
Método de aleatorización		
No mencionado	52	81.2%
Aleatorización simple	6	9.4%
Aleatorización en bloques	5	7.8%
Aleatorización estratificada	1	1.6%
Minimización	0	0.0%
Método de ocultación de la asignación		
No mencionado	40	62.5%
Aleatorización central	5	7.8%
Sobres opacos y sellados numerados secuencialmente	19	29.7%
Cegamiento de las intervenciones		
Cegamiento del paciente	17	26.6%
Cegamiento del operador	0	0.0%
Cegamiento del examinador	43	67.2%
Dirección del ensayo		
No mencionado	62	96.9%
No inferioridad	0	0.0%
Superioridad	2	3.1%
Equivalencia	0	0.0%
Tipo de análisis		
Análisis de intención de tratar (ITT)	35	54.7%
Análisis de intención de tratar modificado (mITT)	29	45.3%
Análisis por protocolo	0	0.0%
Manejo de los datos perdidos		
Datos perdidos como faltantes al azar (MAR) o completamente al azar (MCAR)	28	43.7%
Datos perdidos relevantes como no faltantes al azar (MNAR)	1	1.6%
Análisis de sensibilidad	0	0.0%
Financiación		

© Álvaro Limones Burgos, 2024

No declarada	25	39.10%
Declarada	39	60.9%
Financiación interna	13	20.3%
Financiación externa	26	40.6%
Organización gubernamental	6	9.4%
Organización privada con fines de lucro	20	31.2%
Organización privada sin fines de lucro	0	0.0%
Ambos	6	9.4%
Descripción de la implicación del financiador y de los conflictos de interés	26	40.6%

© Álvaro Limones Burgos, 2024

7.2.6. TABLA RESUMEN DE LAS LIMITACIONES ENCONTRADAS

En la siguiente tabla se resumen las limitaciones intrínsecas y extrínsecas identificadas en las cinco RS originales, utilizando las herramientas mencionadas en las secciones de Material y Métodos, y Resultados de esta disertación de Tesis Doctoral:

Revisión sistemática	Tipo de revisión sistemática	Limitaciones Intrínsecas	Limitaciones Extrínsecas
Limones, Saez-Alcaide, et al. (2020)	RS de efectos adversos	1. No registrar del protocolo del estudio en un registro internacional de acceso público (signo de rigurosidad). 2. No reflejar la no necesidad de análisis de sensibilidad, análisis de subgrupos o metarregresiones (signo de rigurosidad). 3. No poner a disposición del público las plantillas de extracción de datos o cualquier otro material utilizado en la revisión (signo de transparencia). 4. No discutir las limitaciones que pudo tener la no inclusión de estudios en idiomas distintos al inglés, francés o español (signo de exhaustividad) 5. No definir qué es un resultado clínicamente significativo (signo de objetividad).	1. Todos los estudios primarios incluidos estuvieron financiados por la empresa distribuidora y fabricante de la intervención estudiada. Esto representa un riesgo potencial de sesgo y enfatiza la necesidad de manejar con cautela las relaciones contractuales entre investigadores y empresas financiadoras. 2. No se superó el tamaño óptimo de información.
Limones, Molinero-Mourelle, et al. (2020)	RS de intervenciones	1. No se realizaron análisis de sensibilidad (signo de rigurosidad). 2. No se evaluó la certeza de la evidencia científica mediante en enfoque GRADE (signo de rigurosidad). 3. No se presentó un listado explicito de las referencias de los estudios excluidos (signo de transparencia). 4. No se definió qué es un resultado clínicamente significativo para cada variable crítica (signo de objetividad).	1. Se encontró heterogeneidad entre las variables e instrumentos de medición de complicaciones mecánicas/técnicas y biológicas. 2. No se superó el tamaño optimo de información en algunas de las variables críticas. 3. No se informó de todas de variables críticas (i.e. sangrado al sondaje). 4. No se registró el protocolo de los estudios primarios con carácter prospectivo.

© Álvaro Limones Burgos, 2024

Wilensky et al. (2023)	RS de intervenciones	No encontradas	1. Riesgos de sesgo en los estudios primarios por falta de cegamiento, reporte selectivo de resultados o presencia de conflictos de interés 2. Un tamaño muestral insuficiente por la complejidad de ejecutar las intervenciones. 3. Presencia de heterogeneidad no explicada.
Martin et al. (2023)	RS de intervenciones	No encontradas	1. Riesgos de sesgo en los estudios primarios por falta de cegamiento en la asignación/evaluación o reporte selectivo de resultados. 2. Un tamaño muestral insuficiente por falta de consistencia entre las intervenciones de los estudios, características, tiempos de uso de retención y tiempos de seguimiento.
Limones et al. (2022)	RS de instrumentos y variables de evaluación	No encontradas	1. Heterogeneidad de índices de evaluación de complicaciones, mecánicas, biológicas y estéticas. 2. Moderado cumplimiento de las guías CONSORT. 3. Riesgos de sesgo en los estudios primarios por falta de cegamiento en la asignación o evaluación y reporte selectivo de resultados.

© Álvaro Limones Burgos, 2024

8. DISCUSIÓN

La presente Tesis Doctoral pone la RS en el centro de su disertación, la contextualiza dentro de la OBE y explora las limitaciones, fortalezas y evolución de este diseño de estudio desde un enfoque dual: teórico y práctico.

8.1. Componente Teórica – Revisión del estado del arte

8.1.1. Odontología Basada en la Evidencia: de la evidencia a la recomendación

El aumento de la producción científica en el momento actual, genera una sobrecarga informativa o "infoxicación", que hace que sea prácticamente imposible que los profesionales de la salud se mantengan actualizados de todos los hallazgos publicados, discerniendo cuáles son de alta calidad para fundamentar su toma de decisiones con seguridad. Rara vez exploran la evidencia tan minuciosamente como los panelistas expertos de GPC, que hacen un balance neto entre beneficios, riesgos, preferencias de los pacientes y costes, entre otros. Por lo tanto, la realización de RS y el desarrollo de GPC se convierten en herramientas esenciales para filtrar y sintetizar la evidencia, y fundamentar la práctica clínica diaria con eficiencia y el menor riesgo para los pacientes.

Mientras que las RS ofrecen resúmenes precisos, concisos, confiables y completos de la evidencia disponible sobre un tema, son las GPC las que ponen en balanza los beneficios y riesgos presentados y resumidos en RS, y las que brindan recomendaciones clínicas claras y concisas mediante su difusión a través las sociedades científicas relevantes de cada especialidad (i.e. SEPES, SEPA, EFP, ITI, etc.) (Figura 2 del apartado 2.2.1).

© Álvaro Limones Burgos, 2024

8.1.2. Ventajas de las revisiones sistemáticas

Es por ello por lo que las RS se convierten en un pilar central para el desarrollo de GPC de alta calidad (Nivel S3), presentando múltiples ventajas:

1. **Aumentan la generalidad de los resultados (validez externa).** Cuando diferentes características de los estudios incluidos (i.e. diferentes edades o países) coinciden en una magnitud del efecto similar (es decir, no se detecta heterogeneidad).

2. **Pueden lograr resultados de mayor precisión**, al combinar los resultados de los diferentes estudios individuales, disminuyendo el error aleatorio y los intervalos de confianza, y aumentando la potencia estadística. Pudiendo, en ocasiones, revelar diferencias entre dos intervenciones que no habían sido evidentes en estudios individuales por un insuficiente número de participantes.

3. **Detectan inconsistencias entre estudios aparentemente homogéneos** y sugieren hipótesis para explicar la heterogeneidad mediante metarregresiones, análisis de subgrupos o análisis de la sensibilidad.

4. **Presentan de forma visual la dirección del efecto** mediante gráficas, como los *forest* o *Albatross plots*.

5. **Reducen el sesgo de selección de la evidencia o *cherry picking.*** Al realizar una búsqueda sistemática y exhaustiva de la literatura publicada y no publicada, se evalúa toda la evidencia, no solo aquella que respalda las hipótesis de los investigadores (Andrade, 2021).

6. **Permiten evaluar el sesgo de publicación**. Detectar la tendencia de que estudios con resultados positivos o significativos sean publicados con mayor probabilidad que estudios con resultados negativos o no concluyentes, pudiendo sobreestimar el efecto evaluado.

© Álvaro Limones Burgos, 2024

7. **Penalizan la manipulación de hipótesis o *HARKing*** (por sus siglas en inglés *"Hypothesizing After the Results are Known"*), práctica en el que el investigador presenta una hipótesis desarrollada post hoc como a priori, basándose en resultados estadísticamente significativos no previstos, y la presenta como si el diseño, ejecución y análisis del estudio se hubieran orientado desde el inicio a probar tal hipótesis. (Andrade, 2021). Este aspecto disminuye nuestra confianza en el estimativo. Se evidencia mediante las herramientas de evaluación del riesgo de sesgo en los estudios primarios incluidos, las cuales determinan si el protocolo de investigación fue publicado previamente al inicio del estudio o no; y se penaliza en el metaanálisis (i.e. realizando un metaanálisis acumulado en función del riesgo de sesgo de los estudios) y en la interpretación de sus resultados.

8. **Penalizan la manipulación de análisis estadísticos o *P-hacking***, práctica en la cual el investigador de los estudios primarios obtiene el resultado significativo deseado probando distintos test estadísticos, eliminando valores atípicos, definiendo nuevos umbrales, creando nuevos subgrupos, convirtiendo variables continuas en dicotómicas, etc. (Andrade, 2021). Se evidencia y penaliza igual que el *HARKing*.

9. **Permiten calificar la confianza que tenemos en los estimativos** a partir de la evaluación del riesgo de sesgo, la imprecisión, la inconsistencia, la evidencia indirecta y el sesgo de publicación, siguiendo la metodología de evaluación de la certeza de la evidencia propuesta por el grupo GRADE.

10. **Reciben más citaciones que otros diseños de estudio** y**,** por tanto, aumentan el índice-h de los autores y el factor de impacto de las revistas científicas.

8.1.3. Soluciones a las limitaciones de las revisiones sistemáticas

No obstante, las RS no están exentas de limitaciones, las cuales pueden clasificarse en intrínsecas, originadas por la propia metodología de la RS y que son solventables, y extrínsecas, derivadas de las metodologías de los estudios incluidos, y que, por lo tanto, no lo son en el

© Álvaro Limones Burgos, 2024

momento de la revisión. Para abordar estas limitaciones intrínsecas, se han propuesto las siguientes soluciones (Uttley et al., 2023):

Solución a la **Exhaustividad**:

1. Incorporar un experto en bases de datos biomédicas con experiencia en el diseño y la ejecución de búsquedas bibliográficas, es esencial para garantizar que una RS incluya todos los estudios pertinentes.
2. Revisar la bibliografía de los estudios incluidos o de RS publicadas previamente, usar sistemas de alerta de nuevos estudios publicados en la base de datos principal, revisar la literatura gris, revisar los registros internacionales de estudios en curso y contactar a sus investigadores para conocer el estado de sus investigaciones o la cesión de sus datos (sólo recomendable bajo el amparo de instituciones de reconocido prestigio para revisiones por encargo).

Solución a la **Rigurosidad**:

1. Adherirse a las directrices metodológicas y de reporte para protocolos de investigación de RS (i.e. MECIR, PRISMA-P (Moher et al., 2015)), poniéndolos a disposición del publico a través de plataformas de registro internacional (i.e. PROSPERO o OSF), haciendo uso de los manuales Cochrane (Deeks et al., 2023; Higgins et al., 2022) y GRADE (Schünemann et al., 2013) para consultas.
2. Preespecificar las hipótesis a priori, cómo se tratará la heterogeneidad, los análisis de subgrupos, metarregresiones, los análisis de sensibilidad, de sesgo de publicación u otros previstos.

Solución a la **Transparencia**:

1. Cumplir las directrices de reporte (i.e. PRISMA y sus extensiones).
2. Publicar los resultados de las revisiones independientemente de la dirección de los hallazgos.
3. Poner los datos a disposición del publico en plataformas de acceso internacional para datos científicos (i.e. OSF).

© Álvaro Limones Burgos, 2024

Solución a la **Objetividad**:

1. Contar con un equipo adecuadamente cualificado.
2. Confirmar que las preguntas de investigación son prioritarias y no redundantes.
3. Comprobar por duplicado la selección de estudios, la introducción de los datos y el análisis de riesgo de sesgo.
4. Revelar posibles conflictos de intereses (i.e. patrocinios de los investigadores).

8.1.4. Importancia de la selección de marcos de relevancia clínica y el impacto del P-valor en la investigación odontológica

El concepto del "P-valor" se remonta al siglo XVIII, cuando John Arbuthnot realizó los primeros cálculos sobre la probabilidad del género de recién nacidos (Arbuthnot, 1710). No obstante, la adopción de la estadística matemática en la investigación médica se atribuye ampliamente a figuras como Ronald Fisher, Jerzy Neyman y Egon Pearson durante las décadas de 1920 y 1930, quienes realmente impulsaron su uso. Fisher, estableció un umbral rígido para el P-valor de .05, aunque también sugirió flexibilidad en su interpretación (Fisher, 1925). En la actualidad, un P-valor menor o mayor que .05 se ha convertido en sinónimo de resultados positivos o negativos, respectivamente, lo cual ha generado múltiples problemas en la realización y el reporte de investigaciones. Esto ha alimentado el sesgo de publicación y la notificación selectiva de resultados, donde los estudios con hallazgos estadísticamente significativos son más propensos a ser publicados que aquellos con resultados "no significativos". Además, las variables estadísticamente significativas tienden a ser más reportadas que las no significativas en los artículos científicos. Asimismo, la obsesión por alcanzar un P-valor significativo fomenta prácticas metodológicas cuestionables, como el P-hacking, que implica manipular estadísticamente los datos hasta obtener los resultados deseados (P-valor < 0.05), o el HARking, que consiste en crear hipótesis después de obtener los resultados para justificarlos. Esto plantea la pregunta: ¿somos realmente esclavos del P-valor (Goel et al., 2024)?

La falta de contextualización del P-valor por no seleccionar adecuadamente marcos de decisión o de relevancia clínica puede resultar en falsos positivos, es decir, resultados estadísticamente significativos pero no clínicamente relevantes. Estos falsos positivos pueden

© Álvaro Limones Burgos, 2024

llevar a directrices de tratamiento o recomendaciones políticas que no sean beneficiosas e incluso perjudiciales. Por ello, es crucial establecer las diferencias mínimas clínicamente relevantes desde la perspectiva de los investigadores para marcar umbrales del efecto (no efecto, efecto leve, efecto moderado y efecto severo) y realizar interpretaciones más precisas de las hipótesis planteadas y preespecificadas en protocolos registrados en plataformas de acceso público para fomentar la transparencia.

8.1.5. Necesidad de actualización de las revisiones sistemáticas

Garner et al. (2016) definen la actualización de una RS como la creación de una nueva versión que podría incorporar nuevos datos, métodos o análisis, diferentes a los de la edición anterior. Garritty et al. (2010) mencionan que las razones para actualizar pueden ser cuantitativas, como un cambio significativo en la magnitud de un efecto, o cualitativas, debido al surgimiento de nuevas circunstancias o contextos que impactan los resultados previos.

Entidades como la Colaboración Cochrane recomiendan actualizar las RS cada dos años para mantener su relevancia y actualidad (Higgins et al., 2022). Sin embargo, adherirse a un intervalo fijo puede resultar simplista. Factores como la aparición de nuevos estudios significativos, la obsolescencia de metodologías anteriores o la relevancia actual de la pregunta de investigación pueden indicar la necesidad de una actualización. No obstante, la decisión sobre cuándo proceder con estas actualizaciones suele recaer en el equipo de investigación o en organizaciones que gestionan las RS, generalmente, por encargo.

8.1.6. Evolución de las revisiones sistemáticas, un nuevo enfoque de la pirámide de la evidencia científica

De acuerdo con la pirámide tradicional de la evidencia científica, las RS y metaanálisis ocupan un lugar jerárquicamente superior a otros diseños de estudio. Sin embargo, la limitada necesidad de recursos económicos o logísticos requeridos para el desarrollo de estos estudios ha llevado a que la producción de RS y metaanálisis se dispare, alcanzando proporciones epidémicas. Wallach (2019) denominó este fenómeno como "metástasis de metaanálisis". El gráfico de la

© Álvaro Limones Burgos, 2024

Figura 16 exhibe una creciente divergencia entre el número de RS y el número de ECA publicados en las últimas dos décadas. Como se puede observar en el gráfico, en la última década, las curvas se separan drásticamente, siendo cuatro veces mayor el número de RS que el de ECA publicados en el momento actual, no esperando que esta tendencia vaya a cambiar. A pesar de que resulta ilógico pensar que estudios secundarios que sintetizan y evalúan la evidencia retrospectivamente, superen al número de estudios primarios que aportan nueva evidencia.

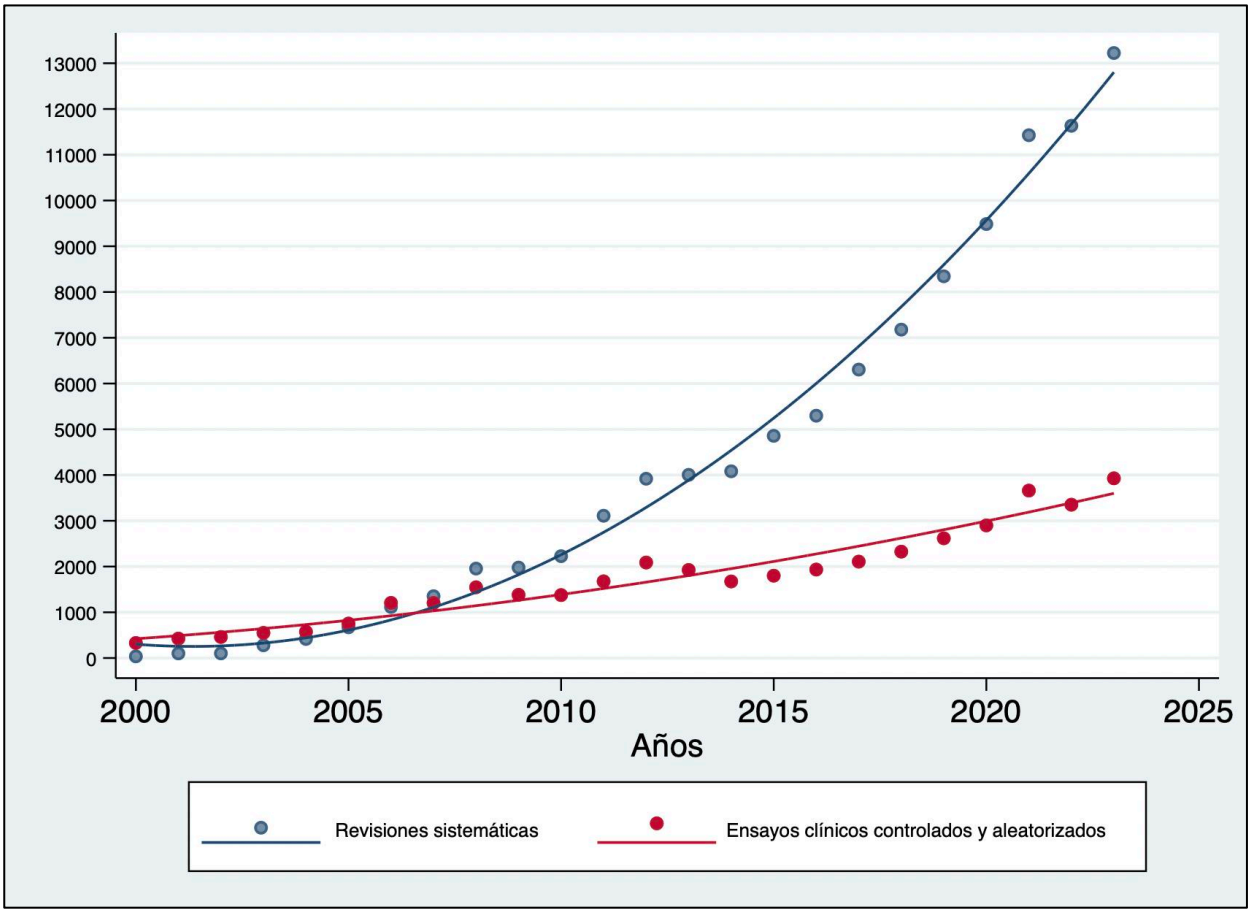


Figura 16. Número de revisiones sistemáticas y ensayos clínicos controlados y aleatorizados publicados en los últimos 20 años.

Recientemente, se ha producido un cambio de paradigma en cuanto a la posición que las RS ocupan en la pirámide de evidencia científica (Murad et al., 2016). El enfoque contemporáneo ve a las RS como una lupa a través de la que visualizar detalladamente la evidencia, reconociendo que es dinámica y variable, como se ilustra en la Figura 17. Esto se debe a que no siempre un ECA

© Álvaro Limones Burgos, 2024

resulta ser el diseño de estudio más óptimo; en ciertas circunstancias, estudios observacionales bien estructurados que involucran un número amplio de participantes pueden ser complementarios.

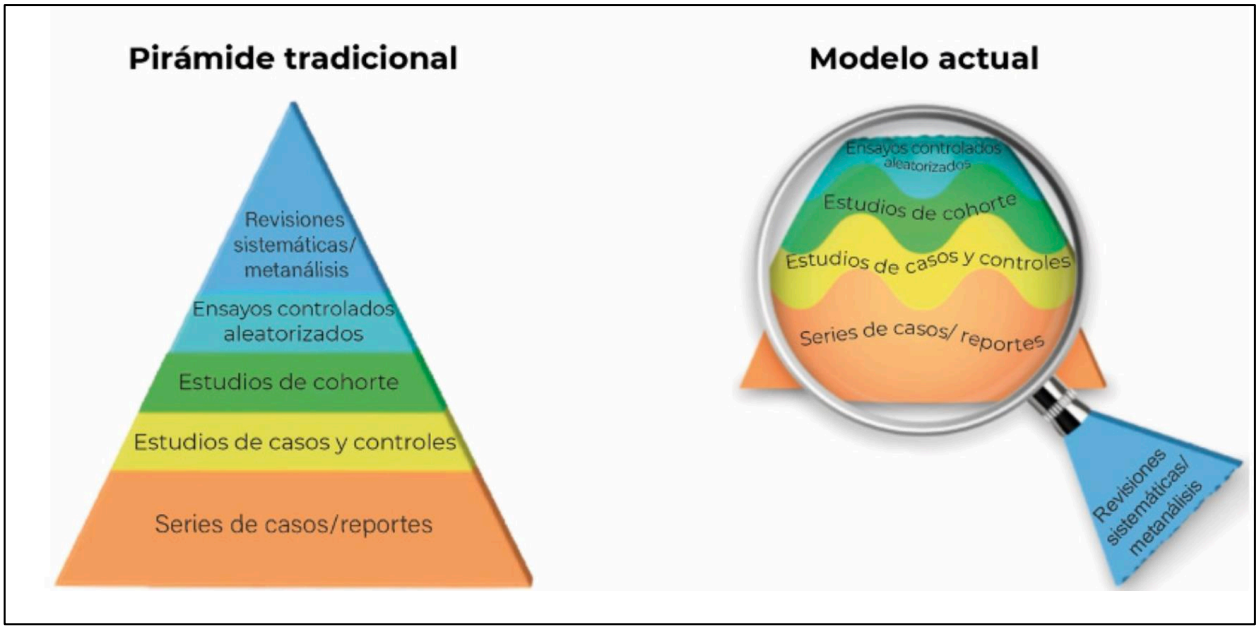


Figura 17. Enfoque contemporáneo de la posición que ocupan las revisiones sistemáticas en la pirámide de la evidencia científica (Murad et al., 2016).

8.1.7. Inteligencia artificial aplicada a la confección de las revisiones sistemáticas

En la actualidad, la confección de RS es un campo que no está exento de la implementación de la Inteligencia Artificial (IA). A través de la automatización, el reconocimiento de patrones y el aprendizaje automático mediante grandes modelos de lenguaje (LLM, *Large Language Models*), se pueden hacer los procesos de revisión más eficientes, objetivos, consistentes y rápidos.

Si bien, estas herramientas (i.e. Abstrackr, RobotAnalyst, EPPI-Reviewer, DistillerSR, etc.) no están validadas. Requieren un entrenamiento más profundo para poder encargarse de forma autónoma y fehaciente de partes críticas del proceso de revisión, como la selección de estudios, la extracción de datos o la evaluación del riesgo de sesgo de los estudios primarios a partir de las complejidades del lenguaje natural (Millard et al., 2016). Se observó entre un 7-29% de pérdidas en selecciones automatizadas de estudios en el trabajo de Feng et al. (2022), cayendo en un sesgo

© Álvaro Limones Burgos, 2024

de selección severo. Por tanto, en el momento actual, las herramientas de IA deben utilizarse como ayudas complementarias y no como sustitutos completos de revisores humanos (Yao et al., 2024).

8.2. Componente Práctica – Revisiones sistemáticas publicadas y evaluadas

8.2.1. Limitaciones globales identificadas que comprometen la formulación de recomendaciones fuertes en guías de práctica clínica

Las recomendaciones fuertes de acuerdo con el enfoque GRADE se emiten cuando un panel de expertos de GPC tiene alta confianza en que los beneficios de una intervención superan significativamente sus efectos negativos (generando una recomendación fuerte a favor de la intervención) o, inversamente, cuando los efectos negativos predominan sobre los beneficios esperados (generando una recomendación fuerte en contra de la intervención). Esto implica que la mayoría de los individuos se beneficiarían notablemente al seguir el curso de acción recomendado. Los factores que influyen en la fuerza de las recomendaciones incluyen el balance entre beneficios y perjuicios significativos, preferencias de los pacientes, costos, el contexto clínico de aplicación, equidad, aceptabilidad y factibilidad. Todos estos, excepto la certeza de la evidencia, dependen de la intervención o exposición específica y, por lo tanto, no se pueden generalizar. No obstante, la certeza de la evidencia sigue un principio distinto: según el enfoque GRADE, cuanto mayor es la certeza, más probable es que se emita una recomendación fuerte. Asimismo, GRADE desaconseja emitir recomendaciones fuertes cuando la certeza en las variables críticas es baja o muy baja.

Un ejemplo real y práctico de todo el proceso presentado en esta disertación sobre "la OBE: de la evidencia a la recomendación" fue la RS de Wilensky et al. (2023) sobre los procedimientos químicos coadyuvantes al tratamiento quirúrgico de la periimplantitis, la cual fue encargada y empleada, junto con otras 12 RS, en el desarrollo de una GPC de nivel S3 (Herrera et al., 2023). Este proceso se llevó a cabo durante el XVIII *Workshop* Europeo de Periodoncia en La Granja de San Ildefonso (Segovia, España) entre el 6 y el 9 de noviembre de 2022. En ella no se pudo emitir una recomendación fuerte a favor del uso de las intervenciones de estudio por disponer

© Álvaro Limones Burgos, 2024

de una muy baja certeza de la evidencia; lo que implica la aceptación de la hipótesis nula primaria de esta disertación. Este aspecto resalta la necesidad de exponer y evaluar minuciosamente las limitaciones de las RS en contextos prácticos.

Con el objetivo de hacer este análisis lo más objetivo posible, se utilizaron herramientas validadas de evaluación de la calidad metodológica (AMSTAR II), el riesgo de sesgo (ROBIS), la certeza de la evidencia científica (GRADE) y guías de reporte (PRISMA). Pudimos clasificar las limitaciones de las RS en intrínsecas, relacionadas con el propio método sistemático de búsqueda y síntesis de la información, y extrínsecas, que provienen de la calidad del material científico disponible. Las limitaciones intrínsecas identificadas incluyeron:

1. No registrar del protocolo de la revisión en un registro internacional de acceso público (signo de rigurosidad). En la revisión llevada a cabo por Limones, Saez-Alcaide, et al. (2020) no se desarrolló un protocolo previo al inicio del estudio. Este escenario subraya la existencia de una curva de aprendizaje asociada a este proceso y la importancia de una formación exhaustiva; así como de la constitución de equipos con experiencia desde la fase inicial de elaboración del protocolo.
2. No incluir estudios en idiomas distintos al inglés (signo de exhaustividad). En la revisión realizada por Wilensky et al. (2023), existió una limitación de tiempo debido a que la RS fue encargada por un comité de expertos con el propósito específico de desarrollar una GPC. Esta restricción temporal obligó a los investigadores a incluir únicamente estudios disponibles en inglés. En el momento actual, existen herramientas de traducción que permiten a los investigadores trabajar con publicaciones en otros idiomas, como ocurrió en la revisión Cochrane realizada por Martin et al. (2023), quienes tradujeron publicaciones en chino para evaluar su elegibilidad.
3. No se presentó un listado explicito de las referencias de los estudios excluidos (signo de transparencia). En las revisiones de Limones, Molinero-Mourelle, et al. (2020) y Limones, Saez-Alcaide, et al. (2020) se presentaron los motivos de exclusión de forma descriptiva y en el propio diagrama de flujo, pero sin ofrecer una lista explicita. En todas la RS se evaluó la concordancia inter-revisor mediante el índice Kappa Cohen obteniendo un nivel superior al 80%.
4. No se definieron los marcos de relevancia clínica para cada variable crítica (signo de objetividad). El enfoque GRADE resalta el valor de presentar umbrales de decisión o

© Álvaro Limones Burgos, 2024

relevancia clínica para poder juzgar adecuadamente la presencia de imprecisión durante la evaluación de la certeza de la evidencia o, posteriormente, los beneficios y riesgos en GPC (Neumann et al., 2022).

5. No informar sobre la necesidad de análisis de sensibilidad, análisis de subgrupos o metarregresiones (signo de transparencia). Aunque la mayoría de las RS no agruparon más de 10 estudios en el metaanálisis, es crucial desde el protocolo establecer las intenciones de los investigadores mediante hipótesis a priori y los métodos para explorar posibles fuentes de heterogeneidad, para no caer en prácticas cuestionables (Guyatt et al., 2023).
6. No se evaluó la certeza de la evidencia científica mediante el enfoque GRADE (signo de rigurosidad). En la revisión de Limones, Molinero-Mourelle, et al. (2020), no se incluyó una "Tabla Resumen de la Evidencia". Esta omisión dificulta la interpretación de la certeza o confianza en cada variable crítica y limita la idoneidad de la RS como base para la elaboración de una GPC.
7. No poner a disposición del publico las plantillas de extracción de datos o cualquier otro material utilizado en la revisión (signo de transparencia). A excepción de la revisión de Limones, Saez-Alcaide, et al. (2020), en las demás RS se proporcionaron datos adicionales como material suplementario publicado en las páginas web de las revistas, así como en plataformas de compartición de datos científicos. Este fue el caso de la revisión de Limones et al. (2022) sobre la evaluación de la investigación clínica en prótesis fija dentosoportada, en la que los datos extraídos fueron compartidos en el repositorio internacional de acceso abierto *Open Science Framework*.

Todas estas limitaciones se fueron reduciendo de acuerdo con:

a) Una curva de aprendizaje del revisor (A.L) mediante asistencia a cursos formativos (Limones et al., 2022; Limones, Molinero-Mourelle, et al., 2020; Limones, Saez-Alcaide, et al., 2020; Martin et al., 2023; Wilensky et al., 2023).
b) Trabajo en equipo con investigadores experimentados pertenecientes a instituciones de reconocido prestigio, como la Colaboración Cochrane (Martin et al., 2023) o la EFP (Wilensky et al., 2023).
c) Aparición de nuevas guías metodológicas y publicaciones por parte de expertos, dirigidas a abordar los problemas emergentes identificados durante los procesos de revisión.

© Álvaro Limones Burgos, 2024

Las limitaciones extrínsecas identificadas incluyeron:

1. Riesgos de sesgo en los estudios primarios por: falta de cegamiento en la asignación o en la evaluación, reporte selectivo de resultados, el no registro del protocolo con carácter prospectivo antes del comienzo de este o la presencia de conflictos de interés. Fue observable en Odontología la dificultad de cegamiento en los estudios de intervención que implicaban procedimientos prácticos. Especialmente, la ocultación durante la asignación aleatoria de los participantes en los grupos o durante el transcurso de las intervenciones se ha estudiado que pueden sobreestimar alrededor del 10% las estimaciones del efecto del tratamiento (Savovic et al., 2018). En la revisión de Limones, Saez-Alcaide, et al. (2020), todos los estudios primarios incluidos, fueron financiados por la empresa distribuidora y fabricante de la intervención estudiada. Esto representa un riesgo potencial de sesgo y enfatiza la necesidad de manejar con cautela las relaciones contractuales entre investigadores y empresas financiadoras. En la revisión de Martin et al. (2023), el sesgo de publicación selectiva de resultados se mitigó contactando a los autores cuando existían datos faltantes o haciendo transformaciones estadísticas con la calculadora de RevMan 5 (RevMan 5, 2020).
2. Inconsistencias en las intervenciones, comparadores y las variables e instrumentos de medición. En la revisión de Martin et al. (2023), a pesar de incluir 48 ECA en la revisión, fueron aparentes las inconsistencias entre las intervenciones y los compradores a lo largo de los estudios incluidos, que conllevó a la imposibilidad de combinar sus resultados en las variables de relevancia clínica, y consigo, llevó a una confianza muy baja de las estimaciones (muy baja certeza de la evidencia). En la revisión de Limones, Molinero-Mourelle, et al. (2020) fueron aparentes las inconsistencias por la disparidad de variables e instrumentos de medición utilizados, lo que nos inspiró a realizar una revisión específica con el fin de poner de manifiesto y dilucidar este aspecto en el campo de la investigación clínica en prótesis fija dentosoportada de la última década (Limones et al., 2022)
3. Imprecisión por un tamaño muestral insuficiente, no superando, en muchas ocasiones, el tamaño óptimo de información (Zeng et al., 2022). Esto se atribuye probablemente a las complejidades inherentes a los estudios de intervención en Odontología, donde se suelen examinar intervenciones de procedimientos prácticos, a diferencia de otras disciplinas

© Álvaro Limones Burgos, 2024

médicas que investigan intervenciones farmacológicas (Zavalis et al., 2024). En el ámbito de la prótesis fija, por ejemplo, en las intervenciones participan al menos dos operadores, el prostodoncista y el técnico de laboratorio, lo que introduce múltiples posibles factores de confusión y hace complejo agrupar grandes tamaños muéstrales en los ECA.

4. Carácter indirecto de la evidencia en los resultados, pues son escasos los resultados a largo plazo, que son precisamente los escenarios de interés clínico.
5. Omisión de variables críticas para proceder a recomendaciones, como efectos adversos, costes o satisfacción del paciente.
6. No registro del protocolo con carácter prospectivo de los estudios primarios antes del comienzo de este, que reduce la credibilidad de la investigaciones pudiendo caer en prácticas de investigación cuestionables como el *HARKing*, el *cherry-picking*, el *P-hacking*, o las *fishing expeditions* (cuando se analizan grandes bases de datos para detectar resultados estadísticamente significativos sin una hipótesis previa, aumentando la probabilidad de falsos positivos).
7. Moderado cumplimiento de las guías CONSORT para ECA.
8. Imposibilidad de evaluar el sesgo de publicación, concretamente, por escasez de ECA, lo que recalca la necesidad de más investigación primaria.

En definitiva, a la luz de los hallazgos presentados, se puede afirmar que la RS es una herramienta metodológica de valor real, que no está exenta de limitaciones. Sin embargo, antes de embarcarse en proyectos de investigación de este tipo, se vuelve crucial cuestionarse: ¿está verdaderamente justificada la realización de la RS? ¿Existe un problema clínico real que requiera la síntesis de la mejor evidencia y el desarrollo de recomendaciones? ¿Es una prioridad? ¿Hay suficiente evidencia disponible para realizar una síntesis o existen RS previas de alta calidad publicadas recientemente? Estos interrogantes encuentran respuesta en el desarrollo de GPC por panelistas expertos, las cuales representan un faro que orienta y dirige la investigación hacia la priorización, la calidad metodológica y la relevancia clínica.

© Álvaro Limones Burgos, 2024

8.2.2. Situación actual de la investigación clínica publicada en prótesis fija dentosoporatada

La RS, cuyo objetivo fue evaluar el riesgo de sesgo, la consistencia de las variables e instrumentos de medición y el reporte de la investigación clínica en el campo de la prótesis fija dentosoportada de la última década, encontró que un total del 20% de los estudios tuvieron un riesgo bajo de sesgo, el 79% presentó algunas preocupaciones, y el 1% tuvo un riesgo alto de sesgo, por lo que se aceptó parcialmente la hipótesis alternativa secundaria de esta disertación. Un total del 72% de los estudios no mencionaron el uso de la declaración CONSORT, el 69% no describieron un cálculo del tamaño de la muestra, y el 50% no describieron la variable primaria. Se observó que la adherencia a la declaración CONSORT estaba asociada con un menor riesgo de sesgo. Se documentó una gran cantidad de variables e instrumentos de medición utilizados, sin mejoras desde la última RS con objetivos homólogos publicada en 2013 (Patel et al., 2014). La necesidad de un conjunto de variables mínimas esenciales para evaluar la investigación clínica en prótesis fija dentosoportada resultó evidente.

La mayoría de los autores utilizaron la tasa de éxito y supervivencia como variables primarias en los ensayos clínicos. La tasa de éxito se ha definido comúnmente como la permanencia de la prótesis en su lugar y sin complicaciones o alteraciones en la satisfacción del paciente. La supervivencia se ha definido rutinariamente como la permanencia de la prótesis en su lugar, incluso con complicaciones que pueden considerarse aceptables y sin necesidad de reemplazo. Es esencial registrar los detalles de estos criterios, o de lo contrario, la comparación entre estudios puede comprometerse.

Tres grandes conjuntos de variables se han descrito hasta la fecha para evaluar complicaciones mecánicas, el sistema de la USPHS (Cvar & Ryge, 2005), el sistema de la CDA (California-Dental-Association, 1977) y el de la FDI (Hickel et al., 2010).

Los estudios en prótesis suelen enfocarse más en evaluar la resistencia de los materiales y el rendimiento mecánico general, en lugar de examinar sus interacciones con los tejidos blandos. Quedando patente la heterogeneidad metodológica en la selección de variables para evaluar clínicamente la relación entre las PFD y el periodonto. Esta diversidad de variables plantea desafíos en la síntesis de datos y la comparación entre estudios (Limones, Molinero-Mourelle, et

© Álvaro Limones Burgos, 2024

al., 2020). Las variables más frecuentemente seleccionadas consistentes con la investigación en periodoncia incluyen: el sangrado al sondaje (BoP, por sus siglas en inglés de *Bleeding on Probing*), la pérdida de inserción clínica (CAL, por sus siglas en inglés de *Clinical attachement loss*), la profundidad de sondaje (PPD, por sus siglas en inglés de *Probing Pocket Depth*) y el índice de control de placa (PCR, por sus siglas en inglés de *Plaque Control Record*). Las principales complicaciones biológicas evaluadas fueron la vitalidad, la caries secundaria y la hipersensibilidad dentinaria.

En la investigación sobre estética dental, el índice objetivo que ha demostrado mayor repetibilidad es el índice de evaluación de la estética rosa y blanca (PES/WES, por sus siglas en inglés *Pink Esthetic Score/White Esthetic Score*) (Chen et al., 2018; Hosseini & Gotfredsen, 2012). Los sistemas de la USPHS/Ryge, CDA y FDI/Hickel también evalúan complicaciones estéticas de manera similar de acuerdo con el color, la forma anatómico-estética o la tinción del material. Actualmente, el instrumento más objetivo y rápido para evaluar el color dental son los espectrofotómetros que utilizan la escala de color de la *Commission Internationale de l'Eclairage* (CIE)Lab o CIECh, que identifica el color según las dimensiones de negro/blanco (valor L*), verde/rojo (valor a*), amarillo/azul (valor b*), luminosidad (valor L*), croma (valor C*) y tono (valor h*). Esta técnica se ha utilizado ampliamente para cuantificar el color dental y sus diferencias (Gómez-Polo et al., 2017).

La satisfacción del paciente es otro aspecto crítico a evaluar, que considera la opinión subjetiva de cada paciente en puntos de tiempo específicos (antes o después del tratamiento a 1, 3, 5, 10 años o más). Tanto la escala analógica visual (VAS, por sus siglas en inglés: *Visual Analogue Scale*) como la evaluación de la satisfacción oral (OSS, por sus siglas en inglés: *Oral Satisfaction Assessment*) se han generalizado para una evaluación rápida y global de la satisfacción del paciente, mientras que los perfiles de impacto de salud oral de 14 y 20 ítems (OHIP-14 y OHIP-20), siendo versiones abreviadas del OHIP-49 original, han sido instrumentos ampliamente utilizados en las últimas décadas para una evaluación más completa (Allen & Locker, 2002; Montero et al., 2011; Slade, 1997). El impacto oral en el rendimiento diario (OIDP, por sus siglas en inglés: *Oral Impacts on Daily Performances*) es otro instrumento frecuentemente utilizado para describir el impacto de las condiciones orales en la calidad de vida (Castillo-Oyagüe, Suárez-García, et al., 2016). Recientemente, entre otros avances en el campo, se ha validado un cuestionario específico, completo y corto (de 10 ítems) denominado QoLFAST-10 (Castillo-

© Álvaro Limones Burgos, 2024

Oyagüe, Perea, et al., 2016; Castillo-Oyagüe, Suárez-García, et al., 2016). El índice QoLFAST-10, que inicialmente fue diseñado para prótesis implantosoportadas, puede ser útil para mejorar la eficiencia de los ensayos clínicos. Según un metaanálisis reciente de Wittneben et al. (2018), la evidencia sobre PFD todavía es escasa, justificando la necesidad de ensayos clínicos adicionales.

Los hallazgos de esta RS mostraron aspectos metodológicos mejorables, como la falta de uso de la declaración CONSORT para mejorar la calidad de los informes en ECA. En consecuencia, nos encontramos con ciertos aspectos críticos en la metodología que no fueron informados. Esta RS demostró que un mayor cumplimiento de sus items, disminuyó el RoB, lo que concuerda con revisiones similares de otros tipos de intervenciones en salud oral (Sarkis-Onofre et al., 2017; Sponchiado-Júnior et al., 2021).

El cálculo del tamaño de la muestra es un requisito fundamental que debe incluirse en el informe de un ECA, a excepción de los nuevos análisis basados en inteligencia artificial que pueden trabajar con grandes tamaños muéstrales, incluso con la totalidad de la población (Liu, Cruz Rivera, et al., 2020; Shan et al., 2021). Sin embargo, por norma general, se selecciona un subconjunto de participantes que represente adecuadamente a la población, a partir del cual se pueden extraer conclusiones. Tan solo un 31% de los ECA informaron de su cálculo del tamaño muestral y su justificación.

Por otro lado, la aleatorización es obligatoria para generar grupos balanceados en sus características iniciales y evitar factores de confusión que puedan sobreestimar o subestimar las conclusiones. En los ensayos clínicos de prótesis, la aleatorización estratificada o la minimización aseguran mejor la equidad de los grupos que la aleatorización simple o por bloques al equilibrar los posibles factores de confusión, ya que la participación de al menos un prostodoncista y un técnico de laboratorio hace que las intervenciones sean complejas y que sea difícil obtener tamaños de muestra grandes. A diferencia de los estudios de intervenciones farmacológicas cuyas intervenciones son más fáciles de aplicar.

El ocultamiento de la asignación, a menudo confundido con el enmascaramiento, protege la posible manipulación de los participantes antes de que se asignen a los grupos de intervención. Los métodos utilizados para ocultar la asignación incluyen la asignación central por un tercero responsable o métodos a prueba de manipulación, como sobres sellados, opacos y numerados secuencialmente que se abren consecutivamente después de que el participante ha sido asignado de manera irreversible.

© Álvaro Limones Burgos, 2024

Las desviaciones del protocolo pueden ocurrir en ensayos clínicos, donde los participantes no siempre siguen las instrucciones, abandonan los estudios o se pierden en el seguimiento. Esto puede llevar a sesgos de selección o sesgos por abandono en los resultados. Por lo tanto, para preservar completamente el beneficio de la aleatorización, todos los participantes aleatorizados deben incluirse en el análisis estadístico, y concretamente en el grupo al que fueron asignados. A este análisis se le conoce como análisis por intención de tratar (ITT, por sus siglas en inglés: *Intention-To-Treat*), un método que se ha recomendado ampliamente como estrategia de análisis preferente. El análisis por intención de tratar requiere identificar y reportar completamente las desviaciones del protocolo y las razones de las pérdidas para clasificar los datos perdidos como: faltantes completamente por azar, faltantes por azar y faltantes no por azar; para establecer un método apropiado para manejar los datos perdidos (como la imputación múltiple, la ponderación por probabilidad inversa o la estimación de máxima verosimilitud) y realizar análisis de sensibilidad. Por el contrario, si los datos faltantes se omiten del análisis estadístico y solo se analizan los participantes que hayan completado el estudio, se realizará un análisis por protocolo (PP), comprometiendo así la aleatorización, reduciendo el tamaño de la muestra y la potencia, introduciendo sesgos y proporcionando una estimación alterada de los resultados. Esto ocurre cuando la frecuencia o las causas de las pérdidas difieren entre los grupos de intervención (Hakoum et al., 2017).

Para cubrir los costes económicos involucrados en los estudios, a menudo, se obtiene financiación de organizaciones gubernamentales o privadas (con o sin fines de lucro), apoyo de casas comerciales a cambio de dar visibilidad a sus marcas en publicaciones (patrocinio), o colaborar con campañas de responsabilidad social corporativa de empresas, entre otros. Sin embargo, los intereses financieros siempre deben ser declarados en el artículo. Es crítico que el protocolo describa explícitamente la participación y responsabilidades de esos financiadores en el diseño, realización, análisis e interpretación de los datos o en la divulgación de los resultados del estudio. Los autores deben tener el control final sobre estos asuntos. La ética de la investigación, la objetividad y el sentido común deben prevalecer en los acuerdos contractuales firmados entre los financiadores y los investigadores.

Los futuros estudios deberían explorar el desarrollo evidente de una estandarización de variables e instrumentos de medición mediante la iniciativa COMET. Los futuros ECA deberían esforzarse por registrar sus estudios de manera prospectiva en registros internacionales, adherirse

© Álvaro Limones Burgos, 2024

a la declaración CONSORT, realizar cálculos de tamaño de la muestra, informar de sus variables primarias, describir la aleatorización, la ocultación de la asignación y el enmascaramiento en detalle, y explicar las razones de los abandonos. Aunque la mayoría de las revistas han adoptado la declaración CONSORT como requisito entre sus normas, aún no se ha utilizado ampliamente en los informes publicados.

© Álvaro Limones Burgos, 2024

9. LIMITACIONES

Esta Tesis Doctoral enfrentó ciertas limitaciones.

En primer lugar, el análisis crítico de las limitaciones de las RS en la componente práctica fue realizado por un investigador únicamente (A.L.), que además estuvo involucrado activamente en el diseño, ejecución y divulgación de las mismas (sesgo de giro). Por lo tanto, es posible que un equipo diferente de investigadores pudiera llegar a interpretaciones distintas a las aquí presentadas.

En segundo lugar, a pesar de que se intentó maximizar la objetividad utilizando herramientas validadas de evaluación o guías de reporte para RS desde diversas consideraciones (AMSTAR-II para evaluar la calidad metodológica, ROBIS para evaluar el riesgo de sesgo, GRADE para evaluar la certeza de la evidencia y PRISMA para comprobar el seguimiento de la guía de reporte), es plausible que existan otras limitaciones no identificadas en la componente práctica.

En tercer lugar, la elección de las publicaciones a analizar en la componente práctica de esta Tesis Doctoral no siguió ningún proceso sistemático, y se realizó fundamentado en la participación como autor del doctorando.

© Álvaro Limones Burgos, 2024

© Álvaro Limones Burgos, 2024

10. CONCLUSIONES

10.1. Conclusión primaria de carácter general en Odontología

Se aceptó la hipótesis nula primaria de esta Tesis Doctoral: las revisiones sistemáticas no generaron un sustrato adecuado, para la consecución de recomendaciones fuertes a través de guías de práctica clínica, cuando presentaron limitaciones, bien en su metodología (intrínsecas al propio proceso de revisión), o bien, en los estudios primarios incluidos (por la evidencia disponible).

1. La componente teórica de la presente disertación, a través de una revisión del estado del arte, identificó fortalezas y limitaciones de las revisiones sistemáticas, sugiriendo soluciones, y enfatizando la importancia de: reconsiderar su posición dentro de la pirámide de la evidencia científica, seleccionar marcos de relevancia clínica, actualizar periódicamente las revisiones, explorar las nuevas aplicaciones de inteligencia artificial para su confección, y la importancia de desarrollar guías de práctica clínica sobre las que fundamentar la práctica odontológica basada en la evidencia, con seguridad y eficiencia.
2. La componente práctica de la presente disertación, a través de cuatro revisiones sistemáticas originales, detectó en contextos reales, y de principio a fin, las siguientes limitaciones extrínsecas: insuficiencia de ensayos clínicos y participantes, debido a las complejidades de las intervenciones en Odontología, que dificultan aunar grandes tamaños muéstrales (generando imprecisión); heterogeneidad metodológica en las intervenciones, comparadores, variables e instrumentos de medición (generando inconsistencias entre estudios); falta de resultados a largo plazo (generando cierto carácter indirecto de la evidencia), y falta de consideración de efectos adversos (limitando la representatividad clínica); y otras intrínsecas que son solventables trabajando en equipo con investigadores experimentados.

10.2. Conclusión secundaria de carácter específico, en el campo de la Odontología Conservadora y la Prótesis Bucofacial

Se aceptó parcialmente la hipótesis alternativa secundaria de esta Tesis Doctoral: los ensayos clínicos en prótesis fija dentosoportada de la última década, demostraron una calidad metodológica aceptable, aunque fue notoria la falta de adherencia a la declaración CONSORT y la presencia de preocupaciones en el análisis del riesgo de sesgo, especialmente, por no describir de forma explícita los métodos de aleatorización, o no publicar los protocolos de forma prospectiva en plataformas de registro internacional. Los futuros ensayos podrían mejorar, aún más, sus metodologías y la calidad de sus reportes. Se hizo evidente la necesidad de estandarización de las variables e instrumentos de medición en este campo de conocimiento, contemplando (I) el éxito, (II) la supervivencia, (III) las complicaciones técnicas, (IV) biológicas, (V) y estéticas, (VI) los costes, (VII) la eficiencia y (VIII) la satisfacción de los pacientes, a través de la iniciativa COMET mediante grupos de consenso con expertos de relevancia internacional.

© Álvaro Limones Burgos, 2024

© Álvaro Limones Burgos, 2024

11. ANEXOS

ANEXO I: Manuscrito de la primera revisión sistemática

(Limones, Saez-Alcaide, et al., 2020)

© Álvaro Limones Burgos, 2024

Journal section: Oral Medicine and Pathology
Publication Types: Review

doi:10.4317/medoral.23324
http://dx.doi.org/doi:10.4317/medoral.23324

Medication-related osteonecrosis of the jaws (MRONJ) in cancer patients treated with denosumab VS. zoledronic acid: A systematic review and meta-analysis

Alvaro Limones [1,2], Luis Miguel Sáez-Alcaide [3], Santiago Angulo Díaz-Parreño [4], Alexandra Helm [3], Michael M. Bornstein [5,6], Pedro Molinero-Mourelle [1]

[1] Department of Conservative Dentistry and Orofacial Prosthesis. Faculty of Dentistry, Complutense University of Madrid, Spain
[2] Department of Dentistry. Universidad San Pablo CEU, Madrid, Spain
[3] Department of Dental Clinical Specialties. Faculty of Dentistry, Complutense University of Madrid, Spain
[4] Department of Statistics and Applied Mathematics. San Pablo CEU University, Madrid, Spain
[5] Applied Oral Sciences & Community Dental Care. Faculty of Dentistry, The University of Hong Kong, Hong Kong SAR, China
[6] Department of Oral Health & Medicine, University Center for Dental Medicine Basel UZB, University of Basel, Basel, Switzerland

Correspondence:
Department of Conservative Dentistry and Orofacial Prosthesis
Plaza Ramón y Cajal S/N, 28040, Madrid, Spain
Faculty of Dentistry, Complutense University of Madrid, Spain
alimones@ucm.es

Received: 14/07/2019
Accepted: 17/02/2020

Limones A, Sáez-Alcaide LM, Díaz-Parreño SA, Helm A, Bornstein MM, Molinero-Mourelle P. Medication-related osteonecrosis of the jaws (MRONJ) in cancer patients treated with denosumab VS. zoledronic acid: A systematic review and meta-analysis. Med Oral Patol Oral Cir Bucal. 2020 May 1;25 (3):e326-36.

Article Number: 23324 http://www.medicinaoral.com/
© Medicina Oral S. L. C.I.F. B 96689336 - pISSN 1698-4447 - eISSN: 1698-6946
eMail: medicina@medicinaoral.com
Indexed in:
Science Citation Index Expanded
Journal Citation Reports
Index Medicus, MEDLINE, PubMed
Scopus, Embase and Emcare
Indice Médico Español

Abstract

Background: The aim of the present study was to analyse the incidence, risk ratio (RR) and prognoses of two types of medication-related osteonecrosis of the jaws (MRONJ): denosumab-related osteonecrosis of the jaws (DRONJ) and Bisphosphonate-Related Osteonecrosis of the Jaws (BRONJ) in cancer patients under treatment with denosumab or zoledronic acid (ZA).

Material and Methods: An electronic and manual search was conducted for randomized controlled trials (RCTs) until May 2019. Assessment of the identified studies, risk of bias and data extraction were performed independently by two reviewers. The incidence of DRONJ and BRONJ and the RR to develop MRONJ were calculated at 1 year, 2 years and 3 years of exposure. It was also calculated the odds ratio (OR) of their respective prognoses. They were calculated normalizing the values of the individual studies to 1 year, 2 years or 3 years when necessary through robust regression models using a statistical program.

Results: From 1.277 references identified, 8 RCTs were included, which comprised a total of 13.857 patients with a variety of neoplasms. The incidence of DRONJ in cancer patients under treatment with denosumab ranged from 0.5 to 2.1% after 1 year, 1.1 to 3.0% after 2 years, and 1.3 to 3.2% after 3 years of exposure. The incidence of BRONJ in cancer patients under treatment with ZA ranged from 0.4 to 1.6% after 1 year of exposure, 0.8 to

2.1% after 2 years, and 1.0 to 2.3% after 3 years of exposure. Statistically significant differences were found between denosumab and ZA in the risk of developing MRONJ after 1, 2 and 3 years of exposure. Nevertheless, there were no significant differences in terms of patient prognosis.
Conclusions: Denosumab is associated with a significantly higher risk of developing MRONJ compared to ZA. Nevertheless, no differences were found in its prognoses.

Key words: *Denosumab, zoledronic acid, bisphosphonate-associated osteonecrosis of the Jaws, medication-related osteonecrosis of the jaws, neoplasms.*

Introduction

The increasing aging population goes hand in hand with a growing prevalence of disabling disease along with the use of medication to prevent and treat metabolic bone diseases (1). The bone is the most common site for metastasis, mostly associated with malignant tumours of the breast (73%), prostate (68%) or lung (36%) (2). Bone metastases can cause skeletal-related events (SREs) such as pain, pathological fractures, hypercalcemia and spinal cord compression, requiring radiation and surgery. They are also linked to an overall increase in mortality.

In 2009, denosumab was approved by the Food and Drug Administration of the United States (FDA) and the European Medicines Agency (EMA) for the treatment and prevention of bone metastases. Numerous case reports and case series have been published since then (3-6). In 2014, the American Association of Oral and Maxillofacial Surgeons (AAOMS) changed the term "Bisphosphonate-Related Osteonecrosis of the Jaws" (BRONJ) to "Medication-Related Osteonecrosis of the Jaws" (MRONJ) (7), as it is not only triggered by bisphosphonates, but also by other antiresorptive and antiangiogenic drugs such as monoclonal antibodies (MABs), tyrosine kinase inhibitors (TKI), mammalian target of rapamycin inhibitors (mTORi), selective estrogen receptor modulators (SERMs) and immunosuppressants (8). MRONJ can be the cause of serious functional and masticatory disorders with an important influence on patient quality of life and may even result in death (9).

To date, the pathophysiology of MRONJ has not been fully elucidated. It is believed to be multifactorial, due to a decrease in physiological bone remodelling, inflammation, infection, inhibition of angiogenesis, and innate or acquired immunity dysfunction (10,11). However, there are two emerging theories on the aetiology behind MRONJ. The first one, named "inside-outside", is based on the inhibition of osteoclastic activity, resulting in a decrease of bone turnover. Due to this, jaw microdamage is not repaired and may lead to bone tissue necrosis and then to bone exposure over time. The second theory, termed "outside-inside", is based on a local depression of the immune system, leading to local infection or inflammation inducing osteonecrosis (12).

The use of denosumab is expected to increase in the near future, because of its favourable profile in terms of avoiding adverse effects and renal toxicity compared to zoledronic acid (ZA) in the treatment and prevention of SREs in patients with advanced solid tumours (13,14). Several meta-analyses have already reported the incidence of DRONJ (15,16). Nevertheless, several new randomized-controlled clinical trials have been published recently. Therefore, the aim of this updated systematic review and meta-analysis is to compare the incidence, risk ratio (RR) and prognoses of DRONJ and BRONJ in cancer patients under treatment with denosumab and ZA.

Material and Methods

This review was focused on answering the following three PICO questions: "In cancer patients under treatment with denosumab or ZA, do exist differences in the incidence of BRONJ (due to ZA) and DRONJ? If so, "what is the RR of MRONJ in patients treated with denosumab compared to patients treated with ZA?" and "do exist differences in the prognosis of BRONJ (due to ZA) compared to DRONJ?"

1) Study type: randomized clinical trials (RCTs).
2) Population: adult patients (> 18 years old) who were diagnosed with a solid tumour or with bone metastasis.
3) Intervention: subcutaneous denosumab in 120 mg doses every 4 weeks.
4) Comparison: intravenous ZA in 4 mg doses every 4 weeks.
5) Outcome: the primary outcome was the incidence of denosumab and zoledronic acid-related MRONJ; the secondary outcome was the RR of DRONJ compared to ZA-related BRONJ; and the third outcome was the OR of their respective prognoses.

- Eligibility criteria

Only double-blinded, ZA-control randomized clinical trials (RCTs) including patients followed up for at least 8 months were selected. RCTs without ZA-control, and non-randomized controlled clinical trials, retrospective and prospective studies were excluded. All studies were limited to research in humans published in English, French and Spanish.

- Search strategy

An electronic search for articles published from 2003 onwards was performed by entering the combination of the following search terms and Boolean operators: ((Cancer [All Fields] OR Neoplasms [MeSH Terms]) AND (Denosumab [MeSH Terms]) AND ("Zoledronic Acid" [MeSH Terms]) AND ("Bisphosphonate-Associated Osteonecrosis of the Jaw" [MeSH Terms] OR Osteonecrosis [MeSH Terms] OR MRONJ [All Fields] OR DRONJ [All Fields] OR ONJ [All Fields] OR ARONJ [All Fields])). An additional manual search was performed in selected journals of the field: "Journal of Dental Research", "Oral Oncology", "Clinical Oral Implant Research", "International Journal of Oral Science", "Oral Diseases", "Journal of Oral Pathology & Medicine", "International Journal of Oral and Maxillofacial Surgery", "Journal of Cranio-Maxillofacial Surgery", "Journal of Oral and Maxillofacial Surgery", "Oral Surgery Oral Medicine Oral Pathology Oral Radiology", "Medicina Oral Patología Oral y Cirugía Oral" and "International Journal of Clinical Oncology".

- Information sources

The following five electronic databases were screened between January and May 2019: SCOPUS, MEDLINE (via OvidSP); Web of Science (WOS); the Central Registry of Controlled Clinical Trials of the Cochrane Library (CENTRAL), and the International Clinical Trials Registry Platform of the World Health Organization (WHO ICTRP) (Fig. 1).

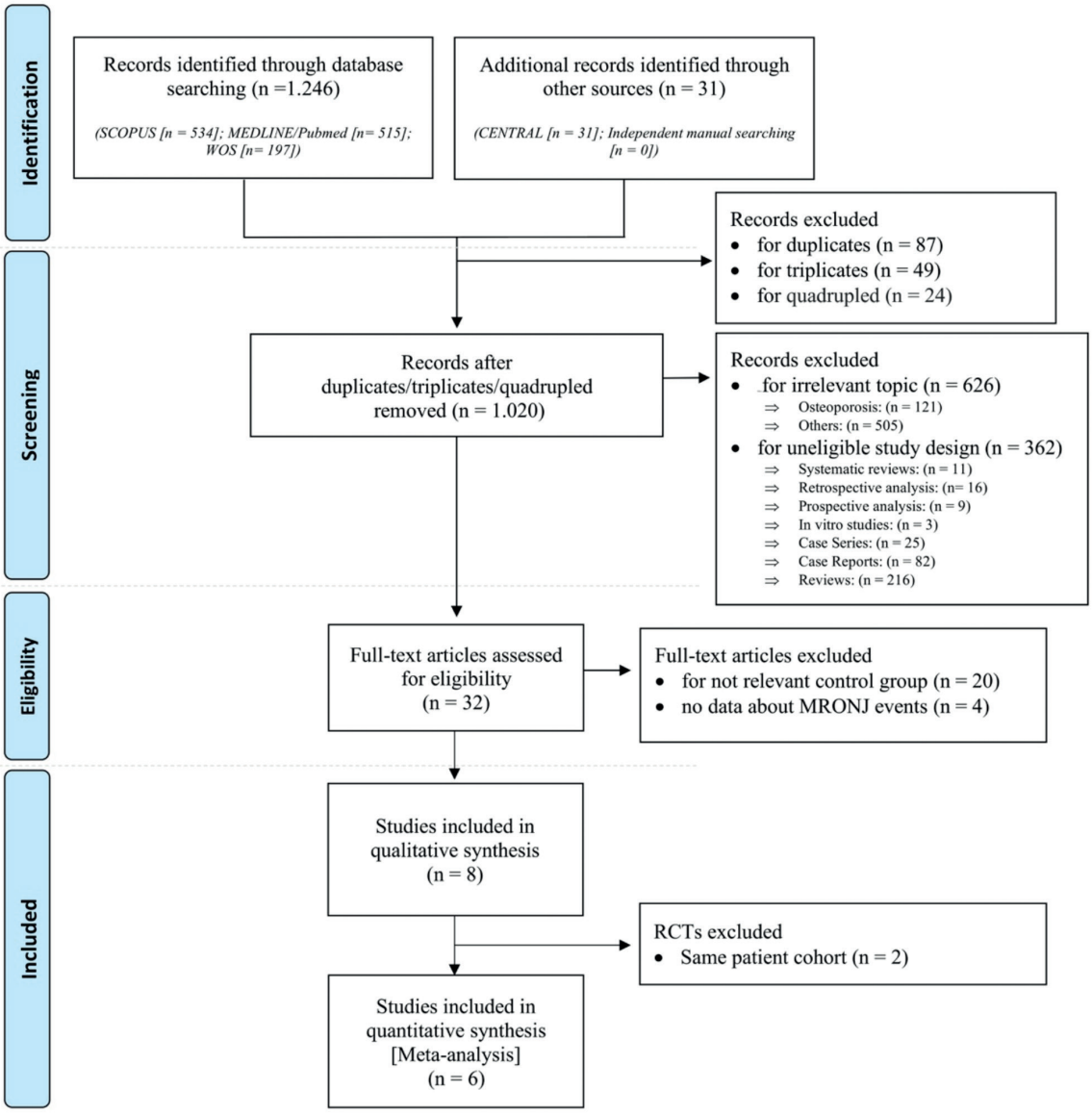


Fig. 1: PRISMA flow chart for the present systematic review and meta-analysis.

- Study selection & Data collection and items

Study selection and data extraction were performed independently by two reviewers (A.L. & P.M.M.). In the case of disagreement on inclusion or exclusion, a consensus was reached by discussion with a third researcher (L.M.S.). The level of agreement between the reviewers was estimated using Cohen's Kappa coefficient at title/abstract selection and at full-text selection. A Kappa value of more than 0.80 was considered as substantial agreement between the reviewers. The following items were entered into a Microsoft Excel spreadsheet (version 15.17, Microsoft Inc. 2015) (Table 1) (Table 2):

name of the first author, year of publication; country of origin; type of study; study population; number of participants; age; gender; follow-up; number of cases of MRONJ; cases of resolved MRONJ; type of drug (denosumab or ZA); dosage; frequency of administration, duration of drug administration, and route of administration.

Table 1: Characteristics of the included studies.

Author	Year	Country	Study design	Phase	Population	Mean age	Gender	Follow up	Drugs	Dose	Frecuency of administration	Route of administration
Stopeck *et al.* (24)	2010	Multicenter (USA)	RCT	Phase III	Breast Cancer	56 (49-65)	Female	34 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						57 (48-65)			Denosumab	120 mg	4 weeks	Subcutaneous
Fizazi *et al.* (25)	2011	Multicenter (France)	RCT	Phase III	Prostate Cancer	71 (66-77)	Male	41 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						71 (64-77)			Denosumab	120 mg	4 weeks	Subcutaneous
Henry *et al.* (26)	2011	Multicenter (USA)	RCT	Phase III	Multiple Myoeloma Bone metastasis	61 (22-89)	Male (66%)	34 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						60 (18-89)	Male (62%)		Denosumab	120 mg	4 weeks	Subcutaneous
Lipton *et al.* (27)	2012	Multicenter (USA)	RCT	Phase III	Bone metastasis	63	Male (53%)	8. 21 months	Zoledronic Acid	4 mg	4 weeks	I.V.
							Male (54%)		Denosumab	120 mg	4 weeks	Subcutaneous
Scagliotti *et al.* (28)	2012	Multicenter (USA)	RCT	Phase III	Lung Cancer	61 (54-69)	Male (68%)	NM	Zoledronic Acid	4 mg	4 weeks	I.V.
						60 (54-68)	Male (74%)		Denosumab	120 mg	4 weeks	Subcutaneous
Henry *et al.* (29)	2014	Multicenter (USA)	RCT	Phase III	Bone metastasis (except Breast & Prostate Cancers)	61 (22-87)	Male (62%)	24-30 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						59 (19-89)	Male (66%)		Denosumab	120 mg	4 weeks	Subcutaneous
Stopeck *et al.* (30,31)	2015	Multicenter (USA)	RCT	Phase III	A: Breast Cancer	56 (48-65)	Female	34 + 18.4 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						56 (48-64)			Denosumab	120 mg	4 weeks	Subcutaneous
					B: Prostate Cancer	71 (66-75.5)	Male	41 + 10.2 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						70 (66-77)			Denosumab	120 mg	4 weeks	Subcutaneous
Raje *et al.* (32)	2018	Multicenter (USA)	RCT	Phase III	Multiple Myeloma	63 (31-89)	Male (55%)	24 months	Zoledronic Acid	4 mg	4 weeks	I.V.
						63 (29-91)	Male (54%)		Denosumab	120 mg	4 weeks	Subcutaneous

RCT: Randomized Control Trial; NM: Not mentioned; I.V.: Intravenous Via.

Table 2: The incidence rate and prognosis of MRONJ reported by the included studies in the quantitative analysis and its normalization to 1 year, 2 years and 3 years of drug exposure.

Author	Year	Drugs	Sample size	MRONJ Cases					Associated Risk factors		Healed cases
				1 year of exposure	Interim period	2 year of exposure	Interim period	3 year of exposure	Dental Extraction, Poor Oral Higine Dental appliences	Chemo-therapy	
Stopeck *et al.* (24)	2010	Zoledronic Acid	1013	5 (0.5%)		12 (1.2%)		14 (1.4%)	10 (71%)	11 (79%)	6 (43%)
		Denosumab	1020	8 (0.8%)		19 (1.9%)		20 (2.0%)	18 (90%)	15 (75%)	10 (50%)
Fizazi *et al.* (25)	2011	Zoledronic Acid	945	5 (0.53%)		12 (1%)		NM - 12* (1,27%)*	10 (83%)	9 (75%)	1 (8%)
		Denosumab	943	10 (1%)		22 (2%)		NM - 22* (2,33%)*	17 (77%)	14 (64%)	4 (18%)
Henry *et al.* (26)	2011	Zoledronic Acid	878	5 (0.6%)		8 (0.9%)		11 (1.3%)	7 (70%)	7 (64%)	NM
		Denosumab	878	4 (0.5%)		9 (1.1%)		10 (1.1%)	10 (91%)	6 (60%)	
Lipton *et al.* (27)	2012	Zoledronic Acid	2836	NM - 18* (0,63%)*		NM - 31* (1,09%)*		37 (1.3%)	NM	NM	NM
		Denosumab	2841	NM - 28* (0,98%)*		NM - 34* (1,55%)*		52 (1.8%)	NM		
Scaglio-tti *et al.* (28)	2012	Zoledronic Acid	395	NM - 2* (0,38%)*		NM - 3* (0,84%)*	3 (0.8%) - *[21 months]*	NM - 4* (1,05%)*	NM	NM	NM
		Denosumab	406	NM - 1* (0,26%)*		NM - 3* (0,84%)*	3 (0.7%) - *[21 months]*	NM - 5* (1,11%)*	NM		
Raje *et al.* (32)	2018	Zoledronic Acid	852	NM - 23* (2,67%)*	24 (3%) - *[14,8 months]*	NM - 27* (3,13%)*		NM - 28* (3,34%)*	13 (54%)	NM	6 (25%)
		Denosumab	850	NM - 33* (3,87%)*	35 (4%) - *[15.8 months]*	NM - 38* (4,46%)*		NM - 40* (4,73%)*	19 (54%)		12 (35%)

NM: Not mentioned; (*) normalized-value.

- Statistical analysis

The incidence of MRONJ associated with denosumab and ZA was normalized when was necessary to 12, 24 and 36 months through a regression model using the statistical program IBM SPSS Statistics 25.0. The function S(t): exp (b0 + (b1/t)) was selected since it has the best fit with the course of the disease according to the incidence of each study and their follow-ups. The incidence was calculated analysing the MRONJ events among all participants involved in the included studies with a confidence interval (CI) of 95% using fixed or random-effect models depending on the heterogeneity of the included trials. Cochran's Q test and I^2 were used to determine statistical heterogeneity (18). If the I^2 value was between 0 and 50% and p-value of the Q test was > 0.05, the level of heterogeneity was interpreted to be within acceptable limits, and therefore, a fixed-effect model would be applied.

The RR of MRONJ and the OR of the respective prognoses was calculated by comparing denosumab vs. ZA with the same meta-analytic methodology as described above. The analyses and forest plots were performed using Comprehensive Meta-Analysis version 3 software (Biostat, Englewood, NJ, USA). The overall quality of evidence for outcomes addressed by direct evidence (analyses with RR or OR) was rated using the Grades of Recommendations, Assessment, Development and Evaluation (GRADE) approach grouped into very low, low, moderate or high (19).

- Risk of bias of included studies

The risk of bias for each study was determined using the Cochrane Collaboration Tool, described in the "Cochrane Handbook for Systematic Reviews of Interventions" version 5.1.0 (20). The articles were assessed according to 7 domains: selection bias, allocation bias, blinding of participants and staff, blinding of the outcome assessors, incomplete data, and selective notification of results. Furthermore, one category was added: conflict of interest, as it may be an important factor to take into consideration for pharmaceutically related studies. Each category was graded as low risk with a point, represented by (+) in green, high risk with zero represented by (-) in red, and uncertain risk with half a point represented by (?) in yellow (Fig. 2) (21). Studies with 5 points or more have a lower risk of bias, whereas studies with less than five points exhibited a higher risk of bias (Table 3).

- Study quality

Publication bias was determined visually using funnel plots (22) using the statistical program Mix 2.0 (Fig. 2) (23).

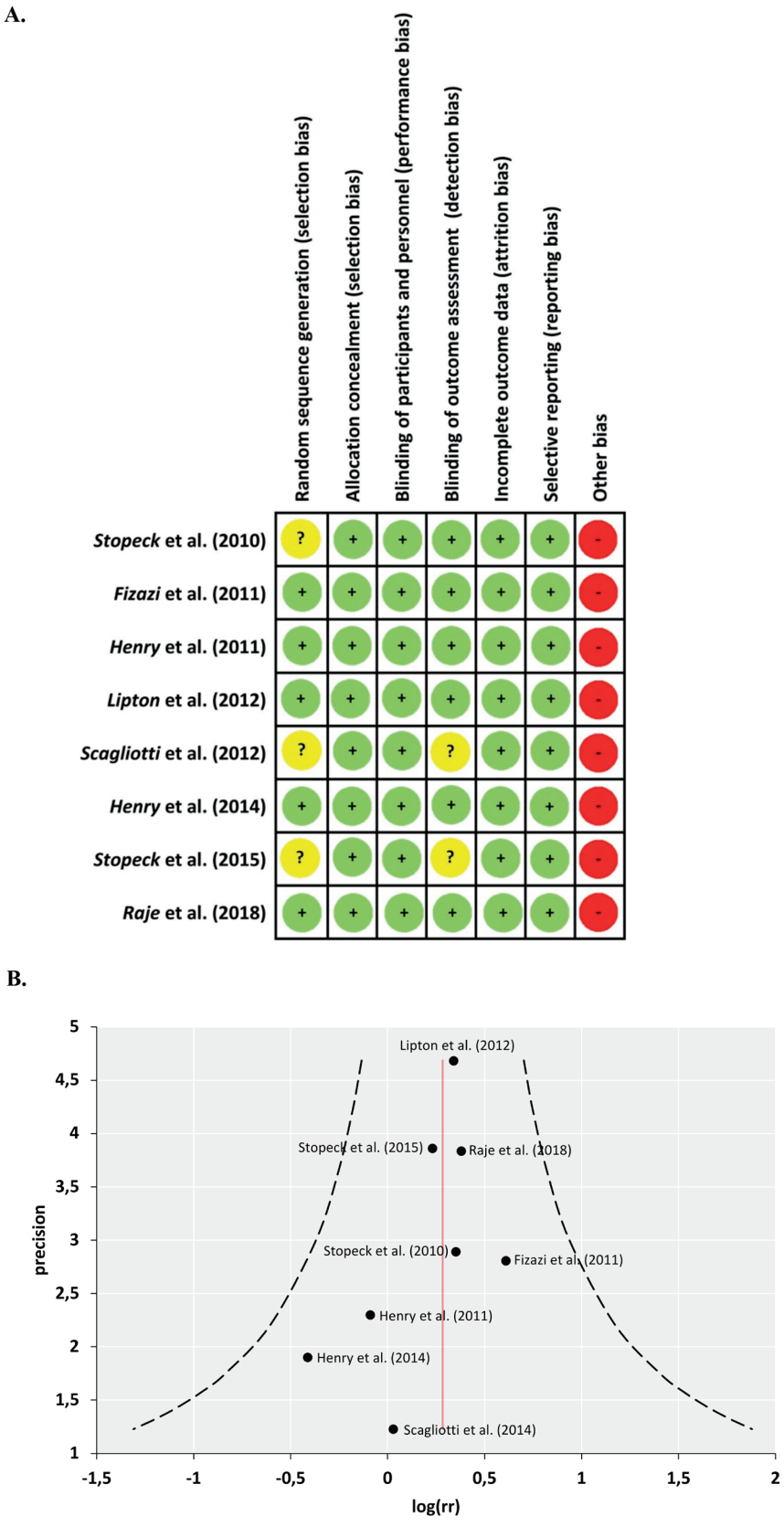


Fig. 2: Assessment of the risk of bias. A) The summary of the Cochrane Collaboration's tool for assessing risk of bias for the included studies. B) Evaluation of publication bias through Funnel plot.

Table 3: Tool for assessing risk of reporting biases.

	Level of Evidence	Overall Bias	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other bias
Stopeck *et al*. 2010	1	5,5	0,5	1	1	1	1	1	0
Fizazi *et al*. 2011	1	6	1	1	1	1	1	1	0
Henry *et al*. 2011	1	6	1	1	1	1	1	1	0
Lipton *et al*. 2011	1	6	1	1	1	1	1	1	0
Scagliotti *et al*. 2012	1	5	0,5	1	1	0,5	1	1	0
Henry *et al*. 2014	1	6	1	1	1	1	1	1	0
Stopeck *et al*. 2015	1	5	0,5	1	1	0,5	1	1	0
Raje *et al*. 2018	1	6	1	1	1	1	1	1	0
TOTAL		5,62							

Results

- Study selection

The electronic database search yielded 1.277 references (515 in MEDLINE, 534 in SCOPUS, 197 in WOS, 31 in CENTRAL and 0 in WHO ICTRP) and duplicates (n = 87), triplicates (n = 49) and quadruplicates (n = 24) were removed. After screening by title and abstract, 32 papers were included for full-text assessment. No articles from the manual search were considered for inclusion. 20 articles did not comply with inclusion criteria due to a different control group, and 4 articles did not provide sufficient information on MRONJ. Eight RCTs met our inclusion criteria (Fig. 1) (24-32). Nevertheless, two papers were excluded from the quantitative analysis for reporting on the same patient cohort (29,30,31). A Kappa value of 0.874 was obtained for title/abstract selection, 0.841 for full-text selection and more than 0.90 for data extraction, indicating a high degree of inter-rater reliability.

- Characteristics of the studies

A summary of the general information of the included studies is shown in Table 1. The dosage, frequency and route of administration of denosumab and ZA was concurrent for all included studies. As some studies presented different follow-up times, it was necessary to normalize the incidences to specific moments of the follow-up period. A 12, 24 and 36-month normalization was selected as intermediate follow-up times. Table 2 shows a summary of MRONJ incidences reported by the studies and their normalizations when it was necessary to apply.

- Risk of bias of included studies

The risk of bias for each study was determined using the Cochrane Collaboration Tool.

The studies showed an average risk of bias score of 5.62 out of 7, indicating a low risk of bias (Fig. 2) (Table 3). However, all studies were funded by the pharmaceutical company that manufactures denosumab, which may result in some kind of bias.

- Publication bias

Funnel plot asymmetry should be used when there are at least 10 studies included in a meta-analysis. As seen in Fig. 2, there is moderate symmetry between studies. However, publication bias may not be detected since more studies are needed so that the power of evidence is strong enough to distinguish chance from real asymmetry (22).

- Results of the effect

The quantitative analysis included 6 RCTs comprising a total of 13.857 patients, of whom 6.938 belonged to the denosumab group and 6.919 to the ZA group, allowing for analyses with great statistical power. The estimates of the individual studies were normalized to 12-, 24- and 36-months using a regression model according to MRONJ incidence rates and the follow-up periods reported in the individual studies included. This allowed to draw a graph of the course of the disease that represents DRONJ and BRONJ incidences over time (Fig. 3). When the pooled normalized incidence was calculated, statistical heterogeneity was detected for both groups. Therefore, a random-effects model was chosen. Nevertheless, a robust homogeneity was found when the RR or OR were calculated. A summary of the results of the meta-analysis is shown in Table 4. This analysis shows statistically significant differences between denosumab and ZA to develop MRONJ after 1 year (P = 0.030), 2 years (P = 0.006) and 3 years of exposure (P = 0.007). Nevertheless, no differences were found between the prognosis of DRONJ cases compared to cases due to ZA (P = 0.163). The individual studies reported a favourable prognosis varying from 18% to 50% for cases due to denosumab and 8% to 43% for ZA-related cases, within their respective observation periods.

Table 4: Meta-analysis results of the incidences and prognoses of MRONJ associated with zoledronic acid and Denosumab.

MRONJ Incidence	**Hererogeneity test**	**Model**	**Relative effect (95% CI)**	***p*-Value**	**Absolute Effect (95% CI)**	**Quality of the evidence (GRADE)**
BRONJ after 1 year with ZA	Q (df = 5) = 33.486 *p = 0.001* I^2 = 85.069 %	*Random-Effects*	0.008 (0.004 - 0.016)	*p = 0.001*		
DRONJ after 1 year with DB	Q (df = 5) = 48.701 *p = 0.001* I^2 = 89.733 %	*Random-Effects*	0.010 (0.005 - 0.021)	*p = 0.001*		
Comparison (measured with RR)	Q (df = 5) = 1.029 *p = 0.960* I^2 = 00.000 %	*Fixed-Effects*	RR 1.448 (1.037 - 2.023)	*p = 0.030*	4 more per 1.000 (1 - 9 more)	⊕⊕⊕∅ MODERATE due to risk of bias[a] and imprecision[b]
BRONJ after 2 years with ZA	Q (df = 5) = 22.891 *p = 0.001* I^2 = 78.158 %	*Random-Effects*	0.013 (0.008 - 0.021)	*p = 0.001*		
DRONJ after 2 years with DB	Q (df = 5) = 34.630 *p = 0.001* I^2 = 85.562 %	*Random-Effects*	0.019 (0.011 - 0.030)	*p = 0.001*		
Comparison (measured with RR)	Q (df = 5) = 1.029 *p = 0.960* I^2 = 00.000 %	*Fixed-Effects*	RR 1.448 (1.114 - 1.881)	*p = 0.006*	4 more per 1.000 (1 - 9 more)	⊕⊕⊕∅ MODERATE due to risk of bias[a] and imprecision[b]
BRONJ after 3 years with ZA	Q (df = 5) = 18.664 *p = 0.002* I^2 = 73.210 %	*Random-Effects*	0.015 (0.010 - 0.023)	*p = 0.001*		
DRONJ after 3 years with DB	Q (df = 5) = 31.193 *p = 0.001* I^2 = 83.971 %	*Random-Effects*	0.021 (0.013 - 0.032)	*p = 0.001*		
Comparison (measured with RR)	Q (df = 5) = 1.626 *p = 0.898* I^2 = 00.000 %	*Fixed-Effects*	RR 1.402 (1.095 - 1.795)	*p = 0.007*	3 more per 1.000 (1 - 7 more)	⊕⊕⊕∅ MODERATE due to risk of bias[a] and imprecision[b]
MRONJ Prognosis	**Hererogeneity test**	**Model**	**Values (95% CI)**	***p*-Value**		
Comparison (measured with OR)	Q (df = 4) = 0.631 *p = 0.889* I^2 = 00.000 %	*Fixed-Effects*	OR 1.730 (0.801 - 3.739)	*p = 0.163*	-	⊕∅∅∅ VERY LOW due to imprecision[c]

GRADE: Grading of Recommendations Assessment, Development and Evaluation; RR: Risk Ratio; OR: Odds Ratio; ZA: zoledronic acid; DB: denosumab; Downgraded initially from "high" to "moderate" due to: a) All included studies have been sponsored, supported or funded by Amgen, which may result in some kind of bias. b) The normalization of the different follow-up times provided approximated values. c) Insufficient information relative to prognosis.

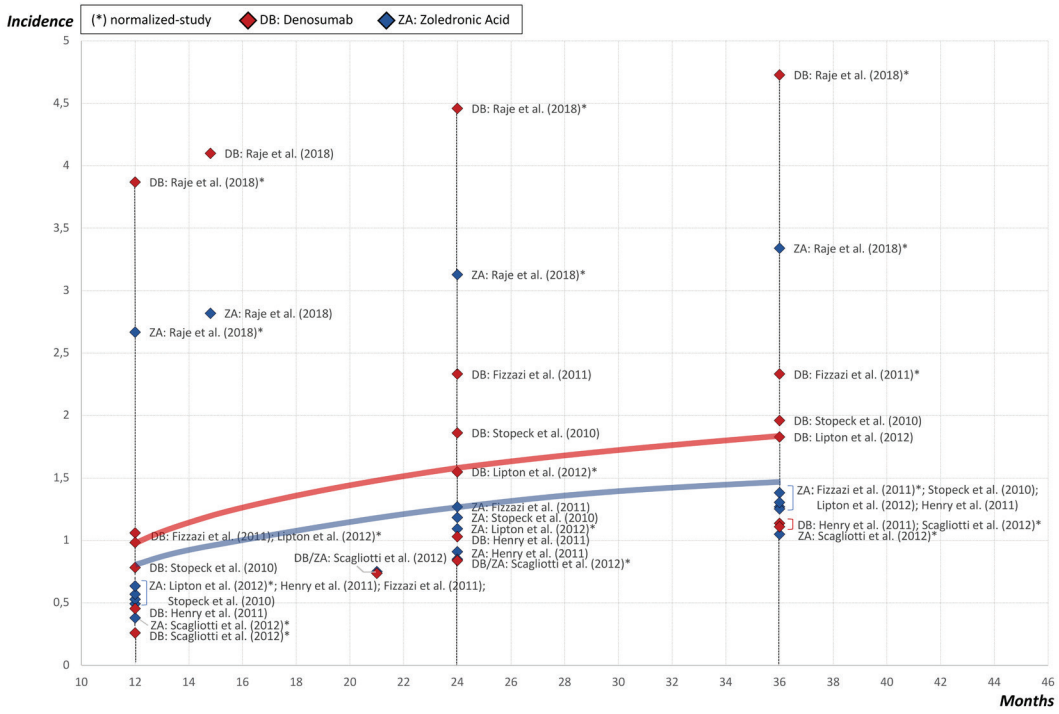


Fig. 3: The summary of the normalized values of the individual studies and diagram of the course of MRONJ during the first 3 years.

Discussion

MRONJ is an uncommon but emerging complication of antiresorptive and antiangiogenic therapy. This updated systematic review and meta-analysis, which to our knowledge, includes all published randomized-controlled clinical trials to date comparing denosumab with ZA as control, aims to determine the incidence of MRONJ due to denosumab and ZA, the RR of developing MRONJ either due to denosumab or ZA, as well as the prognosis of DRONJ and ZA-related BRONJ.

This study has determined that the incidence of DRONJ in cancer patients under treatment with denosumab ranged from 0.5 to 2.1% (50 to 210 cases per 10.000 patients) after 1 year, 1.1 to 3.0% (110 to 300 cases per 10.000 patients) after 2 years, and 1.3 to 3.2% (130 to 320 cases per 10.000 patients) after 3 years of exposure. After longer periods of exposure, Stopeck *et al*. reported DRONJ incidences of 6.9% (30,31). In this sense, as it can be seen in Fig. 3, the MRONJ incidence is expected to increase over the exposure time. The incidence of BRONJ in cancer patients under treatment with ZA ranged from 0.4 to 1.6% (40 to 160 cases per 10.000 patients) after 1 year, 0.8 to 2.1% (80 to 210 cases per 10.000 patients) after 2 years, and 1.0 to 2.3% (100 to 230 cases per 10.000 patients) after 3 years of exposure. Some authors, such as Ruggiero *et al*. reported that in cancer patients under treatment with ZA, the risk of developing BRONJ reaches 1% (100 cases per 10.000 patients) and in patients under treatment with denosumab the risk of developing DRONJ ranges from 0.7 to 1.9% (70 to 190 cases per 10.000 patients) (7). Others, in contrast to these results, have reported a higher MRONJ incidence linked to denosumab (10 %, P = 0.21) with a median follow-up time of 22.0 months, and also due to bisphosphonates (6.7%, P = 0.21) (33). This higher incidence could be associated with the fact that the study of Loyson *et al*. is specifically aimed at evaluating MRONJ cases, contrary to the studies included in this review that have other priority objectives.

Many clinicians switch therapy from bisphosphonates to denosumab for their patients due to its multiple advantages, such as superior prevention of SREs, subcutaneous administration instead of intravenous administration, and no dosage adjustment in case of renal insufficiency (33). Loyson *et al*. reported that the switch from bisphosphonates to denosumab can be considered as safe as having an equivalent exposure to denosumab from the start (33).

According to the results of this analysis, the use of denosumab is associated with a significantly higher risk of developing MRONJ compared to ZA at 1 year (RR 1.448 CI: 1.037-2.023), 2 years (RR 1.488 CI: 1.114-1.881) and 3 years of treatment (RR 1.402 CI: 1.095-1.795). Thus, clinicians should be aware of this risk, promoting preventive measures such as comprehensive oral examinations with appropriate radiographs, oral hygiene instructions, maintenance of good oral health, completion of necessary dental treatments before initi-

ating antiresorptive therapy etc. (34). Similarly, Owosho *et al*. reported that denosumab was associated with an earlier occurrence of MRONJ compared to ZA and pamidronate (35). Stavropoulos *et al*. consider that in patients under high doses of antiresorptive therapy, implant placement, explantation or their presence per se may trigger MRONJ (36). Most authors reported that MRONJ was strongly linked to dental extractions (45%) or poor oral hygiene (between 54 and 93%), simultaneous chemotherapy (between 60 and 75%), or the use of removable dentures with mucosal support (23,25,29-31). All of the causes mentioned above result in bacteria being in contact with susceptible bone or also in a decrease of the immune system in its role of protecting the oral mucosa from a bacterial-contaminated ecosystem.

No statistically significant differences were found in the prognosis of MRONJ cases due to denosumab or ZA (P = 0.163). However, individual studies reported that DRONJ cases have a slightly higher tendency to resolve, ranging from 18 to 50% compared to ZA, which ranges from 8% to 43%. This might be related to the reversibility mechanism inherent to denosumab, which is not found in bisphosphonates in general.

This systematic review and meta-analysis present certain limitations, such as the inclusion of different types of tumours, subcutaneously administered denosumab against intravenous ZA, the normalization of the different follow-up times providing approximated values or that all included randomized trials received some kind of funding by the pharmaceutical company that manufactures denosumab. On the other hand, the strengths of this study lie in the strict inclusion criteria used for homogeneous study selection, in the level of evidence provided by the selected RCTs, the low risk of bias of the individual studies, and in the sample sizes that allow for analyses with great statistical power.

Future RCTs should be carried out in more detail, measuring the real incidence per year, as well as the years after the exposure to medication for more clarity. Specifying further details could be of great value, such as whether oral examinations were carried out at enrolment or periodically, the intake of other concomitant drugs, the stage of MRONJ cases, and the number of patients undergoing dental surgical procedures along with the reason for surgery.

Conclusions

Based on the findings of the present study, the use of denosumab is associated with a significantly higher risk of developing MRONJ compared to ZA. Thus, clinicians should be aware of this risk, and adapt or plan for preventive measures such as thorough dental examinations and interventions prior to the initiation of medication, as well as oral health maintenance and avoidance of surgical procedures during active therapy.

References

1. Vos T, Abajobir AA, Abbafati C, Abbas KM, Abate KH, Abd-Allah F, *et al*. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390:1211-59.
2. Coleman RE, Roodman, Smith, Body, Suva, Vessella. Clinical features of metastatic bone disease and risk of skeletal morbidity. Clin Cancer Res. 2006;12:6243s-9s.
3. Aghaloo TL, Felsenfeld AL, Tetradis S. Osteonecrosis of the Jaw in a Patient on Denosumab. J Oral Maxillofac Surg. 2010;68:959-63.
4. Taylor KH, Middlefell LS, Mizen KD. Osteonecrosis of the jaws induced by anti-RANK ligand therapy. Br J Oral Maxillofac Surg. 2010;48:221-3.
5. Malan J, Ettinger K, Naumann E, Beirne OR. The relationship of denosumab pharmacology and osteonecrosis of the jaws. Oral Surg Oral Med Oral Pathol Oral Radiol. 2012;114:671-6.
6. Bagan L, Jimenez Y, Leopoldo M, Murillo-Cortes J, Bagan J. Exposed necrotic bone in 183 patients with bisphosphonate-related osteonecrosis of the jaw: Associated clinical characteristics. Med Oral Patol Oral Cir Bucal. 2017;22:e582-5.
7. Ruggiero SL, Dodson TB, Fantasia J, Goodday R, Aghaloo T, Mehrotra B, *et al*. American association of oral and maxillofacial surgeons position paper on medication-related osteonecrosis of the jaw -- 2014 update. J Oral Maxillofac Surg. 2014;72:1938-56.
8. King R, Tanna N, Patel V. Medication-related osteonecrosis of the jaw unrelated to bisphosphonates and denosumab - a review. Oral Surg Oral Med Oral Pathol Oral Radiol. 2019;127:289-99.
9. Yamagata K, Nagai H, Baba O, Uchida F, Kanno N, Hasegawa S, *et al*. A Case of Brain Abscess Caused by Medication-Related Osteonecrosis of the Jaw. Case Rep Dent. 2016;2016:7038618.
10. Aghaloo T, Hazboun R, Tetradis S. Pathophysiology of Osteonecrosis of the Jaws. Oral Maxillofac Surg Clin North Am. 2015;27:489-96.
11. McGowan K, McGowan T, Ivanovski S. Risk factors for medication-related osteonecrosis of the jaws: A systematic review. Oral Dis. 2018;24:527-36.
12. Lombard T, Neirinckx V, Rogister B, Gilon Y, Wislet S. Medication-Related Osteonecrosis of the Jaw: New Insights into Molecular Mechanisms and Cellular Therapeutic Approaches. Stem Cells Int. 2016;2016:8768162.
13. Zheng GZ, Chang B, Lin FX, Xie D, Hu QX, Yu GY, *et al*. Meta-analysis comparing denosumab and zoledronic acid for treatment of bone metastases in patients with advanced solid tumours. Eur J Cancer Care (Engl). 2017;26:1-7.
14. Hortobagyi GN, Zheng M, Mohanlal R. Indirect Evaluation of Bone Saturation with Zoledronic Acid After Long-Term Dosing. Oncologist. 2019;24:178-84.
15. Van Den Wyngaert T, Wouters K, Huizing MT, Vermorken JB. RANK ligand inhibition in bone metastatic cancer and risk of osteonecrosis of the jaw (ONJ): Non bis in idem? Support Care Cancer. 2011;19:2035-40.
16. Qi W-X, Tang L-N, He A-N, Yao Y, Shen Z. Risk of osteonecrosis of the jaw in cancer patients receiving denosumab: A meta-analysis of seven randomized controlled trials. Int J Clin Oncol. 2014;19:403-10.
17. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Int J Surg. 2010;8:336-41.
18. Zintzaras E, Ioannidis JPA. Heterogeneity testing in meta-analysis of genome searches. Genet Epidemiol. 2005;28:123-37.
19. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64:380-2.
20. Higgins JPT, Altman DG, Gotzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

21. Ockerman A, Bornstein MM, Leung YY, Li SKY, Politis C, Jacobs R. Incidence of bleeding after minor oral surgery in patients on dual antiplatelet therapy: a systematic review and meta-analysis. Int J Oral Maxillofac Surg. 2019;49:90-8.
22. Anzures-Cabrera J, Higgins JPT. Graphical displays for meta-analysis: An overview with suggestions for practice. Res Synth Methods. 2010;1:66-80.
23. Bax L, Yu L-M, Ikeda N, Tsuruta H, Moons KGM. Development and validation of MIX: Comprehensive free software for meta-analysis of causal research data. BMC Med Res Methodol. 2006;6:50.
24. Stopeck AT, Lipton A, Body J-J, Steger GG, Tonkin K, De Boer RH, *et al*. Denosumab compared with zoledronic acid for the treatment of bone metastases in patients with advanced breast cancer: A randomized, double-blind study. J Clin Oncol. 2010;28:5132–9.
25. Fizazi K, Carducci M, Smith M, Damião R, Brown J, Karsh L, *et al*. Denosumab versus zoledronic acid for treatment of bone metastases in men with castration-resistant prostate cancer: A randomised, double-blind study. Lancet. 2011;377:813-22.
26. Henry DH, Costa L, Goldwasser F, Hirsh V, Hungria V, Prausova J, *et al*. Randomized, double-blind study of denosumab versus zoledronic acid in the treatment of bone metastases in patients with advanced cancer (excluding breast and prostate cancer) or multiple myeloma. J Clin Oncol. 2011;29:1125-32.
27. Lipton A, Fizazi K, Stopeck AT, Henry DH, Brown JE, Yardley DA, *et al*. Superiority of denosumab to zoledronic acid for prevention of skeletal-related events: A combined analysis of 3 pivotal, randomised, phase 3 trials. Eur J Cancer. 2012;48:3082-92.
28. Scagliotti G V, Hirsh V, Siena S, Henry DH, Woll PJ, Manegold C, *et al*. Overall survival improvement in patients with lung cancer and bone metastases treated with denosumab versus zoledronic acid: Subgroup analysis from a randomized phase 3 study. J Thorac Oncol. 2012;7:1823-9.
29. Henry D, Vadhan-Raj S, Hirsh V, Von Moos R, Hungria V, Costa L, *et al*. Delaying skeletal-related events in a randomized phase 3 study of denosumab versus zoledronic acid in patients with advanced cancer: An analysis of data from patients with solid tumors. Support Care Cancer. 2014;22:679-87.
30. Stopeck AT, Fizazi K, Body J-J, Brown JE, Carducci M, Diel I, *et al*. Safety of long-term denosumab therapy: results from the open label extension phase of two phase 3 studies in patients with metastatic breast and prostate cancer. Support Care Cancer. 2016;24:447-55.
31. Stopeck AT, Fizazi K, Body J-J, Brown JE, Carducci M, Diel I, *et al*. Erratum to: Safety of long-term denosumab therapy: results from the open label extension phase of two phase 3 studies in patients with metastatic breast and prostate cancer. Support Care Cancer. 2016;24:447-55.
32. Raje N, Terpos E, Willenbacher W, Shimizu K, García-Sanz R, Durie B, *et al*. Denosumab versus zoledronic acid in bone disease treatment of newly diagnosed multiple myeloma: an international, double-blind, double-dummy, randomised, controlled, phase 3 study. Lancet Oncol. 2018;19:370-81.
33. Loyson T, Van Cann T, Schöffski P, Clement PM, Bechter O, Spriet I, *et al*. Incidence of osteonecrosis of the jaw in patients with bone metastases treated sequentially with bisphosphonates and denosumab. Acta Clin Belgica Int J Clin Lab Med. 2018;73:100-9.
34. Karna H, Gonzalez J, Radia HS, Sedghizadeh PP, Enciso R. Risk-reductive dental strategies for medication related osteonecrosis of the jaw among cancer patients: A systematic review with meta-analyses. Oral Oncol. 2018;85:15-23.
35. Owosho AA, Liang STY, Sax AZ, Wu K, Yom SK, Huryn JM, *et al*. Medication-related osteonecrosis of the jaw: An update on the memorial sloan kettering cancer center experience and the role of premedication dental evaluation in prevention. Oral Surg Oral Med Oral Pathol Oral Radiol. 2018;125:440-5.
36. Stavropoulos A, Bertl K, Pietschmann P, Pandis N, Schiødt M, Klinge B. The effect of antiresorptive drugs on implant therapy: Systematic review and meta-analysis. Clin Oral Implants Res. 2018;29:54-92.

Acknowledgments
The authors express their gratitude to Mr. Santiago Cano, Centre of Data Processing, Computing Service for Research Support, University Complutense of Madrid, for his invaluable collaboration and expertise for the statistics provided for this review and Dr. Ruth Sánchez Sánchez, Department of Oral Surgery and Implantology, San Pablo CEU University, for her support in this study.

Funding
This research did not receive any grant support from public, commercial or non-profit funding agencies. The study was supported by departmental funds only.

Conflicts of interest
The authors received no financial support and declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

ANEXO 1.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II).

PRIMER ARTÍCULO: Limones, Saez-Alcaide, et al., 2020			
Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]	**Respuesta**
Pregunta 1	-	Pregunta de investigación y criterios de inclusión: ¿Se establecieron preguntas de investigación y criterios de inclusión que incluyeron los componentes de PICO (Población, Intervención, Grupo comparador, Resultado)?	Si
Pregunta 2	**CRÍTICA**	**Protocolo registrado antes del inicio de la revisión:** ¿Se registró o publicó el protocolo de la revisión antes de empezar su realización y se han justificado las desviaciones significativas del protocolo?	**No** (No se registró el protocolo con anterioridad al comienzo del estudio, que reflejase la existencia de posibles desviaciones)
Pregunta 3	-	Diseño de estudios seleccionados: ¿Se explica por los autores cual es el diseño de estudio que se ha seleccionado para confeccionar la revisión?	Si
Pregunta 4	**CRÍTICA**	**Búsqueda exhaustiva de la literatura:** ¿Los autores de la revisión utilizaron una estrategia de búsqueda exhaustiva?	**Si**
Pregunta 5	-	Duplicación en la selección de estudios: ¿Los autores de la revisión realizaron la selección de estudios por duplicado?	Si
Pregunta 6		Duplicación en la extracción de datos: ¿Los autores de la revisión realizaron la extracción de datos por duplicado?	Si
Pregunta 7	**CRÍTICA**	**Justificación para excluir estudios individuales:** ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?	**Parcialmente** (Incluida en el diagrama de flujo y descriptivamente en la sección de resultados)
Pregunta 8	-	Descripción de estudios incluidos: ¿Los autores de la revisión describieron los estudios incluidos con suficiente detalle?	Si
Pregunta 9	**CRÍTICA**	**Evaluación del riesgo de sesgo de los estudios primarios incluidos:** ¿Utilizaron los autores de la revisión una técnica satisfactoria para evaluar el riesgo de sesgo (RoB) en los estudios individuales que se incluyeron en la revisión?	**Si**

© Álvaro Limones Burgos, 2024

Pregunta 10	-	Fuentes de financiación: ¿Los autores de la revisión informaron sobre las fuentes de financiación de los estudios incluidos en la revisión?	Si
Pregunta 11	**CRÍTICA**	**Adecuados métodos analíticos:** Si se realizó un metaanálisis, ¿los autores de la revisión utilizaron métodos apropiados para la combinación estadística de resultados?	**Si**
Pregunta 12	-	Impacto del RoB en la metaanálisis: Si se realizó un metaanálisis, ¿los autores de la revisión evaluaron el impacto potencial de RoB de los estudios individuales sobre los resultados del metaanálisis u otra síntesis de evidencia?	Si
Pregunta 13	**CRÍTICA**	**Impacto del RoB en los resultados y discusión:** ¿Los revisores tuvieron en cuenta el RoB de los estudios primarios al interpretar/discutir los resultados de la revisión?	**Si**
Pregunta 14	-	Evaluación de la Heterogeneidad: ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?	**No aplica** (No fue detectada la presencia de heterogeneidad en las comparaciones)
Pregunta 15	**CRÍTICA**	**Evaluación de la presencia e impacto del sesgo de publicación:** Si se realizó una síntesis cuantitativa, ¿se llevó a cabo una investigación adecuada del sesgo de publicación (sesgo de estudios pequeños) y se discutió su probable impacto en los resultados de la revisión?	**Si**
Pregunta 16	-	Conflictos de interés: ¿Se informaron las posibles fuentes de conflicto de interés, incluyendo cualquier financiamiento recibido para llevar a cabo la revisión?	Si
VALORACIÓN FINAL DEL RIESGO DE SESGO: **BAJA CONFIANZA**			

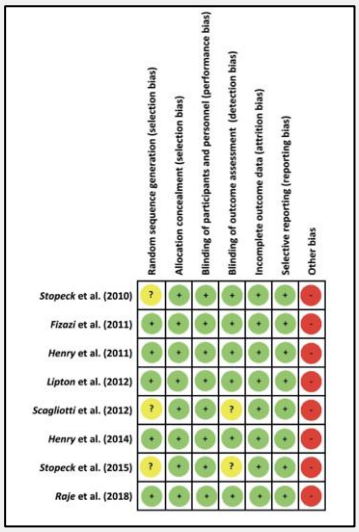
© Álvaro Limones Burgos, 2024

ANEXO 1.2: Resultados de la evaluación del riesgo de sesgo (ROBIS).

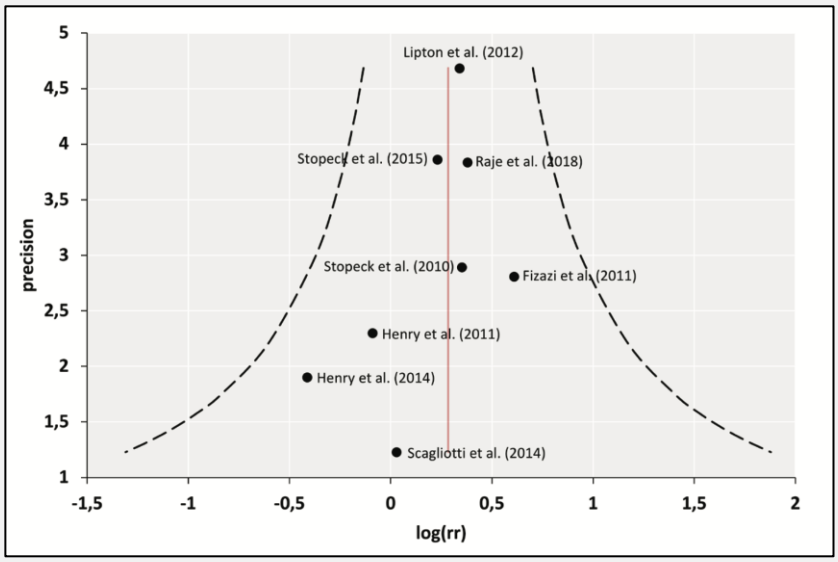
1ª FASE	2ª FASE				3ª FASE
Identificación de la pregunta de investigación y sus componentes	**Criterios de elegibilidad (Dominio 1)**	**Identificación y selección de estudios (Dominio 2)**	**Recopilación de datos y evaluación de estudios (Dominio 3)**	**Síntesis y conclusiones (Dominio 4)**	**Riesgo de sesgo global**
¿La pregunta abordada en la revisión está correctamente definida de acuerdo a sus objetivos? S/PS/PN/N/NI* Revisiones de intervención: *P: Población* *I: Intervención* *C: Comparador* *O: Variables* Revisiones de pruebas de precisión diagnóstica: *P: Población* *I: Test diagnóstico* *C: Estándar de referencia* *O: Enfermedad objetivo* Revisiones de asociaciones etiológicas o factores de riesgo: *P: Población* *I: Exposición* *C: Comparador* *O: Variables* Revisiones de estudios pronóstico: *P: Población* *I: Factor pronóstico o modelo predictivo* *O: Desenlace a predecir* Revisiones de variables e instrumentos de medición: *P: Población/Campo de conocimiento* *O: Variables e instrumentos de medición*	1.1 ¿Se ajustó la revisión a los objetivos y criterios de elegibilidad predefinidos? S/PS/PN/N/NI* 1.2 ¿Los criterios de elegibilidad eran apropiados para la pregunta de la revisión? S/PS/PN/N/NI* 1.3 ¿Eran inequívocos los criterios de elegibilidad? S/PS/PN/N/NI* 1.4 ¿Eran apropiadas las restricciones de los criterios de elegibilidad basadas en las características del estudio (por ejemplo, fecha, tamaño de la muestra, calidad del estudio, resultados medidos, etc.)? S/PS/PN/N/NI* 1.5 ¿Era adecuada alguna restricción en los criterios de elegibilidad basada en las fuentes de fuentes de información (por ejemplo, estado o formato de publicación, idioma, disponibilidad de datos)? S/PS/PN/N/NI*	2.1 ¿Incluyó la búsqueda una gama apropiada de bases de datos/fuentes electrónicas para informes publicados y no publicados? S/PS/PN/N/NI* 2.2 ¿Se utilizaron métodos adicionales a la búsqueda en bases de datos para identificar informes relevantes? S/PS/PN/N/NI* 2.3 ¿Los términos y la estructura de la estrategia de búsqueda permitieron recuperar el mayor número posible de estudios elegibles? S/PS/PN/N/NI* 2.4 ¿Fueron apropiadas las restricciones basadas en la fecha, el formato de publicación o el idioma? S/PS/PN/N/NI* 2.5 ¿Se realizaron esfuerzos para minimizar los errores en la selección de los estudios? S/PS/PN/N/NI*	3.1 ¿Se hicieron esfuerzos para minimizar los errores en la recogida de datos? S/PS/PN/N/NI* 3.2 ¿Se dispuso de suficientes características de los estudios para que tanto los autores de la revisión como los lectores pudieran interpretar los resultados? S/PS/PN/N/NI* 3.3 ¿Se recogieron todos los resultados relevantes de los estudios para su uso en la síntesis? S/PS/PN/N/NI* 3.4 ¿Se evaluó formalmente el riesgo de sesgo (o la calidad metodológica) utilizando criterios apropiados? S/PS/PN/N/NI* 3.5 ¿Se hicieron esfuerzos para minimizar los errores en la evaluación del riesgo de sesgo? S/PS/PN/N/NI*	4.1 ¿Incluyó la síntesis todos los estudios que debía? S/PS/PN/N/NI* 4.2 ¿Se comunicaron todos los análisis predefinidos o se explicaron las desviaciones? S/PS/PN/N/NI* 4.3 ¿Fue adecuada la síntesis dada la naturaleza y la similitud de las preguntas de investigación, los diseños de los estudios y los resultados de los estudios incluidos? S/PS/PN/N/NI* 4.4 ¿La variación entre estudios (heterogeneidad) fue mínima o se abordó en la síntesis? S/PS/PN/N/NI* 4.5 ¿Fueron sólidos los hallazgos, por ejemplo, como se demostró mediante el gráfico en embudo o los análisis de sensibilidad? S/PS/PN/N/NI* 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? S/PS/PN/N/NI*	A. ¿Se han abordado en la interpretación de los resultados todos los problemas señalados en los dominios del uno al cuatro? S/PS/PN/N/NI* B. ¿Se consideró adecuadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? S/PS/PN/N/NI* C. ¿Evitaron los revisores destacar los resultados en función de su significación estadística? S/PS/PN/N/NI*
¿Coincide la pregunta abordada por la revisión con la pregunta objetivo? SI/INCIERTO/NO	**Preocupación por la especificación de los criterios de elegibilidad del estudio** ALTA/INCIERTA/BAJA	**Preocupación por los métodos utilizados para identificar y/o seleccionar los estudios** ALTA/INCIERTA/BAJA	**Preocupación por los métodos de recogida de datos y evaluación de los estudios** ALTA/INCIERTA/BAJA	**Preocupaciones sobre la síntesis y los resultados** ALTA/INCIERTA/BAJA	**Riesgo de sesgo global considerado** ALTO/INCIERTO/BAJO/INEXISTENTE

© Álvaro Limones Burgos, 2024

ANEXO 1.3: Resultados de la evaluación de la certeza de la evidencia científica (GRADE).

Dominios GRADE	Consideraciones	Variación del nivel de certeza de la evidencia
Riesgo de sesgo [Dominio 1]	La herramienta original para evaluar el riesgo de sesgo (Higgins et al., 2011) no reveló riesgos significativos de sesgo en las dimensiones evaluadas. Sin embargo, se identificó una excepción notable: la financiación de todos los estudios clínicos analizados provino de la compañía que distribuye el tratamiento en estudio (Denosumab). Este hecho podría haber afectado a los resultados, especialmente para una RS que está centrada en la detección de efectos adversos del tratamiento. Los resultados de esta evaluación se presentan en la siguiente figura: 	↓ 1 nivel a ambas variables críticas evaluadas en la revisión
Inconsistencia [Dominio 2]	No se detectó heterogeneidad no explicada en el metaanálisis para ninguna de las variables estudiadas: Incidencia de MRONJ – Test de heterogeneidad: $X^2 = 1.029$; $P = 0.960$; $I^2 = 0.00\ \%$) Resolución de MRONJ – Test de heterogeneidad: $X^2 = 0.631$; $P = 0.889$; $I^2 = 0.00\ \%$) Por lo que no fue necesario realizar análisis de sensibilidad, análisis de subgrupos o metarregresiones.	No aplica sobre ninguna de las variables
Carácter indirecto de la evidencia [Dominio 3]	La variable primaria y secundaria evaluó la evidencia de la manera más directa posible, mediante el recuento de casos diagnosticados o resueltos de MRONJ.	No aplica sobre ninguna de las variables
Imprecisión [Dominio 4]	A pesar de que los intervalos de confianza de los estudios primarios se superpusieron unos con los otros en el metaanálisis de incidencia de MRONJ, y el efecto combinado del efecto y su intervalo de confianza no cruza la línea de no efecto ni los umbrales de decisión o relevancia clínica del efecto, no se superó el tamaño optimo de información (OIS) establecido en 9.581 participantes por grupo, y calculado a partir de los resultados del estudio de Lipton et al., 2012, con una proporción de muestreo de 1:1, un nivel de significancia (alfa) del 5%, un poder estadístico del 80% y una prueba de hipótesis de χ^2 para muestras independientes. Por esta razón se degradó un nivel por imprecisión. De acuerdo con el enfoque GRADE en variables dicotómicas una muestra inferior a 2000 participantes podría requerir degradar la certeza de la evidencia científica. La variable resolución de los casos de MRONJ reunió un total de 127 pacientes, por lo que se decidió degrar 3 niveles por imprecisión.	↓ 1 nivel en la variable primaria (incidencia de MRONJ) ↓ 3 niveles para la variable secundaria (resolución de MRONJ)

© Álvaro Limones Burgos, 2024

Sesgo de Publicación [Dominio 5]	No se detectaron asimetrías en el diagrama de embudo para la variable principal. Los resultados de esta evaluación se presentan en la siguiente figura:  Tampoco se detectó presencia de sesgo de publicación para la variable principal en la Prueba de Egger (P=0.704). No fue posible estudiar el sesgo de publicación en la variable secundaria por ineficiente número de ECA (menor a cinco estudios).	No aplica sobre ninguna de las variables
Gradiente dosis-respuesta [Dominio 6]	La incidencia de MRONJ tiene un gradiente dosis-respuesta, soportado por evidencia indirecta (Yan et al., 2022)y consistencia en estudios sobre pacientes oncológicos (mayor dosis, mayor incidencia) y osteoporóticos (menor dosis, menor incidencia) (Ruggiero et al., 2022).	↑ 1 nivel en la variable primaria (incidencia de MRONJ)
Gran magnitud del efecto [Dominio 7]	La magnitud combinada del efecto no presento una gran magnitud con un RR inferior 2 en ambas variables.	No aplicable
Factores de confusión plausibles [Dominio 8]	No se detectaron factores de confusión plausible que aumentase nuestra confianza en las estimaciones, en ninguna de las variables estudiadas	No aplicable
Computo total de la evaluación	Tras la disminución de la certeza en dos niveles (debido al riesgo de sesgo e imprecisión) y el incremento de la certeza en un nivel (por la presencia de un gradiente dosis-respuesta) se optó por reducir la certeza de la evidencia para la variable primaria (incidencia de MRONJ) en 1 nivel. Se redujo la certeza de la evidencia para la variable secundaria (resolución de MRONJ) 3 niveles.	↓ 1 nivel en la variable principal (incidencia de MRONJ) ↓ 3 niveles en la variable secundaria (resolución de MRONJ)

© Álvaro Limones Burgos, 2024

ANEXO 1.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA).

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
TÍTULO			
Título	1	Identificar en el título del manuscrito que se trata de una revisión sistemática.	p. 326
RESUMEN			
Objetivos	2.1	Proporcionar una declaración explícita de los principales objetivos de la revisión.	p.326
Criterios de Inclusión y exclusion	2.2	Especificar los criterios de inclusión y exclusión de la revisión.	/
Fuentes de información	2.3	Especificar las fuentes de información (por ejemplo, las bases de datos o registros) utilizados para identificar los estudios y la fecha de la última búsqueda.	/
Riesgo de sesgo	2.4	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos.	/
Métodos de la síntesis de los resultados	2.5	Especifique los métodos utilizados para presentar y sintetizar los resultados.	/
Estudios incluidos	2.6	Dar el número total de estudios y participantes incluidos y resumir las características pertinentes de los estudios.	p.326
Síntesis de los resultados	2.7	Indicar las variables de estudio. Si se ha realizado un metaanálisis, notifique la estimación del efecto o exposición y su IC. Si se comparan dos o más grupos, indíquen la dirección del efecto (es decir, que grupo es favorecido).	p. 327 *(Parcialmente – no informa de IC)*
Degradación de la evidencia	2.8	Proporcionar un breve resumen de la degradación de la evidencia (por ejemplo, por riesgo de sesgo, incoherencia o impresión.	/
Interpretación	2.9	Proporcionar una interpretación general de los resultados e implicaciones importantes.	p.327
Financiación	2.10	Especificar la principal fuente de financiación de la revisión.	/
Registro	2.11	Proporcionar el nombre del registro y el número de registro.	/
INTRODUCIÓN			
Fundamento	3	Describir la justificación de la revisión en el contexto de los conocimientos existentes.	p.327
Objetivos	4	Proporcionar una declaración explícita de los objetivos de la revisión.	p.327
MÉTODOS			
Criterios de elegibilidad	5	Especificar los criterios de inclusión y exclusión de la revisión y cómo se agruparon los estudios para la síntesis.	p.327-8
Fuentes de información	6	Especificar todas las bases de datos, registros, sitios web, organizaciones, listas de referencia y otras fuentes consultadas para identificar estudios. Especificar la fecha de la última revisión.	p.328
Estrategia de búsqueda	7	Presentar las estrategias de búsqueda completa para al menos la base de datos principal, indicar filtros u otras estrategias de búsqueda utilizadas.	p.327
Proceso de selección	8	Especificar los métodos utilizados para la selección de estudios.	p.329

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Recogida de datos	9	Especifique los métodos utilizados para recopilar datos de los estudios, incluyendo cuántos revisores recogieron datos de cada informe.	p. 329
Datos extraidos	10a	Listar y definir todas las variables críticas.	p.329
	10b	Listar y definir todas las demás variables secundarias. Describa cualquier suposición sobre cualquier información que falta o no esté clara.	p.330
Métodos de evaluación del riesgo de sesgo	11	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos, incluyendo detalles de la herramienta o herramientas utilizadas, cuántos revisores evaluaron cada estudio y si trabajaban independientemente, y si se utilizaron procesos de automatización.	p.330
Medidas del efecto	12	Especifícar para cada variable la medida(s) del efecto (por ejemplo. relación de riesgo, diferencia media) utilizada en la síntesis o presentación de resultados.	p.330
Métodos de síntesis	13a	Describir los procesos utilizados para decidir qué estudios eran elegibles para cada síntesis (por ejemplo: tabular las características de intervención del estudio y comparar con los grupos previstos para cada síntesis).	p.329
	13b	Describir cualquier método necesario para preparar los datos de presentación o síntesis, como el tratamiento de las estadísticas de datos faltantes o las conversiones de datos.	p.330
	13c	Describir cualquier método utilizado para tabular o visualizar los resultados de los estudios y la síntesis individuales.	p.330
	13d	Describir cualquier método utilizado para sintetizar los resultados y proporcionar una justificación. Si se realizó un metaanálisis, describa el(s) modelo(s), método(s) para identificar la presencia y extensión de la heterogeneidad estadística, y paquetes de software utilizado.	p.334
	13e	Describa cualquier método utilizado para explorar posibles causas de heterogeneidad entre los resultados del estudio (por ejemplo, análisis de subgrupos, metaregresión).	/
	13f	Describir cualquier análisis de sensibilidad realizado para evaluar la solidez de los resultados sintetizados.	/
Implicaciones del riesgo de sesgo por datos faltantes	14	Describir cualquier método han utilizado para evaluar el riesgo de sesgo debido a la falta de datos (a partir de sesgos informativos).	p. 332
Evaluación de la certeza de la evidencia	15	Describir cualquier método utilizado para evaluar la certeza (o la confianza) de la evidencia.	p.333
RESULTADOS			
Selección de estudios	16a	Describir los resultados del proceso de búsqueda y selección, idealmente utilizando un diagrama de flujo.	p. 332
	16b	Presentar los estudios excluidos y sus razones.	p. 332
Características de los estudios incluidos	17	Presentar cada estudio incluido y sus características.	p. 329
Riesgo de sesgo en los estudios incluidos	18	Presentar el riesgo de sesgo para cada estudio incluido.	p. 331

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Resultados de los estudios individuales	19	Para todas las variables presentar la estimación del efecto y su IC.	p. 333
Resultados de la síntesis	20a	Resumir brevemente las características y el riesgo de sesgo de los estudios incluidos.	p. 331-2
	20b	Si se ha realizado un metaanálisis, presente para cada una de las estimaciones medidas de heterogeneidad estadística. Si se comparan los grupos, describa la dirección del efecto.	p. 332-3
	20c	Informar de posibles causas de heterogeneidad.	/
	20d	Informar de los análisis de sensibilidad realizados para evaluar la solidez de los resultados sintetizados.	/
Evaluación del riesgo de sesgo por datos faltantes	21	Informar de la evaluación del riesgo de sesgo debido a datos faltantes.	p.332
Certeza de la evidencia	22	Evaluaciones de la certeza (o confianza) de la evidencia para cada variable evaluada.	p.333
DISCUSIÓN			
Discusión	23a	Proporcionar una interpretación general de los resultados en el contexto científico actual.	p.334
	23b	Examinar cualquier limitación de la evidencia incluida en la revisión.	p.335
	23c	Debatir las limitaciones de los procesos de revisión utilizados.	/
	23d	Debatir las implicaciones de los resultados de la práctica, la política y la futura investigación.	p.335
OTRA INFORMACIÓN			
Registro y protocolo	24a	Proporcionar información del registro de la revisión, incluyendo el nombre del registro y el número de registro, o afirmar que la revisión no está registrada.	/
	24b	Indicar dónde puede accederse al protocolo de revisión o afirmar que no se ha elaborado un protocolo.	/
	24c	Describir y explicar las modificaciones de la información facilitada en el registro o en el protocolo.	/
Apoyo	25	Describa las fuentes de ayuda financiera o no financiera para la revisión, y el papel de los patrocinadores en la revisión.	p.336
Intereses de competencia	26	Declare cualquier interés competitivo de los autores de revisión.	p.336
Disponibilidad de datos, código y otros materiales	27	Informe sobre cuál de los siguientes está disponible públicamente y dónde se pueden encontrar: plantillas de recogida de datos; datos extraídos de estudios incluidos; datos utilizados para todos los análisis; o cualquier otro material utilizado en la revisión.	/

© Álvaro Limones Burgos, 2024

ANEXO II: Manuscrito de la segunda revisión sistemática

(Limones, Molinero-Mourelle, et al., 2020)

© Álvaro Limones Burgos, 2024

JADA
THE JOURNAL OF THE AMERICAN DENTAL ASSOCIATION
APRIL 2020

CSA CORNER:
Human Papillomavirus Vaccine 303

ORIGINAL CONTRIBUTIONS

Dental Treatment and Elevated Blood Pressure 239

Powered Toothbrush Efficacy 265

Radiation Exposure Reduction Methods 287



COVER STORY

Posterior Multiunit Tooth-Supported Fixed Dental Prostheses: A Systematic Review 230

April 2020
Volume 151, Issue 4

Cover Story

Check for updates

ADA CE Online
ADA American Dental Association

Supplemental material is available online.

Zirconia-ceramic versus metal-ceramic posterior multiunit tooth-supported fixed dental prostheses

A systematic review and meta-analysis of randomized controlled trials

Alvaro Limones, DMD; Pedro Molinero-Mourelle, DMD, MSc; Luis Azevedo, DMD, MSc; Marta Romeo-Rubio, DMD, PhD; André Correia, DMD, PhD; Miguel Gómez-Polo, DMD, PhD

ABSTRACT

Background. The authors aimed to compare the survival and complication rates of zirconia-ceramic (ZC) versus metal-ceramic (MC) restorative material in multiunit tooth-supported posterior fixed dental prostheses (FDP).

Types of Studies Reviewed. The authors conducted a systematic search of randomized controlled trials (RCTs), with no time or language restrictions, up to May 2019 using the MEDLINE (PubMed), Scopus, Web of Science, and Cochrane Central Register of Controlled Trials databases, followed by a manual search.

Results. The authors included 7 RCTs in the review and 5 RCTs in the meta-analysis. All studies had a low risk of bias. The authors included 330 participants (177 ZC and 173 MC tooth-supported FDP) in the meta-analysis, which revealed a medium-term survival rate of 95.4% (95% confidence interval [CI], 90.5% to 99.1%) for ZC FDP compared with 96.9% (95% CI, 94.3% to 99.4%) for MC FDP, with no significant differences ($P = .364$). The biological or technical complications did not show statistically significant differences, except in the global ceramic veneering chipping analysis ($P = .023$; risk difference [RD], 22.3%; 95% CI, 3.0% to 41.6%) and their subanalysis: minor chipping or chipping that can be solved with polishing ($P = .044$; RD, 19.5%; 95% CI, 0.5% to 38.4%), and major chipping or chipping that needs repair in the laboratory ($P = .023$; RD, 6.0%; 95% CI, 0.8% to 11.3%).

Conclusions and Practical Implications. Posterior multiunit ZC restorations are considered a predictable treatment in the medium term, although they are slightly more susceptible to chipping of the veneering ceramic than MC restorations.

Key Words. Dental prosthesis; zirconium oxide; yttria-stabilized tetragonal zirconia; metal-ceramic alloys; survival rate; biological complications; technical complications.

JADA 2020:151(4):230-238
https://doi.org/10.1016/j.adaj.2019.12.013



This article has an accompanying online continuing education activity available at: http://jada.ada.org/ce/home.

Copyright © 2020 American Dental Association. All rights reserved.

Restorations based on ceramic systems have been gaining in importance because of their good esthetic and biological results, as well as an increase in the esthetic demands of patients and dentists. These restorations, as well as the metal-ceramic (MC) ones, have shown good survival rates after more than 10 years of use and gradually have become a preferred alternative and a predictable treatment compared with MC restorations.[1-3] This demand promoted the development of different ceramic systems for FDP, of which the yttria-stabilized tetragonal zirconia polycrystal is 1 of the most used owing to its good esthetic, biological, and mechanical properties.[4,5] However, these restorations are not free of complications. Chipping of the veneering ceramic is the most common complication, followed by the fracture of the framework.[6]

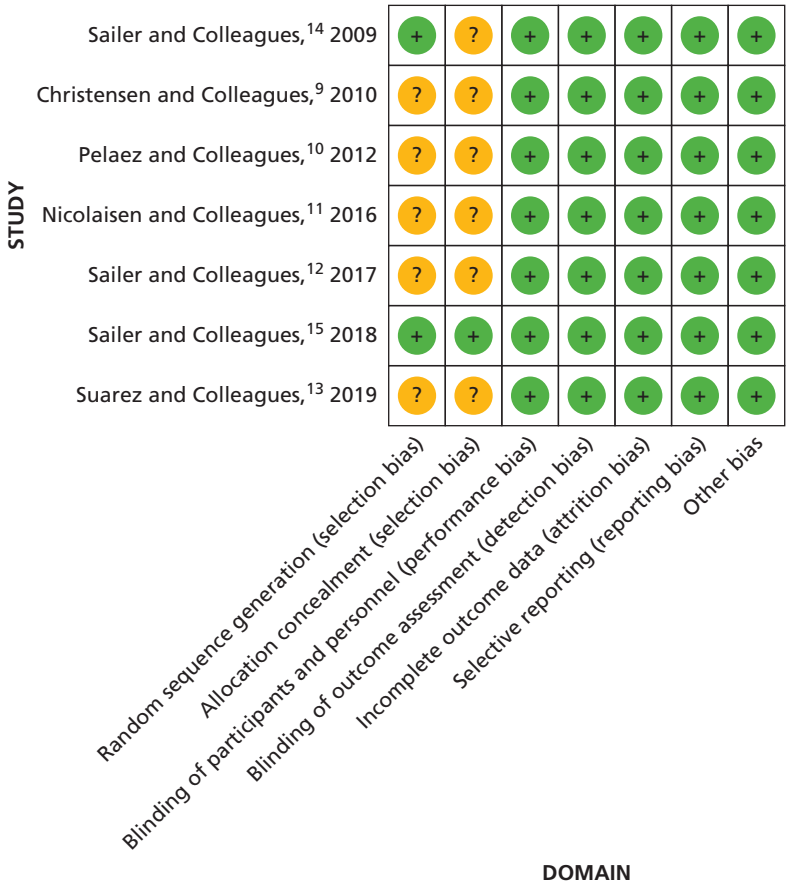


Figure 1. Risk of bias summary. Review authors' judgment about each risk of bias item for each included study. Red (-) corresponds to high risk, green (+) to low risk and yellow (?) to uncertain risk of bias.

Although monolithic zirconia restorations have gained in popularity since 2010, veneered zirconia FDP have been widely used. Therefore, clinical studies on these restoration with medium- and long-term outcomes should be evaluated.[7] Accordingly, our aim in this article was to systematically review the existing randomized controlled clinical trials (RCTs) for evidence on the clinical survival and the biological, technical, and esthetic complications of zirconia-ceramic (ZC) multiunit tooth-supported posterior FDP.

METHODS

Study registration

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines[8] and registered the review protocol at PROSPERO International Prospective Register of Systematic Reviews (CRD42019126096) hosted by the Centre for Reviews and Dissemination, University of York, National Institute for Health Research, United Kingdom.

Focused question

We developed the review protocol according to the patient, intervention, comparison, and outcomes form as follows:

- population: patients who are partially edentulous in the posterior sectors;
- intervention: tooth-supported posterior FDP of 3 through 5 units with ZC as a restoration material;
- comparison: tooth-supported posterior FDP of 3 through 5 units with MC as a restorative material;
- outcome: technical, biological, and esthetic complications and survival rates;
- studies: RCTs.

ABBREVIATION KEY

CENTRAL: Cochrane Central Register of Controlled Trials.
FDP: Fixed dental prostheses.
MC: Metal-ceramic.
RCT: Randomized clinical trial.
USPHS: US Public Health Service.
WOS: Web of Science.
ZC: Zirconia-ceramic.

The focused question of our review was as follows: In partially edentulous patients with tooth-supported posterior FDP, are the survival and complications rates of ZC restorative material different from those of MC restorative material?

Inclusion criteria

The inclusion criteria were RCTs comparing ZC and MC restorative material in 3-unit to 5-unit posterior FDPs with a follow-up of at least 3 years that reported data on survival rates and technical, biological, and esthetic complication rates.

Exclusion criteria

We excluded RCTs that did not meet the inclusion criteria, observational studies, case series, case reports, finite element analyses, in vitro studies, and animal studies.

Sources of information

We conducted the electronic search up through May 2019 in 4 databases: National Library of Medicine (MEDLINE [PubMed]) via Ovid, Web of Science (WOS), Scopus, and Cochrane Central Register of Controlled Trials (CENTRAL). We did not apply any time or language restrictions. The search strategy is shown in Appendix 1, available online at the end of this article. We contacted the researchers of the included studies to obtain relevant information when necessary.

Study selection

After the electronic search in the databases without using any filters, we saved the references in a Zotero library. Two reviewers (A.L., P.M.-M.) independently performed the study selection through the title for each database, and the level of agreement between the reviewers was estimated using κ statistics (Figure 1).[9-15] An agreement higher than 0.80 corresponded to a substantial agreement between reviewers. After removing the duplicates and triplicates, the reviewers checked the titles and abstracts of the remaining articles for adequacy and excluded irrelevant studies (for example, observational studies). If the relevance of a study was not clear from the title and abstract, it was included for full-text screening.

Data collection and items

Two reviewers (A.L., L.A.) working together collected data from each included article and recorded them in an Excel spreadsheet (Version 15.17, Microsoft). Any potential disagreement was resolved by consensus (A.L., L.A., P.M.-M., M.R.-R., M.G.-P.). The following items were recorded: name of the first author, year of publication, number of participants, sex, age, follow-up, number of dropouts, number of posterior units, restoration material, veneering ceramic used, manufacturing technique, cement used, biological complications, technical complications, and survival rates (eTables 1-3b, available online at the end of this article).[9-15] Additional information was obtained by contacting the investigators when indicated.

Summary measures and synthesis of results

The biological complications measured in the abutment teeth included probing pocket depth, probing attachment level, absence or presence of plaque (plaque control record), and bleeding on probing. We considered these complications to be continuous outcomes and calculated them using standard mean difference values and visualized them by means of a forest plot. We considered the loss of tooth vitality a dichotomous outcome and evaluated it according to risk difference. The data extraction regarding biological complications is shown in eTable 2, available online at the end of this article.[9-15]

The technical complications examined were core fracture, ceramic veneering chipping, occlusal wear, marginal adaptation, anatomic form, and loss of retention. We considered all of these as dichotomous outcomes and evaluated them according to risk difference, as this can be calculated for any study even when there are no events in the intervention and control groups.[16] We performed a subanalysis for technical complications on the groups "Alfa and 4," "Bravo and 3," "Charlie and 2," and "Delta and 1," according to US Public Health Service (USPHS) and California Dental Association criteria (Appendix 2, available online at the end of this article).

The data extraction regarding technical complications is shown in eTables 3a and 3b, available online at the end of this article.[9-15]

Statistical analysis

We calculated standard mean difference and risk difference by means of comparing the estimates of MC and ZC restorations with a 95% confidence interval (CI) using a fixed-effects or random-effects model depending on the homogeneity of the trials included. We used the Cochran Q test and I^2 to determine the statistical heterogeneity. If the I^2 value was from 0% through 50% and the P value was higher than .05, we interpreted the level of heterogeneity as within acceptable limits and, therefore, applied a fixed-effects model. We applied the analyses using Comprehensive Meta-Analysis software, Version 3 (Biostat).

Risk of bias in person studies

Two reviewers (A.L., P.M.-M.) determined the risk of bias of individual studies using the Cochrane Collaboration Tool to assess the risk of bias described in the Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0.[10] The articles were assessed in 7 categories: selection bias, allocation bias, blinding of participants and staff, blinding of outcome assessors, incomplete data, selective notification of results, and other bias. Each category was classified as low risk, high risk, or uncertain risk (Figure 1).[9-15]

RESULTS

Search results

Via the search strategy, we identified 783 references: 316 from MEDLINE (PubMed), 267 from Scopus, 200 from WOS, 0 from CENTRAL, and 12 from an independent manual search. Of these, we removed 90 duplicates and 58 triplicates. We downloaded the full texts of 51 articles. Finally, we conducted a qualitative synthesis on 7 RCTs comparing ZC and MC tooth-supported FDPs within the same study. We excluded 2 of the 3 articles reporting on the same patient cohort from the quantitative analysis (meta-analysis).[12,14,15] We decided to consider exclusively the study with 5 years of follow-up for reasons of homogeneity.[12] We estimated the level of agreement between the reviewers via κ statistics with values of 0.85 on Scopus, 0.82 on MEDLINE, 0.87 on WOS, and 1.0 on CENTRAL, indicating a high degree of interrater reliability. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses systematic search flow chart is shown in the eFigure.

Characteristics of the studies

The total number of participants in the included studies was 330, with a mean age of 40.36 years (range, 24-70 years). On average, women were 56.31% of the study populations. A total of 177 ZC and 184 MC FDPs were included. All studies evaluated the posterior area of the maxilla and mandible over 3 units (97.62%), 4 units (1.84%), or 5 units (0.54%). The mean dropout rate within the observation period was low (7.89%), and the mean observation period was 4.09 years (range, 3-5.28 years).

Tooth preparation

All studies were carried out at universities except for 1.[9] All preparations of the abutment teeth followed some requirements: circumferentially rounded shoulder (0.8-1.0 millimeters), axial reduction (1.5 mm), occlusal reduction (1.5-2 mm), and convergence angle (6-15 degrees) by 1, 2, or several experienced clinicians. Most authors airborne-particle abraded the intaglio surfaces of the FDPs and cemented with a resin-based cement. Sailer and colleagues[15] pretreated the abutment teeth with a dentin primer (ED Primer, Kuraray) before cementing the FDPs and pretreated the intaglio surfaces of the metal-based FDPs with an alloy primer (Alloy Primer, Kuraray) compatible with the resin cement.

Restoration material

Different metal alloys for casting were used for the metal framework; the authors in 3 articles used gold alloy,[12,15,16] those in 1 article used gold-platinum-palladium alloy,[9] those in 1 article used high-

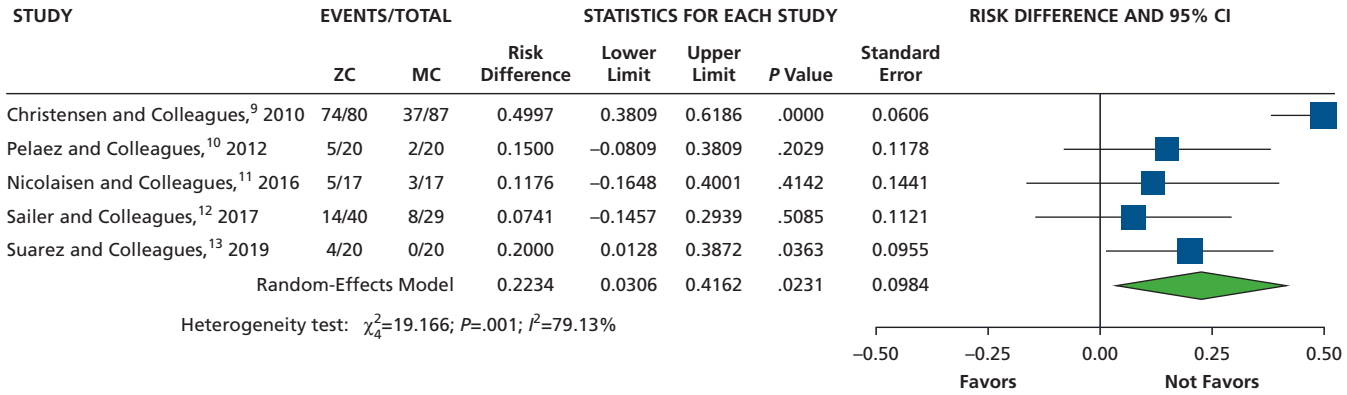


Figure 2. Forest plot for meta-analysis of risk difference of zirconia-ceramic (ZC) versus metal-ceramic (MC) fixed dental prostheses in terms of global ceramic chipping. CI: Confidence interval.

noble gold-platinum alloy,[11] and those in 2 articles used cobalt-chromium alloy.[10,13] The zirconia framework was milled from presintered blocks. All frameworks were manufactured by 1 experienced technician or 2 technicians working together.

Biological complications

The analysis of biological complications showed no significant differences between ZC and MC restorations for the following parameters: probing pocket depth, probing attachment level, and loss of tooth vitality. We could not obtain the plaque control record and perform bleeding on probing analysis because these measures were reported by the authors in only 1 study.[15] The results of these analysis are shown in the table and Appendix 3, available online at the end of this article.

Technical complications

The analysis of technical complications showed significant differences between ZC and MC FDP regarding global veneer chipping ($P = .023$) with a 22.3% higher risk (95% CI, 3% to 41.6%) (Figure 2).[9-13] The subanalyses showed minor chipping or chipping resolved with polishing ($P = .044$) with a 19.5% higher risk (95% CI, 0.5% to 38.4%) and major chipping or chipping down to the framework ($P = .023$) with a 6% higher risk (95% CI, 0.8% to 11.3%) when ZC was chosen as the restoration material. We found no statistically significant differences in occlusal wear or its subanalysis, marginal adaptation or its subanalysis, and anatomic form or its subanalysis. The results of this analysis are shown in the table and Appendix 3, available online at the end of this article.

Survival rate

We defined survival as the FDP remaining in situ with or without modifications. The meta-analysis revealed a 95.4% survival rate (95% CI, 90.5% to 99.1%) for ZC and 96.9% (95% CI, 94.3% to 99.4%) for MC with no significant differences between materials ($P = .364$) (Table). Sailer and colleagues[15] reported survival rates of 91.3% of ZC FDP 10 years after placement and 100% for MC FDP.

Risk of bias in included studies

The Cochrane Collaboration Tool[17] showed a low risk of bias in all of the included studies. In some studies, the authors did not report sufficient information about the sequence-generation process to allow an evaluation of "low risk" or "high risk."[9-13] In others, the authors did not describe the allocation concealment or did not describe it in sufficient detail.[9-14] The assessment of the risk of bias for each study is summarized in Figure 1.[9-15]

Statistical analysis results

The length of the follow-up in the included studies was chosen rigorously for the medium-term (3-5 years) for reasons of homogeneity. The heterogeneity test of the meta-analysis was favorable in most estimates (Table; Appendix 3, available online at end of this article).

Table. Meta-analysis results of technical and biological complications and survival rates.

ESTIMATED RD,* MANTEL-HAENZEL METHOD				
Restoration Materials	**Heterogeneity Test**	**Model**	**Values (95% CI†)**	***P* Value**
Zirconia ceramic	$\chi^2_4 = 12.794$; $P = .012$; $I^2 = 68.735\%$	Random-effects	0.954 (0.905 to 1.000)	.001
Metal ceramic	$\chi^2_4 = 1.602$; $P = .808$; $I^2 = 0.000\%$	Fixed-effects	0.969 (0.943 to 0.994)	.001
Comparison (measure with RD)	$\chi^2_4 = 4.865$; $P = .301$; $I^2 = 17.778\%$	Fixed effects	0.017 (−0.020 to 0.054)	.364
Technical Complications				
Core fractures	$\chi^2_4 = 0.739$; $P = .946$; $I^2 = 0.000\%$	Fixed-effects	0.0122 (−0.016 to 0.041)	.401
Loss of retention	$\chi^2_2 = 1.581$; $P = .454$; $I^2 = 0.000\%$	Fixed-effects	0.023 (−0.031 to 0.078)	.398
Global ceramic veneering chipping	$\chi^2_4 = 19.166$; $P = .001$; $I^2 = 79.130\%$	Random-effects	0.223 (0.030 to 0.416)	.023
Bravo and 3 ceramic veneering chipping	$\chi^2_3 = 11.624$; $P = .009$; $I^2 = 74.191\%$	Random-effects	0.195 (0.005 to 0.384)	.044
Charlie and 2 ceramic veneering chipping	$\chi^2_3 = 4.867$; $P = .182$; $I^2 = 38.357\%$	Fixed-effects	0.060 (0.008 to 0.113)	.023
Delta and 1 ceramic veneering chipping	$\chi^2_3 = 2.626$; $P = .453$; $I^2 = 0.000\%$	Fixed-effects	0.017 (−0.022 to 0.056)	.399
Global occlusal wear	$\chi^2_4 = 36.119$; $P = .001$; $I^2 = 88.925\%$	Random-effects	0.094 (−0.171 to 0.359)	.485
Bravo and 3 occlusal wear	$\chi^2_2 = 9.304$; $P = .010$; $I^2 = 78.503\%$	Random-effects	−0.024 (−0.337 to 0.289)	.881
Charlie and 2 occlusal wear	$\chi^2_2 = 2.201$; $P = .333$; $I^2 = 9.118\%$	Fixed-effects	0.028 (−0.048 to 0.104)	.470
Delta and 1 occlusal wear	$\chi^2_2 = 0.000$; $P = .999$; $I^2 = 0.000\%$	Fixed-effects	0.000 (0.000 to 0.044)	.999
Global marginal adaptation	$\chi^2_3 = 38.120$; $P = .000$; $I^2 = 92.130\%$	Random-effects	0.173 (−0.230 to 0.576)	.399
Bravo and 3 marginal adaptation	$\chi^2_2 = 32.199$; $P = .000$; $I^2 = 93.789\%$	Random-effects	0.239 (−0.298 to 0.777)	.383
Charlie and 2 marginal adaptation	$\chi^2_2 = 0.788$; $P = .674$; $I^2 = 0.000\%$	Fixed-effects	0.014 (−0.070 to 0.099)	.742
Delta and 1 marginal adaptation	$\chi^2_2 = 0.000$; $P = .999$; $I^2 = 0.000\%$	Fixed-effects	0.000 (−0.044 to 0.044)	.999
Global anatomic form	$\chi^2_3 = 0.972$; $P = .808$; $I^2 = 0.000\%$	Fixed-effects	−0.039 (−0.161 to 0.083)	.530
Bravo and 3 anatomic form	$\chi^2_2 = 1.756$; $P = .416$; $I^2 = 0.000\%$	Fixed-effects	−0.036 (−0.166 to 0.094)	.588
Charlie and 2 anatomic form	$\chi^2_2 = 1.004$; $P = .605$; $I^2 = 0.000\%$	Fixed-effects	0.006 (−0.048 to 0.059)	.837
Delta and 1 anatomic form	$\chi^2_2 = 0.000$; $P = .999$; $I^2 = 0.000\%$	Fixed-effects	0.000 (−0.044 to 0.044)	.999
Biological Complications				
Probing pocket depth	$\chi^2_1 = 0.000$; $P = .999$; $I^2 = 0.000\%$	Fixed-effects	0.000 (−0.390 to 0.390)	.999
Probing attachment level	$\chi^2_1 = 0.216$; $P = .642$; $I^2 = 0.000\%$	Fixed-effects	0.066 (−0.324 to 0.456)	.741
Plaque control record	Not calculable	–	–	–
Bleeding on probing	Not calculable	–	–	–
Loss of tooth vitality (measure with RD)	$\chi^2_3 = 1.937$; $P = .586$; $I^2 = 0.000\%$	Fixed-effects	0.020 (−0.033 to 0.073)	.463

* RD: Risk difference. † CI: Confidence interval. ‡ –: Not applicable.

DISCUSSION

Systematic reviews and meta-analyses have been used extensively in evidence-based dentistry for the past 2 decades.[18-20] Two systematic reviews have reported on this topic.[18,21] Nevertheless, we are unaware of previous meta-analyses that have used only RCTs to obtain comparable results with the greatest possible evidence, given that 2 treatments can be compared rigorously only when they have been performed under the same controlled conditions within the same study.[22] In our meta-analysis, we found 5 medium-term RCTs[9-13] reporting on the survival and complication rates of ZC compared with MC tooth-supported posterior FDP, with a median follow-up of 4.07 years. We did not include 2 of the 3 studies reporting on the same patient cohort in the quantitative analysis.[14,15]

Zirconia is 1 of the most developed core materials for FDP and exhibits excellent mechanical, biocompatibility, and esthetic properties that satisfy the growing esthetic demands for avoiding dark

materials in the copings and making restorations increasingly similar to natural teeth with a profile better integrated with soft tissues.

According to the literature, the survival rates of ZC FDP after 5 years are slightly lower than those of MC FDP,[18] and our meta-analysis found similar results with a 95.4% survival rate (95% CI, 90.5% to 99.1%) for ZC and 96.9% (95% CI, 94.3% to 99.4%) for MC with no significant differences between the materials ($P = .364$, Table). No framework fractures were found in either group. Sailer and colleagues[15] reported a 91.3% survival rate for the ZC FDP after 10 years from placement. Conversely, they did not report any framework fractures in MC FDP, with a 100% survival rate in all periods studied.

The major biological complications reported were secondary caries, loss of vitality, abutment tooth fractures, and periodontal disease.[23] Concerning the biological complications, no significant differences were found between the 2 materials. However, the evidence is weaker in this aspect, as not all studies provided data or used the same indexes.

Regarding the technical complications assessed, the only significant differences observed were in global ceramic veneering chipping ($P = .023$) (Figure 2) and their subanalyses: minor chipping or "Bravo and 3 group" ($P = .044$) and major chipping or "Charlie and 2 group" ($P = .023$).[9-13] We estimated that, in global terms, the ceramic veneering chipping was 22.3% (95% CI, 3.0% to 41.6%) more frequent in ZC restorations after 4 years. On the basis of the subanalyses, we estimated that chipping when polishing is possible (minor chipping or "Bravo and 3 group") appeared 19.5% (95% CI, 0.5% to 38.4%) more frequently in the ZC group and that chipping that needs repair in the laboratory (major chipping or "Charlie and 2 group") appeared 6.0% (95% CI, 0.8% to 11.3%) more frequently than in MC restorations. These results reveal that most of the chipping that appear in ZC restorations can be addressed by means of polishing and that only a small proportion of catastrophic chipping requires repair in the laboratory.

Special care is required with the zirconia framework design, the veneering-ceramic application technique (layered or pressed),[24] and the firing procedures (slow cooling and heating protocols) to reduce chipping rates,[25,26] which are expected to decrease in the near future.[13,27] However, insufficient clinical trials are available to completely support the reduction of chipping rates in ZC restorations compared with previous chipping rates. Another solution to veneering-ceramic chipping may be monolithic zirconia restorations, which require minimal tooth reduction because there is no need for space clearance for the veneering material.[28,29] Monolithic zirconia FDP may be a reliable clinical solution, especially in molar areas with large occlusal forces. However, this concept presents some worries that need to be measured clinically, such as matching the color to the other teeth, the long-term chemical stability, or the clinical wear behavior. More investigations are required in the medium term and long term to allow for evidence-based decision making.

Our analysis did not find statistical differences in the loss of retention between the 2 materials. However, Sailer and colleagues[15] reported a higher prevalence in ZC restorations after 10 years of follow-up. This can be explained by the fact that zirconia frameworks exhibited internal gaps larger than those of frameworks fabricated with conventional MCc techniques. The internal gap values of zirconia FDP in the literature ranged from a mean (standard deviation [SD]) of 140 (26) micrometers in laboratory studies through 130 (56) μm in clinical studies.[15,30-32] However, in the past 5 years, internal gap assessment values of zirconia FDP in the literature have ranged from a mean (SD) of 76.2 (36.3) μm in laboratory studies through 88.27 (41.49) μm and 92.13 (49.87) μm in clinical studies.[15,27,28] Hence, ZC FDP manufacturing methods have improved, achieving values of marginal discrepancy similar to those of MC FDP[33,34] whose values range from a mean (SD) of 48.9 (24) μm through 128.8 (29.3) μm depending on the fabrication technique of the core: casting, milling, selective laser melting, or milling and sintering.[35] Another factor that may influence the loss of retention is the cementation technique, which should achieve a reliable bond between the dentin and the zirconia copings to increase marginal adaptation and prevent leakage. Several techniques have been proposed to overcome the limitations of the bond to zirconia through airborne-particle abrasion to create micromechanical retention in combination with the use of zirconia primers to increase the chemical bond.[36] From a macroscopic point of view, no statistically significant differences were found in the marginal adaptation when a clinical examination with an explorer was performed in different studies ($P = .399$).

Esthetics seems to be key in the selection of 1 type of restoration over another. Zirconia cores could be better accepted than dark metallic cores owing to their white color and better integration with soft tissues and esthetics over time. A study of implant-supported restorations that assessed patients' opinions on the esthetic result did not show statistically significant differences between the ZC and MC restorations.[37] However, most studies have not reported on patient satisfaction with esthetics.

Until more evidence becomes available, the use of MC FDP should be considered first for the posterior sector, in which esthetics do not seem to be critical, because ZC FDP present more technical complications in terms of chipping than MC FDP. However, ZC FDP are a reliable treatment option for patients with high esthetic expectations.

The limitations of our analysis are related to the potential differences between the 2 technical complication assessment systems (CDA and modified USPHS). Also, the results of our analysis might not represent fully the procedures available because the software and hardware of the computer-aided design and computer-aided manufacturing technology are constantly evolving and improving in accuracy. Another limitation is that we drew conclusions from only 7 studies for the qualitative assessment and 5 studies for the quantitative assessment.

Authors of future RCTs should use the same standardized assessment system (ideally, the modified USPHS criteria as it is more accurate) and measure the same specific parameters to provide comparable results for the scientific community. Esthetic complications should be measured more thoroughly with objective methods, and more long-term studies are needed.

CONCLUSIONS

In the medium term, ZC FDP showed rates of clinical survival and biological complications similar to those of MC FDP. The only technical complication with statistically significant differences found was chipping of the ceramic veneer, which was mostly resolved via polishing. On this basis, although MC FDP seem to remain the reference standard for multiunit tooth-supported posterior FDP, ZC FDP are considered a predictable treatment. ■

SUPPLEMENTAL DATA

Supplemental data related to this article can be found at: https://doi.org/10.1016/j.adaj.2019.12.013

Dr. Limones is a master of science candidate, Department of Conservative Dentistry and Orofacial Prosthodontics, Faculty of Dentistry, Complutense University of Madrid, Madrid, Spain.

Dr. Molinero-Mourelle is a PhD candidate, Department of Conservative Dentistry and Orofacial Prosthodontics, Complutense University of Madrid, Plaza Ramón y Cajal S/N, 28040, Madrid, Spain, e-mail pedromol@ucm.es. Address correspondence to Dr. Molinero-Mourelle.

Dr. Azevedo is a PhD candidate, Department of Conservative Dentistry and Orofacial Prosthodontics, Faculty of Dentistry, Complutense University of Madrid, Madrid, Spain.

Dr. Romeo-Rubio is an associate professor, Department of Conservative Dentistry and Orofacial Prosthodontics, Faculty of Dentistry, Complutense University of Madrid, Madrid, Spain.

Dr. Correia is an associate professor, Institute of Health Sciences of Viseu, Universidade Católica Portuguesa, Porto, Portugal, and a researcher, Center for Interdisciplinary Research in Health, Porto, Portugal.

Dr. Gómez-Polo is an associate professor, Department of Conservative Dentistry and Orofacial Prosthodontics, Faculty of Dentistry, Complutense University of Madrid, Madrid, Spain.

Disclosure. None of the authors reported any conflicts of interest.

The authors state that this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector. Alvaro Limones and Pedro Molinero-Mourelle are joint first authors and contributed equally to this work.

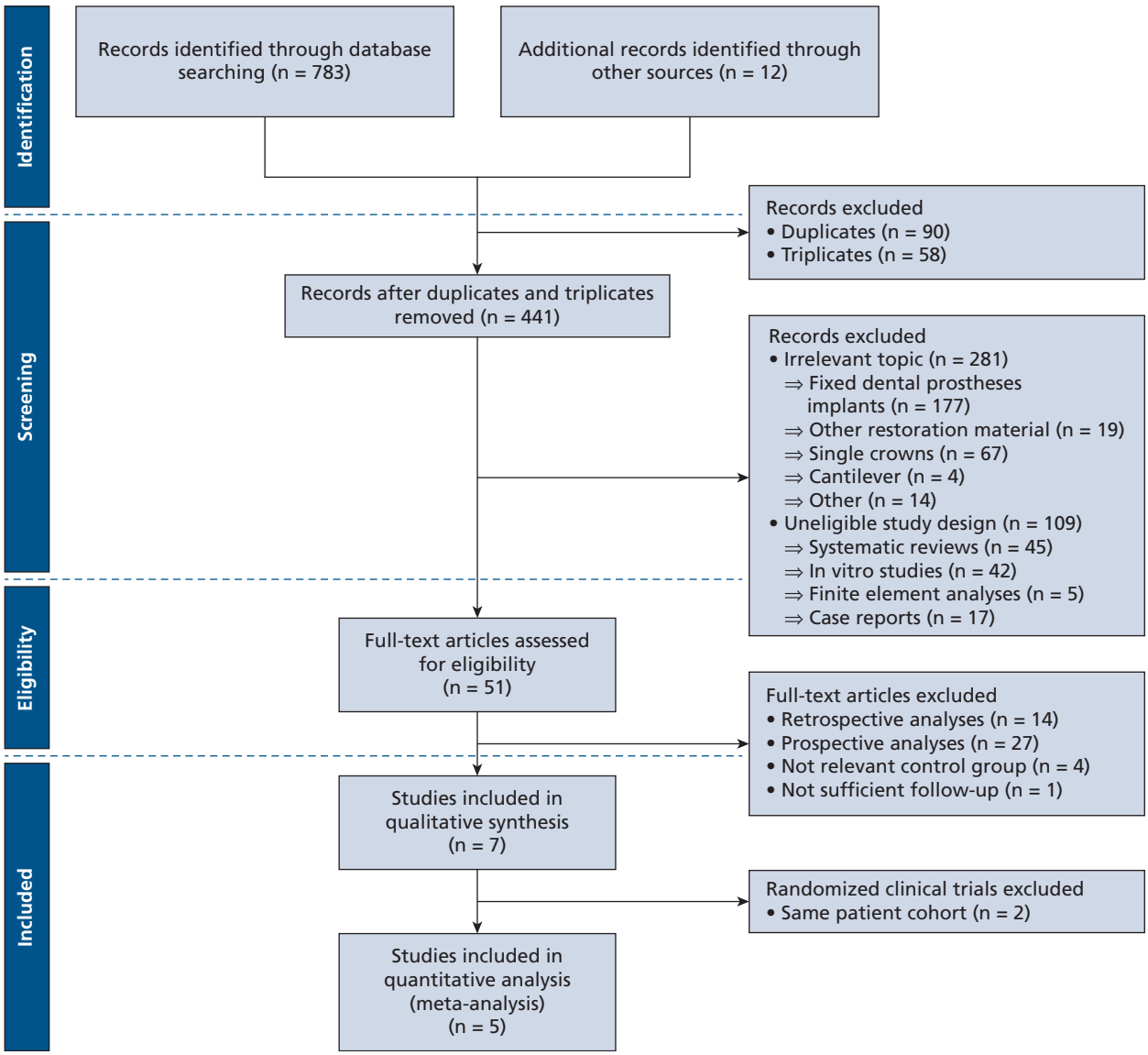
1. Edelhoff D, Sorensen JA. Tooth structure removal associated with various preparation designs for posterior teeth. *Int J Periodontics Restorative Dent*. 2002;22(3):241-249.
2. McLean JW. Evolution of dental ceramics in the twentieth century. *J Prosthet Dent*. 2001;85(1):61-66.
3. Kelly JR, Benetti P. Ceramic materials in dentistry: historical evolution and current practice. *Aust Dent J*. 2011;56(suppl 1):84-96.
4. Manicone PF, Rossi Iommetti P, Raffaelli L. An overview of zirconia ceramics: basic properties and clinical applications. *J Dent*. 2007;35(11):819-826.
5. Zarone F, Russo S, Sorrentino R. From porcelain-fused-to-metal to zirconia: clinical and experimental considerations. *Dent Mater*. 2011;27(1):83-96.
6. Choi B-K, Han J-S, Yang J-H, Lee J-B, Kim S-H. Shear bond strength of veneering porcelain to zirconia and metal cores. *J Adv Prosthodont*. 2009;1(3):129-135.
7. Malkondu Ö, Tinastepe N, Akan E, Kazazoğlu E. An overview of monolithic zirconia in dentistry. *Biotechnol Equip*. 2016;30(4):644-652.
8. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
9. Christensen RP, Ploeger BJ. A clinical comparison of zirconia, metal and alumina fixed-prosthesis frameworks veneered with layered or pressed ceramic: a three-year report. *JADA*. 2010;141(11):1317-1329.
10. Pelaez J, Cogolludo PG, Serrano B, Serrano JFL, Suarez MJ. A four-year prospective clinical evaluation of zirconia and metal-ceramic posterior fixed dental prostheses. *Int J Prosthodont*. 2012;25(5):451-458.
11. Nicolaisen MH, Bahrami G, Schropp L, Isidor F. Comparison of metal-ceramic and all-ceramic three-unit posterior fixed dental prostheses: a 3-year randomized clinical trial. *Int J Prosthodont*. 2016;29(3):259-264.
12. Sailer I, Balmer M, Husler J, Hammerle CHF, Kanel S, Thoma DS. Comparison of fixed dental prostheses with zirconia and metal frameworks: five-year results of a randomized controlled clinical trial. *Int J Prosthodont*. 2017;30(5):426-428.
13. Suarez MJ, Perez C, Pelaez J, Lopez-Suarez C, Gonzalo E. A randomized clinical trial comparing zirconia and metal-ceramic three-unit posterior fixed partial dentures: a 5-year follow-up. *J Prosthodont*. 2019;28(7):750-756.

14. Sailer I, Gottnerb J, Kanelb S, Hammerle CHF. Randomized controlled clinical trial of zirconia-ceramic and metal-ceramic posterior fixed dental prostheses: a 3-year follow-up. *Int J Prosthodont*. 2009;22(6):553-560.
15. Sailer I, Balmer M, Husler J, Hammerle CHF, Kanel S, Thoma DS. 10-year randomized trial (RCT) of zirconia-ceramic and metal-ceramic fixed dental prostheses. *J Dent*. 2018;76:32-39.
16. Cheng J, Pullenayegum E, Marshall JK, Iorio A, Thabane L. Impact of including or excluding both-armed zero-event studies on using standard meta-analysis methods for rare event outcome: a simulation study. *BMJ Open*. 2016;6(8):e010983.
17. Higgins JPT, Altman DG, Gotzsche PC, et al; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
18. Pjetursson BE, Sailer I, Makarov NA, Zwahlen M, Thoma DS. All-ceramic or metal-ceramic tooth-supported fixed dental prostheses (FDPs)? A systematic review of the survival and complication rates, part II—multiple-unit FDPs [correction appears in *Dent Mater*. 2017;33(1):e48-e51]. *Dent Mater*. 2015;31(6):624-639.
19. Bidra AS. Evidence-based prosthodontics: fundamental considerations, limitations, and guidelines. *Dent Clin North Am*. 2014;58(1):1-17.
20. Egger M, Smith GD, Sterne JA. Uses and abuses of meta-analysis. *Clin Med*. 2001;1(6):478-484.
21. Heintze SD, Rousson V. Survival of zirconia- and metal-supported fixed dental prostheses: a systematic review. *Int J Prosthodont*. 2010;23(6):493-502.
22. Every-Palmer S, Howick J. How evidence-based medicine is failing due to biased trials and selective publication. *J Eval Clin Pract*. 2014;20(6):908-914.
23. Larsson C, Wennerberg A. The clinical success of zirconia-based crowns: a systematic review. *Int J Prosthodont*. 2014;27(1):33-43.
24. Naenni N, Bindl A, Sax C, Hammerle C, Sailer I. A randomized controlled clinical trial of 3-unit posterior zirconia-ceramic fixed dental prostheses (FDP) with layered or pressed veneering ceramics: 3-year results. *J Dent*. 2015;43(11):1365-1370.
25. Tan JP, Sederstrom D, Polansky JR, McLaren EA, White SN. The use of slow heating and slow cooling regimens to strengthen porcelain fused to zirconia. *J Prosthet Dent*. 2012;107(3):163-169.
26. Rodrigues C da S, Aurelio IL, Kaizer M da R, Zhang Y, May LG. Do thermal treatments affect the mechanical behavior of porcelain-veneered zirconia? A systematic review and meta-analysis. *Dent Mater*. 2019;35(5):807-817.
27. Kanat B, Comlekoglu EM, Dundar-Comlekoglu M, Hakan Sen B, Ozcan M, Ali Gungor M. Effect of various veneering techniques on mechanical strength of computer-controlled zirconia framework designs. *J Prosthodont*. 2014;23(6):445-455.
28. Konstantinidis I, Trikka D, Gasparatos S, Mitsias ME. Clinical outcomes of monolithic zirconia crowns with CAD/CAM technology: a 1-year follow-up prospective clinical study of 65 patients. *Int J Environ Res Public Health*. 2018;15(11):2523.
29. Nakamura K, Harada A, Inagaki R, et al. Fracture resistance of monolithic zirconia molar crowns with reduced thickness. *Acta Odontol Scand*. 2015;73(8):602-608.
30. Buchi DL, Ebler S, Hammerle CHF, Sailer I. Marginal and internal fit of curved anterior CAD/CAM-milled zirconia fixed dental prostheses: an in-vitro study. *Quintessence Int*. 2014;45(10):837-846.
31. Brawek PK, Wolfart S, Endres L, Kirsten A, Reich S. The clinical accuracy of single crowns exclusively fabricated by digital workflow: the comparison of two systems. *Clin Oral Investig*. 2013;17(9):2119-2125.
32. Ahrberg D, Lauer HC, Ahrberg M, Weigl P. Evaluation of fit and efficiency of CAD/CAM fabricated all-ceramic restorations based on direct and indirect digitalization: a double-blinded, randomized clinical trial. *Clin Oral Investig*. 2016;20(2):291-300.
33. Lopez-Suarez C, Gonzalo E, Pelaez J, Serrano B, Suarez MJ. Marginal vertical discrepancies of monolithic and veneered zirconia and metal-ceramic three-unit posterior fixed dental prostheses. *Int J Prosthodont*. 2016;29(3):256-258.
34. Freire Y, Gonzalo E, Lopez-Suarez C, Suarez MJ. The marginal fit of CAD/CAM monolithic ceramic and metal-ceramic crowns. *J Prosthodont*. 2019;28(3):299-304.
35. Kim E-H, Lee D-H, Kwon S-M, Kwon T-Y. A microcomputed tomography evaluation of the marginal fit of cobalt-chromium alloy copings fabricated by new manufacturing techniques and alloy systems. *J Prosthet Dent*. 2017;117(3):393-399.
36. Steiner R, Heiss-Kisielewsky I, Schwarz V, et al. Zirconia primers improve the shear bond strength of dental zirconia [published online ahead of print January 9, 2019]. *J Prosthodont*.
37. Hosseini M, Worsaae N, Schiodt M, Gotfredsen K. A 3-year prospective study of implant-supported, single-tooth restorations of all-ceramic and metal-ceramic materials in patients with tooth agenesis. *Clin Oral Implants Res*. 2013;24(10):1078-1087.

APPENDIX 1 - Search Strategy

Studies were identified by using the following search terms:

((Posterior [All Fields] OR "3-unit" [All Fields] OR "Fixed-dental Prosthesis" [All Fields] OR "Dental Prosthesis" [MeSH Terms]) AND ("Metal-ceramic Alloy" [MeSH Terms]) AND ("Zirconia-ceramic" [All Fields] OR Zirconia [All Fields] OR "Yttria stabilized tetragonal zirconia" [MeSH] OR "Zirconium oxide" [All Fields]) AND ("Survival Rate" [MeSH Terms] OR "Biological complications" [All Fields] OR "Technical complications" [All Fields] OR Complications [MeSH Terms] OR Chipping [All fields])). The following journals were hand-searched (A.L.) from year 2008 to 2019: *"Journal of Dental Research," "Journal of Dentistry," "Journal of Prosthodontic Research", "Clinical Oral Investigation,s, "Journal of Prosthetic Dentistry, "Journal of Oral Rehabilitation," "Journal of Prosthodontic - Implant Esthetic and Restorative Dentistry," "International Journal of Prosthodontics," "British Dental Journal," and "The Journal of Advanced Prosthodontics".*



eFigure. Flowchart of the study selection process.

eTable 1. Data extraction of included studies, general information.*

STUDY	PARTICIPANTS, NO. (EXAMINED)	DROPOUTS, NO. (%)	SEX, NO.		AGE, Y, MEDIAN	FOLLOW-UP, MO, Median (SD†)
			Female	Male		
Sailer and Colleagues,[14] 2009	59 (53)	6 participants, 9 FDP (8.96)	27	32	54.4 (SD, 12)	ZC§ and MC¶: 40.3 (2.8)
Christensen and Colleagues,[9] 2010	193 (167)††	26 participants, 26 FDP (13.47)	71	122	50 (range, 16-89)	ZC and MC: 36
Pelaez and Colleagues,[10] 2012	37 (37)	0 participants, 0 FDP (0.00)	22	15	Range, 23-65	ZC and MC: 50 (2.4)
Nicolaisen and Colleagues,[11] 2016	34 (34)	NR	21	13	51 (female, 53; male, 49)	ZC and MC: 36
Sailer and Colleagues,[12] 2017	58 (52)	6 participants, 7 FDP (7.89)	27	31	54.8 ZC: 52.7 (13.0), MC:57.0 (12.1)	ZC: 61.2 (3.6) MC: 60 (4.8)
Sailer and Colleagues,[15] 2018	58 (44)	13 participants, 17 FDP (24.53)	NR		60.9	ZC: 123 MC: 120
Suarez and Colleagues,[13] 2018	40 (40)	0 participants, 0 FDP (0.00)	23	17	Range, 24-70	ZC and MC: 63.3 (2.4)

* All the included studies were randomized clinical trials. † SD: Standard deviation. ‡ FDP: Fixed dental prostheses. § ZC: Zirconia ceramic. ¶ MC: Metal ceramic. # Mx: Maxilla. ** Md: Mandible. †† Only the results at 36 months of follow-up were included, per the inclusion criteria. ‡‡ NR: Not reported.

eTable 1. Continued

POSTERIOR UNITS FDP‡				RESTORATION MATERIAL		CEMENT USED
Total	**3 Units (%)**	**4 Units (%)**	**5 Units (%)**	**Ceramic Framework (n)**	**Manufacturing Technique**	
76 (67)	ZC: 11 Mx#, 18 Md** MC: 13 Mx, 16Md (86.57)	ZC: 4 Mx, 2 Md MC: 1 Mx (10.44)	ZC: 1 Mx MC: 1 Md (2.9)	ZC (36) (Cercon Ceram S, Ceramco) MC (31) (Duceram Plus, DeguDent)	Zirconia milling, metal casting (gold alloy)	Resin cement (Panavia 21 TC, Kuraray)
193 (167)	ZC: 80 MC: 87 (100)	NR‡‡	NR	ZC (80) (Ceramco, Dentsply) (CZR Press, Noritake Dental) (Lava Ceram, 3M ESPE) MC (87) (Pulse interface , Jensen Dental) (Creation, Jensen Dental) (Ceramco SWE, Dentsply)	Zirconia milling, metal casting (silver-free alloy or gold-platinum-palladium alloy)	Resin cement (RelyX Unicem, 3M ESPE)
37	ZC: 11 Mx, 9 Md MC: 6 Mx, 14 Md (100)	NR	NR	ZC (20) (Lava Ceram) MC (20) (VITA VM 13, VITA Zahnfabrik)	Zirconia milling, metal casting (chrome-cobalt alloy)	Resin cement (RelyX Unicem), glass ionomer cement (Ketac Cem Plus, 3M ESPE)
34	ZC: 11 Mx, 6 Md MC: 9 Mx and 8 Md (100)	NR	NR	ZC (17) (VITA VM 9, VITA Zahnfabrik) MC (17) (VITA VM 13)	Zirconia milling metal casting (gold-platinum alloy)	Glass ionomer cement (Ketac Cem Plus)
76 (69)	ZC: 33 MC: 34 (88.16)	ZC: 6 MC: 1 (9.21)	ZC: 1 MC: 1 (2.63)	ZC (40) (Cercon Ceram S) MC (29) (Duceram Plus)	Zirconia milling, metal casting (gold alloy)	Resin cement (Panavia 21 TC)
76 (53)	ZC: 20 MC: 29 (85.52)	ZC: 6 MC: 1 (10.71)	ZC: 1 MC: 1 (3.77)	ZC (29) (Cercon Ceram S) MC (24) (Duceram Plus)	Zirconia milling, metal casting (gold alloy)	Resin cement (Panavia 21 TC)
40	ZC: 10 Mx, 10 Md MC: 10 Mx, 10 Md (100)	NR	NR	ZC (20) (IPS e.max Ceram, Ivoclar Vivadent) MC (20) (VITA VM 13)	Zirconia milling, metal casting (chrome-cobalt alloy)	Resin cement (RelyX Unicem)

eTable 2. Data extraction of included studies, biological complications.

STUDY	TYPE OF FDP*	PROBING POCKET DEPTH, MILLIMETERS (SD†)		PROBING ATTACHMENT LEVEL, mm (SD)		PLAQUE CONTROL RECORD, % (SD)		BLEEDING ON PROBING, % (SD)		LOSS OF TOOTH VITALITY, NO. (%)
		Test	Control	Test	Control	Test	Control	Test	Control	
Sailer and Colleagues,[14] 2009	ZC‡ MC§	2.4 (0.3) 2.4 (0.3)	2.3(0.3) 2.4 (0.4)	2.5 (0.2) 2.3 (0.2)	2.0 (0.4) 2.6 (0.6)	0.1 (0.1) 0.1 (0.1)	0.2 (0.2) 0.3 (0.3)	0.3 (0.2) 0.3 (0.2)	0.2 (0.2) 0.2 (0.2)	1 (NR) 1 (NR)
Christensen and Colleagues,[9] 2010	ZC MC	NR¶		NR		NR		NR		NR
Pelaez and Colleagues,[10] 2012	ZC	Gingiva index			0: 6 (25%) 1: 14 (70 %) 2: 0 (5%) 3: 0 (0%)					1 (NR)
	MC				0: 4 (5%) 1: 13 (80%) 2: 3 (15 %) 3: 0 (0%)					0 (0)
Nicolaisen and Colleagues,[11] 2016	ZC MC	< 4 mm (0.2) < 4 mm (0.1)	< 4 mm < 4 mm	0.7 (0.8) 0.9 (1.2)	NR NR	0 0	NR NR	0 0	NR NR	0 (0) 1 (NR)
Sailer and Colleagues,[12] 2017	ZC MC	2.5 (0.4) 2.5 (0.4)	NR NR	0.0 (0.2) 0.0 (0.2)	NR NR	13.8 (24.5) 12.9 (17.8)	NR NR	32.8 (26.7) 29.8 (24.0)	NR NR	2 (NR) 0 (0)
Sailer and Colleagues,[15] 2018	ZC MC	2.5 (0.3) 2.6 (0.5)	NR NR	−0.1 (0.4) 0.0 (0.3)	NR NR	14 (16) 10 (17)	NR NR	26 (29) 26 (27)	NR NR	NR (3.6) NR (4.2)
Suarez and Colleagues,[13] 2018	ZC	Gingiva index			0: 5 (25%) 1: 4 (70 %) 2: 1 (5 %) 3: 0 (0 %)					0 (0)
	MC				0: 1 (5%) 1: 16 (80 %) 2: 3 (15 %) 3: 0 (0 %)					0 (0)

* FDP: Fixed dental prostheses. † SD: Standard deviation. ‡ ZC: Zirconia-ceramic. § MC: Metal-ceramic. ¶ NR: Not reported.

eTable 3a. Data extraction of included studies, technical complications.

STUDY	TYPE OF FDP*	ASSESSMENT SYSTEM	FRAMEWORK FRACTURE, NO (%)	CERAMIC VENEERING CHIPPING, NO. (%)			
				Bravo (3 or Acceptable)	Charlie (2 or Unacceptable)	Delta (1 or Replacement)	Global Analysis
Sailer and Colleagues,[14] 2009	ZC[†] MC[‡]	USPHS[§]	0 (0) 0 (0)	9 (25) 6 (19.4)	2 (5.6) 0 (0)	1 (2.8) 0 (0)	12 (33.34) 6 (19.4)
Christensen and Colleagues,[9] (2010)	ZC MC	Own criteria (adapted)	2 (2) 0 (0)	50 (NR[¶]) 19 (NR)	16 (NR) 5 (NR)	10 (NR) 4 (NR)	74 (92.5) 37 (42.53)
Pelaez and Colleagues,[10] 2012	ZC MC	CDA[#]	0 (0) 0 (0)	NR	NR	NR	5 (25) 2 (10)
Nicolaisen and Colleagues,[11] 2016	ZC MC	CDA	0 (0) 0 (0)	4 (NR) 2 (NR)	1 (NR) 1 (NR)	0 (0) 0 (0)	5 (29) 3 (18)
Sailer and Colleagues,[12] 2017	ZC MC	USPHS	0 (0) 0 (0)	NR (27.5) NR (27.6)	NR (7.5) 0 (0)	NR (0) 0 (0)	NR (35) NR (27.6)
Sailer and Colleagues,[15] 2018	ZC MC	USPHS	2 (5.9) 0 (0)	NR (37.9) NR (33.3)	NR (13.8) 0 (0)	0 (0) 0 (0)	NR (51.7) NR (33.3)
Suarez and Colleagues,[13] 2018	ZC MC	CDA	0 (0) 0 (0)	NR (20) 0 (0)	0 (0) 0 (0)	0 (0) 0 (0)	NR (20) 0 (0)

* FDP: Fixed dental prostheses. † ZC: Zirconia ceramic. ‡ MC: Metal ceramic. § USPHS: US Public Health Service. ¶ NR: Not reported. # CDA: California Dental Association criteria.

eTable 3a. Continued

OCCLUSAL WEAR, N (%)			
Bravo (3 or Acceptable)	**Charlie (2 or Unacceptable)**	**Delta (1 or Replacement)**	**Global Analysis**
23 (63.90) 18 (58.10)	1 (2.8) 1 (3.2)	0 (0%) 0 (0)	24 (66.67) 19 (61.29)
NR	NR	NR	45 (46.39) 10 (9.6)
NR	NR	NR	10 (50) 12 (60)
8 (47) 14 (82)	2 (12) 2 (12)	0 (0) 0 (0)	10 (59) 16 (94)
NR (72.5) NR (72.4)	NR (25) NR (10.3)	0 (0) NR (0)	NR (97.5) NR (82.7)
NR (64.3) NR (70.8)	NR (28.6) NR (25)	NR (3.6) 0 (0)	NR (96.5) NR (95.8)
7 (35) 2 (10)	0 (0) 0 (0)	0 (0) 0 (0)	NR (35) NR (10)

eTable 3b. Data extraction of included studies, technical complications (continued).

STUDY	TYPE OF FDP*	MARGINAL ADAPTATION, NO. (%)				ANATOMIC FORM, NO. (%)				LOSS OF RETENTION, NO. (%)
		Bravo (3 or Acceptable)	Charlie (2 or Unacceptable)	Delta (1 or Replacement)	Global Analysis	Bravo (3 or Acceptable)	Charlie (2 or Unacceptable)	(Delta (1 or Replacement)	Global Analysis	
Sailer and Colleagues,[14] 2009	ZC† MC‡	23 (63.90) 20 (64.50)	6 (16.7) 2 (6.5)	0 (0) 0 (0)	29 (80.56) 22 (70.97)	2 (5.6) 6 (19.4)	0 (0) 0 (0)	0 (0) 0 (0)	2 (5.6) 6 (19.4)	NR
Christensen and Colleagues,[9] 2010	ZC MC	NR¶	NR	NR	NR	NR	NR	NR	NR	NR
Pelaez and Colleagues,[10] 2012	ZC MC	NR	NR	NR	NR (15) NR (25)	NR	NR	NR	10 (50) 12 (60)	0 (0) 0 (0)
Nicolaisen and Colleagues,[11] 2016	ZC MC	14 (82) 1 (6)	0 (0) 0 (0)	0 (0) 0 (0)	14 (82) 1 (6)	7 (41) 5 (29)	0 (0) 1 (6)	0 (0) 0 (0)	7 (41) 6 (35)	NR
Sailer and Colleagues,[12] 2017	ZC MC	NR (65) NR (65.5)	NR (17.5) NR (10.3)	0 (0) 0 (0)	NR (82.5) NR (75.8)	NR (7.5) NR (17.2)	NR (2.5) 0 (0)	0 (0) 0 (0)	NR (10) NR (17.2)	3 (7.5) 0 (0)
Sailer and Colleagues,[15] 2018	ZC MC	NR (51.7) NR (58.33)	NR (24.14) NR (20.83)	NR (10.3) NR (0)	NR (86.14) NR (79.16)	NR (7.1) NR (0)	NR (0) NR (0)	NR (0) NR (0)	NR(7.1) NR (0%)	5 (NR) 1 (NR)
Suarez and Colleagues,[13] 2018	ZC MC	4 (20) 5 (25)	NR (0) NR (0)	NR (0) NR (0)	NR (20) NR (25)	11 (55) 10 (50)	NR (0) NR (0)	NR (0) NR (0)	NR (55) NR (50)	0 (0) 0 (0)

* FDP: Fixed dental prostheses. † ZC: Zirconia ceramic. ‡ MC: Metal ceramic. § USPHS: US Public Health Service. ¶ NR: Not reported.

Appendix 3 - Forest Plots of Meta-analyses
(Supplementary data)

1. Technical complications:

Fig. A. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of survival rates.

Study name	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Christensen et al. (2010)	14 / 80	5 / 87	0,1175	0,0210	0,2141	0,0171	0,0493
Pelaez et al. (2012)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Nicolaisen et al. (2016)	0 / 17	0 / 17	0,0000	-0,1074	0,1074	1,0000	0,0548
Sailer et al. (2017)	0 / 40	0 / 29	0,0000	-0,0568	0,0568	1,0000	0,0290
Suarez et al. (2018)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Fixed-Effect Model			0,0171	-0,0197	0,0538	0,3635	0,0188

Heterogeneity: Q value (df = 4) = 4.865, P = 0.301; I^2 = 17.778%; T^2 = 0.000

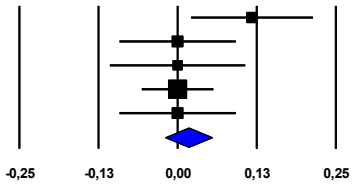


Fig. B. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of core fractures.

Study name	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Christensen et al. (2010)	2 / 80	0 / 87	0,0250	-0,0158	0,0658	0,2298	0,0208
Pelaez et al. (2012)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Nicolaisen et al. (2016)	0 / 17	0 / 17	0,0000	-0,1074	0,1074	1,0000	0,0548
Sailer et al. (2017)	0 / 40	0 / 29	0,0000	-0,0568	0,0568	1,0000	0,0290
Suarez et al. (2018)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Fixed-Effect Model			0,0122	-0,0163	0,0407	0,4019	0,0145

Heterogeneity: Q value (df = 4) = 0.739, P = 0.946; I^2 = 0.000%; T^2 = 0.000

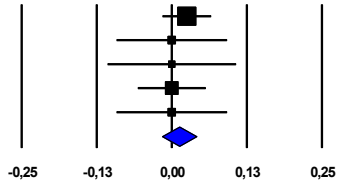


Fig. C. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of loss of retention.

Study name	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Pelaez et al. (2012)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Sailer et al. (2017)	3 / 40	0 / 29	0,0750	-0,0220	0,1720	0,1298	0,0495
Suarez et al. (2018)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Fixed-Effect Model			0,0233	-0,0308	0,0775	0,3981	0,0276

Heterogeneity: Q value (df = 2) = 1.581, P = 0.454; I^2 = 0.000%; T^2 = 0.000

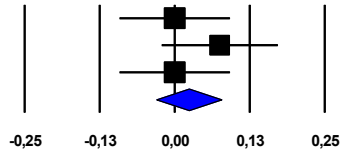
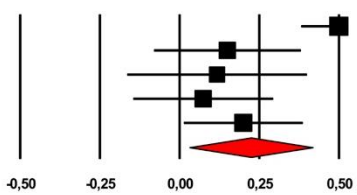


Fig. D.0. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Global Ceramic Veneering Chipping.

Study name	Events / Total		Risk difference	Lower limit	Upper limit	p-Value	Standard error
	Zr-C	M-C					
Christensen et al. (2010)	74 / 80	37 / 87	0,4997	0,3809	0,6186	0,0000	0,0606
Pelaez et al. (2012)	5 / 20	2 / 20	0,1500	-0,0809	0,3809	0,2029	0,1178
Nicolaisen et al. (2016)	5 / 17	3 / 17	0,1176	-0,1648	0,4001	0,4142	0,1441
Sailer et al. (2017)	14 / 40	8 / 29	0,0741	-0,1457	0,2939	0,5085	0,1121
Suarez et al. (2018)	4 / 20	0 / 20	0,2000	0,0128	0,3872	0,0363	0,0955
Random-Effect Model			0,2234	0,0306	0,4162	0,0231	0,0984

Heterogeneity test: Q *value* (df =4) = 19.166; P = 0.001 ; I^2 = 79.130 %.



© Álvaro Limones Burgos, 2024

Fig. D.1. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Beta & 3 Ceramic Veneering Chipping.

Study name	Events / Total		Statistics for each study				
	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Christensen et al. (2010)	50 / 80	19 / 87	0,4066	0,2695	0,5437	0,0000	0,0699
Nicolaisen et al. (2016)	4 / 17	2 / 17	0,1176	-0,1356	0,3709	0,3625	0,1292
Sailer et al. (2017)	11 / 40	8 / 29	-0,0009	-0,2144	0,2127	0,9937	0,1090
Suarez et al. (2018)	4 / 20	0 / 20	0,2000	0,0128	0,3872	0,0363	0,0955
Random-Effect Model			0,1946	0,0053	0,3839	0,0440	0,0966

Heterogeneity: Q *value* (df = 3) = 11.624; P = 0.009 ; I^2 = 74.191%.

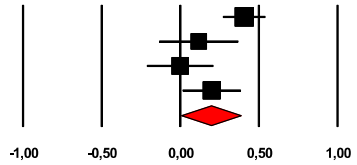


Fig. D.2. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Charlie & 2 Ceramic Veneering Chipping.

Study name			Statistics for each study				
	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Christensen et al. (2010)	16 / 80	5 / 87	0,1425	0,0422	0,2429	0,0054	0,0512
Nicolaisen et al. (2016)	1 / 17	1 / 17	0,0000	-0,1582	0,1582	1,0000	0,0807
Sailer et al. (2017)	3 / 40	0 / 29	0,0750	-0,0220	0,1720	0,1298	0,0495
Suarez et al. (2018)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Fixed-Effect Model			0,0609	0,0084	0,1134	0,0229	0,0268

Heterogeneity: Q *value* (df = 3) = 4.867 ; P = 0.182 ; I^2 = 38.357%.

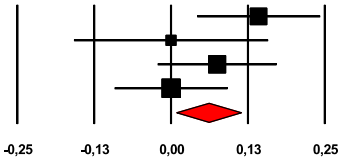


Fig. D.3. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Delta & 1 Ceramic Veneering Chipping.

Study name			Statistics for each study				
	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Christensen et al. (2010)	10 / 80	4 / 87	0,0790	-0,0058	0,1638	0,0677	0,0433
Nicolaisen et al. (2016)	0 / 17	0 / 17	0,0000	-0,1074	0,1074	1,0000	0,0548
Sailer et al. (2017)	0 / 40	0 / 29	0,0000	-0,0568	0,0568	1,0000	0,0290
Suarez et al. (2018)	0 / 20	0 / 20	0,0000	-0,0922	0,0922	1,0000	0,0470
Fixed-Effect Model			0,0168	-0,0223	0,0559	0,3993	0,0200

Heterogeneity: Q value (df = 3) = 2.626, P = 0.453; I^2 = 0.000%; T^2 = 0.000

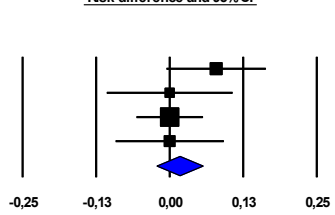
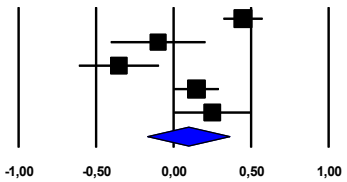


Fig. E.0. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Global Occlusal wear.

Study name	Events / Total		Statistics for each study				
	Zr-C	M-C	Risk difference	Lower limit	Upper limit	p-Value	Standard error
Christensen et al. (2010)	45 / 80	10 / 87	0,4476	0,3199	0,5753	0,0000	0,0652
Pelaez et al. (2012)	10 / 20	12 / 20	-0,1000	-0,4068	0,2068	0,5229	0,1565
Nicolaisen et al. (2016)	10 / 17	16 / 17	-0,3529	-0,6123	-0,0936	0,0076	0,1323
Sailer et al. (2017)	39 / 40	24 / 29	0,1474	0,0017	0,2932	0,0474	0,0744
Suarez et al. (2018)	7 / 20	2 / 20	0,2500	0,0031	0,4969	0,0472	0,1260
Random-Effect Model			0,0944	-0,1706	0,3594	0,4850	0,1352

Heterogeneity: Q value (df = 4) = 36.119, P = 0.000; I^2 = 88.925%; T^2 = 0.078



© Álvaro Limones Burgos, 2024

Fig. E.1. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Beta & 3 Occlusal wear.

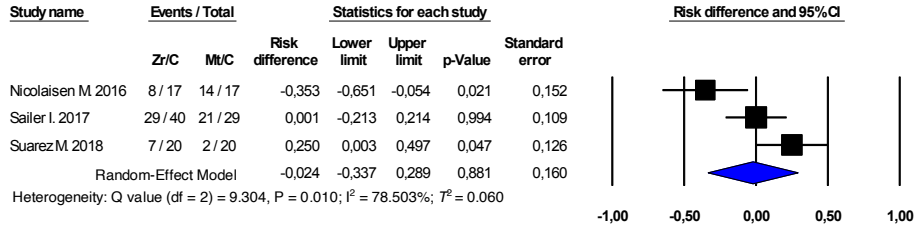


Fig. E.2. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Charlie & 2 Occlusal wear.

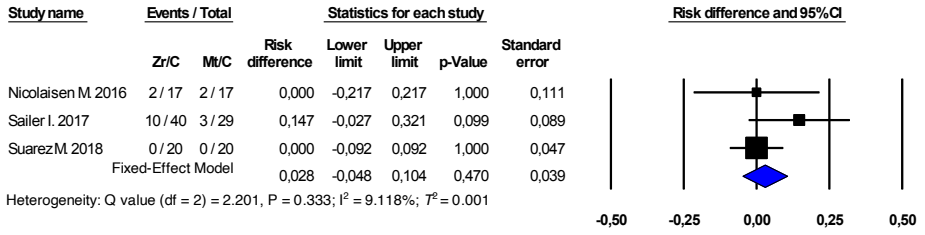


Fig. E.3. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Delta & 1 Occlusal wear.

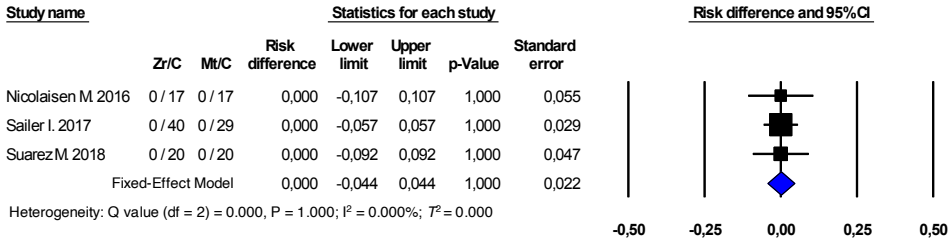


Fig. F.0. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Global Marginal adaptation.

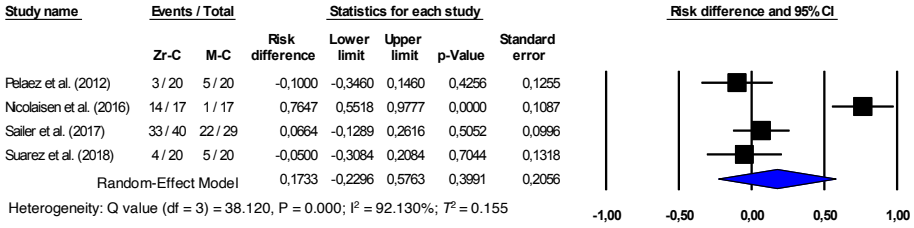
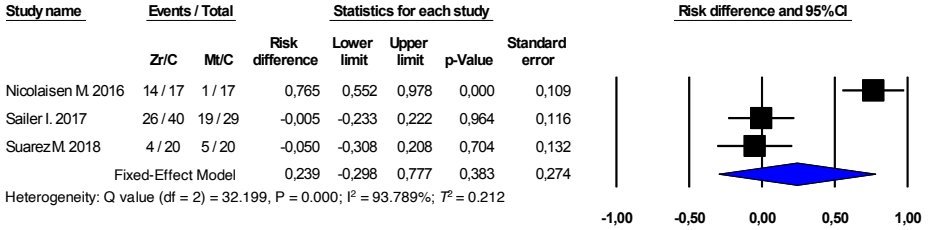


Fig. F.1. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Beta & 3 Marginal adaptation.



© Álvaro Limones Burgos, 2024

Fig. F.2. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Charlie & 2 Marginal adaptation.

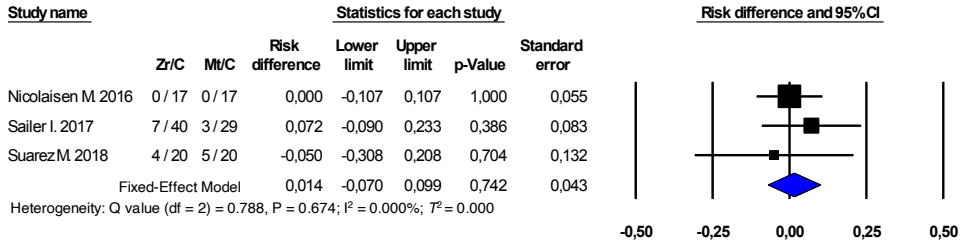


Fig. F.3. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Delta & 1 Marginal adaptation.

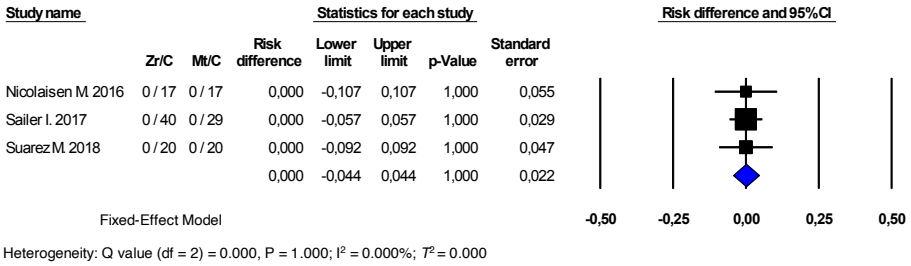


Fig. G.0. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Global Anatomical Form.

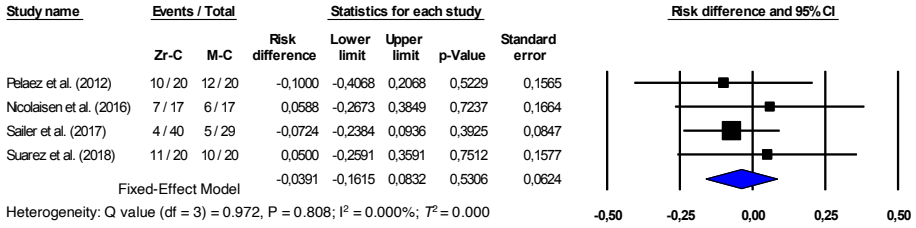
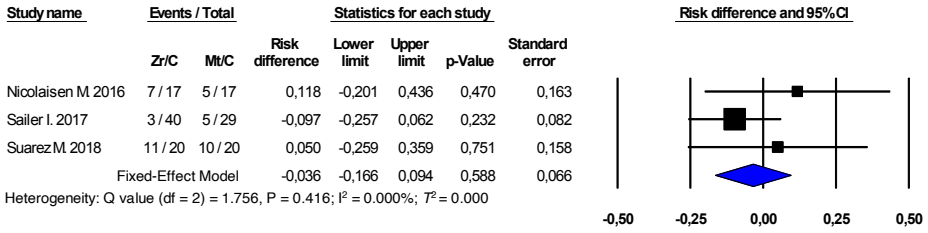


Fig. G.1. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Beta & 3 Anatomical Form.



© Álvaro Limones Burgos, 2024

Fig. G.2. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Charlie & 2 Anatomical Form.

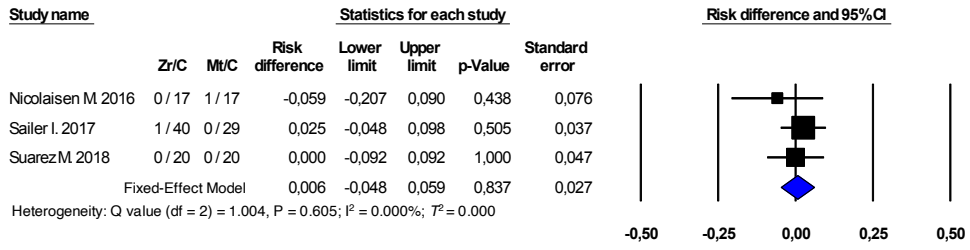
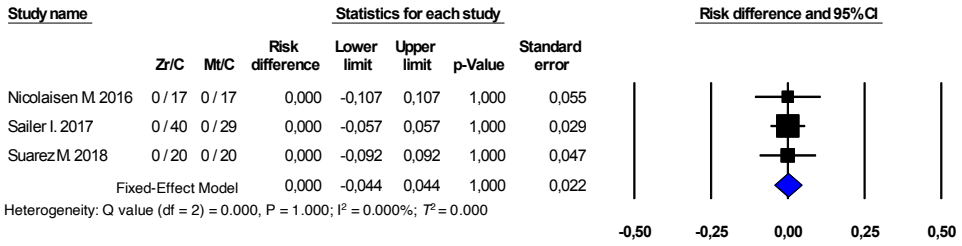


Fig. G.3. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of Delta & 1 Anatomical Form.



2. Biological complications:

Fig. H. Standard mean difference forest plot of zirconia-ceramic vs metal-ceramic in terms of probing pocket depth .

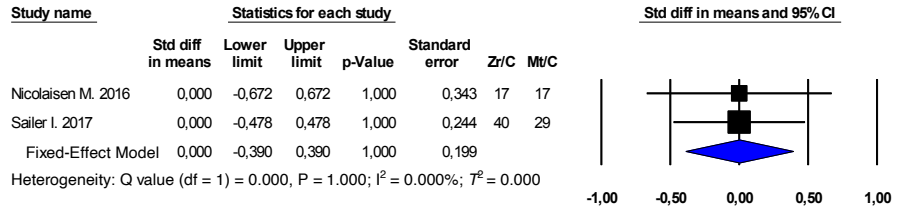


Fig. I. Standard mean difference forest plot of zirconia-ceramic vs metal-ceramic in terms of proving attachment level.

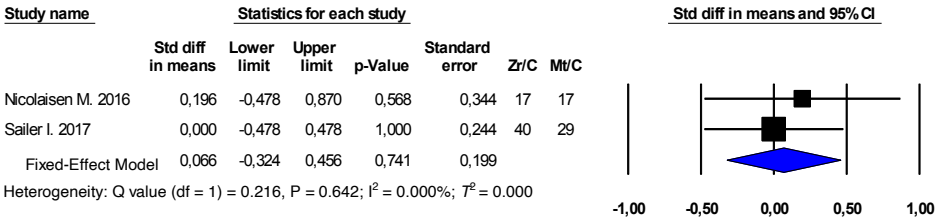
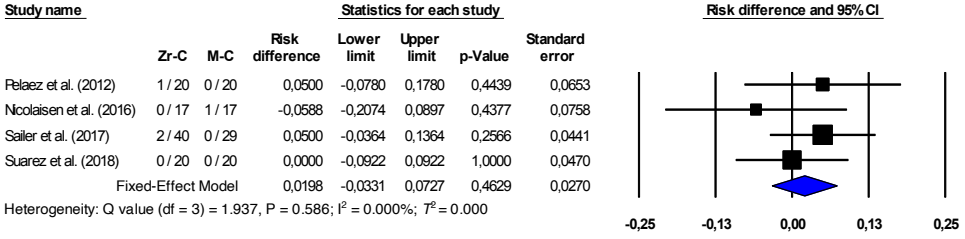


Fig. J. Risk difference forest plot of zirconia-ceramic vs metal-ceramic in terms of loss of tooth vitality.



© Álvaro Limones Burgos, 2024

Appendix 2 – Equivalent Criteria

(Supplementary data)

Equivalence for modified United States Public Health Service (USPHS) criteria and the California Dental Association (CDA) assessment system.

Equivalent Criteria				
	USPHS	**Criteria**	**CDA**	**Criteria**
Framework Fracture	Alfa	No fracture of framework	4	Excellent
	Beta		3	Good
	Charlie		2	Acceptable (repair)
	Delta	Fracture of framework	1	Unacceptable (replacement)
Veneering fracture	Alfa	No fracture	4	Excellent
	Beta	Chipping, but polishing is possible	3	Good
	Charlie	Chipping extending to the framework	2	Acceptable (repair)
	Delta	New restoration needed	1	Unacceptable (replacement)
Occlusal Wear	Alfa	No occlusal wear on restoration or opposite teeth	4	Excellent
	Beta	Occlusal wear on restoration or opposite teeth is < 2 mm	3	Good
	Charlie	Occlusal wear on restoration or opposite teeth is > 2 mm	2	Acceptable (repair)
	Delta	New restoration needed	1	Unacceptable (replacement)
Margical Adaptation	Alfa	No explorer catch	4	Excellent
	Beta	Slight explorer catch, but no gap	3	Good
	Charlie	Gap with some dentin or cement exposure	2	Acceptable (repair)
	Delta	New restoration needed	1	Unacceptable (replacement)
Anatomical form	Alfa	Ideal anatomical shape, good proximal contact	4	Excellent
	Beta	Slightly over or under contoured, weak proximal contact	3	Good
	Charlie	Highly over or under contoured, open proximal contact	2	Acceptable (repair)
	Delta	New restoration needed	1	Unacceptable (replacement)

© Álvaro Limones Burgos, 2024

ANEXO 2.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II).

SEGUNDO ARTÍCULO: (Limones, Molinero-Mourelle, et al., 2020)			
Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]	**Respuesta**
Pregunta 1	-	Pregunta de investigación y criterios de inclusión: ¿Se establecieron preguntas de investigación y criterios de inclusión que incluyeron los componentes de PICO (Población, Intervención, Grupo comparador, Resultado)?	Si
Pregunta 2	**CRÍTICA**	**Protocolo registrado antes del inicio de la revisión:** ¿Se registró o publicó el protocolo de la revisión antes de empezar su realización y se han justificado las desviaciones significativas del protocolo?	**Si**
Pregunta 3	-	Diseño de estudios seleccionados: ¿Se explica por los autores cual es el diseño de estudio que se ha seleccionado para confeccionar la revisión?	Si
Pregunta 4	**CRÍTICA**	**Búsqueda exhaustiva de la literatura:** ¿Los autores de la revisión utilizaron una estrategia de búsqueda exhaustiva?	**Si**
Pregunta 5	-	Duplicación en la selección de estudios: ¿Los autores de la revisión realizaron la selección de estudios por duplicado?	Si
Pregunta 6		Duplicación en la extracción de datos: ¿Los autores de la revisión realizaron la extracción de datos por duplicado?	Si
Pregunta 7	**CRÍTICA**	**Justificación para excluir estudios individuales:** ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?	**Parcialmente** (Incluida en el diagrama de flujo y descriptivamente en la sección de resultados)
Pregunta 8	-	Descripción de estudios incluidos: ¿Los autores de la revisión describieron los estudios incluidos con suficiente detalle?	Si
Pregunta 9	**CRÍTICA**	**Evaluación del riesgo de sesgo de los estudios primarios incluidos:** ¿Utilizaron los autores de la revisión una técnica satisfactoria para evaluar el riesgo de sesgo (RoB) en los estudios individuales que se incluyeron en la revisión?	**Si**

© Álvaro Limones Burgos, 2024

Pregunta 10	-	Fuentes de financiación: ¿Los autores de la revisión informaron sobre las fuentes de financiación de los estudios incluidos en la revisión?	No
Pregunta 11	**CRÍTICA**	**Adecuados métodos analíticos:** Si se realizó un metaanálisis, ¿los autores de la revisión utilizaron métodos apropiados para la combinación estadística de resultados?	**Si**
Pregunta 12	-	Impacto del RoB en la metaanálisis: Si se realizó un metaanálisis, ¿los autores de la revisión evaluaron el impacto potencial de RoB de los estudios individuales sobre los resultados del metaanálisis u otra síntesis de evidencia?	Si (Ninguno de los estudios incluidos reportó alto RoB por lo que no fue necesario hacer una consideración especial al respecto)
Pregunta 13	**CRÍTICA**	**Impacto del RoB en los resultados y discusión:** ¿Los revisores tuvieron en cuenta el RoB de los estudios primarios al interpretar/discutir los resultados de la revisión?	**Si** (Ninguno de los estudios incluidos reportó alto RoB por lo que no fue necesario hacer una mención especial al respecto).
Pregunta 14	-	Evaluación de la Heterogeneidad: ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?	Si (Se realizó un análisis por subgrupos para explorar la heterogeneidad)
Pregunta 15	**CRÍTICA**	**Evaluación de la presencia e impacto del sesgo de publicación:** Si se realizó una síntesis cuantitativa, ¿se llevó a cabo una investigación adecuada del sesgo de publicación (sesgo de estudios pequeños) y se discutió su probable impacto en los resultados de la revisión?	**No aplicable** (Inclusión insuficiente de ECA para analizarlo)
Pregunta 16	-	Conflictos de interés: ¿Se informaron las posibles fuentes de conflicto de interés, incluyendo cualquier financiamiento recibido para llevar a cabo la revisión?	Si
VALORACIÓN FINAL DEL RIESGO DE SESGO: **MODERADA CONFIANZA**			

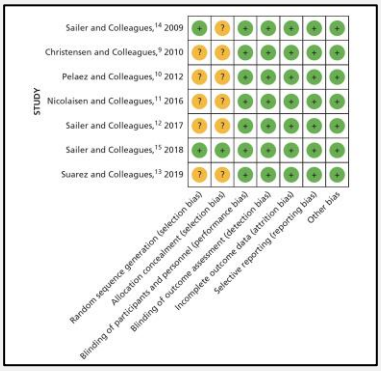
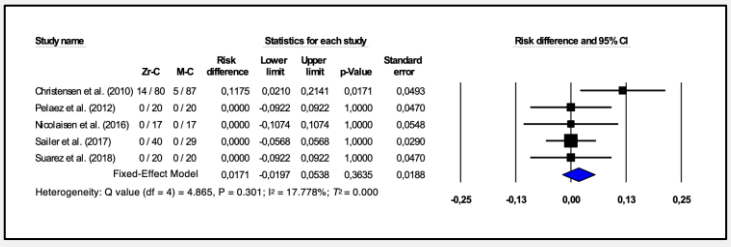
© Álvaro Limones Burgos, 2024

ANEXO 2.2: Resultados de la evaluación del riesgo de sesgo (ROBIS).

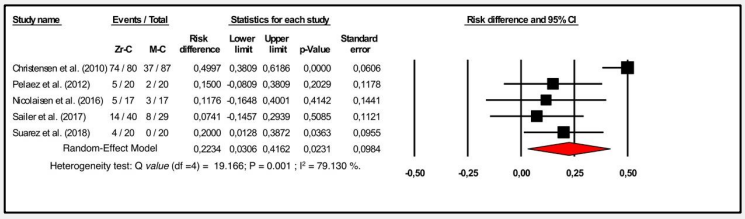
1ª FASE	2ª FASE				3ª FASE
Identificación de la pregunta de investigación y sus componentes	**Criterios de elegibilidad (Dominio 1)**	**Identificación y selección de estudios (Dominio 2)**	**Recopilación de datos y evaluación de estudios (Dominio 3)**	**Síntesis y conclusiones (Dominio 4)**	**Riesgo de sesgo global**
¿La pregunta abordada en la revisión está correctamente definida de acuerdo a sus objetivos? S/PS/PN/N/NI* Revisiones de intervención: ✓ *P: Población* ✓ *I: Intervención* ✓ *C: Comparador* ✓ *O: Variables* ✓ Revisiones de pruebas de precisión diagnóstica: *P: Población* *I: Test diagnóstico* *C: Estándar de referencia* *O: Enfermedad objetivo* Revisiones de asociaciones etiológicas o factores de riesgo: *P: Población* *I: Exposición* *C: Comparador* *O: Variables* Revisiones de estudios pronóstico: *P: Población* *I: Factor pronóstico o modelo predictivo* *O: Desenlace a predecir* Revisiones de variables e instrumentos de medición: *P: Población/Campo de conocimiento* *O: Variables e instrumentos de medición*	1.1 ¿Se ajustó la revisión a los objetivos y criterios de elegibilidad predefinidos? S/PS/PN/N/NI* 1.2 ¿Los criterios de elegibilidad eran apropiados para la pregunta de la revisión? S/PS/PN/N/NI* 1.3 ¿Eran inequívocos los criterios de elegibilidad? S/PS/PN/N/NI* 1.4 ¿Eran apropiadas las restricciones de los criterios de elegibilidad basadas en las características del estudio (por ejemplo, fecha, tamaño de la muestra, calidad del estudio, resultados medidos, etc.)? S/PS/PN/N/NI* 1.5 ¿Era adecuada alguna restricción en los criterios de elegibilidad basada en las fuentes de fuentes de información (por ejemplo, estado o formato de publicación, idioma, disponibilidad de datos)? S/PS/PN/N/NI*	2.1 ¿Incluyó la búsqueda una gama apropiada de bases de datos/fuentes electrónicas para informes publicados y no publicados? S/PS/PN/N/NI* 2.2 ¿Se utilizaron métodos adicionales a la búsqueda en bases de datos para identificar informes relevantes? S/PS/PN/N/NI* 2.3 ¿Los términos y la estructura de la estrategia de búsqueda permitieron recuperar el mayor número posible de estudios elegibles? S/PS/PN/N/NI* 2.4 ¿Fueron apropiadas las restricciones basadas en la fecha, el formato de publicación o el idioma? S/PS/PN/N/NI* 2.5 ¿Se realizaron esfuerzos para minimizar los errores en la selección de los estudios? S/PS/PN/N/NI*	3.1 ¿Se hicieron esfuerzos para minimizar los errores en la recogida de datos? S/PS/PN/N/NI* 3.2 ¿Se dispuso de suficientes características de los estudios para que tanto los autores de la revisión como los lectores pudieran interpretar los resultados? S/PS/PN/N/NI* 3.3 ¿Se recogieron todos los resultados relevantes de los estudios para su uso en la síntesis? S/PS/PN/N/NI* 3.4 ¿Se evaluó formalmente el riesgo de sesgo (o la calidad metodológica) utilizando criterios apropiados? S/PS/PN/N/NI* 3.5 ¿Se hicieron esfuerzos para minimizar los errores en la evaluación del riesgo de sesgo? S/PS/PN/N/NI*	4.1 ¿Incluyó la síntesis todos los estudios que debía? S/PS/PN/N/NI* 4.2 ¿Se comunicaron todos los análisis predefinidos o se explicaron las desviaciones? S/PS/PN/N/NI* 4.3 ¿Fue adecuada la síntesis dada la naturaleza y la similitud de las preguntas de investigación, los diseños de los estudios y los resultados de los estudios incluidos? S/PS/PN/N/NI* 4.4 ¿La variación entre estudios (heterogeneidad) fue mínima o se abordó en la síntesis? S/PS/PN/N/NI* 4.5 ¿Fueron sólidos los hallazgos, por ejemplo, como se demostró mediante el gráfico en embudo o los análisis de sensibilidad? S/PS/PN/N/NI* 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? S/PS/PN/N/NI*	A. ¿Se han abordado en la interpretación de los resultados todos los problemas señalados en los dominios del uno al cuatro? S/PS/PN/N/NI* B. ¿Se consideró adecuadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? S/PS/PN/N/NI* C. ¿Evitaron los revisores destacar los resultados en función de su significación estadística? S/PS/PN/N/NI*
¿Coincide la pregunta abordada por la revisión con la pregunta objetivo? SI/INCIERTO/NO	**Preocupación por la especificación de los criterios de elegibilidad del estudio** ALTA/INCIERTA/BAJA	**Preocupación por los métodos utilizados para identificar y/o seleccionar los estudios** ALTA/INCIERTA/BAJA	**Preocupación por los métodos de recogida de datos y evaluación de los estudios** ALTA/INCIERTA/BAJA	**Preocupaciones sobre la síntesis y los resultados** ALTA/INCIERTA/BAJA	**Riesgo de sesgo global considerado** ALTO/INCIERTO/ BAJO/INEXISTENTE

© Álvaro Limones Burgos, 2024

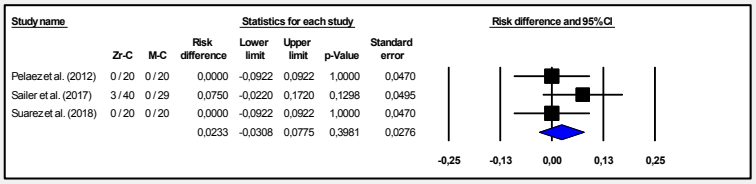
ANEXO 2.3 Resultados de la evaluación de la certeza de la evidencia científica (GRADE).

Dominios	Consideraciones	Variación del nivel de certeza de la evidencia
Riesgo de sesgo [Dominio 1]	La herramienta original para evaluar el riesgo de sesgo (Higgins et al., 2011) no reveló riesgos significativos de sesgo en las dimensiones evaluadas. Los resultados de esta evaluación se presentan en la siguiente figura: 	No aplica
Inconsistencia [Dominio 2]	No se detectó heterogeneidad en ninguna de las variables críticas, excepto en el deslaminado cerámico que fue resuelta moderadamente mediante un análisis de subgrupos (aun así se redujo la evidencia un nivel por este motivo): SUPERVIVENCIA: Q = 4.865; P = 0.301; I^2 = 17.78% DELAMINADO CERÁMICO: Q = 19.166; P = 0.001; I^2 = 79.13%* PÉRDIDA DE RETENCIÓN: Q = 1.581; P = 0.454; I^2 = 00.00% PROFUNDIDAD DE SONDAJE Q = 0.000; P = 0.999; I^2 = 00.00%* PÉRDIDA DE VITALIDAD: Q = 1.937; P = 0.586; I^2 = 00.00% NOTA*: La disparidad de variables e instrumentos de medicación al evaluar complicaciones biológicas y técnicas implicó reducir la certeza de la evidencia otro nivel.	No aplica ↓ 2 niveles No aplica ↓ 1 nivel No aplica
Carácter indirecto de la evidencia [Dominio 3]	Las variables críticas se evaluaron de la manera más directa posible. Sin embargo, el tiempo de seguimiento medio de los estudios incluidos fue de 4.07 años, muy por debajo del tiempo estimado de supervivencia clínicamente razonable, por lo que la extrapolación de conclusiones extraídas merece degradar la evidencia un nivel por presencia de cierto carácter indirecto de la evidencia.	↓ 1 nivel en todas las variables críticas
Imprecisión [Dominio 4]	Los intervalos de confianza de los estudios incluidos se superpusieron unos con los otros en todos los *Forest plot* de las variables críticas como se puede ver en las siguientes figuras: SUPERVIVENCIA: (Umbral de relevancia clínica: diferencia > 10%) 	No aplica a ninguna variable crítica excepto a la Profundidad de Sondaje.y a la Pérdida de Retención que ↓ 1 nivel

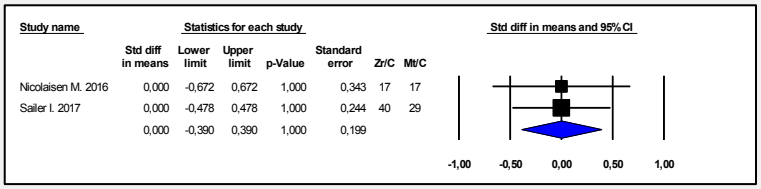
DELAMINADO CERÁMICO: (Umbral de relevancia clínica: diferencia > 10%)



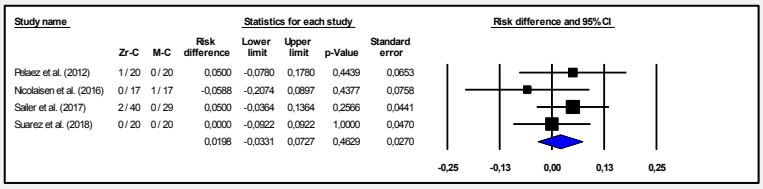
PÉRDIDA DE RETENCIÓN: (Umbral de relevancia clínica: diferencia > 10%)



PROFUNDIDAD DE SONDAJE: (Umbral de relevancia clínica: diferencia > 0.5 mm)



PÉRDIDA DE VITALIDAD: (Umbral de relevancia clínica: diferencia > 10%)

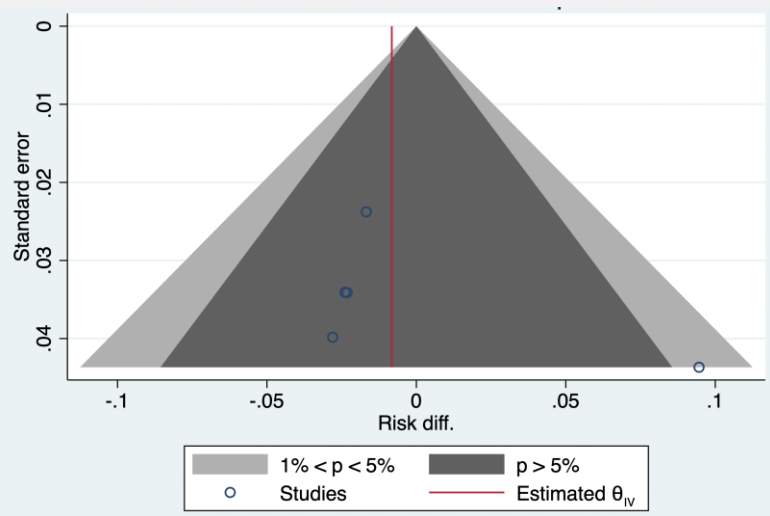


Por otro lado, en la mayoría de las variables críticas la magnitud del efecto y sus IC no cruzaron dos umbrales de decisión o relevancia clínica al mismo tiempo.

En la variable primaria Tasa de Supervivencia, se superó el tamaño optimo de información establecido en 163 participantes por grupo y calculado a partir de los resultados del estudio de Christensen et al., 2010, con una proporción de muestreo de 1:1, un nivel de significancia del 5%, un poder estadístico del 80% y una prueba de hipótesis de χ^2 para muestras independientes).

La excepción fueron las variables – Pérdida de Retención y Profundidad de Sondaje, que agruparon tamaños muéstrales inferiores al tamaño optimo de información de la variable principal por lo que se redujo la evidencia un nivel.

© Álvaro Limones Burgos, 2024

Sesgo de Publicación [Dominio 5]	Los métodos visuales de comprobación del sesgo de publicación con gráficos de embudo basados en asimetrías no son precisos cuando hay pocos estudios incluidos en el análisis (<10 estudios). Los resultados de esta evaluación se presentan en la siguiente figura:  No se detectó presencia de sesgo de publicación con métodos estadísticos mediante la Prueba de Egger (*P=0.273*).	No aplica
Gradiente dosis-respuesta [Dominio 6]	No existe.	No aplica
Gran magnitud del efecto [Dominio 7]	La magnitud combinada del efecto informó de efecto nulo *(P = 0.364)*.	No aplica
Factores de confusión plausibles [Dominio 8]	No se detectaron factores de confusión plausible que aumentase nuestra confianza en las estimaciones.	No aplica
Computo final de la evaluación	Disminución de la certeza de la evidencia fue del siguiente modo: SUPERVIVENCIA: ↓ 1 nivel por carácter indirecto de la evidencia. DELAMINADO CERÁMICO: ↓ 3 nivel por carácter indirecto de la evidencia e inconsistencia (dos niveles). PÉRDIDA DE RETENCIÓN: ↓ 2 nivel por carácter indirecto de la evidencia e imprecisión. PROFUNDIDAD DE SONDAJE: ↓ 3 nivel por carácter indirecto de la evidencia, imprecisión e inconsistencia. PÉRDIDA DE VITALIDAD: ↓ 1 nivel por carácter indirecto de la evidencia.	

ANEXO 2.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA).

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
TÍTULO			
Título	1	Identificar en el título del manuscrito que se trata de una revisión sistemática.	p. 230
RESUMEN			
Objetivos	2.1	Proporcionar una declaración explícita de los principales objetivos de la revisión.	p.230
Criterios de Inclusión y exclusion	2.2	Especificar los criterios de inclusión y exclusión de la revisión.	p.230
Fuentes de información	2.3	Especificar las fuentes de información (por ejemplo, las bases de datos o registros) utilizados para identificar los estudios y la fecha de la última búsqueda.	p.230
Riesgo de sesgo	2.4	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos.	/
Métodos de la síntesis de los resultados	2.5	Especifique los métodos utilizados para presentar y sintetizar los resultados.	/
Estudios incluidos	2.6	Dar el número total de estudios y participantes incluidos y resumir las características pertinentes de los estudios.	p.230
Síntesis de los resultados	2.7	Indicar las variables de estudio. Si se ha realizado un metaanálisis, notifique la estimación del efecto o exposición y su IC. Si se comparan dos o más grupos, indíquen la dirección del efecto (es decir, que grupo es favorecido).	p.230
Degradación de la evidencia	2.8	Proporcionar un breve resumen de la degradación de la evidencia (por ejemplo, por riesgo de sesgo, incoherencia o impresión.	p.230
Interpretación	2.9	Proporcionar una interpretación general de los resultados e implicaciones importantes.	p.230
Financiación	2.10	Especificar la principal fuente de financiación de la revisión.	/
Registro	2.11	Proporcionar el nombre del registro y el número de registro.	/
INTRODUCIÓN			
Fundamento	3	Describir la justificación de la revisión en el contexto de los conocimientos existentes.	p.230-1
Objetivos	4	Proporcionar una declaración explícita de los objetivos de la revisión.	p.231
MÉTODOS			
Criterios de elegibilidad	5	Especificar los criterios de inclusión y exclusión de la revisión y cómo se agruparon los estudios para la síntesis.	p.232
Fuentes de información	6	Especificar todas las bases de datos, registros, sitios web, organizaciones, listas de referencia y otras fuentes consultadas para identificar estudios. Especificar la fecha de la última revisión.	p.232
Estrategia de búsqueda	7	Presentar las estrategias de búsqueda completa para al menos la base de datos principal, indicar filtros u otras estrategias de búsqueda utilizadas.	p.232
Proceso de selección	8	Especificar los métodos utilizados para la selección de estudios.	p.232

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Recogida de datos	9	Especifique los métodos utilizados para recopilar datos de los estudios, incluyendo cuántos revisores recogieron datos de cada informe.	p.232
Datos extraidos	10a	Listar y definir todas las variables críticas.	p.232-3
	10b	Listar y definir todas las demás variables secundarias. Describa cualquier suposición sobre cualquier información que falta o no esté clara.	p.330
Métodos de evaluación del riesgo de sesgo	11	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos, incluyendo detalles de la herramienta o herramientas utilizadas, cuántos revisores evaluaron cada estudio y si trabajaban independientemente, y si se utilizaron procesos de automatización.	p.232-3
Medidas del efecto	12	Especifícar para cada variable la medida(s) del efecto (por ejemplo. relación de riesgo, diferencia media) utilizada en la síntesis o presentación de resultados.	p.233
Métodos de síntesis	13a	Describir los procesos utilizados para decidir qué estudios eran elegibles para cada síntesis (por ejemplo: tabular las características de intervención del estudio y comparar con los grupos previstos para cada síntesis).	p.233
	13b	Describir cualquier método necesario para preparar los datos de presentación o síntesis, como el tratamiento de las estadísticas de datos faltantes o las conversiones de datos.	p.338e2-3
	13c	Describir cualquier método utilizado para tabular o visualizar los resultados de los estudios y la síntesis individuales.	p.233
	13d	Describir cualquier método utilizado para sintetizar los resultados y proporcionar una justificación. Si se realizó un metaanálisis, describa el(s) modelo(s), método(s) para identificar la presencia y extensión de la heterogeneidad estadística, y paquetes de software utilizado.	p.233
	13e	Describa cualquier método utilizado para explorar posibles causas de heterogeneidad entre los resultados del estudio (por ejemplo, análisis de subgrupos, metaregresión).	p.233
	13f	Describir cualquier análisis de sensibilidad realizado para evaluar la solidez de los resultados sintetizados.	/
Implicaciones del riesgo de sesgo por datos faltantes	14	Describir cualquier método han utilizado para evaluar el riesgo de sesgo debido a la falta de datos (a partir de sesgos informativos).	/
Evaluación de la certeza de la evidencia	15	Describir cualquier método utilizado para evaluar la certeza (o la confianza) de la evidencia.	/
RESULTADOS			
Selección de estudios	16a	Describir los resultados del proceso de búsqueda y selección, idealmente utilizando un diagrama de flujo.	p.338e1
	16b	Presentar los estudios excluidos y sus razones.	p.338e1 y p.233
Características de los estudios incluidos	17	Presentar cada estudio incluido y sus características.	p.338e2-3
Riesgo de sesgo en los estudios incluidos	18	Presentar el riesgo de sesgo para cada estudio incluido.	p. 234

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Resultados de los estudios individuales	19	Para todas las variables presentar la estimación del efecto y su IC.	p. 233-4
Resultados de la síntesis	20a	Resumir brevemente las características y el riesgo de sesgo de los estudios incluidos.	p. 234
	20b	Si se ha realizado un metaanálisis, presente para cada una de las estimaciones medidas de heterogeneidad estadística. Si se comparan los grupos, describa la dirección del efecto.	p. 233-4
	20c	Informar de posibles causas de heterogeneidad.	p. 233-4
	20d	Informar de los análisis de sensibilidad realizados para evaluar la solidez de los resultados sintetizados.	/
Evaluación del riesgo de sesgo por datos faltantes	21	Informar de la evaluación del riesgo de sesgo debido a datos faltantes.	p. 234
Certeza de la evidencia	22	Evaluaciones de la certeza (o confianza) de la evidencia para cada variable evaluada.	/
DISCUSIÓN			
Discusión	23a	Proporcionar una interpretación general de los resultados en el contexto científico actual.	p. 235-6
	23b	Examinar cualquier limitación de la evidencia incluida en la revisión.	p. 237
	23c	Debatir las limitaciones de los procesos de revisión utilizados.	/
	23d	Debatir las implicaciones de los resultados de la práctica, la política y la futura investigación.	p. 237
OTRA INFORMACIÓN			
Registro y protocolo	24a	Proporcionar información del registro de la revisión, incluyendo el nombre del registro y el número de registro, o afirmar que la revisión no está registrada.	p. 231
	24b	Indicar dónde puede accederse al protocolo de revisión o afirmar que no se ha elaborado un protocolo.	p. 231
	24c	Describir y explicar las modificaciones de la información facilitada en el registro o en el protocolo.	/
Apoyo	25	Describa las fuentes de ayuda financiera o no financiera para la revisión, y el papel de los patrocinadores en la revisión.	/
Intereses de competencia	26	Declare cualquier interés competitivo de los autores de revisión.	p. 237
Disponibilidad de datos, código y otros materiales	27	Informe sobre cuál de los siguientes está disponible públicamente y dónde se pueden encontrar: plantillas de recogida de datos; datos extraídos de estudios incluidos; datos utilizados para todos los análisis; o cualquier otro material utilizado en la revisión.	p.232

© Álvaro Limones Burgos, 2024

ANEXO III: Manuscrito de la tercera revisión sistemática

(Wilensky et al., 2023)

Check for updates

Received: 7 September 2022 | Revised: 3 January 2023 | Accepted: 12 February 2023

DOI: 10.1111/jcpe.13794

SYSTEMATIC REVIEW

Journal of Clinical Periodontology WILEY

The efficacy of implant surface decontamination using chemicals during surgical treatment of peri-implantitis: A systematic review and meta-analysis

Asaf Wilensky[1] | Lior Shapira[1] | Alvaro Limones[2] | Conchita Martin[3,4]

[1]Department of Periodontology, Hadassah Medical Center, Faculty of Dental Medicine, Hebrew University of Jerusalem, Jerusalem, Israel

[2]Department of Conservative and Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain

[3]Department of Dental Clinical Specialties, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain

[4]BIOCRAN Research group, Complutense University of Madrid (UCM), Madrid, Spain

Correspondence
Asaf Wilensky, Department of Periodontology, Hadassah Medical Center, Faculty of Dental Medicine, Hebrew University of Jerusalem, P.O. Box 12272, Jerusalem 91120, Israel.
Email: asafw@ekmd.huji.ac.il

Abstract

Aim: To answer the following PICOS question: "In adult patients with peri-implantitis, what is the efficacy of surgical therapy with chemical surface decontamination of implant surfaces in comparison with surgical therapy alone or surgery with placebo decontamination, on probing pocket depth (PD) reduction and bleeding on probing (BoP)/suppuration on probing (SoP), in randomized controlled clinical trials (RCTs) and non-RCTs with at least 6 months of follow-up?"

Materials and Methods: Six databases were searched from their inception up to 20 May 2022. Data on clinical outcome variables were pooled and analysed using mean differences (MDs), risk ratios (RRs), or risk differences (RDs) as appropriate, 95% confidence intervals (CIs), and prediction intervals (PIs) in the case of significant heterogeneity. Primary outcomes were determined as changes in PD and BoP/SoP. Secondary outcomes were radiographic marginal bone loss (MBL), implant loss, and disease resolution. PROSPERO registration number: CRD42022325603.

Results: Six RCTs—two with moderate, three with high, and one with low risk of bias (RoB)—were included. These studies test the adjunctive effect of photodynamic therapy (PDT), chlorhexidine (CHX), and administration of local antibiotics (LAbs) during surgery on the clinical outcome. In a single 12-month study, the adjunctive use of local antibiotics showed a clinically relevant reduction of PD [MD = 1.44; 95%CI (0.40 to −2.48)] and MBL [MD = 1.21; 95%CI (0.44–1.98); one trial, 32 participants]. PDT showed a small but significant reduction in BoP [MD = 7.41%; 95%CI (0.81–14.00); $p = 0.028$; two trials; 42 participants]. Treatment with CHX resulted in no significant changes in PD, BoP, or MBL compared to placebo (saline solution). None of the interventions affected disease resolution and implant loss. Certainty of the evidence was very low for all outcome measures assessed.

Conclusions: Within the limitations of this systematic review and the meta-analysis, adjunctive use of chemicals such as PDT, CHX, and LAbs for surface decontamination during surgery of peri-implantitis cannot be recommended as superior to standard debridement procedures (mechanical debridement with or without saline).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2023 The Authors. *Journal of Clinical Periodontology* published by John Wiley & Sons Ltd.

KEYWORDS
dental implants, systematic review, therapy

Clinical Relevance

Scientific rationale for the study: Numerous surgical techniques to treat peri-implantitis have been suggested, most include various approaches of decontamination of implant surfaces. This work systematically reviews the efficacy of chemical decontamination of implant surfaces during surgery.

Principal findings: Twelve months post-treatment, PDT significantly reduced BoP, while LAbs significantly reduced PD and MBL. However, the analysis was based on a small number of studies with very low certainty of the evidence.

Practical implications: There is no decisive evidence (limited and heterogenous results) showing that the adjunctive use of chemical implant surface decontamination improves the outcome of surgical therapy of peri-implantitis.

1 | INTRODUCTION

Peri-implantitis is a chronic inflammatory disease involving the breakdown of soft and hard tissues supporting a dental implant, which may ultimately result in implant and restoration loss (Berglundh et al., 2018). Among the various factors that have been implicated in the aetiology of peri-implantitis, bacterial dysbiosis and the concomitant growth of various Gram-negative anaerobic microflora is considered a major risk factor (Sanz-Martin et al., 2017; Zhuang et al., 2016). The prevalence of peri-implantitis ranges from 10% to 22% at the implant level (Salvi et al., 2019) and between 22% and 56% at the patient level (Derks & Tomasi, 2015; Lindhe & Meyle, 2008). Thus, it is a global health concern affecting millions of people worldwide. Moreover, since dental implants are generally performed in the mature/elderly population, disease burden will further expand owing to the expected increase in life expectancy.

Since the aetiology of peri-implantitis is associated with the peri-implant biofilm, most of the suggested therapeutic strategies are aimed at reducing the biofilm load and shifting its composition. Numerous mechanical treatment protocols have been suggested for the treatment of peri-implantitis, including non-surgical techniques, surgical procedures, and combined approaches. However, non-surgical anti-infective treatments have limited efficacy in most cases of peri-implantitis (Lindhe & Meyle, 2008; Renvert et al., 2008; Schwarz et al., 2015), which is possibly due to the difficulty of completely removing the plaque from the implant surfaces by mechanical debridement alone. Therefore, surgical debridement approaches have been suggested, and the adjunctive use of various chemotherapeutic agents was advocated to allow better decontamination of implant surfaces to increase the overall effectiveness of the surgical mechanical treatment. The most common anti-microbial treatments include the local application of substances such as chlorhexidine (CHX), citric acid, hydrogen peroxide, 35% phosphoric acid gel, tetracycline, minocycline, or saline (Mellado-Valero et al., 2013). Although some studies show that surgical debridement can improve clinical parameters, such as reduced PD and bleeding on probing (BoP), and are capable of slowing or attenuating the progression of peri-implantitis (Heitz-Mayfield & Mombelli, 2014; Renvert et al., 2012), there is still no single "gold standard" treatment protocol for treating peri-implantitis. Furthermore, evidence regarding the efficacy of chemical decontamination of implant surfaces is limited (Patil et al., 2022).

The aim of the present systematic review was to answer the following PICOS question: "In adult patients with peri-implantitis (P), what is the efficacy of surgical therapy with adjunctive chemical implant surface decontamination (I), as compared to mechanical debridement with or without saline (standard therapy) (C), with regard to reduction of PD and BoP/suppuration on probing (SoP) (O), as reported in prospective controlled clinical trials with a follow-up of at least 6 months"?(S) (Table 1).

RCTs and non-RCTs were chosen because these types of studies offer the highest level of testing the efficacy of the selected interventions.

2 | MATERIALS AND METHODS

2.1 | Protocol development and systematic review reporting

The reporting of this systematic review adhered to the 2020 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement and was conducted following the Cochrane's Handbook (Cumpston et al., 2019).

The protocols were presented to the Workshop Committee for the XVIII European Workshop on Periodontology. Protocol approval and registration in the International Prospective Register of Systematic Reviews (PROSPERO registration number: CRD42022325603) were completed prior to the study commencement.

2.2 | Eligibility criteria

2.2.1 | Inclusion criteria

Prospective controlled clinical trials with a minimum of eight randomized/allocated patients per group were included. All patients were

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Components of the PICOS question.

P: Population	Individuals aged above 18 years otherwise healthy with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018) or with other well-defined criteria, based on clear clinical and radiographic parameters.
I: Intervention	Surgical therapy of peri-implantitis, using scalers (hand and ultrasonic) to remove hard deposits with or without saline application, with chemical surface decontamination of the implant surfaces with topical antibiotics, any antibacterial agent [chlorhexidine (CHX), cetylpyridinium chloride (CPC) etc.], acid, ethylenediaminetetraacetic acid (EDTA) and photodynamic therapy (PDT), alone or combined with other mechanical/physical approach.
C: Comparator	Surgical therapy of peri-implantitis using scalers with or without saline application or another placebo solution not aiming at chemical decontamination of the implant surface.
O: Outcomes	*Primary outcomes* • Probing depth reduction measured as mean difference of PD (mm) • Bleeding on probing/suppuration on probing • *Secondary outcomes* • Radiographic marginal bone level (MBL) changes • Composite outcomes (primary outcomes + MBL) • Dental implant/implant-supported prosthesis loss or survival • Adverse effects: gingival recession (mm) and antibiotics resistance (in studies that used local antibiotics) • Patient-reported outcome measures (PROMs): patient discomfort/pain, oral-health-related quality of life (OHIP-CP), and economic factors • Disease resolution
S: Study design and duration	Parallel, cross-over or split-mouth RCTs or CCTs with a minimum of eight randomized/allocated patients per group, with a follow-up of at least 6 months.

>18 years old and were diagnosed with peri-implantitis, as defined by the 2017 World Workshop (Berglundh et al., 2018) or by other well-defined criteria, based on clear clinical and radiographic parameters. Reports included PD and BoP/SoP measurements at 6 months or more of follow-up.

2.2.2 | Exclusion criteria

Studies carried out specifically on patients with diabetes or other known systemic conditions, studies with no supporting periodontal therapy performed at 3–6-month intervals during the first year post therapy and throughout the follow-up period, and pre-clinical in vivo studies were excluded.

2.2.3 | Information sources and search

Two investigators (CM, AL) independently searched MEDLINE (via Pubmed), Embase, Cochrane Library (CENTRAL), Web of Science, Scopus, and clinicaltrials.gov from the inception of each database up to 20 May 2022, to identify published, unpublished, and ongoing studies. The full electronic search strategy for all electronic databases is presented (exactly as run without search filters) in Table S1. To identify newly published studies, an email alert service was used on MEDLINE. Reference lists of all relevant primary studies and systematic reviews were also screened to identify additional data sources. Because of time limitations, only publications in English language were included. Reference lists of retrieved papers and previously published systematic reviews were also manually searched. Grey literature was not excluded from our search. Attempt to contact primary authors was made if additional information was needed.

2.3 | Study selection

Citations identified from the electronic searches were imported into Covidence Software. After reviewer calibration, eligibility assessment was based on a two-step approach: (1) screening of titles and abstracts, and (2) full-text analysis. Discrepancies arising from independent peer review were resolved by discussion and consultation among all authors. During the full-text screening, a list of excluded reports and the reasons for elimination from the study were recorded. Inter-observer agreement for the screening of complete articles was assessed using the kappa (κ) score.

2.4 | Data collection process

Based on the Cochrane recommendations, a standardized, pre-piloted data extraction form was designed for data collection. Data were extracted from eligible studies and recorded by two reviewers (Asaf Wilensky, Lior Shapira). Two additional examiners (Alvaro Limones, Conchita Martin) independently cross-checked the accuracy and validity of all data obtained from the studies. Studies lacking enough data for meta-analyses were kept in the systematic review but excluded from the meta-analyses. Corresponding authors of studies or published protocols were contacted for further information in cases where the data were incomplete or missing. In the event that the same study was reported in multiple publications, the data were collated in order for each study, rather than each report, to be constituted as the unit of interest for the review, giving it a unique identifier based on its primary publication.

2.5 | Risk of bias (RoB) in individual studies

RoB and quality assessment were conducted by two reviewers (Alvaro Limones, Conchita Martin) using the Cochrane RoB 2.0 tool

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(Sterne, Savovic, et al., 2019). Assessments were made based on journal articles and trial protocols, when available on international registry platforms, using PD and BoP as primary outcomes. Funding for the study and perceived independence of the authors were also evaluated.

2.6 | Data analysis

When the differences between (Δ) baseline and end were not reported, they were calculated using the formula: $\Delta Var = Var2 - Var1$, where Var1 is the mean value before treatment and Var2 the mean value after treatment. In addition, the variance ΔVar was estimated using the formula $SVar^2 = SVar1^2 + SVar2^2 - (2*r*SVar1*SVar2)$, where $SVar^2$ is the variance of the difference, $SVar1^2$ is the variance of the mean baseline value, and $SVar2^2$ is the variance of the mean end value. SVar1 and SVar2 are the standard deviations (SDs) of the baseline and end values, respectively. A correlation *r* value of 0.5 was used as previously described (Paraskevas et al., 2008). If treatment effects (mean or SD) were not reported in the publications, the authors were requested to provide this information. Dichotomous outcomes were pooled and analysed using risk difference (RD) or risk ratio (RR), along with their corresponding 95% confidence intervals (CIs). Subgroup analyses were performed based on the chemical procedure type: (1) photodynamic therapy (PDT); (2) antiseptics; and (3) local antibiotics (LAbs). Forest plots were created to illustrate the effects of the global estimation in the meta-analysis and the various sub-analyses.

The extent and impact of heterogeneity between studies was assessed by inspecting the forest plots and by calculating the *Q*-test according to Dersimonian and Laird (DerSimonian & Laird, 1986). As a complement to the *Q*-test, the I^2 index was calculated in order to measure the percentage of variation in the global estimate that was attributable to heterogeneity ($I^2 = 25\%$: low; $I^2 = 50\%$: moderate; $I^2 = 75\%$: high). Study-specific estimates were pooled using both the fixed effects model (Mantel–Haenszel-Peto test) and the random effects model (Dersimonian–Laird test). In the event significant heterogeneity was found, the random effects model results were presented.

In cases of heterogeneity, in addition to the summary estimate (MD) and 95%CI, prediction intervals were reported to allow more informative inferences and illustrate the range of true effects that can be expected in future settings, by presenting heterogeneity in the same metric as the original effect size measure (IntHout et al., 2016).

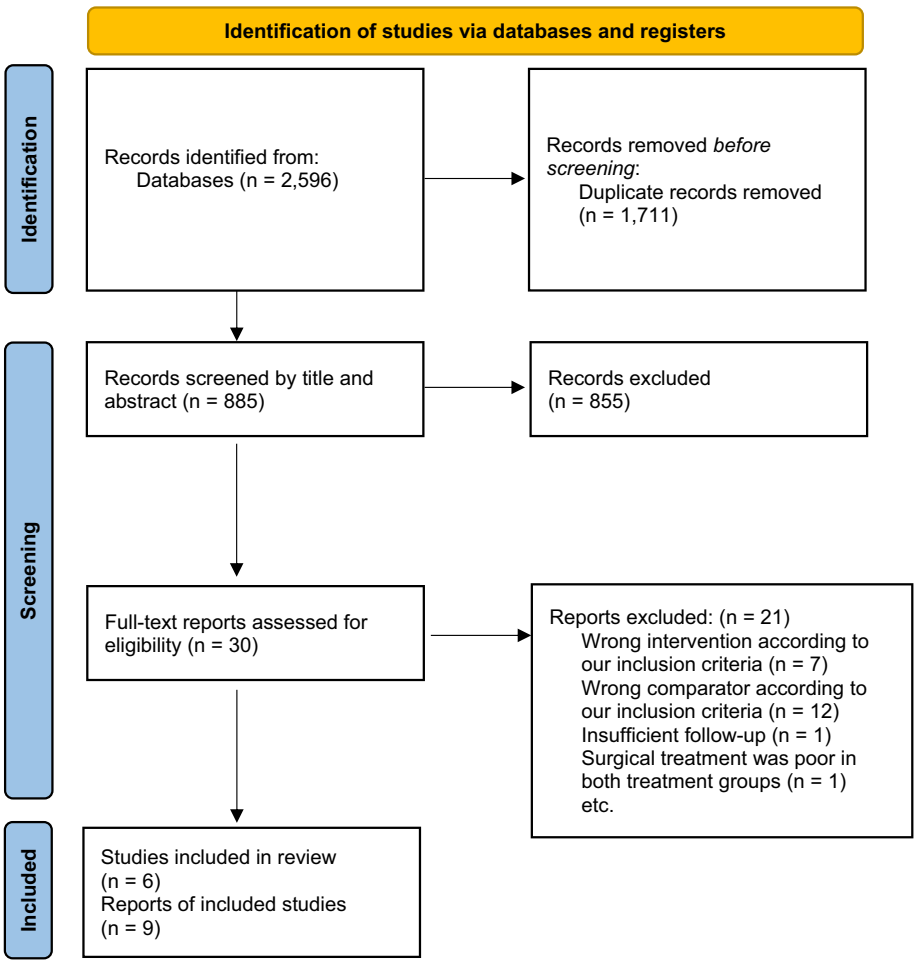


FIGURE 1 Flow diagram: selection of studies evaluating adjunctive chemical surface decontamination during surgical therapy of peri-implantitis. [Colour figure can be viewed at wileyonlinelibrary.com]

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Main characteristics of included studies.

First author and year	Setting	Fundin/support	Study design	Follow-up (months)	Time points (6 months or more)	Author's conclusion
Albaker et al., 2018	Private	University	RCT, Parallel	12	6 and 12	A single application of PDT as an adjunct to OFD does not provide additional benefit in improving clinical and radiographic peri-implant parameters in peri-implantitis.
Carcuac et al., 2016 (without systemic antibiotics)	Public clinics/ University setting	Swedish Research Council, University, Swedish Dental Society	RCT Parallel	12	6 and 12	The local use of chlorhexidine had no overall effect on treatment outcomes.
Carcuac et al., 2016 (with systemic antibiotics)	Public clinics/ University setting	Swedish Research Council, University, Swedish Dental Society	RCT Parallel	12	6 and 12	The local use of chlorhexidine had no overall effect on treatment outcomes. The effect of adjunctive systemic antibiotics depended on implant surface characteristics.
Cha et al., 2019	University	Industry	RCT Parallel	6	6	Repeated local delivery of minocycline would improve the success rate of the surgical treatment of peri-implantitis, in terms of increasing the peri-implant bone gain and improving clinical outcomes in the short healing period.
de Waal et al., 2013	University	Industry	RCT Parallel	12	6 and 12	Surface decontamination with 0.12% CHX + 0.05% CPC in resective surgical treatment of peri-implantitis does not translate to better clinical or radiographical outcomes of the intervention.
Emanuel et al., 2020	University	Industry	RCT Parallel	12	6 and 12	β-Tricalcium phosphate granules with doxycycline hyclate might create favourable conditions that enable implant surface decontamination and soft and hard tissue healing over a prolonged period.
Esposito et al., 2013	University and private	NR	RCT Parallel	12	4 and 12	The use of adjunctive PDT with mechanical cleaning of implants affected by peri-implantitis did not improve any clinical outcomes when compared to mechanical cleaning alone up to 1 year after treatment.

Abbreviations: CHX, chlorhexidine; CPC, cetylpyridinium chloride; NR, not reported; OFD, open flap debridement; PDT, photodynamic therapy; RCT, randomized controlled clinical trial.

16000051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

STATA14 (StataCorp LP, Lakeway Drive, College Station, Texas, USA) intercooled software was used to perform all analyses. Statistical significance was set at $p \leq .05$.

2.7 | RoB across studies: strength of the evidence

The overall quality of evidence was rated using the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) approach (Guyatt et al., 2011) following the summary of findings table format by Carrasco-Labra et al. (Carrasco-Labra et al., 2016). To reach an outcome-level grading of evidence, the authors considered RoB, inconsistency, imprecision, reporting bias, publication bias, large effects, plausible confounding, and dose–response gradient as per the GRADE methodology.

2.8 | Method for addressing the reporting of biases

Publication bias was addressed in several ways: (1) In the case of a sufficient number of included studies (i.e., at least 10 studies; with fewer studies, the power of the tests is too low to distinguish chance from real asymmetry), funnel plots were used, although an asymmetrical funnel plot was not equated with publication bias. (2) Egger's test was used to assess publication bias. (3) Results of tests for funnel plot asymmetry were interpreted by visual inspection of the funnel plot. The contribution made to the totality of the evidence by studies with statistically non-significant results compared to those from studies with statistically significant results was considered.

2.9 | Sensitivity analysis

Post hoc sensitivity analyses were performed by excluding studies, one by one, from the global estimation for primary outcome (PD and BoP changes). Additionally, an overall visual distribution of the specific estimators of each publication was carried out, and studies following an alternative pattern were identified.

3 | RESULTS

3.1 | Study selection

Figure 1 shows the inclusion and exclusion study flow-chart. The electronic search yielded 2596 references, of which 1711 titles were removed after deduplication. Titles and abstracts were screened in the remaining 885 references. After screening, 855 references were discarded (agreement = 95.65%; κ = 0.83), resulting in 30 references. Following the review of the full-text articles, 21 articles were excluded (Table S2) for the following reasons: 7 articles did not meet our inclusion criteria regarding intervention; 12 studies were excluded because the control group deviated from our PICOS question; and 1 study contained an insufficient follow-up period. The results of one study were vastly divergent from results previously reported in other studies on surgical treatment of peri-implantitis, especially for the control groups. Moreover, the effects on BoP reduction seemed inconsistent with changes in PD (Bombeccari et al., 2013). Six RCTs published in nine articles were included in our analysis (agreement = 90%; κ = 0.75).

3.2 | Study characteristics

The methodologies used by the six included studies are shown in Table 2. One study, by Carcuac et al., 2016, included four groups: intervention and control with administration of systemic antibiotics, and intervention and control without administration of systemic antibiotics; hence, this study was added to our analysis as two separate studies, according to the administration of pre-surgical antibiotics (Carcuac et al., 2016). All studies included in the meta-analysis were designed as parallel-arm RCTs. Three studies were conducted in a university setting (Cha et al., 2019; de Waal et al., 2013; Emanuel et al., 2020), one study was carried out at a private clinic (Albaker

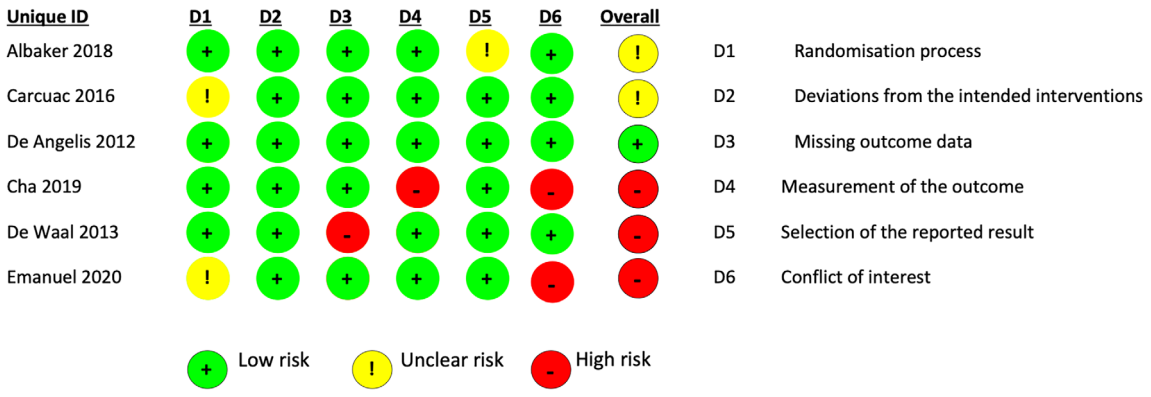


FIGURE 2 Risk-of-bias assessment of the individual studies. [Colour figure can be viewed at wileyonlinelibrary.com]

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Characteristics of the studies and population.

Study	Definition of peri-implantitis	Hygienic phase prior to treatment	Pre-/post-treatment systemic antibiotics	Intervention	Sample size-intention to treat (*n*)	Total no. of implants intention to treat	Overall age (years)			Gender (females)	Smokers	
							Mean	SD	Range	(%)	*n*	Diabetes
Albaker et al., 2018	PD≥5 + BoP/Sup + MBL≥2 mm in 1 year or MBL≥3 mm	Yes	Yes	Debridement + PTD	24	24	NR	NR	NR	NR	12	6 pre-diabetic
Carcuac et al., 2016 (without systemic antibiotics)	PD ≥6 mm + BoP/Sup; MBL >3 mm	Yes	No	Debridement + 0.2% CHX	48	86	65.75	NR	27–88	58.3%	15	3
Carcuac et al., 2016 (with systemic antibiotics)	PD ≥6 mm + BoP/Sup; MBL >3 mm	Yes	Yes	Debridement + 0.2% CHX	52	93	66.3	NR	21–90	71%	18	2
Cha et al., 2019	PD >5, MBL >2 mm + BoP	Yes	Yes	Debridement + minocycline ointment	50	50	61.6	NR	40–84	50%	0	0
de Waal et al., 2013	PD ≥5 mm + BoP/Sup + bone loss ≥2 mm	Yes	No	Debridement + CHX/CPC	30	79	59.4	14	NR	66%	11	NR
Emanuel et al., 2020	PD 6–10 mm + BoP/Sup + bone loss >2 mm	Yes	Yes	Debridement + βTCP sustained-release doxycycline	27	28	64.81	7.61	NR	16%	0	NR
Esposito et al., 2013 de Angelis 2012	MBL ≥5 mm + inflammation (Sup/swelling/redness)	NR	No	PDT	19	19	53	12.19	25–71	42%	4	NR

Abbreviations: BoP, bleeding on probing; βTCP, β-tricalcium phosphate; CHX, chlorhexidine; CPC, cetylpyridinium chloride; MBL, marginal bone level; NR: not reported; PD, pocket depth; PDT, photodynamic therapy; Sup, suppuration.

16000518, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2018), one study was performed in a public clinic within a university setting (Carcuac et al., 2016), and one study was at a university and private clinics (Esposito et al., 2013). Three of the studies were funded or supported by the industry (Cha et al., 2019; de Waal et al., 2013; Emanuel et al., 2020) and two studies by non-industrial public funding (Albaker et al., 2018; Carcuac et al., 2016), whereas in one study the source of funding was not reported (Esposito et al., 2013). Overall, the follow-up periods relevant to the inclusion criteria of this meta-analysis varied between 6 and 12 months. In the study by Esposito et al. (Esposito et al., 2013), the first follow-up was at 4 months (de Angelis et al., 2012).

3.3 | RoB in individual studies

Figure 2 presents the RoB assessed using the Cochrane RoB 2.0 tool (Sterne, Savović, et al., 2019). One study was qualified as having a "low" RoB (de Angelis et al., 2012), two studies were qualified as having "some concerns" (Albaker et al., 2018; Carcuac et al., 2016) due to the randomization process and selection of the reported results, and three studies were qualified as having a "high" RoB due to missing outcome data (multiple implant loss in the control group; de Wall et al., 2013), the manner in which the outcome was measured (questionable blindness; Cha et al., 2019), and possible conflict of interest (Cha et al., 2019; Emanuel et al., 2020). The number of studies that contained a low RoB for specific criteria varied according to the assessed item: randomization process ($n = 4$), deviations from the intended interventions ($n = 6$), missing outcome data ($n = 5$), measurement of the outcome ($n = 5$), and selection of the reported result ($n = 5$). Detailed information about each specific domain and complete evaluation is presented in Table S3. The contribution made to the entire evidence by studies with different RoB assessments was further considered during the sensitivity analysis. No significant impact from studies rated as high or low RoB was found.

TABLE 4 Results of the meta-analyses for the primary and secondary outcomes showing the changes from baseline to 6 months at implant and at patient level. Subgroup analysis according to the interventions in the test group.

	Baseline to 6 months								
					95%CI				
Implant level		**Subgroup by test group**	***n***	**MD**	**Lower**	**Upper**	***p*-value**	**I^2 (%)**	***p*-value**
PPD changes (mm)	Overall		376	0.265	−0.29	0.821	.349	62.6	.014
		PDT	42	−0.179	−0.988	0.629	.664	0	.401
		CHX	256	0.319	−0.655	1.294	.521	83	.002
		L_Antib	78	0.592	0	1.421	.161	9.8	.292
BOP changes (%)	Overall		198	7.30	1.599	13	.012	0	.836
		PDT	42	6.622	0.047	13.197	.048	0	.916
		CHX	78	6.8	−7.607	21.207	.355	-	
		L_Antib	78	13.714	−5.121	32.548	.154	0	.331
MBL changes (mm)	Overall		180	0.481	−0.031	0.992	.065	47.2	.128
		PDT	24	−0.100	−1.221	1.021	.861	-	
		CHX	78	0.100	−0.894	1.094	.844	-	
		L_Antib	78	0.779	−0.096	1.655	.081	75.4	.044
	Baseline to 6 months								
					95%CI				
Patient level		**Subgroup by test group**	***n***	**MD**	**Lower**	**Upper**	***p*-value**	**I^2 (%)**	***p*-value**
PPD changes (mm)	Overall		115	0.146	−0.508	0.800	.663	17.3	.305
		PDT	42	−0.179	−0.988	0.629	.664	0	.410
		L_Antib	73	0.526	−0.558	1.609	.342	39.0	.200
BOP changes (%)	Overall		115	7.626	1.374	13.877	.017	0	.685
		PDT	42	6.622	0.047	13.197	.048	0	.916
		L_Antib	73	17.086	−3.100	37.272	.097	0	.46
MBL changes (mm)	Overall		97	0.698	−0.236	1.632	.143	78.5	.009
		PDT	24	−0.100	−1.221	1.021	.861	-	
		L_Antib	73	1.019	−0.297	2.336	.129	87.6	.004

Abbreviations: BOP, bleeding on probing; CI, confidence interval; CHXc Chlorhexidine; I^2, heterogeneity; L_Antib, local antibiotics; MBL, radiographic bone loss; MD, mean difference; mm, millimetre; PPD, probing pocket depth; PDT, photodynamic therapy.

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.4 | Study and participants characteristics

Information related to the general characteristics of the studies and participants included in this systematic review and meta-analysis is shown in Table 3. Overall, this review includes data from 250 patients in the age range 21–90, mainly non-smoking women, and most were not diabetic. In most studies, the inclusion criteria for peri-implantitis was based on the definition of the 2017 World Workshop (Berglundh et al., 2018). In all but one study (Esposito et al., 2013), the participants were enrolled in a hygiene regimen phase prior to surgery, while pre- or post-systemic antibiotic treatment was given in 50% of the studies. The interventions included PTD, CHX application, minocycline ointment, cetylpyridinium chloride (CPC) together with CHX, sustained-release doxycycline from a β-tricalcium phosphate (β-TCP) carrier, or combinations of the above treatments.

3.5 | Study outcomes

All included studies were designed as parallel-arm RCTs with a follow-up of 6 and 12 months following surgery. The primary outcomes of this systematic review and meta-analysis (PD, BoP/SoP) are presented as values at the 6- and 12-month follow-up. The results of the meta-analysis were divided according to the type of intervention administered in each study and is presented in the forest plots at the implant or patient level when possible.

3.5.1 | Probing depth (PD)

The results of the meta-analysis regarding PD at the implant and patient level at 6- and 12-month follow-ups are presented in Tables 4 and 5 and in forest plots in Figure 3. At 6 months, treatment with

TABLE 5 Results of the meta-analyses for the primary and secondary outcomes showing the changes from baseline to 12 months at implant and at patient level.

	Baseline to 12 months								
					95%CI				
Implant level		Subgroup by test group	*n*	MD	Lower	Upper	*p*-value	I^2 (%)	*p*-value
PPD changes (mm)	Overall		331	0.249	−0.352	0.850	.417	59.3	.031
		PDT	42	−0.073	−0.881	0.735	.860	0	.698
		CHX	257	0.089	−0.678	0.856	.819	65.5	.055
		LAbs	32	1.440	0.401	2.479	.007	-	
BOP changes (%)	Overall		152	5.686	−1.309	12.681	.111	10.4	.341
		PDT	42	7.406	0.808	14.004	.028	0	.39
		CHX	78	−2.600	−15.527	10.327	0.693	-	
		LAbs	32	21.100	−13.335	55.535	0.230	-	
MBL changes (mm)	Overall		331	0.155	−0.383	0.694	.572	62.3	.021
		PDT	42	0.052	−0.932	1.035	.918	0	.861
		CHX	257	−0.137	−0.558	0.284	0.523	31.3	0.233
		LAbs	32	1.210	0.437	1.983	0.002	-	
	Baseline to 12 months								
					95%CI				
Patient level		Subgroup by test group	*n*	MD	Lower	Upper	*p*-value	I^2 (%)	*p*-value
PPD changes (mm)	Overall		69	0.417	−0.649	1.483	.443	53.2	.118
		PDT	42	−0.073	−0.883	0.737	.860	0.0	.698
		LAbs	27	1.410	0.230	2.590	.019	-	
BOP changes (%)	Overall		69	7.976	1.476	14.475	.016	0	.426
		PDT	42	7.406	0.808	14.004	.028	0	.39
		LAbs	27	26.600	−11.122	64.322	.167	-	
MBL changes (mm)	Overall		69	0.708	−0.443	1.860	.228	61.5	.075
		PDT	42	0.052	−0.932	1.035	.918	0	.861
		LAbs	27	1.570	0.705	2.435	<.001	-	

Note: Subgroup analysis according to the interventions in the test group.

Abbreviations: BOP, bleeding on probing; CHX, chlorhexidine; CI, confidence interval; I^2, heterogeneity; LAbs, local antibiotics; MBL, radiographic bone loss; MD, mean difference; mm, millimetre; PDT, photodynamic therapy; PPD, probing pocket depth.

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

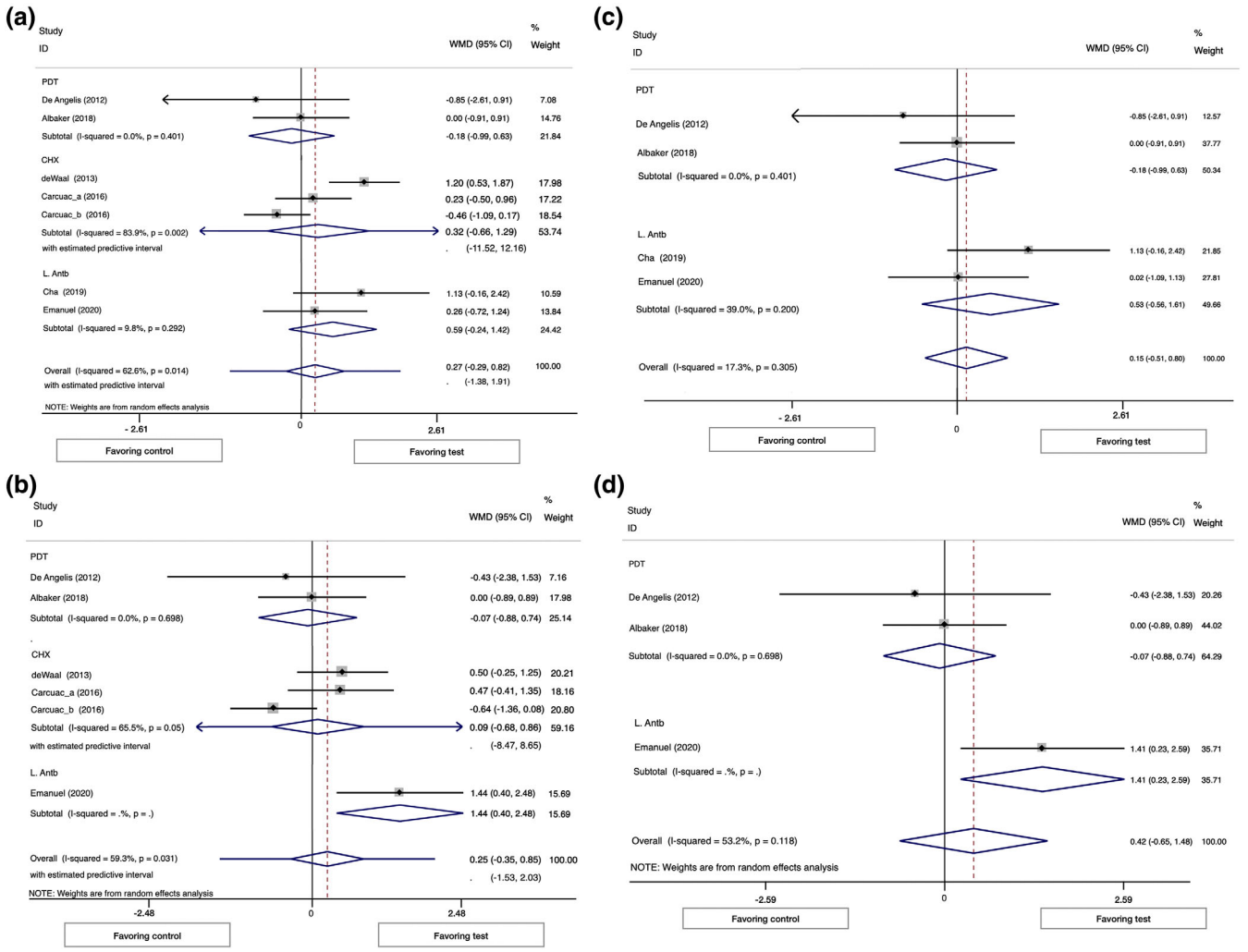


FIGURE 3 Forest plots showing comparisons of probing depth (PD) changes (in mm) between groups. Subgroup analysis according to the type of adjunctive chemical surface decontamination used: photodynamic therapy (PDT), chlorhexidine (CHX), and local antibiotic (L. Antb). (a) Implant-level 6-month follow-up results. (b) Implant-level 12-month follow-up results. (c) Patient-level 6-month follow-up results. (d) Patient-level 12-month follow-up results. Data presented are also available in Tables 4 and 5. 95%CI, 95% confidence interval; MD, mean difference. [Colour figure can be viewed at wileyonlinelibrary.com]

PDT resulted in no significant changes in PD reduction as compared to the control group at both the implant [$n = 2$; MD = −0.179; 95%CI (−0.988 to 0.629); $p = .664$; no heterogeneity: $I^2 = 0\%$] and patient levels [$n = 2$; MD = −0.179; 95%CI (−0.988 to 0.629); $p = .664$; no heterogeneity: $I^2 = 0\%$]. The use of CHX also showed no significant improvement in PD at the implant level, with or without the use of systemic antibiotics [$n = 3$; MD = 0.319; 95%CI (−0.655 to 1.294); $p = .521$; high heterogeneity: $I^2 = 83\%$]. The same pattern was observed for LAbs treatment: no significant improvement was observed in PD at the implant [$n = 2$; MD = 0.592; 95%CI (0–1.421); $p = .161$; very low heterogeneity: $I^2 = 9.8\%$] or patient level [$n = 2$; MD = 0.526; 95%CI (−0.558 to 1.609); $p = .342$; low heterogeneity: $I^2 = 39\%$]. The overall analysis for all treatments at the 6-month follow-up showed no additive effect of chemical treatment on PD reduction either at the implant level [$n = 7$; MD = 0.265; 95%CI (−0.29 to 0.821); $p = .349$; moderate heterogeneity: $I^2 = 62.6\%$] or the patient level [$n = 4$; MD = 0.146; 95%CI (−0.508 to 0.800); $p = .663$; low heterogeneity: $I^2 = 38.6\%$]. The 12-month analysis revealed significant improvement in PD following administration of LAbs only ($p = 0.007$, one study) (Table 5 and Figure 3). The overall analysis for all treatments at 12 months post surgery showed no significant change in PD at the implant level [$n = 6$; MD = 0.249; 95%CI (−0.352 to 0.85); $p = .417$; moderate heterogeneity: $I^2 = 59.3\%$] or the patient level [$n = 3$; MD = 0.417; 95%CI (−0.649 to 1.483); $p = .443$; moderate heterogeneity: $I^2 = 53.2\%$].

3.5.2 | Bleeding on probing (BoP)

Figure 4 and Tables 4 and 5 present the effect of each chemical treatment on BoP at 6 and 12 months post treatment at the implant and patient levels. Six months post treatment with PDT, BoP was

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

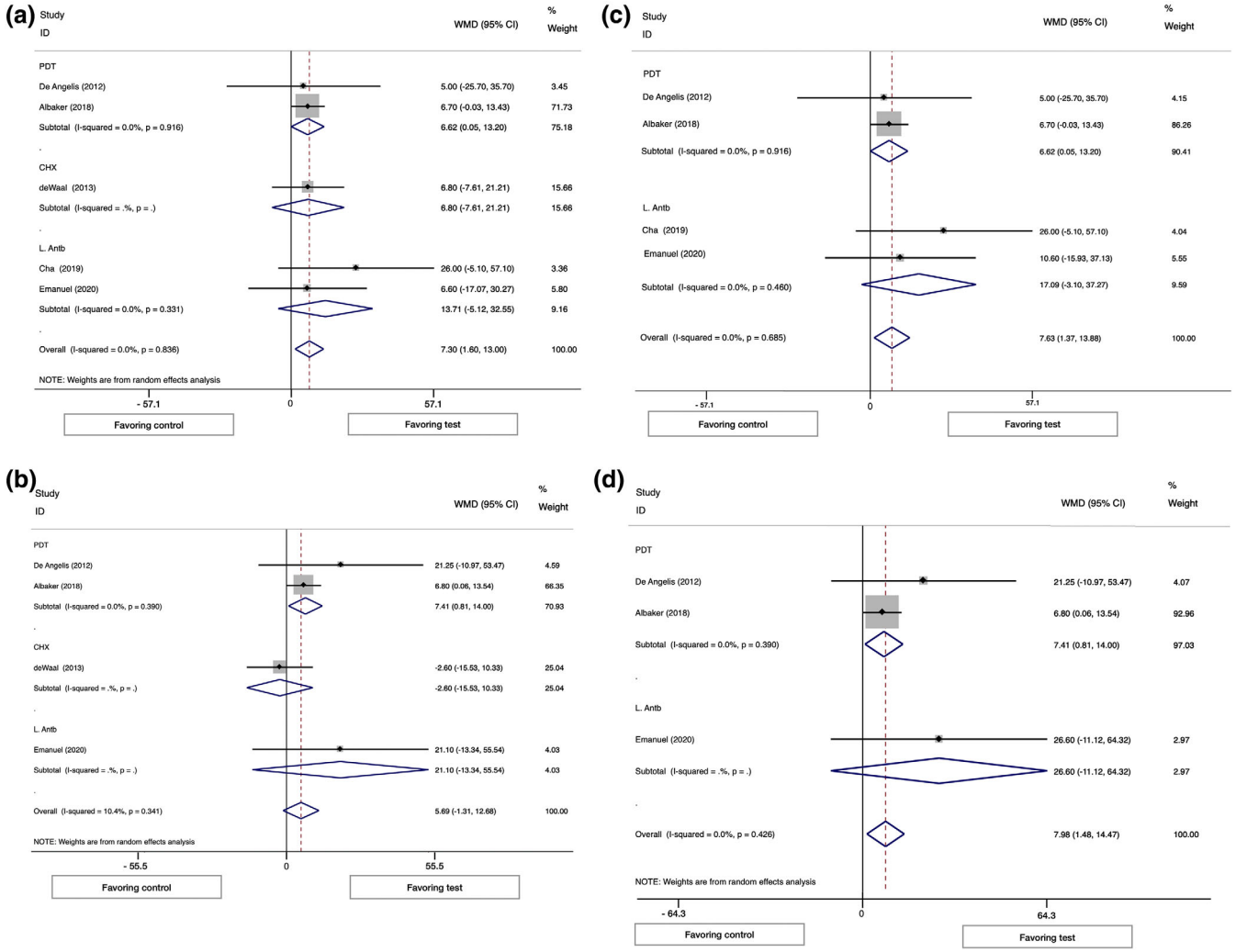


FIGURE 4 Forest plots showing comparisons of bleeding on probing (BoP) changes (in %) between groups. Subgroup analysis according to the type of adjunctive chemical surface decontamination used: photodynamic therapy (PDT), chlorhexidine (CHX), and local antibiotic (L. Antb). (a) Implant-level 6-month follow-up results. (b) Implant-level 12-month follow-up results. (c) Patient-level 6-month follow-up results. (d) Patient-level 12-month follow-up results. Data presented are also available in Tables 4 and 5. 95%CI, 95% confidence interval; MD, mean difference. [Colour figure can be viewed at wileyonlinelibrary.com]

significantly reduced at the implant level [$n = 2$; MD = 6.622; 95%CI (0.047–13.197); $p = .048$; no heterogeneity: $I^2 = 0\%$] and the patient level [$n = 2$; MD = 6.622; 95%CI (0.047–13.197); $p = .048$; no heterogeneity: $I^2 = 0\%$]. While treatments with CHX and LAbs did not result in significant changes in BoP, the overall analysis showed that chemical treatment with surgical debridement significantly reduced BoP 6 months post treatment at the implant level [$n = 5$; MD = 7.3; 95%CI (1.599–13); $p = .012$; no heterogeneity: $I^2 = 0\%$] and the patient level [$n = 4$; MD = 7.626; 95%CI (1.374–13.877); $p = .017$; no heterogeneity: $I^2 = 0\%$]. At 12 months post treatment, only PDT was found to be effective in reducing BoP at both the implant and patient levels [$n = 2$; MD = 7.406; 95%CI (0.808–14.004); $p = .028$; no heterogeneity: $I^2 = 0\%$]. The overall analysis at 12 months following surgery showed that chemical treatment significantly improved BoP levels at the patient level [$n = 3$; MD = 7.976; 95%CI (1.476–14.475); $p = .016$; no heterogeneity: $I^2 = 0\%$] but not at the implant level.

3.5.3 | Marginal bone level (MBL)

The results of the meta-analysis regarding MBL are shown in Tables 4 and 5 and Figure 5. At 6 months post surgery, none of the individual treatments or the overall analysis resulted in significant improvement in MBL at either the implant [$n = 4$; MD = 0.481; 95%CI (−0.031 to 0.992); $p = .065$; moderate heterogeneity: $I^2 = 47.2\%$] or patient levels [$n = 3$; MD = 0.698; 95% CI (−0.236 to 1.632); $p = .143$; high heterogeneity: $I^2 = 78.5\%$]. At 12 months post surgery, administration of LAbs was the only effective treatment that improved MBL at both the implant and patient levels (Table 5 and Figure 5b,d). Of note, the results of the effect of LAbs on MBL were based exclusively on one study by Emanuel et al. (Emanuel et al., 2020). The overall results demonstrated no effect for chemical treatment on MBL for both the implant [$n = 6$; MD = 0.155; 95%CI (−0.383 to 0.694); $p = .572$; moderate heterogeneity: $I^2 = 62.3\%$] and patient levels [$n = 3$;

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

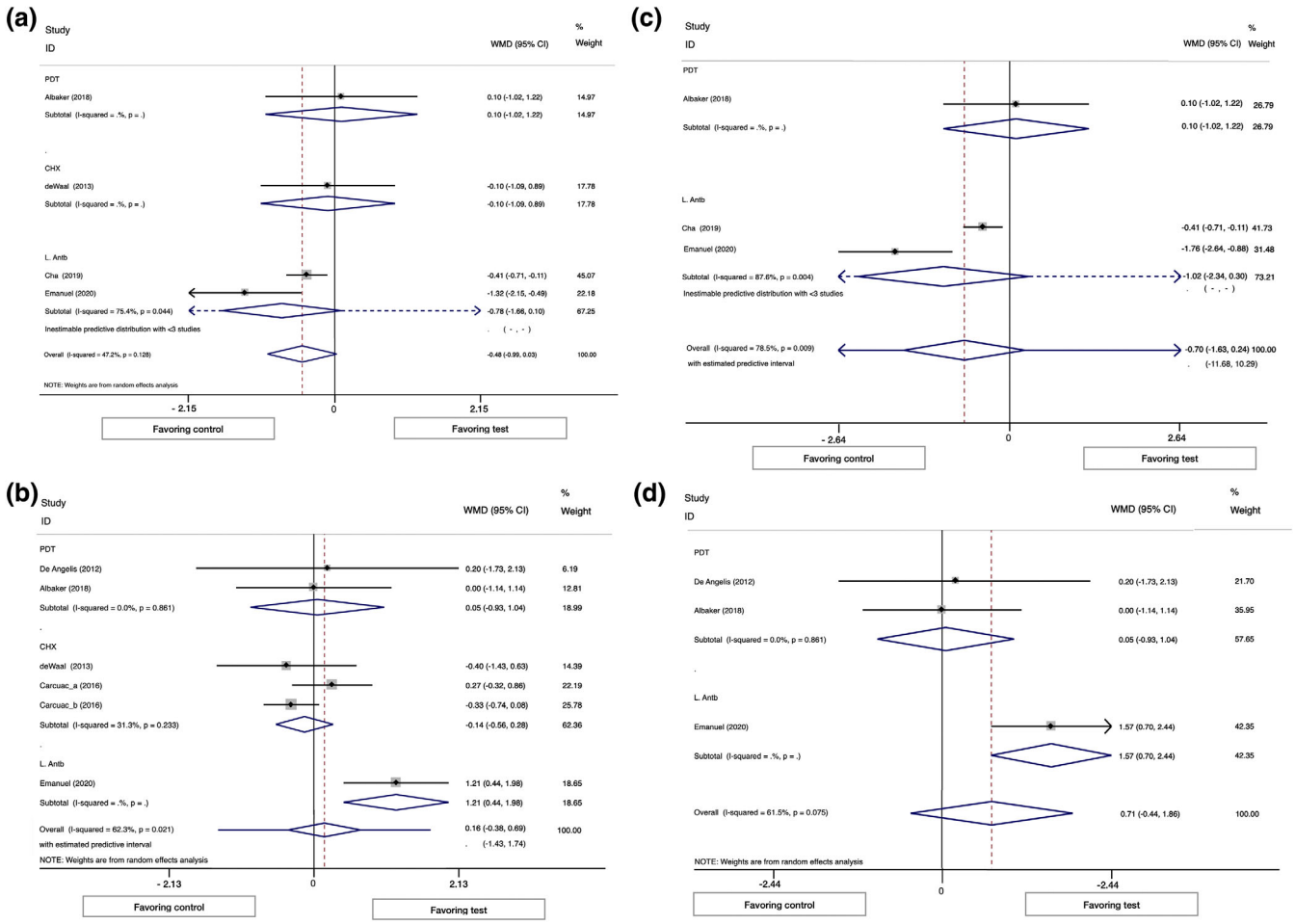


FIGURE 5 Forest plots showing comparisons of MBL (marginal bone loss) changes (in mm) between groups. Subgroup analysis according to the type of adjunctive chemical surface decontamination used: photodynamic therapy (PDT), chlorhexidine (CHX), and local antibiotic (L. Antb). (a) Implant-level 6-month follow-up results. (b) Implant-level 12-month follow-up results. (c) Patient-level 6-month follow-up results. (d) Patient-level 12-month follow-up results. Data presented are also available in Tables 4 and 5. 95%CI, 95% confidence interval; MD, mean difference. [Colour figure can be viewed at wileyonlinelibrary.com]

MD = 0.708; 95%CI (−0.443 to 1.860); $p = .228$; moderate heterogeneity: $I^2 = 61.5\%$].

3.5.4 | Implant loss

Implant loss rate (RD RI) for the different treatments at 6 and 12 months is presented in Table 6 and Figure 6. None of the individual treatments or the overall analysis resulted in reduced risk of implant loss at the implant [$n = 6$: RD = 0.–0.029; 95%CI (−0.087 to 0.029); $p = .332$; moderate heterogeneity: $I^2 = 48.2\%$] or patient level [$n = 6$: RD = 0.0.011; 95%CI (−0.035 to 0.057); $p = .641$; no heterogeneity: $I^2 = 0\%$].

3.5.5 | Disease resolution

Table 6 and Figure 7 show the results of the meta-analysis for disease resolution at 12 months following the various treatments. None of the individual treatments or the overall results was found to affect disease resolution at the implant level [$n = 4$: RR = 0.948; 95%CI (0.718–1.265); $p = .715$; moderate heterogeneity: $I^2 = 48.8\%$] or patient level [$n = 3$: RR = 0.765; 95%CI (0.466–1.257); $p = .291$; no heterogeneity: $I^2 = 0\%$]. It should be noted that the analysis at the patient level included a single study by Carcuac et al. (2016).

3.6 | RoB across studies (publication bias) and sensitivity analyses

Owing to the small number of studies (<10), publication bias could not be calculated using a statistical test. Post hoc sensitivity analysis was performed by excluding studies, one at a time, from the global estimation. Sensitivity analysis for each treatment subgroup could not be calculated because of the small number of studies in each subgroup. Considering the overall results, the global estimators for each of the reported outcomes did not change significantly after omitting each of

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 Results of the meta-analyses for failure rate (risk difference) and disease resolution (risk ratio), at implant and at patient level. Subgroup analysis according to the interventions in the test group.

Failure rate (at 6 and 12 months)		Subgroup by test group	n	RD (risk diff.)	95%CI		p-value	I^2 (%)	p-value
					Lower	Upper			
Implant level	Overall		376	−0.029	−0.087	0.029	.332	48.2	.072
		PDT	42	0	−0.118	0.118	1	0	1
		CHX	256	−0.044	−0.160	0.072	.55	79.7	.007
		L_Antib	78	−0.040	−0.166	0.086	.531	39.6	.198
Patient level	Overall		187	0.011	−0.035	0.057	.641	0	.99
		PDT	42	0	−0.118	0.118	1.00	0	1.00
		CHX	99	0.038	−0.047	0.124	.382	0	.964
		L_Antib	46	0	−0.085	0.085	1000	-	

Disease resolution (at 12 months)		Subgroup by test group	n	RR (relat risk)	95% CI		p-value	I^2 (%)	p-value
					Lower	Upper			
Implant level	Overall		274	0.948	0.718	1265	.715	48.8	.118
		PDT	18	1094	0.824	1452	.536	-	
		CHX	256	0.884	0.590	1325	.552	54,3	.112
Patient level		L_Antib	99	0.765	0.466	1257	.291	0	.401

Abbreviations: CHX, chlorhexidine; CI, confidence interval; I^2, heterogeneity; L_Antib, local antibiotics; PDT, photodynamic therapy; RD, risk difference; RR, relative risk.

the contributing studies, oscillating within the limits of the original 95%CI.

The certainty of the evidence was assessed using the GRADE approach. A summary of the key findings (SoF) from this systematic review is presented in Tables 7–9 (6 months) and Tables 10–12 (12 months) using the GRADE method for each intervention: PDT (Tables 7 and 10), CHX (Tables 8 and 11), and LAbs (Tables 9 and 12). The following outcome measures were included and assessed in the SoF table: PD, BoP, MBL, disease resolution, and implant loss at the 6- and 12-month follow-up. According to the GRADE guidelines, the certainty of the evidence was extremely low for every measure in all comparisons.

4 | DISCUSSION

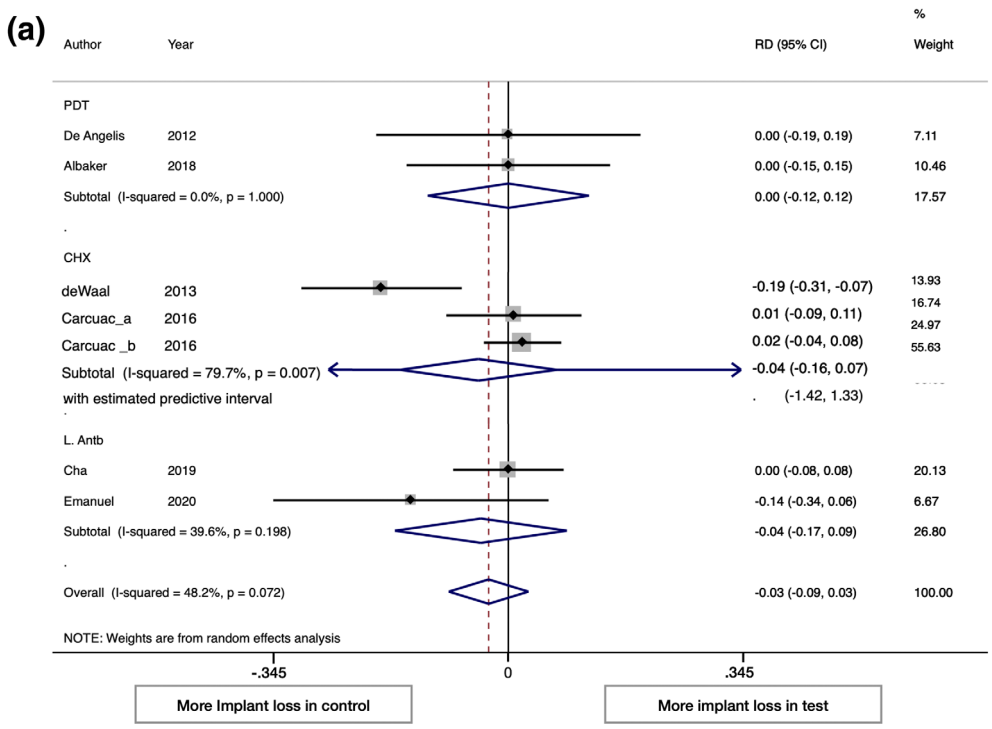
The present systematic review and meta-analysis evaluated the efficacy of adjunctive chemical surface decontamination during surgical therapy of peri-implantitis. Although the present results show that the use of LAbs during surgery may have a positive effect on PD reduction at 12 months post treatment, this result is based on a single study with a very low certainty of evidence. With regard to the second primary outcome, BoP, the overall results at 6 months show that chemical treatment significantly improved this parameter at the implant and patient levels. At 12 months post surgery, the overall results showed reduced BoP at the patient level only. Of note, the reduction observed in BoP following chemical treatment at both time points was derived mainly from the PDT subgroup, which included only two studies. These studies carried heavy weight in the final analysis because of their narrow SD. In this context, the composite parameter of "disease resolution", which includes PD + BoP, may be more accurate for forecasting implant prognosis (Koldsland et al., 2018; Sanz, Chapple, & Working Group 4, 2012).

Since the integration of implants is a prerequisite to their functioning, improvement in MBL following treatment could be considered an important goal when treating peri-implantitis. The overall results of this meta-analysis suggest that chemical treatment of the implant surface has no effect on the marginal bone height. It should be noted, however, that the one study in which LAbs were used showed improved MBL (Emanuel et al., 2020), but has a very low certainty of evidence (β-tricalcium phosphate was used as a carrier in the test group only, the study was industry-funded, and the corresponding author holds a senior position in the company).

Two additional secondary outcome variables, namely implant loss and disease resolution, were also analysed. These are two end-point parameters important for verifying treatment success by practitioners. None of the chemical decontamination procedures was superior to saline with regard to these outcome variables at the 12-month follow-up. Thus, we cannot conclude that any of the chemical decontamination protocols included in this meta-analysis is an essential step during surgical treatment of peri-implantitis for determining a better outcome.

To formulate clinical guidelines, it is necessary to evaluate the effect of each treatment on the primary and secondary outcomes, and only then can we decide on the recommendations based on each specific treatment. The present meta-analysis includes too few controlled studies for evaluating the efficacy of each chemical intervention, and all results obtained are with low certainty of evidence (Tables 7–12).

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



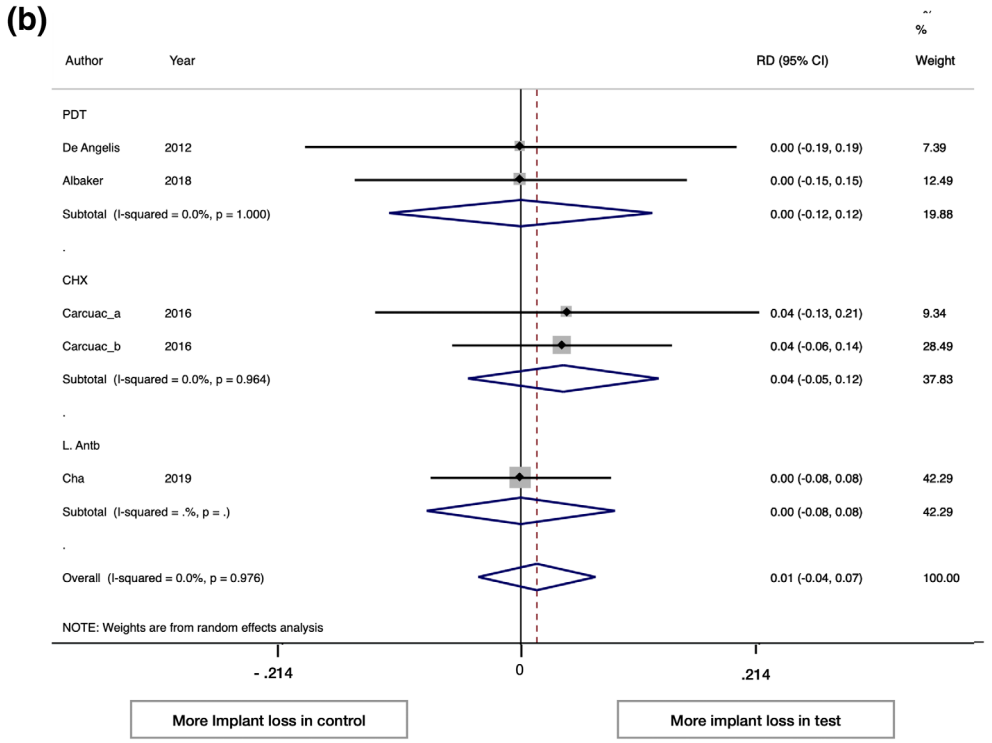


FIGURE 6 Forest plots showing comparisons of implant loss between groups. Subgroup analysis according to the type of adjunctive chemical surface decontamination used: photodynamic therapy (PDT), chlorhexidine (CHX), and local antibiotic (L. Antb). (a) Implant-level 6 and 12-month follow-up results. (b) Patient-level 6- and 12-month follow-up results. Data presented are also available in Tables 4 and 5. 95%CI, 95% confidence interval; MD, mean difference. [Colour figure can be viewed at wileyonlinelibrary.com]

Therefore, it is difficult to reach any definitive conclusions. These findings are in agreement with previous in vitro pre-clinical and other systematic reviews. Almohandes et al. examined the efficacy of various treatment modalities using a canine model of experimental peri-implantitis. They concluded that decontamination procedures, including citric acid gel or a rotating titanium brush, did not improve the outcome following surgical treatment of experimental peri-implantitis (Almohandes et al., 2022). Dostie et al. tested the efficacy of the commonly used antimicrobial agents (CHX, tetracycline, phosphoric acid, etc.) in the decontamination of multiple species of mature oral biofilm on sand-blasted, large-grit, acid-etched (SLA) titanium implants and found that rinsing the surfaces with 0.9% NaCl removed most of the biofilm. However, bacteria persisted in all specimens and none of the disinfectants tested was superior to the saline group (Dostie et al., 2017). These findings may explain the results of the aforementioned studies. In addition, other systematic reviews support our findings, showing no significant

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

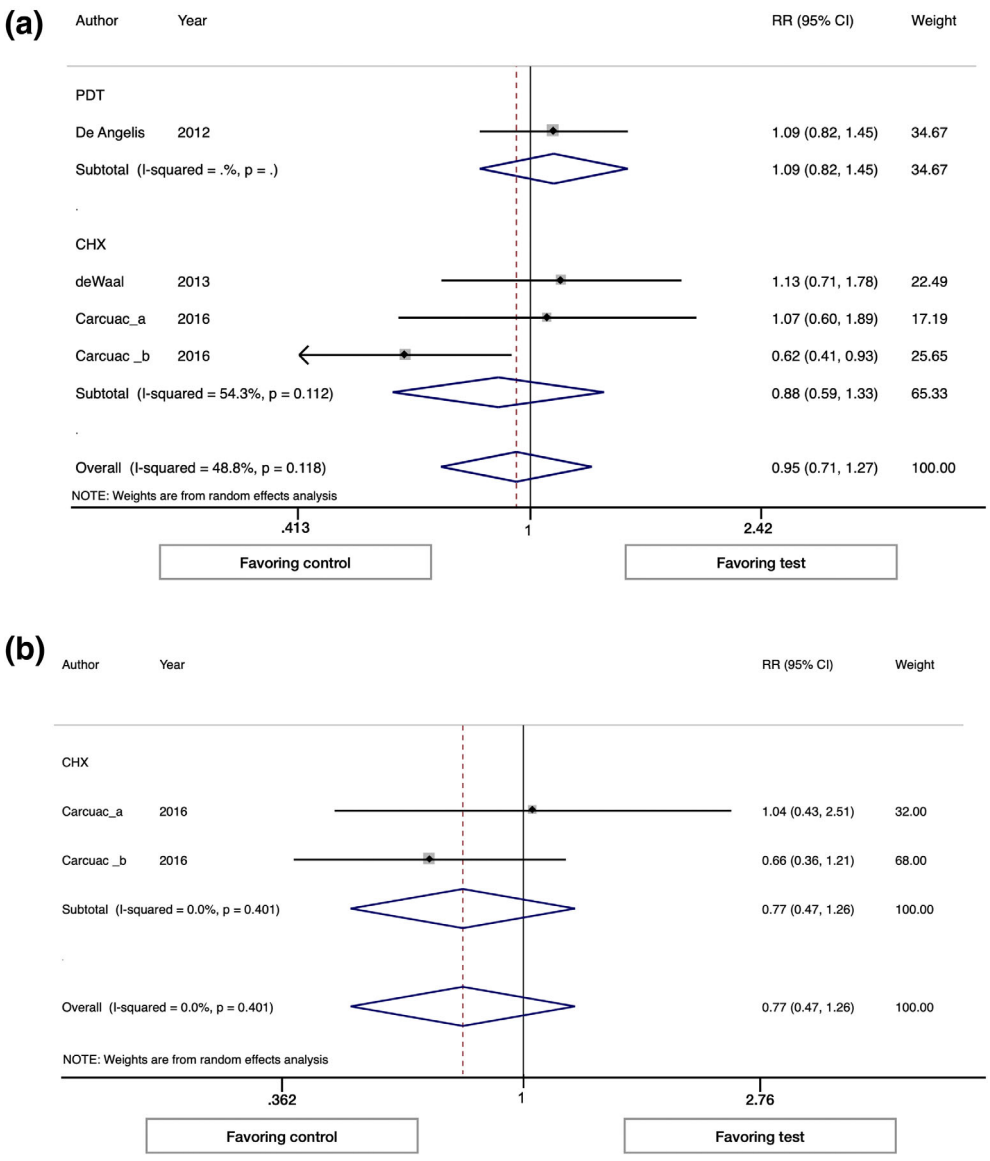


FIGURE 7 Forest plots showing comparisons of disease resolution between groups. Subgroup analysis according to the type of adjunctive chemical surface decontamination used: photodynamic therapy (PDT), chlorhexidine (CHX). (a) Implant-level 6- and 12-month follow-up results. (b) Patient-level 6- and 12-month follow-up results. Data presented are also available in Tables 4 and 5. 95%CI, 95% confidence interval; MD, mean difference. [Colour figure can be viewed at wileyonlinelibrary.com]

additive effect for the use of chemical decontamination of the implant surface during surgery (Faggion Jr. et al., 2013; Koo et al., 2019; Li et al., 2022; Ramanauskaite et al., 2021). Importantly, the present study and previous systematic reviews are based on a limited number of studies with various chemical agents. Therefore, we conclude that, upon analysing the currently available literature, chemical decontamination of the implant surface, in addition to surgical debridement, does not provide any additional benefit over surgical debridement.

4.1 | Limitations

The present systematic review has several limitations that should be considered: (1) Scarcity of existing evidence: In the present systematic review and meta-analysis, we included RCTs and controlled clinical trials (CCTs) with a follow-up of at least 6 months, which yielded only six studies; one study only had a follow-up of 12 months with the initial follow-up period at 4 months. (2) Owing to the small number of studies, some of the treatments were evaluated from only one or two RCTs. (3) The RCTs included only 10–27 subjects per treatment arm, and half of those studies contained less than 20 subjects per arm. This may have led to a large CI that affected the results. (4) Variability in the definition of peri-implantitis: Initially, we intended to include only studies that defined peri-implantitis according to the World Workshop on the Classification of Periodontal and Peri-implant Diseases and Conditions (2017), yet because of the limited amount of information using this definition, we included studies with similar definitions. (5) Heterogeneity in inclusion criteria between RCTs: Some of the studies included pre-diabetic patients, smokers, and patients with a history of periodontitis, while others included only healthy patients. (6) Surgical techniques and measurement of the outcomes varied between studies. (7) In half of the studies, systemic antibiotic treatment was administered before and after surgical treatment. This could affect the intra-subgroup comparisons. (8) Although LAbs showed

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 7 Summary of findings using the grading of recommendations assessment, development and evaluation (grade) approach at 6 months: Photodynamic therapy

Patient or population: Individuals aged above 18 years, otherwise healthy, with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018)
Setting: University/private clinics
Intervention: Surgical treatment + photodynamic therapy (PDT)
Control: Surgical treatment + placebo
Outcomes: Change in peri-implant outcomes

		Illustrative comparative risks[a] (95%CI)					
Outcomes[a]	Unit of interest	Assumed risk[a] (surgical treatment + placebo − comparator)	Corresponding risk[a] (surgical treatment + PDT − intervention)	Relative effect (95%CI)	No. (studies)	Certainty of the evidence (GRADE)	Comments
Pocket depth (PD) change (mm) assessed with periodontal probe Follow-up: 6 months	Implant level or patient level	Mean PD was 1.65 mm reduction	MD 0.18 mm lower reduction (0.99 lower to 0.63 higher)	-	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d] Due to risk of bias, imprecision, and indirectness	PD was not affected by PDT in addition to the surgical treatment of peri-implantitis
Bleeding on probing (BoP) change (%) assessed with periodontal probe Follow-up: 6 months	Implant level or patient level	Mean BoP was 31.2% reduction	MD 6.62% higher reduction (0.05% higher to 13.20% higher)	-	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d] Due to risk of bias, imprecision, and indirectness	BoP showed slightly higher reduction when PDT was used in addition to surgical treatment of peri-implantitis, but this reduction was not clinically relevant
Marginal bone loss (MBL) change (mm) assessed with periapical X-rays Follow-up: 6 months	Implant level or patient level	Mean MBL was 0.5 mm gain	MD 0.10 mm lower gain (1.22 mm loss to 1.02 gain)	-	24 (1 RCT)	⨁◯◯◯ VERY LOW[b,c,d] Due to risk of bias, imprecision, and indirectness	MBL was not affected by PDT in addition to the surgical treatment of peri-implantitis
Disease resolution (DR) assessed with clinical examination Follow-up: 6 months	Implant level or patient level	NR	NR	NR	NR	NR	NR
Implant loss assessed with clinical examination Follow-up: 6 and12 months	Implant level or patient level	0 implant loss per 1000 implants	0 per 1000	RD 0.00 (−0.12 to 0.12)	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d,e] Due to risk of bias, imprecision, and indirectness	Implant loss was not affected by PDT in addition to the surgical treatment of peri-implantitis

Note: GRADE Working Group grades of evidence. High certainty—We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty—We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Very low certainty—We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Abbreviations: CI, confidence interval; MD, mean difference; RD, risk difference; RR, risk ratio.

[a]"Assumed risk" is obtained from the median control group risk across studies. "Corresponding risk" (and its 95% confidence interval) is based on the "Assumed risk" in the comparison group and the relative effect of the intervention (and its 95%CI).

[b]Downgraded one level for study limitations ("unclear risk of bias" due to some concerns about reporting bias – Albaker et al., 2018).

[c]Downgraded two levels for very serious imprecision due to the very small sample size.

[d]Downgraded one level for severe indirectness due to the partial consideration of the intervention, restricted to the surgical approach – De Angelis et al. 2012.

[e]Downgraded one level for study limitations ("unclear risk of bias" due to some concerns about reporting and selection bias – Bombeccari et al., 2013).

16000051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 8 Summary of findings using the grading of recommendations assessment, development and evaluation (grade) approach at 6 months: Clorhexidine

Patient or population: Individuals aged above 18 years, otherwise healthy, with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018)
Setting: University/private clinics
Intervention: Surgical treatment + chlorhexidine solution (CHX)
Control: Surgical treatment + placebo
Outcomes: Change in peri-implant outcomes

		Illustrative comparative risks[a] (95%CI)					
Outcomes[a]	Unit of interest	Assumed risk[a](Surgical treatment + placebo − comparator)	Corresponding risk[a] (Surgical treatment + chlorhexidine − intervention)	Relative effect (95%CI)	No. (studies)	Certainty of the evidence (GRADE)	Comments
Pocket depth (PD) change (mm) assessed with periodontal probe Follow-up: 6 months	Implant level	Mean PD was 2.28 mm reduction	MD 0.32 mm higher reduction (0.66 lower to 1.29 higher)	-	256 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	PD was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis
	Patient level	NR	NR	NR	NR	NR	NR
Bleeding on Probing (BoP) change (%) assessed with periodontal probe Follow-up: 6 months	Implant level	Mean BoP was 19.6% reduction	MD 6.80% higher reduction (7.61% lower to 21.21% higher)	-	78 (1 RCT)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	BoP was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis
	Patient level	NR	NR	NR	NR		
Marginal bone loss (MBL) change (mm) assessed with periapical X-rays Follow-up: 6 months	Implant level	Mean MBL was 0.7 mm loss	MD 0.10 mm lower loss (0.89 mm higher loss to 1.09 gain)	-	78 (1 RCT)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	MBL was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	NR	NR	NR	NR	NR	NR
Disease resolution (DR) assessed with clinical examination Follow-up: 6 months	Implant level or patient level	NR	NR	NR	NR	NR	NR
Implant loss assessed with clinical examination Follow-up: 6 &12 months	Implant level	85 implant loss per 1000 implants	4 less per 1000 (14 less to 6 more)	RD −0.04 (−0.16 to 0.07)	256 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	Implant loss was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	40 implant loss per 1000 participants	2 more per 1000 (2 less to 5 more)	RD 0.04 (−0.05 to 0.12)	99 (1 RCTs)		

Note: GRADE Working Group grades of evidence. High certainty—We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty—We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Very low certainty—We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Abbreviations: CI, onfidence interval; MD, mean difference; NR, not reported; RR, risk ratio; RD, risk difference.

[a]"Assumed risk" is obtained from the median control group risk across studies. "Corresponding risk" (and its 95% confidence interval) is based on the "Assumed risk" in the comparison group and the relative effect of the intervention (and its 95%CI).

[b]Downgraded one level for study limitations ("high risk of bias" due to attrition bias – De Waal et al. 2013; "unclear risk of bias" due to some concerns about selection bias – Carcuac et al., 2016).

[c]Downgraded two levels for very serious imprecision due to the very small sample size.

16000051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 9 Summary of findings using the grading of recommendations assessment, development and evaluation (grade) approach at 6 months: Local antibiotics

Patient or population: Individuals aged above 18 years, otherwise healthy, with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018)
Setting: University/private clinics
Intervention: Surgical treatment + local antibiotics
Control: Surgical treatment + placebo
Outcomes: Change in peri-implant outcomes

		Illustrative comparative risks[a] (95%CI)					
Outcomes[a]	Unit of interest	Assumed risk[a] (surgical treatment + placebo − comparator)	Corresponding risk[a] (surgical treatment + local antibiotics − intervention)	Relative effect (95%CI)	No. (studies)	Certainty of the evidence (GRADE)	Comments
Pocket depth (PD) change (mm) assessed with periodontal probe Follow-up: 6 months	Implant level	Mean PD was 1.89 mm reduction	MD 0.59 mm higher reduction (0.24 lower to 1.42 higher)	-	78 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	PD was not affected by the local application of antibiotics adjunct to the surgical treatment of peri-implantitis
	Patient level	Mean PD was 1.95 mm reduction	MD 0.53 mm higher reduction (0.56 lower to 1.61 higher)	-	73 (2 RCTs)		
Bleeding on Probing (BoP) change (%) assessed with periodontal probe Follow-up: 6 months	Implant level	Mean BoP was 22.5% reduction	MD 13.71% higher reduction (5.12% lower to 32.55% higher)	-	78 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	BoP was not affected by the local application of antibiotics adjunct to the surgical treatment of peri-implantitis
	Patient level	Mean BoP was 22.95% reduction	MD 17.09% higher reduction (3.10% lower to 37.27% higher)	-	73 (2 RCTs)		
Marginal bone loss (MBL) change (mm) assessed with periapical X-rays Follow-up: 6 months	Implant level	Mean MBL was 0.035 mm gain	MD 0.78 mm higher gain (0.10 mm loss to 1.66 higher gain)	-	78 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	MBL was not affected by the local application of antibiotics adjunct to the surgical treatment of peri-implantitis
	Patient level	Mean MBL was 0.005 mm loss	MD 1.02 mm higher gain (0.30 mm loss to 2.34 higher gain)	-	73 (2 RCTs)		
Disease resolution (DR) assessed with clinical examination Follow-up: 6 months	Implant level or patient level	NR	NR	NR	NR	NR	NR
Implant loss assessed with clinical examination Follow-up: 6 &12 months	Implant level	56 implant loss per 1000 implants	3 less per 1000 (10 less to 5 more)	RD - 0.04 (−0.17 to 0.09)	78 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	Implant loss was not affected by the local application of antibiotics adjunct to the surgical treatment of peri-implantitis
	Patient level	0 implant loss per 1000 participants	0 more per 1000	RD 0.00 (−0.08 to 0.08)	46 (1 RCTs)		

Note: GRADE Working Group grades of evidence. High certainty—We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty—We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Very low certainty—We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Abbreviations: CI, confidence interval; MD, mean difference; NR, not reported; RD, risk difference.

[a]"Assumed risk" is obtained from the median control group risk across studies."Corresponding risk" (and its 95% confidence interval) is based on the "Assumed risk" in the comparison group and the relative effect of the intervention (and its 95%CI).

[b]Downgraded one level for study limitations ("high risk of bias" due to detection bias – Cha et al., 2019; "unclear risk of bias" due to some concerns about selection bias – Emanuel et al., 2020).

[c]Downgraded two levels for very serious imprecision due to the very small sample size.

16000051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 10 Summary of findings using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach at 12 months: Photodynamic therapy

Patient or population: Individuals aged above 18 years otherwise healthy with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018)
Setting: University/Private clinics
Intervention: Surgical treatment + Photodynamic therapy (PDT)
Control: Surgical treatment + Placebo
Outcomes: Change in peri-implant outcomes

		Illustrative comparative risks (95% CI)					
Outcomes[a]	Unit of interest	Assumed risk[a] (Surgical treatment + Placebo − Comparator]	Corresponding risk[a] (Surgical treatment +PDT− Intervention)	Relative effect (95% CI)	No. (studies)	Certainty of the evidence (GRADE)	Comments
Pocket depth (PD) change (mm) assessed with periodontal probe Follow-up: 12 months	Implant level or Patient level	Mean PD was 1.65 mm reduction	MD 0.07 mm lower reduction (0.88 lower to 0.74 higher)	-	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d] Due to risk of bias, imprecision, and indirectness	PD was not affected by PDT in addition to the surgical treatment of peri-implantitis.
Bleeding on Probing (BoP) change (%) assessed with periodontal probe Follow-up: 12 months	Implant level or Patient level	Mean BoP was 23.35 % reduction	MD 7.41% higher reduction (0.81% higher to 14% higher)	-	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d] Due to risk of bias, imprecision, and indirectness	BoP is reduced when PDT is used in addition to surgical treatment of peri-implantitis, but this reduction was not clinically relevant.
Marginal bone loss (MBL) change (mm) assessed with periapical X-rays Follow-up: 12 months	Implant level or Patient level	Mean MBL was 0.75 mm gain	MD 0.05 mm higher gain (0.93mm loss to 1.04 gain)	-	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d] Due to risk of bias, imprecision, and indirectness	MBL was not affected by PDT in addition to the surgical treatment of peri-implantitis.
Disease resolution (DR) assessed with clinical examination Follow-up: 12 months	Implant level or Patient level	900 resolved per 1000 implants or participants	9 more per 1000 (162 less to 405 more)	RR 1.09 (0.82 to 1.45)	18 (1 RCT)	⨁◯◯◯ VERY LOW[b,c,d] Due to imprecision and indirectness	DR was not affected by PDT in addition to the surgical treatment of peri-implantitis.
Implant loss assessed with clinical examination Follow-up: 6 &12 months	Implant level or Patient level	0 implant loss per 1000 implants	0 per 1000	RD 0.00 (−0.12 to 0.12)	42 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c,d,e] Due to risk of bias, imprecision, and indirectness	Implant loss was not affected by PDT in addition to the surgical treatment of peri-implantitis.

Note: GRADE Working Group grades of evidence. High certainty—We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty—We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Very low certainty—We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Abbreviations: CI, Confidence interval; MD, Mean difference; RD, Risk difference; RR, Risk ratio.

[a]The 'Assumed risk' is obtained from the median control group risk across studies. The 'Corresponding risk' (and its 95% confidence interval) is based on the 'Assumed risk' in the comparison group and the relative effect of the intervention (and its 95% CI).

[b]Downgraded one level for study limitations ('unclear risk of bias' due to some concerns about reporting bias – Albaker et al., 2018).

[c]Downgraded two levels for very serious imprecision due to the very small sample size.

[d]Downgraded one level for severe indirectness due to the partial consideration of the intervention, restricted to the surgical approach – De Angelis et al. 2012.

[e]Downgraded one level for study limitations ('unclear risk of bias' due to some concerns about reporting and selection bias – Bombeccari et al., 2013).

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 11 Summary of findings using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach at 12 months: Clorhexidine.

Patient or population: Individuals aged above 18 years otherwise healthy with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018)
Setting: University/Private clinics
Intervention: Surgical treatment + Chlorhexidine solution (CHX)
Control: Surgical treatment + Placebo
Outcomes: Change in peri-implant outcomes

		Illustrative comparative risks (95% CI)					
Outcomes[a]	Unit of interest	Assumed risk[a] (Surgical treatment + Placebo − Comparator)	Corresponding risk[a] (Surgical treatment + Chlorhexidine − Intervention)	Relative effect (95% CI)	No. (studies)	Certainty of the evidence (GRADE)	Comments
Pocket depth (PD) change (mm) assessed with periodontal probe Follow-up: 12 months	Implant level	Mean PD was 2.31 mm reduction	MD 0.09 mm higher reduction (0.68 lower to 0.86 higher)	-	257 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision.	PD was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	NR	NR	NR	NR	NR	NR
Bleeding on Probing (BoP) change (%) assessed with periodontal probe Follow-up: 12 months	Implant level	Mean BoP was 22.5 % reduction	MD 2.60 % lower reduction (15.53% lower to 10.33% higher)	-	78 (1 RCT)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	BoP was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	NR	NR	NR	NR	NR	NR
Marginal bone loss (MBL) change (mm) assessed with periapical X-rays Follow-up: 12 months	Implant level	Mean MBL was 0.75 mm loss	MD 0.14 mm higher loss (0.56 mm higher loss to 0.28 lower loss)	-	257 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	MBL was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	NR	NR	NR	NR	NR	NR
Disease resolution (DR) assessed with clinical examination Follow-up: 12 months	Implant level	500 resolved per 1000 implants	60 less per 1000 (205 less to 165 more)	RR 0.88 (0.59 to 1.33)	256 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	DR was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	429 resolved per 1000 participants	99 less per 1000 (227 less to 111 more)	RR 0.77 (0.47 to 1.26)	99 (1 RCT)		
Implant loss assessed with clinical examination Follow-up: 6 &12 months	Implant level	85 implant loss per 1000 implants	4 less per 1000 (14 less to 6 more)	RD −0.04 (−0.16 to 0.07)	256 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	Implant loss was not affected by local application of chlorhexidine adjunct to the surgical treatment of peri-implantitis.
	Patient level	40 implant loss per 1000 participants	2 more per 1000 (2 less to 5 more)	RD 0.04 (−0.05 to 0.12)	99 (1 RCTs)		

Note: GRADE Working Group grades of evidence. High certainty—We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty—We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Very low certainty—We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Abbreviations: CI, Confidence interval; MD, Mean difference; NR, Not reported; RD, Risk difference; RR, Risk ratio.

[a]The 'Assumed risk' is obtained from the median control group risk across studies. The 'Corresponding risk' (and its 95% confidence interval) is based on the 'Assumed risk' in the comparison group and the relative effect of the intervention (and its 95% CI).

[b]Downgraded one level for study limitations ('high risk of bias' due to attrition bias – De Waal et al. 2013; 'unclear risk of bias' due to some concerns about selection bias – Carcuac et al., 2016)

[c]Downgraded two levels for very serious imprecision due to the very small sample size.

16000051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 12 Summary of findings using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach at 12 months: Local antibiotics.

Patient or population: Individuals aged above 18 years otherwise healthy with peri-implantitis as defined by the 2017 World Workshop (Berglundh et al., 2018)
Setting: University/Private clinics
Intervention: Surgical treatment + Local antibiotics
Control: Surgical treatment + Placebo
Outcomes: Change in peri-implant outcomes

		Illustrative comparative risks (95% CI)					
Outcomes[a]	**Unit of interest**	**Assumed risk**[a] **(Surgical treatment + Placebo − Comparator)**	**Corresponding risk**[a] **(Surgical treatment + Local antibiotics − Intervention)**	**Relative effect (95% CI)**	**No. (studies)**	**Certainty of the evidence (GRADE)**	**Comments**
Pocket depth (PD) change (mm) assessed with periodontal probe Follow-up: 12 months	Implant level	Mean PD was 0.96 mm reduction	MD 1.44 mm higher reduction (0.40 higher to 2.48 higher)	-	32 (1 RCT)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision.	PD is reduced when local antibiotics are employed as an adjunct to surgical treatment of peri-implantitis. This reduction was clinically relevant.
	Patient level	Mean PD was 1.04 mm reduction	MD 1.41 mm higher reduction (0.23 higher to 2.59 higher)	-	27(1 RCT)		
Bleeding on Probing (BoP) change (%) assessed with periodontal probe Follow-up: 12 months	Implant level	Mean BoP was 15.2 % reduction	MD 21.10% higher reduction (13.34 % lower to 55.54 % higher)	-	32 (1 RCT)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	BoP was not affected by the local application of antibiotics adjunct to the surgical treatment of peri-implantitis.
	Patient level	Mean BoP was 18.6 % reduction	MD 26.60 % higher reduction (11.12 % lower to 64.32 % higher)	-	27 (1 RCT)		
Marginal bone loss (MBL) change (mm) assessed with periapical X-rays Follow-up: 12 months	Implant level	Mean MBL was 0.33 mm loss	MD 1.21 mm higher gain (0.44 mm higher to 1.98 higher)	-	32 (1 RCT)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	MBL is slightly restored when topical antibiotics are used in the surgical treatment of peri-implantitis. This gain was clinically relevant.
	Patient level	Mean MBL was 0.37 mm loss	MD 1.57 mm higher gain (0.70 mm gain to 2.44 gain)	-	27 (1 RCT)		
Disease resolution (DR) assessed with clinical examination Follow-up: 12 months	Implant level or Patient level	NR	NR	NR	NR	NR	NR
Implant loss assessed with clinical examination Follow-up: 6 &12 months	Implant level	56 implant loss per 1000 implants	3 less per 1000 (10 less to 5 more)	RD - 0.04 (−0.17 to 0.09)	78 (2 RCTs)	⨁◯◯◯ VERY LOW[b,c] Due to risk of bias and imprecision	Implant loss was not affected by the local application of antibiotics adjunct to the surgical treatment of peri-implantitis.
	Patient level	0 implant loss per 1000 participants	0 more per 1000	RD 0.00 (−0.08 to 0.08)	46(1 RCTs)		

Note: GRADE Working Group grades of evidence. High certainty—We are very confident that the true effect lies close to that of the estimate of the effect. Moderate certainty—We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Low certainty—Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Very low certainty—We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Abbreviations: CI, Confidence interval; MD, Mean difference; NR, Not reported; RD: Risk difference.

[a]The 'Assumed risk' is obtained from the median control group risk across studies. The 'Corresponding risk' (and its 95% confidence interval) is based on the 'Assumed risk' in the comparison group and the relative effect of the intervention (and its 95% CI).

[b]Downgraded one level for study limitations ('high risk of bias' due to detection bias – Cha et al., 2019; 'unclear risk of bias' due to some concerns about selection bias – Emanuel et al., 2020)

[c]Downgraded two levels for very serious imprecision due to the very small sample size.

16000051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

significant advantages in two studies, these studies were very different and have a high RoB because of industry support and strong conflicts of interest of some authors.

4.2 | Implications for research

In light of the above limitations, further parallel-arm RCTs with low RoB, a common definition of peri-implantitis, strict inclusion criteria, and similar treatment protocols need to be carried out.

5 | CONCLUSIONS

Because of the small number of studies included, the high RoB, and very low certainty of evidence, the adjunctive use of chemicals such as PDT, CHX, and LAbs for surface decontamination during surgery of peri-implantitis cannot be advised as superior to standard debridement procedures (mechanical debridement with/without saline).

AUTHOR CONTRIBUTIONS

Asaf Wilensky: designed the methodology, performed the search, analyzed the data, and drafted the manuscript; Lior Shapira: performed the search, analysed the data, and drafted the manuscript. Alvaro Limones, Conchita Martin: analyzed the data, drafted the manuscript, and critically revised the manuscript.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to all colleagues who responded to our requests for additional information, particularly Olivier Carcuac and Marco Esposito.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Lior Shapira https://orcid.org/0000-0001-9145-5155
Alvaro Limones https://orcid.org/0000-0002-1214-8286
Conchita Martin https://orcid.org/0000-0003-3997-6900

REFERENCES

Albaker, A. M., ArRejaie, A. S., Alrabiah, M., Al-Aali, K. A., Mokeem, S., Alasqah, M. N., Vohra, F., & Abduljabbar, T. (2018). Effect of antimicrobial photodynamic therapy in open flap debridement in the treatment of peri-implantitis: A randomized controlled trial. *Photodiagnosis Photodyn Ther*, *23*, 71–74. https://doi.org/10.1016/j.pdpdt.2018.05.003

Almohandes, A., Abrahamsson, I., Dionigi, C., & Berglundh, T. (2022). Surgical treatment of experimental peri-implantitis using mechanical and chemical decontamination procedures: A pre-clinical in vivo study. *J Clin Periodontol*, *49*(5), 518–525. https://doi.org/10.1111/jcpe.13607

Berglundh, T., Armitage, G., Araujo, M. G., Avila-Ortiz, G., Blanco, J., Camargo, P. M., Chen, S., Cochran, D., Derks, J., Figuero, E., Hämmerle, C. H. F., Heitz-Mayfield, L. J. A., Huynh-Ba, G., Iacono, V., Koo, K.-T., Lambert, F., McCauley, L., Quirynen, M., Renvert, S., … Zitzmann, N. (2018). Peri-implant diseases and conditions: Consensus report of workgroup 4 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J Clin Periodontol*, *45*(S20), S286–S291. https://doi.org/10.1111/jcpe.12957

Bombeccari, G. P., Guzzi, G., Gualini, F., Gualini, S., Santoro, F., & Spadari, F. (2013). Photodynamic therapy to treat periimplantitis. [Comparative Study Randomized Controlled Trial]. *Implant Dent*, *22*(6), 631–638. https://doi.org/10.1097/01.id.0000433592.18679.91

Carcuac, O., Derks, J., Charalampakis, G., Abrahamsson, I., Wennstrom, J., & Berglundh, T. (2016). Adjunctive Systemic and Local Antimicrobial Therapy in the Surgical Treatment of Peri-implantitis: A Randomized Controlled Clinical Trial. [Comparative Study Randomized Controlled Trial Research Support, Non-U.S. Gov't]. *J Dent Res*, *95*(1), 50–57. https://doi.org/10.1177/0022034515601961

Carrasco-Labra, A., Brignardello-Petersen, R., Santesso, N., Neumann, I., Mustafa, R. A., Mbuagbaw, L., Ikobaltzeta, I. E., De Stio, C., McCullagh, L. J., Alonso-Coello, P., Meerpohl, J. J., Vandvik, P. O., Brozek, J. L., Akl, E. A., Bossuyt, P., Churchill, R., Glenton, C., Rosenbaum, S., Tugwell, P., … Schünemann, H. J. (2016). Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. *Journal of clinical epidemiology*, *74*, 7–18. https://doi.org/10.1016/j.jclinepi.2015.12.007

Cha, J. K., Lee, J. S., & Kim, C. S. (2019). Surgical Therapy of Peri-Implantitis with Local Minocycline: A 6-Month Randomized Controlled Clinical Trial. [Randomized Controlled Trial Research Support, Non-U.S. Gov't]. *J Dent Res*, *98*(3), 288–295. https://doi.org/10.1177/0022034518818479

Cumpston, M., Li, T., Page, M. J., Chandler, J., Welch, V. A., Higgins, J. P., & Thomas, J. (2019). Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *Cochrane Database Syst Rev*, *10*, ED000142. https://doi.org/10.1002/14651858.ED000142

De Angelis, N., Felice, P., Grusovin, M. G., Camurati, A., & Esposito, M. (2012). The effectiveness of adjunctive light-activated disinfection (LAD) in the treatment of peri-implantitis: 4-month results from a multicentre pragmatic randomised controlled trial. Eur. *J Oral Implantol.*, *5*(4), 321–331.

de Waal, Y. C., Raghoebar, G. M., Huddleston Slater, J. J., Meijer, H. J., Winkel, E. G., & van Winkelhoff, A. J. (2013). Implant decontamination during surgical peri-implantitis treatment: a randomized, double-blind, placebo-controlled trial. [Randomized Controlled Trial]. *J Clin Periodontol*, *40*(2), 186–195. https://doi.org/10.1111/jcpe.12034

Derks, J., & Tomasi, C. (2015). Peri-implant health and disease. A systematic review of current epidemiology. *J Clin Periodontol*, *42*(Suppl 16), S158–S171. https://doi.org/10.1111/jcpe.12334

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials*, *7*(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Dostie, S., Alkadi, L. T., Owen, G., Bi, J., Shen, Y., Haapasalo, M., & Larjava, H. S. (2017). Chemotherapeutic decontamination of dental implants colonized by mature multispecies oral biofilm. *J Clin Periodontol*, *44*(4), 403–409. https://doi.org/10.1111/jcpe.12699

Emanuel, N., Machtei, E. E., Reichart, M., & Shapira, L. (2020). D-PLEX500: a local biodegradable prolonged release doxycycline-formulated bone graft for the treatment for peri-implantitis. A randomized controlled clinical study. [Multicenter Study Randomized Controlled Trial]. *Quintessence Int*, *51*(7), 546–553. https://doi.org/10.3290/j.qi.a44629

Esposito, M., Grusovin, M. G., De Angelis, N., Camurati, A., Campailla, M., & Felice, P. (2013). The adjunctive use of light-activated disinfection (LAD) with FotoSan is ineffective in the treatment of peri-implantitis: 1-year results from a multicentre pragmatic randomised controlled

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

trial. [Comparative Study Multicenter Study Randomized Controlled Trial Research Support, Non-U.S. Gov't]. *Eur J Oral Implantol*, *6*(2), 109–119.

Faggion, C. M., Jr., Chambrone, L., Listl, S., & Tu, Y. K. (2013). Network meta-analysis for evaluating interventions in implant dentistry: the case of peri-implantitis treatment. *Clin Implant Dent Relat Res*, *15*(4), 576–588. https://doi.org/10.1111/j.1708-8208.2011.00384.x

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., DeBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., & Schunemann, H. J. (2011). GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*, *64*(4), 383–394. https://doi.org/10.1016/j.jclinepi.2010.04.026

Heitz-Mayfield, L. J., & Mombelli, A. (2014). The therapy of peri-implantitis: a systematic review. *Int J Oral Maxillofac Implants*, *29*(Suppl), 325–345. https://doi.org/10.11607/jomi.2014suppl.g5.3

IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, *6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Koldsland, O. C., Wohlfahrt, J. C., & Aass, A. M. (2018). Surgical treatment of peri-implantitis: Prognostic indicators of short-term results. *J Clin Periodontol*, *45*(1), 100–113. https://doi.org/10.1111/jcpe.12816

Koo, K. T., Khoury, F., Keeve, P. L., Schwarz, F., Ramanauskaite, A., Sculean, A., & Romanos, G. (2019). Implant Surface Decontamination by Surgical Treatment of Periimplantitis: A Literature Review. *Implant Dent*, *28*(2), 173–176. https://doi.org/10.1097/ID.0000000000000840

Li, Z. B., Li, K., Ding, T., Yang, H. Q., Du, M., & Yu, Y. (2022). The Effect of Nonaugmentative Approaches in the Surgical Treatment of Peri-implantitis: A Systematic Review and Network Meta-analysis. *Int J Oral Maxillofac Implants*, *37*(2), 270–282. https://doi.org/10.11607/jomi.9211

Lindhe, J., Meyle, J., & Group, D. o. E. W. o. P. (2008). Peri-implant diseases: Consensus Report of the Sixth European Workshop on Periodontology. *J Clin Periodontol*, *35*(8 Suppl), 282–285. https://doi.org/10.1111/j.1600-051X.2008.01283.x

Mellado-Valero, A., Buitrago-Vera, P., Sola-Ruiz, M. F., & Ferrer-Garcia, J. C. (2013). Decontamination of dental implant surface in peri-implantitis treatment: a literature review. [Review]. *Med Oral Patol Oral Cir Bucal*, *18*(6), e869–e876. https://doi.org/10.4317/medoral.19420

Paraskevas, S., Huizinga, J. D., & Loos, B. G. (2008). A systematic review and meta-analyses on C-reactive protein in relation to periodontitis. *J Clin Periodontol*, *35*(4), 277–290. https://doi.org/10.1111/j.1600-051X.2007.01173.x

Patil, C., Agrawal, A., Abullais, S. S., Arora, S., Khateeb, S. U., & Fadul, A. E. M. (2022). Effectiveness of Different Chemotherapeutic Agents for Decontamination of Infected Dental Implant Surface: A Systematic Review. *Antibiotics (Basel)*, *11*(5). https://doi.org/10.3390/antibiotics11050593

Ramanauskaite, A., Fretwurst, T., & Schwarz, F. (2021). Efficacy of alternative or adjunctive measures to conventional non-surgical and surgical treatment of peri-implant mucositis and peri-implantitis: a systematic review and meta-analysis. *Int J Implant Dent*, *7*(1), 112. https://doi.org/10.1186/s40729-021-00388-x

Renvert, S., Polyzois, I., & Claffey, N. (2012). Surgical therapy for the control of peri-implantitis. *Clin Oral Implants Res*, *23*(Suppl 6), 84–94. https://doi.org/10.1111/j.1600-0501.2012.02554.x

Renvert, S., Roos-Jansaker, A. M., & Claffey, N. (2008). Non-surgical treatment of peri-implant mucositis and peri-implantitis: a literature review. *J Clin Periodontol*, *35*(8 Suppl), 305–315. https://doi.org/10.1111/j.1600-051X.2008.01276.x

Salvi, G. E., Cosgarea, R., & Sculean, A. (2019). Prevalence of Periimplant Diseases. *Implant Dent*, *28*(2), 100–102. https://doi.org/10.1097/ID.0000000000000872

Sanz, M., Chapple, I. L., & Working Group 4 of the, V. E. W. o. P. (2012). Clinical research on peri-implant diseases: consensus report of Working Group 4. *J Clin Periodontol*, *39*(Suppl 12), 202–206. https://doi.org/10.1111/j.1600-051X.2011.01837.x

Sanz-Martin, I., Doolittle-Hall, J., Teles, R. P., Patel, M., Belibasakis, G. N., Hammerle, C. H. F., Jung, R. E., & Teles, F. R. F. (2017). Exploring the microbiome of healthy and diseased peri-implant sites using Illumina sequencing. *J Clin Periodontol*, *44*(12), 1274–1284. https://doi.org/10.1111/jcpe.12788

Schwarz, F., Schmucker, A., & Becker, J. (2015). Efficacy of alternative or adjunctive measures to conventional treatment of peri-implant mucositis and peri-implantitis: a systematic review and meta-analysis. *Int J Implant Dent*, *1*(1), 22. https://doi.org/10.1186/s40729-015-0023-1

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, *366*, l4898. https://doi.org/10.1136/bmj.l4898

Zhuang, L. F., Watt, R. M., Mattheos, N., Si, M. S., Lai, H. C., & Lang, N. P. (2016). Periodontal and peri-implant microbiota in patients with healthy and inflamed periodontal and peri-implant tissues. *Clin Oral Implants Res*, *27*(1), 13–21. https://doi.org/10.1111/clr.12508

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wilensky, A., Shapira, L., Limones, A., & Martin, C. (2023). The efficacy of implant surface decontamination using chemicals during surgical treatment of peri-implantitis: A systematic review and meta-analysis. *Journal of Clinical Periodontology*, 1–23. https://doi.org/10.1111/jcpe.13794

1600051x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jcpe.13794 by Cochrane Netherlands, Wiley Online Library on [20/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ANEXO 3.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II).

TERCER ARTÍCULO: (Wilensky et al., 2023)			
Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]	**Respuesta**
Pregunta 1	-	Pregunta de investigación y criterios de inclusión: ¿Se establecieron preguntas de investigación y criterios de inclusión que incluyeron los componentes de PICO (Población, Intervención, Grupo comparador, Resultado)?	Si
Pregunta 2	**CRÍTICA**	**Protocolo registrado antes del inicio de la revisión:** ¿Se registró o publicó el protocolo de la revisión antes de empezar su realización y se han justificado las desviaciones significativas del protocolo?	**Si**
Pregunta 3	-	Diseño de estudios seleccionados: ¿Se explica por los autores cual es el diseño de estudio que se ha seleccionado para confeccionar la revisión?	Si
Pregunta 4	**CRÍTICA**	**Búsqueda exhaustiva de la literatura:** ¿Los autores de la revisión utilizaron una estrategia de búsqueda exhaustiva?	**Si** (debido a la limitación de tiempo para confeccionar esta revisión por encargo, solo se incluyeron estudios en inglés)
Pregunta 5	-	Duplicación en la selección de estudios: ¿Los autores de la revisión realizaron la selección de estudios por duplicado?	Si
Pregunta 6		Duplicación en la extracción de datos: ¿Los autores de la revisión realizaron la extracción de datos por duplicado?	Si
Pregunta 7	**CRÍTICA**	**Justificación para excluir estudios individuales:** ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?	**Si**
Pregunta 8	-	Descripción de estudios incluidos: ¿Los autores de la revisión describieron los estudios incluidos con suficiente detalle?	Si
Pregunta 9	**CRÍTICA**	**Evaluación del riesgo de sesgo de los estudios primarios incluidos:** ¿Utilizaron los autores de la revisión una técnica satisfactoria	**Si**

		para evaluar el riesgo de sesgo (RoB) en los estudios individuales que se incluyeron en la revisión?	
Pregunta 10	-	Fuentes de financiación: ¿Los autores de la revisión informaron sobre las fuentes de financiación de los estudios incluidos en la revisión?	Si
Pregunta 11	**CRÍTICA**	**Adecuados métodos analíticos:** Si se realizó un metaanálisis, ¿los autores de la revisión utilizaron métodos apropiados para la combinación estadística de resultados?	**Si**
Pregunta 12	-	Impacto del RoB en la metaanálisis: Si se realizó un metaanálisis, ¿los autores de la revisión evaluaron el impacto potencial de RoB de los estudios individuales sobre los resultados del metaanálisis u otra síntesis de evidencia?	Si
Pregunta 13	**CRÍTICA**	**Impacto del RoB en los resultados y discusión:** ¿Los revisores tuvieron en cuenta el RoB de los estudios primarios al interpretar/discutir los resultados de la revisión?	**Si**
Pregunta 14	-	Evaluación de la Heterogeneidad: ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?	SI (Se realizó un análisis por subgrupos para explorar la heterogeneidad)
Pregunta 15	**CRÍTICA**	**Evaluación de la presencia e impacto del sesgo de publicación:** Si se realizó una síntesis cuantitativa, ¿se llevó a cabo una investigación adecuada del sesgo de publicación (sesgo de estudios pequeños) y se discutió su probable impacto en los resultados de la revisión?	**Si**
Pregunta 16	-	Conflictos de interés: ¿Se informaron las posibles fuentes de conflicto de interés, incluyendo cualquier financiamiento recibido para llevar a cabo la revisión?	Si
VALORACIÓN FINAL DEL RIESGO DE SESGO: **ALTA CONFIANZA**			

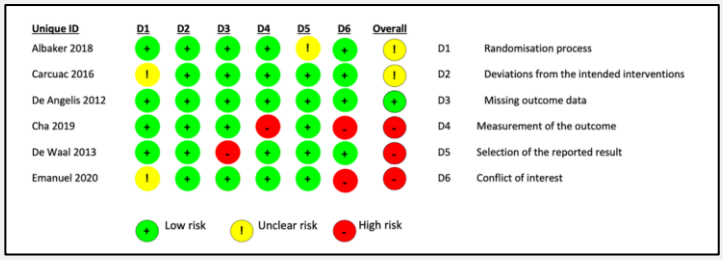
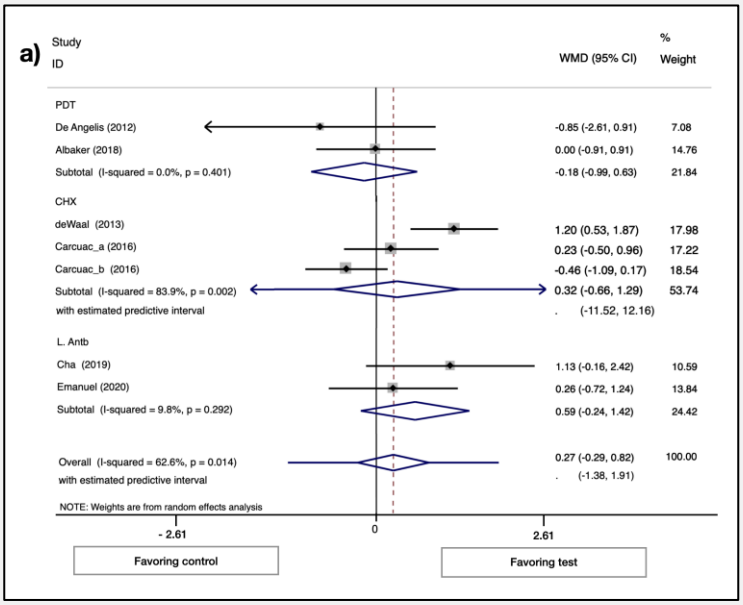
© Álvaro Limones Burgos, 2024

ANEXO 3.2: Resultados de la evaluación del riesgo de sesgo (ROBIS).

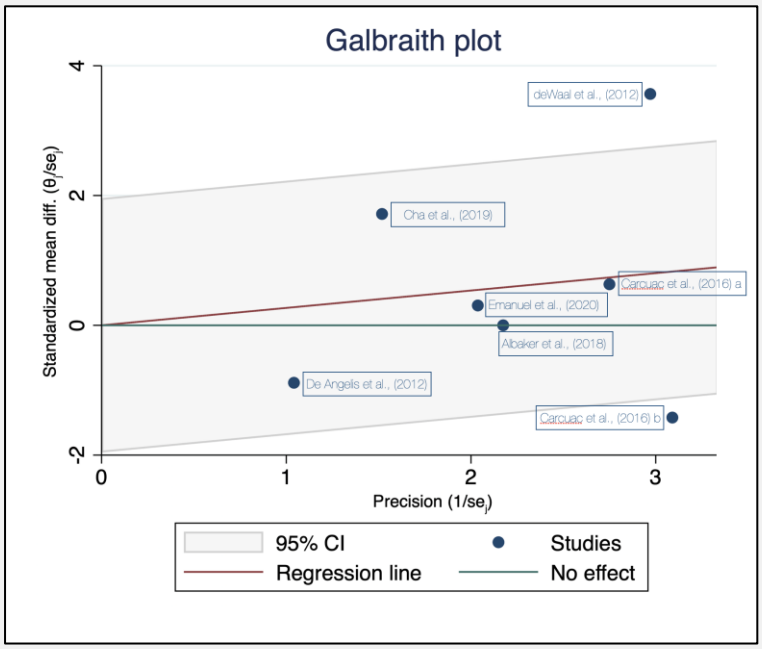
1ª FASE	2ª FASE				3ª FASE
Identificación de la pregunta de investigación y sus componentes	**Criterios de elegibilidad (Dominio 1)**	**Identificación y selección de estudios (Dominio 2)**	**Recopilación de datos y evaluación de estudios (Dominio 3)**	**Síntesis y conclusiones (Dominio 4)**	**Riesgo de sesgo global**
¿La pregunta abordada en la revisión está correctamente definida de acuerdo a sus objetivos? S/PS/PN/N/NI* Revisiones de intervención: *P: Población* *I: Intervención* *C: Comparador* *O: Variables* Revisiones de pruebas de precisión diagnóstica: *P: Población* *I: Test diagnóstico* *C: Estándar de referencia* *O: Enfermedad objetivo* Revisiones de asociaciones etiológicas o factores de riesgo: *P: Población* *I: Exposición* *C: Comparador* *O: Variables* Revisiones de estudios pronóstico: *P: Población* *I: Factor pronóstico o modelo predictivo* *O: Desenlace a predecir* Revisiones de variables e instrumentos de medición: *P: Población/Campo de conocimiento* *O: Variables e instrumentos de medición*	1.1 ¿Se ajustó la revisión a los objetivos y criterios de elegibilidad predefinidos? S/PS/PN/N/NI* 1.2 ¿Los criterios de elegibilidad eran apropiados para la pregunta de la revisión? S/PS/PN/N/NI* 1.3 ¿Eran inequívocos los criterios de elegibilidad? S/PS/PN/N/NI* 1.4 ¿Eran apropiadas las restricciones de los criterios de elegibilidad basadas en las características del estudio (por ejemplo, fecha, tamaño de la muestra, calidad del estudio, resultados medidos, etc.)? S/PS/PN/N/NI* 1.5 ¿Era adecuada alguna restricción en los criterios de elegibilidad basada en las fuentes de fuentes de información (por ejemplo, estado o formato de publicación, idioma, disponibilidad de datos)? S/PS/PN/N/NI*	2.1 ¿Incluyó la búsqueda una gama apropiada de bases de datos/fuentes electrónicas para informes publicados y no publicados? S/PS/PN/N/NI* 2.2 ¿Se utilizaron métodos adicionales a la búsqueda en bases de datos para identificar informes relevantes? S/PS/PN/N/NI* 2.3 ¿Los términos y la estructura de la estrategia de búsqueda permitieron recuperar el mayor número posible de estudios elegibles? S/PS/PN/N/NI* 2.4 ¿Fueron apropiadas las restricciones basadas en la fecha, el formato de publicación o el idioma? S/PS/PN/N/NI* 2.5 ¿Se realizaron esfuerzos para minimizar los errores en la selección de los estudios? S/PS/PN/N/NI*	3.1 ¿Se hicieron esfuerzos para minimizar los errores en la recogida de datos? S/PS/PN/N/NI* 3.2 ¿Se dispuso de suficientes características de los estudios para que tanto los autores de la revisión como los lectores pudieran interpretar los resultados? S/PS/PN/N/NI* 3.3 ¿Se recogieron todos los resultados relevantes de los estudios para su uso en la síntesis? S/PS/PN/N/NI* 3.4 ¿Se evaluó formalmente el riesgo de sesgo (o la calidad metodológica) utilizando criterios apropiados? S/PS/PN/N/NI* 3.5 ¿Se hicieron esfuerzos para minimizar los errores en la evaluación del riesgo de sesgo? S/PS/PN/N/NI*	4.1 ¿Incluyó la síntesis todos los estudios que debía? S/PS/PN/N/NI* 4.2 ¿Se comunicaron todos los análisis predefinidos o se explicaron las desviaciones? S/PS/PN/N/NI* 4.3 ¿Fue adecuada la síntesis dada la naturaleza y la similitud de las preguntas de investigación, los diseños de los estudios y los resultados de los estudios incluidos? S/PS/PN/N/NI* 4.4 ¿La variación entre estudios (heterogeneidad) fue mínima o se abordó en la síntesis? S/PS/PN/N/NI* 4.5 ¿Fueron sólidos los hallazgos, por ejemplo, como se demostró mediante el gráfico en embudo o los análisis de sensibilidad? S/PS/PN/N/NI* 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? S/PS/PN/N/NI*	A. ¿Se han abordado en la interpretación de los resultados todos los problemas señalados en los dominios del uno al cuatro? S/PS/PN/N/NI* B. ¿Se consideró adecuadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? S/PS/PN/N/NI* C. ¿Evitaron los revisores destacar los resultados en función de su significación estadística? S/PS/PN/N/NI*
¿Coincide la pregunta abordada por la revisión con la pregunta objetivo? SI/PARCIERTO/NO	**Preocupación por la especificación de los criterios de elegibilidad del estudio** ALTA/INCIERTA/BAJA	**Preocupación por los métodos utilizados para identificar y/o seleccionar los estudios** ALTA/INCIERTA/BAJA	**Preocupación por los métodos de recogida de datos y evaluación de los estudios** ALTA/INCIERTA/BAJA	**Preocupaciones sobre la síntesis y los resultados** ALTA/INCIERTA/BAJA	**Riesgo de sesgo global considerado** ALTO/[illegible]/BAJA/INEXISTENTE

© Álvaro Limones Burgos, 2024

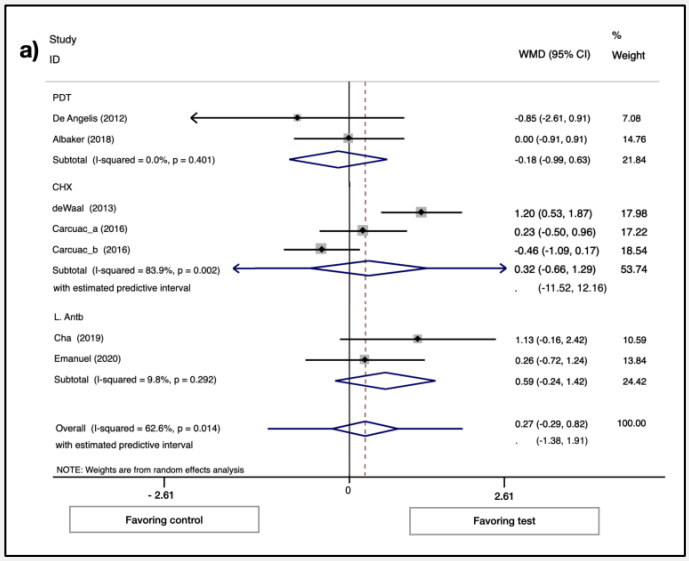
ANEXO 3.3: Resultados de la evaluación de la certeza de la evidencia científica (GRADE).

<table>
<tr><th>Dominios</th><th>Consideraciones
[Para los ejemplos de esta evaluación fue utilizada la variable principal – Profundidad de sondaje, evaluada a 6 meses de seguimiento a nivel de implante]</th><th>Variación del nivel de certeza de la evidencia</th></tr>
<tr><td>Riesgo de sesgo
[Dominio 1]</td><td>Se evaluó con la herramienta de evaluación de riesgo de sesgo RoB-2.0 (Sterne et al., 2019). Se identifico alto riesgo de sesgo por el no informe de todos los resultados, la falta de cegamiento y la presencia de conflictos de interés.
 
</td><td>↓ 1 nivel en todas las variables críticas</td></tr>
<tr><td>Inconsistencia
[Dominio 2]</td><td>Se detectó heterogeneidad en diversos forest plots del metaanálisis, como por ejemplo en el siguiente forest plot extraído del artículo publicado (ANEXO III; Fig. 3a del artículo):
 

Esta figura muestra que la heterogeneidad se atenuó parcialmente mediante un análisis de subgrupos (habiendo aplicado previamente la herramienta ICEMAN). No obstante, en el subgrupo CHX persistió una heterogeneidad inexplicada, evidenciada por:
1. Métodos estadísticos:
1.1. La prueba de heterogeneidad ratificó su continuidad: (P = 0.002; I² = 83.9%).
2. Métodos visuales:
2.1. En el mismo forest plot se puede observar como:
a. Intervalos que se solapan moderadamente entre sí.</td><td>↓ 1 nivel en todas las variables críticas</td></tr>
</table>

	b. El efecto combinado del efecto cruza la línea de no efecto, mientras que también cruza el umbral de diferencia clínicamente significativa (definido en 1 mm para la profundidad de sondaje). 2.2. <u>Gráfico radial de *Galbraith*</u>: a. Podemos ver como los estudios de deWaal y Carcuac_b quedaron fuera de las bandas de confianza de la línea de la ecuación de regresión ajustada, lo que refleja su contribución a la heterogeneidad.  Esto implicó reducir la evidencia un nivel por heterogeneidad no explicada (inconsistencias).	
Carácter indirecto de la evidencia [Dominio 3]	No se detectaron caracteres indirectos de la evidencia en ninguno de los aspectos del formato PICO: 1. La población de estudio reflejo exactamente la población de interés sobre la que se realizaron inferencias. 2. Las intervenciones fueron reproducibles en los entornos clínicos sobre los que se realizaron las inferencias. 3. Los resultados fueron evaluados a través de variables que tenían un significado directo y práctico para los pacientes (profundidad de sondaje, sangrado al sondaje y pérdida ósea marginal). 4. Los estudios incluidos en el metaanálisis eran controlados, ofreciendo comparaciones directas entre las intervenciones.	No aplica
Imprecisión [Dominio 4]	Refiriéndonos nuevamente a la figura 3a del artículo publicado (ANEXO III), identificamos varios indicadores de imprecisión:	↓ 1 nivel en todas las variables críticas

© Álvaro Limones Burgos, 2024



1. Un rango excesivamente amplio en la estimación combinada del efecto y la intersección del intervalo de confianza tanto con los umbrales de relevancia clínica (por ejemplo, 1 mm para la profundidad de sondaje) como con la línea de no efecto de manera simultánea.

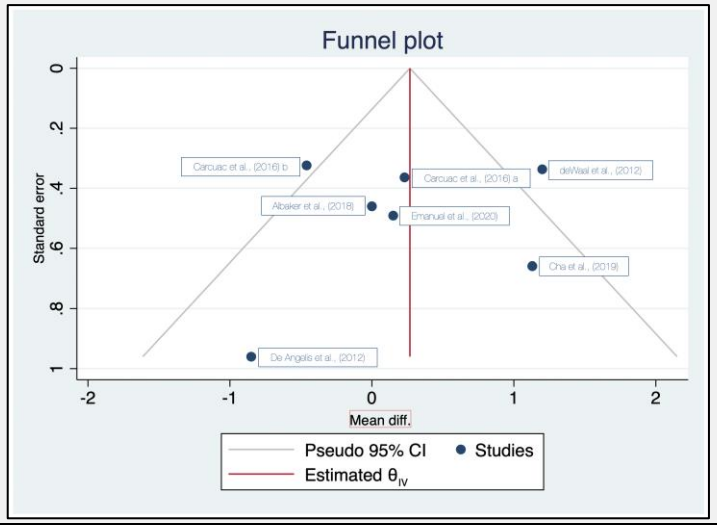
2. El estudio no alcanzó el Tamaño Óptimo de Información (OIS) fijado en 300 participantes por grupo. Esto se calculó basándose en los resultados del estudio de Carcuac et al., 2016, que encontró una diferencia de medias de -0.46 ± 0.63, con una relación de muestreo de 1:1, un nivel de significancia (alfa) del 5%, y un poder estadístico del 80%. Debido a esto, se redujo un nivel por imprecisión.

Sesgo de Publicación [Dominio 5]

Las pruebas visuales y estadísticas para evaluar el sesgo de publicación pueden no ser fiables cuando el número de estudios es menor a 10. Sin embargo, buscando transparencia en esta disertación, se aplicaron los siguientes métodos:

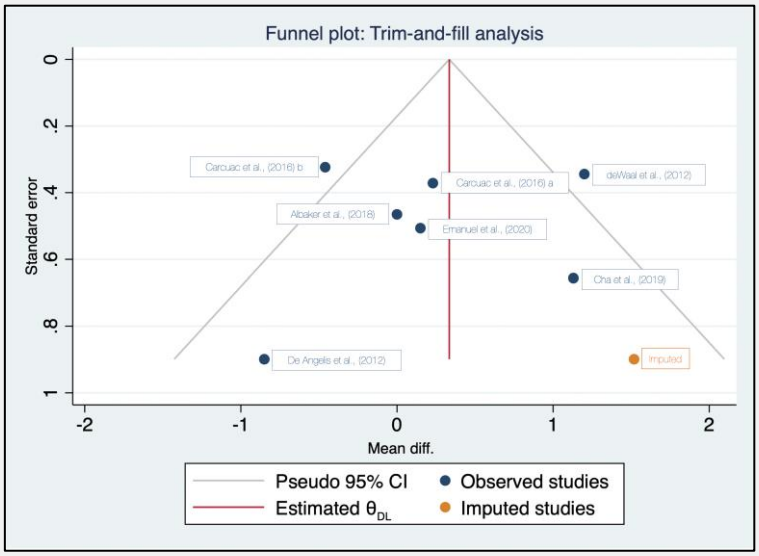
1. Métodos Estadísticos:
   No se encontró evidencia de sesgo de publicación usando métodos estadísticos, como la Prueba de Egger o el test para estudios pequeños (P=0.685).

2. Métodos Visuales:
   Se detectó una leve asimetría atribuible al estudio de deWall et al., 2013, en el diagrama de embudo (*funnel plot*) mostrado en la figura siguiente:



No aplica

© Álvaro Limones Burgos, 2024

	3. Métodos Combinados: Trim-and-fill (T&F): a. La técnica de cortado-y-relleno no modificó de manera significativa la magnitud del efecto: MD (sin imputación) = 0.27 mm (IC 95%: -0.29, 0.82) MD (imputada por T&F) = 0.34 mm (IC 95%: -0.21, 0.88) b. La visualización de la corrección asimétrica no mostró aspectos especialmente alarmantes:  Por lo tanto, se descartó la presencia de sesgo de publicación, a pesar de las limitaciones en su análisis (<10 estudios).	
Gradiente dosis-respuesta [Dominio 6]	No aplicable.	No aplicable
Gran magnitud del efecto [Dominio 7]	La magnitud combinada del efecto no presento una gran magnitud del efecto.	No aplicable
Factores de confusión plausibles [Dominio 8]	No se detectaron factores de confusión plausible que aumentase nuestra confianza en las estimaciones.	No aplicable
Computo total de la evaluación	Disminución de la certeza en tres niveles debido al riesgo de sesgo, inconsistencia e imprecisión.	↓3 nivel en todas las variables críticas ⊕○○○ MUY BAJA

© Álvaro Limones Burgos, 2024

ANEXO 3.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA).

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
TÍTULO			
Título	1	Identificar en el título del manuscrito que se trata de una revisión sistemática.	p. 1
RESUMEN			
Objetivos	2.1	Proporcionar una declaración explícita de los principales objetivos de la revisión.	p.1
Criterios de Inclusión y exclusion	2.2	Especificar los criterios de inclusión y exclusión de la revisión.	p.1
Fuentes de información	2.3	Especificar las fuentes de información (por ejemplo, las bases de datos o registros) utilizados para identificar los estudios y la fecha de la última búsqueda.	/
Riesgo de sesgo	2.4	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos.	/
Métodos de la síntesis de los resultados	2.5	Especifique los métodos utilizados para presentar y sintetizar los resultados.	p. 1
Estudios incluidos	2.6	Dar el número total de estudios y participantes incluidos y resumir las características pertinentes de los estudios.	p. 1
Síntesis de los resultados	2.7	Indicar las variables de estudio. Si se ha realizado un metaanálisis, notifique la estimación del efecto o exposición y su IC. Si se comparan dos o más grupos, indíquen la dirección del efecto (es decir, que grupo es favorecido).	p. 1
Degradación de la evidencia	2.8	Proporcionar un breve resumen de la degradación de la evidencia (por ejemplo, por riesgo de sesgo, incoherencia o impresión.	p. 1
Interpretación	2.9	Proporcionar una interpretación general de los resultados e implicaciones importantes.	p. 1-2
Financiación	2.10	Especificar la principal fuente de financiación de la revisión.	/
Registro	2.11	Proporcionar el nombre del registro y el número de registro.	p. 1
INTRODUCIÓN			
Fundamento	3	Describir la justificación de la revisión en el contexto de los conocimientos existentes.	p. 2
Objetivos	4	Proporcionar una declaración explícita de los objetivos de la revisión.	p. 2
MÉTODOS			
Criterios de elegibilidad	5	Especificar los criterios de inclusión y exclusión de la revisión y cómo se agruparon los estudios para la síntesis.	p. 2-3
Fuentes de información	6	Especificar todas las bases de datos, registros, sitios web, organizaciones, listas de referencia y otras fuentes consultadas para identificar estudios. Especificar la fecha de la última revisión.	p. 3
Estrategia de búsqueda	7	Presentar las estrategias de búsqueda completa para al menos la base de datos principal, indicar filtros u otras estrategias de búsqueda utilizadas.	Protocolo
Proceso de selección	8	Especificar los métodos utilizados para la selección de estudios.	p. 2-3

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Recogida de datos	9	Especifique los métodos utilizados para recopilar datos de los estudios, incluyendo cuántos revisores recogieron datos de cada informe.	p. 3
Datos extraidos	10a	Listar y definir todas las variables críticas.	p. 4
	10b	Listar y definir todas las demás variables secundarias. Describa cualquier suposición sobre cualquier información que falta o no esté clara.	p. 4
Métodos de evaluación del riesgo de sesgo	11	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos, incluyendo detalles de la herramienta o herramientas utilizadas, cuántos revisores evaluaron cada estudio y si trabajaban independientemente, y si se utilizaron procesos de automatización.	p. 6
Medidas del efecto	12	Especifícar para cada variable la medida(s) del efecto (por ejemplo. relación de riesgo, diferencia media) utilizada en la síntesis o presentación de resultados.	p. 4
Métodos de síntesis	13a	Describir los procesos utilizados para decidir qué estudios eran elegibles para cada síntesis (por ejemplo: tabular las características de intervención del estudio y comparar con los grupos previstos para cada síntesis).	p. 5 y 7
	13b	Describir cualquier método necesario para preparar los datos de presentación o síntesis, como el tratamiento de las estadísticas de datos faltantes o las conversiones de datos.	p. 4
	13c	Describir cualquier método utilizado para tabular o visualizar los resultados de los estudios y la síntesis individuales.	p. 4
	13d	Describir cualquier método utilizado para sintetizar los resultados y proporcionar una justificación. Si se realizó un metaanálisis, describa el(s) modelo(s), método(s) para identificar la presencia y extensión de la heterogeneidad estadística, y paquetes de software utilizado.	p. 4
	13e	Describa cualquier método utilizado para explorar posibles causas de heterogeneidad entre los resultados del estudio (por ejemplo, análisis de subgrupos, metaregresión).	p. 4
	13f	Describir cualquier análisis de sensibilidad realizado para evaluar la solidez de los resultados sintetizados.	p. 6
Implicaciones del riesgo de sesgo por datos faltantes	14	Describir cualquier método han utilizado para evaluar el riesgo de sesgo debido a la falta de datos (a partir de sesgos informativos).	p. 6
Evaluación de la certeza de la evidencia	15	Describir cualquier método utilizado para evaluar la certeza (o la confianza) de la evidencia.	p. 6
RESULTADOS			
Selección de estudios	16a	Describir los resultados del proceso de búsqueda y selección, idealmente utilizando un diagrama de flujo.	p. 6
	16b	Presentar los estudios excluidos y sus razones.	p. 4 y 6
Características de los estudios incluidos	17	Presentar cada estudio incluido y sus características.	p. 5 y 7
Riesgo de sesgo en los estudios incluidos	18	Presentar el riesgo de sesgo para cada estudio incluido.	p. 12-3

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Resultados de los estudios individuales	19	Para todas las variables presentar la estimación del efecto y su IC.	p. 9-12
Resultados de la síntesis	20a	Resumir brevemente las características y el riesgo de sesgo de los estudios incluidos.	p. 12-3
	20b	Si se ha realizado un metaanálisis, presente para cada una de las estimaciones medidas de heterogeneidad estadística. Si se comparan los grupos, describa la dirección del efecto.	p. 8, 9 y 13
	20c	Informar de posibles causas de heterogeneidad.	p. 8, 9 y 13
	20d	Informar de los análisis de sensibilidad realizados para evaluar la solidez de los resultados sintetizados.	p. 12-3
Evaluación del riesgo de sesgo por datos faltantes	21	Informar de la evaluación del riesgo de sesgo debido a datos faltantes.	p. 12-3
Certeza de la evidencia	22	Evaluaciones de la certeza (o confianza) de la evidencia para cada variable evaluada.	p. 12-3
DISCUSIÓN			
Discusión	23a	Proporcionar una interpretación general de los resultados en el contexto científico actual.	p. 13-4
	23b	Examinar cualquier limitación de la evidencia incluida en la revisión.	p. 15
	23c	Debatir las limitaciones de los procesos de revisión utilizados.	p. 15
	23d	Debatir las implicaciones de los resultados de la práctica, la política y la futura investigación.	p. 15 y 22
OTRA INFORMACIÓN			
Registro y protocolo	24a	Proporcionar información del registro de la revisión, incluyendo el nombre del registro y el número de registro, o afirmar que la revisión no está registrada.	p. 1 y 2
	24b	Indicar dónde puede accederse al protocolo de revisión o afirmar que no se ha elaborado un protocolo.	p. 2
	24c	Describir y explicar las modificaciones de la información facilitada en el registro o en el protocolo.	p. 2
Apoyo	25	Describa las fuentes de ayuda financiera o no financiera para la revisión, y el papel de los patrocinadores en la revisión.	p. 22
Intereses de competencia	26	Declare cualquier interés competitivo de los autores de revisión.	p. 22
Disponibilidad de datos, código y otros materiales	27	Informe sobre cuál de los siguientes está disponible públicamente y dónde se pueden encontrar: plantillas de recogida de datos; datos extraídos de estudios incluidos; datos utilizados para todos los análisis; o cualquier otro material utilizado en la revisión.	p. 22

ANEXO IV: Manuscrito de la cuarta revisión sistemática

(Martin et al., 2023)

© Álvaro Limones Burgos, 2024

Cochrane Library

Cochrane Database of Systematic Reviews

Retention procedures for stabilising tooth position after treatment with orthodontic braces (Review)

Martin C, Littlewood SJ, Millett DT, Doubleday B, Bearn D, Worthington HV, Limones A

Martin C, Littlewood SJ, Millett DT, Doubleday B, Bearn D, Worthington HV, Limones A.
Retention procedures for stabilising tooth position after treatment with orthodontic braces.
Cochrane Database of Systematic Reviews 2023, Issue 5. Art. No.: CD002283.
DOI: 10.1002/14651858.CD002283.pub5.

www.cochranelibrary.com

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

WILEY

TABLE OF CONTENTS

ABSTRACT 1
PLAIN LANGUAGE SUMMARY 3
SUMMARY OF FINDINGS 5
BACKGROUND 15
OBJECTIVES 15
METHODS 15
RESULTS 18
Figure 1. 19
Figure 2. 24
Figure 3. 26
DISCUSSION 40
AUTHORS' CONCLUSIONS 45
ACKNOWLEDGEMENTS 45
REFERENCES 47
CHARACTERISTICS OF STUDIES 56
DATA AND ANALYSES 137
Analysis 1.1. Comparison 1: Removable versus fixed retainers, Outcome 1: Stability: Irregularity Index 146
Analysis 1.2. Comparison 1: Removable versus fixed retainers, Outcome 2: Stability: intercanine width 147
Analysis 1.3. Comparison 1: Removable versus fixed retainers, Outcome 3: Stability: intermolar width 148
Analysis 1.4. Comparison 1: Removable versus fixed retainers, Outcome 4: Stability: arch length 149
Analysis 1.5. Comparison 1: Removable versus fixed retainers, Outcome 5: Stability: overjet 150
Analysis 1.6. Comparison 1: Removable versus fixed retainers, Outcome 6: Stability: overbite 151
Analysis 1.7. Comparison 1: Removable versus fixed retainers, Outcome 7: Stability: maximum rotation 151
Analysis 1.8. Comparison 1: Removable versus fixed retainers, Outcome 8: Failure of retainers 152
Analysis 1.9. Comparison 1: Removable versus fixed retainers, Outcome 9: Mean survival time (days) 153
Analysis 1.10. Comparison 1: Removable versus fixed retainers, Outcome 10: Patient satisfaction: acceptance 153
Analysis 1.11. Comparison 1: Removable versus fixed retainers, Outcome 11: Patient satisfaction: acceptance (dichotomous) 154
Analysis 1.12. Comparison 1: Removable versus fixed retainers, Outcome 12: Patient satisfaction: compliance (dichotomous) 154
Analysis 1.13. Comparison 1: Removable versus fixed retainers, Outcome 13: Patient satisfaction: efficiency 154
Analysis 1.14. Comparison 1: Removable versus fixed retainers, Outcome 14: Patient satisfaction: cleaning/hygiene (continuous) 155
Analysis 1.15. Comparison 1: Removable versus fixed retainers, Outcome 15: Patient satisfaction: cleaning/hygiene (dichotomous) 155
Analysis 1.16. Comparison 1: Removable versus fixed retainers, Outcome 16: Patient satisfaction: discomfort/pain 155
Analysis 1.17. Comparison 1: Removable versus fixed retainers, Outcome 17: Patient satisfaction: discomfort/pain (dichotomous) 156
Analysis 1.18. Comparison 1: Removable versus fixed retainers, Outcome 18: Patient satisfaction: speech (dichotomous) 156
Analysis 1.19. Comparison 1: Removable versus fixed retainers, Outcome 19: Oral health: evidence of caries (dichotomous) 156
Analysis 1.20. Comparison 1: Removable versus fixed retainers, Outcome 20: Oral health: Plaque Index 157
Analysis 1.21. Comparison 1: Removable versus fixed retainers, Outcome 21: Oral health: Gingival Index 158
Analysis 1.22. Comparison 1: Removable versus fixed retainers, Outcome 22: Oral health: Calculus Index 159
Analysis 1.23. Comparison 1: Removable versus fixed retainers, Outcome 23: Oral health: probing depth 160
Analysis 1.24. Comparison 1: Removable versus fixed retainers, Outcome 24: Oral health: periodontal pocketing (dichotomous) 160
Analysis 1.25. Comparison 1: Removable versus fixed retainers, Outcome 25: Oral health: bleeding on probing 161
Analysis 1.26. Comparison 1: Removable versus fixed retainers, Outcome 26: Oral health: gingival bleeding (dichotomous) 161
Analysis 2.1. Comparison 2: Fixed versus fixed retainers, Outcome 1: Stability: Irregularity Index 174
Analysis 2.2. Comparison 2: Fixed versus fixed retainers, Outcome 2: Stability: intercanine width 176
Analysis 2.3. Comparison 2: Fixed versus fixed retainers, Outcome 3: Stability: interpremolar width 178
Analysis 2.4. Comparison 2: Fixed versus fixed retainers, Outcome 4: Stability: intermolar width 179
Analysis 2.5. Comparison 2: Fixed versus fixed retainers, Outcome 5: Stability: anterior arch length 180
Analysis 2.6. Comparison 2: Fixed versus fixed retainers, Outcome 6: Stability: total arch length 181

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.7. Comparison 2: Fixed versus fixed retainers, Outcome 7: Stability: overjet 182
Analysis 2.8. Comparison 2: Fixed versus fixed retainers, Outcome 8: Stability: overbite 182
Analysis 2.9. Comparison 2: Fixed versus fixed retainers, Outcome 9: Stability: maximum rotation 182
Analysis 2.10. Comparison 2: Fixed versus fixed retainers, Outcome 10: Failure of retainers 183
Analysis 2.11. Comparison 2: Fixed versus fixed retainers, Outcome 11: Mean survival time (months) 186
Analysis 2.12. Comparison 2: Fixed versus fixed retainers, Outcome 12: Failure risk (Hazard Ratio) 187
Analysis 2.13. Comparison 2: Fixed versus fixed retainers, Outcome 13: Patient satisfaction: efficiency 188
Analysis 2.14. Comparison 2: Fixed versus fixed retainers, Outcome 14: Patient satisfaction: cleaning/hygiene 188
Analysis 2.15. Comparison 2: Fixed versus fixed retainers, Outcome 15: Patient satisfaction: comfort 189
Analysis 2.16. Comparison 2: Fixed versus fixed retainers, Outcome 16: Patient satisfaction: speech 189
Analysis 2.17. Comparison 2: Fixed versus fixed retainers, Outcome 17: Patient satisfaction: appointment needed 190
Analysis 2.18. Comparison 2: Fixed versus fixed retainers, Outcome 18: Patient satisfaction: aesthetics 190
Analysis 2.19. Comparison 2: Fixed versus fixed retainers, Outcome 19: Patient satisfaction: preference for type of retainer 190
Analysis 2.20. Comparison 2: Fixed versus fixed retainers, Outcome 20: Oral health: Plaque Index 191
Analysis 2.21. Comparison 2: Fixed versus fixed retainers, Outcome 21: Oral health: Gingival Index 193
Analysis 2.22. Comparison 2: Fixed versus fixed retainers, Outcome 22: Oral health: Calculus Index 195
Analysis 2.23. Comparison 2: Fixed versus fixed retainers, Outcome 23: Oral health: probing depth 196
Analysis 2.24. Comparison 2: Fixed versus fixed retainers, Outcome 24: Oral health: bleeding on probing 197
Analysis 2.25. Comparison 2: Fixed versus fixed retainers, Outcome 25: Oral health: marginal recession 197
Analysis 3.1. Comparison 3: Removable versus removable retainers, Outcome 1: Stability: Irregularity Index 210
Analysis 3.2. Comparison 3: Removable versus removable retainers, Outcome 2: Stability: intercanine width 212
Analysis 3.3. Comparison 3: Removable versus removable retainers, Outcome 3: Stability: interpremolar width 213
Analysis 3.4. Comparison 3: Removable versus removable retainers, Outcome 4: Stability: intermolar width 214
Analysis 3.5. Comparison 3: Removable versus removable retainers, Outcome 5: Stability: arch length 216
Analysis 3.6. Comparison 3: Removable versus removable retainers, Outcome 6: Stability: overjet 217
Analysis 3.7. Comparison 3: Removable versus removable retainers, Outcome 7: Stability: overbite 217
Analysis 3.8. Comparison 3: Removable versus removable retainers, Outcome 8: Stability: maintaining corrected rotations in the upper 218
Analysis 3.9. Comparison 3: Removable versus removable retainers, Outcome 9: Stability: maintaining quality of finish (PAR) 218
Analysis 3.10. Comparison 3: Removable versus removable retainers, Outcome 10: Stability: maintaining quality of finish (ABO Grading) 219
Analysis 3.11. Comparison 3: Removable versus removable retainers, Outcome 11: Failures of retainers: total number of failures 220
Analysis 3.12. Comparison 3: Removable versus removable retainers, Outcome 12: Failure of retainers: broken retainers 221
Analysis 3.13. Comparison 3: Removable versus removable retainers, Outcome 13: Failure of retainers: lost retainers 222
Analysis 3.14. Comparison 3: Removable versus removable retainers, Outcome 14: Failure of retainers: nonfitting retainers 223
Analysis 3.15. Comparison 3: Removable versus removable retainers, Outcome 15: Failure of retainers: serious abrasion 224
Analysis 3.16. Comparison 3: Removable versus removable retainers, Outcome 16: Mean survival time (days) 224
Analysis 3.17. Comparison 3: Removable versus removable retainers, Outcome 17: Failure risk (Hazard Ratio) 225
Analysis 3.18. Comparison 3: Removable versus removable retainers, Outcome 18: Patient satisfaction: acceptance 225
Analysis 3.19. Comparison 3: Removable versus removable retainers, Outcome 19: Patient satisfaction: acceptance (dichotomous) 226
Analysis 3.20. Comparison 3: Removable versus removable retainers, Outcome 20: Patient satisfaction: compliance (dichotomous) 226
Analysis 3.21. Comparison 3: Removable versus removable retainers, Outcome 21: Patient satisfaction: cleaning/hygiene 226
Analysis 3.22. Comparison 3: Removable versus removable retainers, Outcome 22: Patient satisfaction: comfort 227
Analysis 3.23. Comparison 3: Removable versus removable retainers, Outcome 23: Patient satisfaction: discomfort/pain (dichotomous) 227
Analysis 3.24. Comparison 3: Removable versus removable retainers, Outcome 24: Patient satisfaction: speech 227
Analysis 3.25. Comparison 3: Removable versus removable retainers, Outcome 25: Patient satisfaction: speech (dichotomous) 228
Analysis 3.26. Comparison 3: Removable versus removable retainers, Outcome 26: Patient satisfaction: durability 228
Analysis 3.27. Comparison 3: Removable versus removable retainers, Outcome 27: Oral health: Plaque Index 229
Analysis 3.28. Comparison 3: Removable versus removable retainers, Outcome 28: Oral health: Gingival Index 229

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.29. Comparison 3: Removable versus removable retainers, Outcome 29: Oral health: probing depth 230
Analysis 3.30. Comparison 3: Removable versus removable retainers, Outcome 30: Oral health: bleeding on probing 230
Analysis 4.1. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 1: Stability: upper irregularity index 235
Analysis 4.2. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 2: Stability: upper intercanine width 236
Analysis 4.3. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 3: Stability: upper intermolar width 237
Analysis 4.4. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 4: Stability: upper arch length 238
Analysis 4.5. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 5: Stability: lower irregularity index 239
Analysis 4.6. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 6: Stability: lower intercanine width 240
Analysis 4.7. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 7: Stability: lower intermolar width 241
Analysis 4.8. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 8: Stability: lower arch length 242
Analysis 4.9. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 9: Stability: overjet 243
Analysis 4.10. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 10: Stability: overbite 244
Analysis 5.1. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 1: Stability: upper irregularity index 245
Analysis 5.2. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 2: Stability: upper intercanine width 245
Analysis 5.3. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 3: Stability: upper intermolar width 245
Analysis 5.4. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 4: Stability: upper arch length 246
Analysis 5.5. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 5: Stability: upper maximum rotation 246
ADDITIONAL TABLES 246
APPENDICES 248
WHAT'S NEW 252
HISTORY 252
CONTRIBUTIONS OF AUTHORS 253
DECLARATIONS OF INTEREST 253
SOURCES OF SUPPORT 253
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 254
INDEX TERMS 254

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

[Intervention Review]

Retention procedures for stabilising tooth position after treatment with orthodontic braces

Conchita Martin[1], Simon J Littlewood[2], Declan T Millett[3], Bridget Doubleday[4], David Bearn[5], Helen V Worthington[6], Alvaro Limones[7]

[1]Orthodontic Department, Faculty of Dentistry, Complutense University of Madrid, Madrid, Spain. [2]Orthodontic Department, St Luke's Hospital, Bradford, UK. [3]Oral Health and Development, Cork University Dental School and Hospital, Cork, Ireland. [4]Orthodontic Department, Forth Valley Royal Hospital, Falkirk, UK. [5]School of Dentistry, University of Dundee, Dundee, UK. [6]Cochrane Oral Health, Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. [7]Faculty of Dentistry, Complutense Univesity of Madrid, Madrid, Spain

Contact: Simon J Littlewood, simonjlittlewood@aol.com.

Editorial group: Cochrane Oral Health Group.
Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 5, 2023.

Citation: Martin C, Littlewood SJ, Millett DT, Doubleday B, Bearn D, Worthington HV, Limones A. Retention procedures for stabilising tooth position after treatment with orthodontic braces. *Cochrane Database of Systematic Reviews* 2023, Issue 5. Art. No.: CD002283. DOI: 10.1002/14651858.CD002283.pub5.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Without a phase of retention after successful orthodontic treatment, teeth tend to 'relapse', that is, to return to their initial position. Retention is achieved by fitting fixed or removable retainers to provide stability to the teeth while avoiding damage to teeth and gums. Removable retainers can be worn full- or part-time. Retainers vary in shape, material, and the way they are made. Adjunctive procedures are sometimes used to try to improve retention, for example, reshaping teeth where they contact ('interproximal reduction'), or cutting fibres around teeth ('percision').

This review is an update of one originally published in 2004 and last updated in 2016.

Objectives

To evaluate the effects of different retainers and retention strategies used to stabilise tooth position after orthodontic braces.

Search methods

An information specialist searched Cochrane Oral Health Trials Register, CENTRAL, MEDLINE, Embase and OpenGrey up to 27 April 2022 and used additional search methods to identify published, unpublished and ongoing studies.

Selection criteria

Randomised controlled trials (RCTs) involving children and adults who had retainers fitted or adjunctive procedures undertaken to prevent relapse following orthodontic treatment with braces. We excluded studies with aligners.

Data collection and analysis

Two review authors independently screened eligible studies, assessed risk of bias and extracted data. Outcomes were stability or relapse of tooth position, retainer failure (i.e. broken, detached, worn out, ill-fitting or lost), adverse effects on teeth and gums (i.e. plaque, gingival and bleeding indices), and participant satisfaction. We calculated mean differences (MD) for continuous data, risk ratios (RR) or risk differences (RD) for dichotomous data, and hazard ratios (HR) for survival data, all with 95% confidence intervals (CI). We conducted meta-analyses when similar studies reported outcomes at the same time point; otherwise results were reported as mean ranges. We prioritised

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

reporting of Little's Irregularity Index (crookedness of anterior teeth) to measure relapse, judging the minimum important difference to be 1 mm.

Main results

We included 47 studies, with 4377 participants. The studies evaluated: removable versus fixed retainers (8 studies); different types of fixed retainers (22 studies) or bonding materials (3 studies); and different types of removable retainers (16 studies). Four studies evaluated more than one comparison. We judged 28 studies to have high risk of bias, 11 to have low risk, and eight studies as unclear.

We focused on 12-month follow-up.

The evidence is low or very low certainty. Most comparisons and outcomes were evaluated in only one study at high risk of bias, and most studies measured outcomes after less than a year.

Removable versus fixed retainers

Removable (part-time) versus fixed

One study reported that participants wearing clear plastic retainers part-time in the lower arch had more relapse than participants with multistrand fixed retainers, but the amount was not clinically significant (Little's Irregularity Index (LII) MD 0.92 mm, 95% CI 0.23 to 1.61; 56 participants). Removable retainers were more likely to cause discomfort (RR 12.22; 95% CI 1.69 to 88.52; 57 participants), but were associated with less retainer failure (RR 0.44, 95% CI 0.20 to 0.98; 57 participants) and better periodontal health (Gingival Index (GI) MD −0.34, 95% CI −0.66 to −0.02; 59 participants).

Removable (full-time) versus fixed

One study reported that removable clear plastic retainers worn full-time in the lower arch did not provide any clinically significant benefit for tooth stability over fixed retainers (LII MD 0.60 mm, 95% CI 0.17 to 1.03; 84 participants). Participants with clear plastic retainers had better periodontal health (gingival bleeding RR 0.53, 95% CI 0.31 to 0.88; 84 participants), but higher risk of retainer failure (RR 3.42, 95% CI 1.38 to 8.47; 77 participants). The study found no difference between retainers for caries.

Different types of fixed retainers

Computer-aided design/computer-aided manufacturing (CAD/CAM*) nitinol versus conventional/analogue multistrand*

One study reported that CAD/CAM nitinol fixed retainers were better for tooth stability, but the difference was not clinically significant (LII MD −0.46 mm, 95% CI −0.72 to −0.21; 66 participants). There was no evidence of a difference between retainers for periodontal health (GI MD 0.00, 95% CI -0.16 to 0.16; 2 studies, 107 participants), or retainer survival (RR 1.29, 95% CI 0.67 to 2.49; 1 study, 41 participants).

Fibre-reinforced composite versus conventional multistrand/spiral wire

One study reported that fibre-reinforced composite fixed retainers provided better stability than multistrand retainers, but this was not of a clinically significant amount (LII MD −0.70 mm, 95% CI −1.17 to −0.23; 52 participants). The fibre-reinforced retainers had better patient satisfaction with aesthetics (MD 1.49 cm on a visual analogue scale, 95% CI 0.76 to 2.22; 1 study, 32 participants), and similar retainer survival rates (RR 1.01, 95% CI 0.84 to 1.21; 7 studies; 1337 participants) at 12 months. However, failures occurred earlier (MD −1.48 months, 95% CI −1.88 to −1.08; 2 studies, 103 participants; 24-month follow-up) and more gingival inflammation at six months, though bleeding on probing (BoP) was similar (GI MD 0.59, 95% CI 0.13 to 1.05; BoP MD 0.33, 95% CI −0.13 to 0.79; 1 study, 40 participants).

Different types of removable retainers

Clear plastic versus Hawley

When worn in the lower arch for six months full-time and six months part-time, clear plastic provided similar stability to Hawley retainers (LII MD 0.01 mm, 95% CI −0.65 to 0.67; 1 study, 30 participants). Hawley retainers had lower risk of failure (RR 0.60, 95% CI 0.43 to 0.83; 1 study, 111 participants), but were less comfortable at six months (VAS MD -1.86 cm, 95% CI -2.19 to -1.53; 1 study, 86 participants).

Part-time versus full-time wear of Hawley

There was no evidence of a difference in stability between part-time and full-time use of Hawley retainers (MD 0.20 mm, 95% CI −0.28 to 0.68; 1 study, 52 participants).

Authors' conclusions

The evidence is low to very low certainty, so we cannot draw firm conclusions about any one approach to retention over another. More high-quality studies are needed that measure tooth stability over at least two years, and measure how long retainers last, patient satisfaction and negative side effects from wearing retainers, such as tooth decay and gum disease.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

PLAIN LANGUAGE SUMMARY

What is the best method for maintaining the correct position of teeth after orthodontic treatment?

Key messages

- We cannot draw firm conclusions about any one approach to retention over another because we have little confidence in the evidence.

- More well-planned studies are needed. They should measure tooth stability over at least two years, how long the retainers last, patient satisfaction and any unwanted effects from the retainers, such as tooth decay and gum disease.

What is the problem?

Teeth can become crooked again after being straightened with orthodontic braces ('relapse'). Orthodontists try to prevent this using different ways of holding the teeth straight ('retention procedures'). Retention procedures can include wearing custom-made fixed or removable retainers, usually made of wires or clear plastic, after orthodontic braces have been removed. Removable retainers can be worn full- or part-time. Retainers fit over or around teeth, or stick onto the back of teeth. Orthodontists may also use extra treatments ('adjunctive procedures') such as smoothing contacts between the teeth ('interproximal enamel reduction'), or cutting fibres that connect the gum and the neck of the tooth ('pericision').

What did we want to know?

We wanted to find out the best way to maintain teeth in their new position after the end of treatment with orthodontic braces and whether there are any unwanted effects.

What did we do?

We searched for studies that compared different types of retainers following treatment with orthodontic braces, and that measured the effects for at least three months after the orthodontic treatment. We summarised study results and rated our confidence in the evidence.

What did we find?

We found 47 studies with 4377 adults and children. Most studies took place in hospitals, universities or specialist practices, NHS clinics, or a mixture of places. Most of the studies tested: removable retainers versus fixed retainers; different types of fixed retainers or bonding materials; or different types of removable retainers. Most measured the effects of treatment for less than one year.

Main results

Removable versus fixed retainers (8 studies)

Teeth may return to their previous position more with part-time wear of a removable retainer made of clear plastic than with a fixed retainer, but the difference was so small that it may not be important. If the removable retainer is worn full-time, there may be no difference between the retainers in tooth movement. Clear plastic retainers worn part-time may fail less than fixed retainers, but when worn full-time may fail more. Better gum health may be achieved with clear plastic removable retainers than with fixed retainers, but no difference was found in tooth decay.

Different types of fixed retainers and glues to stick them on (21 studies)

CAD/CAM (computer-aided design/computer-aided manufacturing) nitinol fixed retainers compared to conventional multistrand wire fixed retainers may not differ in terms of how well they hold the teeth straight, gum health, or how long they last.

Fibre-reinforced composite retainers are less obvious on the teeth therefore participants are more satisfied with them because of their appearance. Fibre-reinforced composite retainers may be slightly better at keeping the teeth straight, but the difference was so small, it might not be important. Fibre-reinforced composite retainers might be more likely to fail earlier and be associated with more gum problems than multistrand fixed retainers.

Different types of removable retainers (16 studies)

The effects of Hawley removable retainers compared with clear plastic retainers on movement of the teeth may be similar, whether they are worn full- or part-time. Clear plastic retainers may provide better patient satisfaction but be less likely to last than Hawley retainers.

What are the limitations of the evidence?

Our confidence in the evidence is low because the studies were small and did not use the best methods. Very few studies looked at more than one of the things we were interested in. Most studies measured results less than a year following treatment. Also, we do not know

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

how much study results may be affected by the age of the participants and whether they are still growing, what kind of problems they had with their teeth before their first braces, and what other treatment they had, such as extractions.

How up to date is this evidence

The evidence is up to date to 27 April 2022. This review updates one originally published in 2004 and last updated in 2016.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

SUMMARY OF FINDINGS

Summary of findings 1. Removable retainer (part-time) versus fixed retainer to stabilise tooth position

Population: people who had received fixed appliance treatment
Setting: specialised orthodontic practice and hospital or university orthodontic department
Intervention: part-time clear plastic removable retainers
Comparison: any type of fixed retainer

Follow-up: 12 months

Arch: lower

Outcome	Outcome measure	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of evidence (GRADE)	Comments
		Assumed risk	Corresponding risk				
		Any type of fixed retainer	Part-time clear plastic removable retainer				
Stability Follow-up: 12 months	**Little's Irregularity Index (LII)** in lower arch (ideally 0 mm)	Mean LII in the fixed group was 0.77 mm	**MD 0.92 mm higher** (0.23 to 1.61 higher)	-	56 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c,d]	1 RCT at high risk of detection bias suggested that there was**slightly more relapse with the clear plastic retainer, but the mean difference was less than 1 mm, which is not clinically significant.**
Survival Follow-up: 12 months	**Failure of retainer** in lower arch	500 per 1000	**280 fewer per 1000** (400 to 10 less)	**RR 0.44** (0.20 to 0.98)	57 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c,d]	1 RCT at high risk of detection bias suggested that there was a**lower risk of retainer failure with the part-time clear plastic retainer.**
Patient satisfaction Follow-up: 12 months	**Discomfort/pain** in lower arch	33 per 1000	**374 more per 1000** (23 to 1000 more)	**RR 12.22** (1.69 to 88.52)	57 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c,d]	1 RCT at high risk of detection bias suggested that there was a **higher risk of discomfort with the part-time clear plastic retainer.**

Cochrane Library
Trusted evidence.
Informed decisions.
Better health.

Adverse effects on oral health Follow-up: 12 months	**Gingival Index (GI)** in lower arch (ideally 0)	Mean GI in the fixed group was 0.92	**MD 0.34 lower** (0.66 lower to 0.02 lower)	-	59 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c,d]	1 RCT at high risk of detection bias suggested that there was**better periodontal health with the part-time clear plastic retainer**.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **LII:** Little's Irregularity Index; **MD:** mean difference; **RCT:** randomised controlled trial; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level due to study limitations (high risk of detection bias).
[b]Downgraded two levels for very serious imprecision due to the small sample size of a single study.
[c]Downgraded one level since there is only one study for this comparison.
[d]Only one study, which made it difficult to ascertain the likelihood of publication bias.

Summary of findings 2. Removable retainer (full time) versus fixed retainer to stabilise tooth position

Population: people who had received fixed appliance treatment
Setting: specialist orthodontic practice and hospital orthodontic department
Intervention: full-time clear plastic removable retainer
Comparison: any type of fixed retainer

Follow-up: 12 months

Arch: lower

Outcome	Outcome measure	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of evidence (GRADE)	Comments
		Assumed risk	Corresponding risk				
		Any type of fixed retainer	Full-time clear plastic removable retainer				

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Stability Follow-up: 12 months	**Little's Irregularity Index (LII)** in lower arch (ideally 0 mm)	Mean LII in the fixed group was 0.44 mm	**MD 0.60 mm higher** (0.17 to 1.03 higher)	-	84 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at high risk of reporting bias suggested that there was more**relapse with the clear full-time clear plastic removable retainer, but the mean difference was less than 1 mm, which is not clinically significant.**
Survival Follow-up: 12 months	**Failure of retainer** in lower arch	122 per 1000	**295 more per 1000** (46 to 911 more)	**RR 3.42** (1.38 to 8.47)	77 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,c,d]	1 RCT at high risk of reporting bias suggested that there was a **higher risk of failure with clear plastic retainers**.
Patient satisfaction Follow-up: 12 months	**Discomfort/pain** in lower arch	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome.
Adverse effects on oral health Follow-up: 12 months	**Caries** in lower arch	See comments	See comments	**RD 0.00** (−0.05 to 0.05)	84 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,c,d]	1 RCT at high risk of reporting bias suggested that there was **no difference between retainers**.
	Gingival bleeding in lower arch (ideally < 0.10)	617 per 1000	**290 less per 1000** (426 to 74 less)	**RR 0.53** (0.31 to 0.88)	84 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,c,d]	1 RCT at high risk of reporting bias suggested that there was **better periodontal health with the full-time clear plastic retainers**.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **LII:** Little's Irregularity Index; **MD:** mean difference; **RCT:** randomised controlled trial; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect
Moderate certainty: we are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low certainty: our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low certainty: we have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[a]Downgraded one level due to study limitations (serious/unclear risk of reporting bias).

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

[b]Downgraded one level for imprecision due to the small sample size.
[c]One study - difficult to ascertain the likelihood of publication bias.
[d]Downgraded two levels for imprecision due to the very small sample size.

Summary of findings 3. Fixed retainer (CAD/CAM nitinol) versus fixed retainer (conventional/analogue multistrand) to stabilise tooth position

Population: people who had received fixed appliance treatment
Setting: specialised orthodontic practice and hospital or university orthodontic department
Intervention: CAD/CAM nitinol fixed retainer
Comparison: conventional/analogue multistrand fixed retainer

Follow-up: 12-months

Arch: lower

Outcome	Outcome measure	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of evidence (GRADE)	Comments
		Assumed risk	Corresponding risk				
		Multistrand fixed retainer	CAD/CAM nitinol fixed retainer				
Stability Follow-up: 12 months	**Little's Irregularity Index (LII)** in lower arch (ideally 0 mm)	Mean LII in the fixed group was 0.56 mm	**MD 0.46 mm lower** (0.72 lower to 0.21 lower)	-	66 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at high risk of detection bias and unclear risk of selection and reporting bias suggested that there was **less relapse with CAD/CAM nitinol fixed retainer, but the mean difference was less than 1 mm, which is not clinically significant.**
Survival Follow-up: 12 months	**Failure of retainer** in lower arch	409 per 1000	**119 more per 1000** (135 less to 610 more)	**RR 1.29** (0.67 to 2.49)	41 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at high risk of detection bias and unclear risk of selection bias suggested that there was **no difference** between retainers.
Patient satisfaction Follow-up: 12 months	**Discomfort/pain** in lower arch	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Adverse effects on oral health Follow-up: 12 months	**Bleeding on probing** in lower arch (ideally < 0.10)	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome at 12-month follow-up. 1 RCT (40 participants) at high risk of detection bias and 1 RCT (81 participants) at unclear risk of detection bias suggested that there was **no difference between retainers at 6-month follow-up** (low-certainty evidence).
	Gingival Index (GI) in lower arch (ideally 0)	Mean GI ranged from 0.23 to 0.33 across control groups	**MD 0.00** (0.16 lower to 0.16 higher)	-	107 (2 RCTs)	⊕⊕⊖⊖ **Low**[a,b]	At 12 months, 2 RCTs at high risk of detection bias and unclear risk of selection bias suggested that there was **no difference** between retainers. At 6 months, 4 RCTs at high or unclear risk of detection bias suggested that**periodontal health was better with the CAD/CAM nitinol retainer** (GI MD −0.14, 95% CI −0.27 to −0.01; 198 participants), but the amount was not clinically significant.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **GI:** Gingival Index; **LII:** Little's Irregularity Index; **MD:** mean difference; **RCT:** randomised controlled trial; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level due to study limitations (serious/unclear risk detection bias).
[b]Only one or two studies - difficult to ascertain the likelihood of publication bias.
[c]Downgraded one level for imprecision due to the small sample size.

Summary of findings 4. Fixed retainer (fibre-reinforced composite) versus fixed retainer (multistrand/spiral wire) to stabilise tooth position

Population: people who had received fixed appliance treatment
Setting: specialised orthodontic practice and hospital or university orthodontic department
Intervention: fibre-reinforced composite retainer
Comparison: multistrand/spiral wire fixed retainer

Follow-up: 12 months

Arch: lower

Outcome	Outcome measure	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of evidence (GRADE)	Comments
		Assumed risk	Corresponding risk				
		Multi-strand/spiral wire fixed retainers	Fibre-reinforced composite retainer				
Stability Follow-up: 12 months	**Little's Irregularity Index(LII)** in lower arch (ideally 0 mm)	Mean LII was 0.97 mm	**MD 0.70 mm lower** (1.17 lower to 0.23 lower)	-	52 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at high risk of detection bias and unclear risk of selection bias suggested that **there was better stability with the fibre-reinforced composite retainer, but the mean difference was less than 1 mm, which is not clinically significant.**
Survival Follow-up: 12 months	**Failure of retainer** in lower arch	247 per 1000	**2 more per 1000** (39 less to 52 more)	**RR 1.01** (0.84 to 1.21)	1337 (7 RCTs)	⊕⊕⊖⊖ **Low**[a,d]	4 RCTs at high risk of detection, attrition, and reporting bias and 3 RCTs at unclear selection, detection, and attrition bias suggested that there was **no difference between the retainer types**.
	Mean survival time in lower arch (months)	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome at 12-month follow-up. 1 RCT at high risk of detection and reporting bias and 1 RCT at unclear risk of selection and detection bias suggested that **failures occurred earlier when a fibre-reinforced composite retainer was used** at **24-month follow-up** (MD −1.48, 95% CI −1.88 to −1.08; 103 participants; very low-certainty evidence).
Patient satisfaction Follow-up: 12 months	**Aesthetic** in lower arch VAS 0–10 cm	Mean VAS score was 8.24 cm	**MD 1.49 cm higher** (0.76 higher to 2.22 higher)	-	32 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at unclear selection, detection and attrition bias suggested that **satisfaction with aesthetics was higher with the fibre-reinforced composite fixed retainer**.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

	(10 cm is very satisfied)						
Adverse effects on oral health Follow-up: 12 months	**Bleeding on probing** in lower arch (ideally < 0.10)	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome at 12-month follow-up. 1 RCT at high risk of detection and attrition bias suggested that there was **no difference between retainers at 6-month-follow-up** (40 participants; very low-certainty evidence).
	Gingival Index (GI) in lower arch (ideally 0)	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome at 12-month follow-up. 1 RCT at high risk of detection and attrition bias suggested that there was**higher gingival inflammation of moderate clinical relevance with the fibre-reinforced composite fixed retainer at 6-month follow-up** (MD 0.59, 95% CI 0.13 to 1.05; 1 RCT, 40 participants; very low-certainty evidence).

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **GI:** Gingival Index; **LII:** Little's Irregularity Index; **MD:** mean difference; **RCT:** randomised controlled trial; **RR:** risk ratio; **VAS**: visual analogue scale

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level due to study limitations (serious/unclear risk of bias).
[b]Downgraded two levels for imprecision due to the small sample size.
[c]One study - difficult to ascertain the likelihood of publication bias.
[d]Downgraded one level for inconsistency due to the presence of unexplained heterogeneity.

Summary of findings 5. Removable retainer (clear plastic) versus removable retainer (Hawley) to stabilise tooth position

Population: people who had received fixed appliance treatment
Setting: specialist orthodontic practice and hospital orthodontic department
Intervention: Hawley removable retainer (any coincident wearing protocol between treatment arms)
Comparison: clear plastic removable retainer (any coincident wearing protocol between treatment arms)

Follow-up: 12 months

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Arch: lower

Outcome	Outcome measure	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of evidence (GRADE)	Comments
		Assumed risk	Corresponding risk				
		Clear plastic removable retainer	Hawley removable retainer				
Stability Follow-up: 12 months	**Little's Irregularity Index(LII)** in lower arch (ideally 0 mm)	Mean LII was 0.72 mm	**MD 0.01 mm higher** (0.65 lower to 0.67 higher)	-	30 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at high risk of selection bias suggested that there was **no difference between the removable retainer types** when worn 6 months full time and 6 months part time.
Survival Follow-up: 12 months	**Failure of retainer** in upper and lower arches	745 per 1000	**298 fewer per 1000** (425 to 127 fewer)	**RR 0.60** (0.43 to 0.83)	111 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,c,d]	1 RCT at high risk of detection and reporting bias suggested that there were**lower risk of failures with the Hawley** removable retainer.
Patient satisfaction Follow-up: 12 months	**Comfort** in upper and lower arches VAS: 0–10 cm (10 cm is very satisfied)	-	-	-	0 (0 RCTs)	-	No studies measured or reported this outcome at 12-month follow-up. In 1 RCT at 6-month follow-up, participants found the **Hawley retainer less comfortable** (VAS MD −1.86 cm, 95% CI −2.19 to −1.53; 86 participants; very low-certainty evidence).
Adverse effects on oral health Follow-up: 12 months	Not measured						

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **LII:** Little's Irregularity Index; **MD:** mean difference; **RCT:** randomised controlled trial; **RR:** risk ratio; **VAS**: visual analogue scale

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level due to study limitations (serious/unclear risk of bias).
[b]Downgraded one level for imprecision due to the small sample size.
[c]One study - difficult to ascertain the likelihood of publication bias.
[d]Downgraded two levels for imprecision due to the very small sample size of a single study.

Summary of findings 6. Removable (full-time) versus removable (part-time) retainers to stabilise tooth position

Population: people who had received fixed appliance treatment
Setting: specialised orthodontic practice and hospital or university orthodontic department
Intervention: full-time wear of any type of removable retainer
Comparison: part-time wear of any type of removable retainer

Follow-up: 12 months

Arch: lower

Outcome	Outcome measure	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of evidence (GRADE)	Comments
		Assumed risk	Corresponding risk				
		Part-time removable retainers	Full-time removable retainers				
Stability Follow-up: 12 months	**Little's Irregularity Index (LII)** in lower arch (ideally 0 mm)	Mean LII was 2 mm	**MD 0.20 mm higher** (0.28 lower to 0.68 higher)	-	52 (1 RCT)	⊕⊖⊖⊖ **Very low**[a,b,c]	1 RCT at high risk of selection bias suggested that there was **no differencebetween full-time and part-time wearing** protocol with Hawley removable retainer.
Survival Follow-up:	Not measured						

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

12 months	
Patient satisfaction Follow-up: 12 months	Not measured
Adverse effects on oral health Follow-up: 12 months	Not measured

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **LII:** Little's Irregularity Index; **MD:** mean difference; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level due to study limitations (serious risk of bias).
[b]Downgraded one level for imprecision due to the small sample size.
[c]One study - difficult to ascertain the likelihood of publication bias.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

BACKGROUND

Description of the condition

Retention is the phase of orthodontic treatment that attempts to keep teeth in the corrected positions after treatment with orthodontic (dental) braces. Without retention there is a tendency for the teeth to return to their initial position. This unfavourable change from the corrected position is known as relapse, which is a lifetime risk. The causes of relapse are not fully understood but are felt to relate to recoil of the fibres that hold the teeth in the jaw bone, pressures from the lips, cheeks and tongue, further growth and the way the teeth meet together (Melrose 1998). Relapse may be more frequent in the lower jaw due to occlusal forces. Sometimes, orthodontists use adjunctive procedures to smooth the contacts between the teeth or to cut the fibres that connect the gum and the tooth, or strategies to increase compliance with removable retainer wear, such as monitoring with mobile phone apps. To minimise relapse, almost every person who has orthodontic treatment will require some type of retention.

Description of the intervention

Retention can be achieved by placing appliances, called retainers, on the teeth, or by undertaking additional or 'adjunctive' procedures to the teeth or the surrounding structures.

The retainers can be fixed to the teeth, or they can be removable, so that the person can take them out to clean (Atack 2007). Bonded ('fixed') retainers are usually glued or 'bonded' on the back (inside) of the front teeth. There are various types of fixed retainers. Some consist of braided or solid metallic wires; others are made of polyethylene and glass-fibres impregnated with composite, which provide better aesthetics. There are different bonding protocols and manufacturing procedures. There are different types of removable retainers too, the Hawley retainer (made of acrylic and metal) and the clear plastic retainers being the most common ones. Some clinicians use a removable appliance called a positioner, that articulates the upper and lower arches, as a retainer after treatment is completed: this helps to 'fine-tune' the result and can then be worn part-time to help reduce relapse.

Although there is no recognised duration that retainers need to be worn, patient compliance with the prescribed wear of a removable retainer is required to maintain correction. It has been shown that if people stop wearing retainers after one to two years, there is a risk of long-term relapse of the teeth (Little 1981; Little 1988). Therefore, some clinicians prefer to retain for longer periods, sometimes indefinitely. It is also not clear how many hours a day removable retainers need to be worn: some people are asked to wear their retainer full-time (24 hours a day), while other people are only asked to wear them part-time.

Clinicians may try to reduce relapse by using 'adjunctive' procedures that make changes to the teeth (hard tissues) or the surrounding gum (soft tissues). The two main adjunctive procedures used are reshaping of the contact points between teeth, which is known as interproximal reduction (Aasen 2005), and cutting the fibres around the neck of the tooth that connect the tooth to the gum, which is known as pericision (Edwards 1988).

Appendix 1 contains a glossary of terms to help describe some of the common types of retainers and adjunctive procedures that are described in this review.

How the intervention might work

Factors such as periodontal disease, bruxism (tooth grinding), tooth loss and ageing can cause teeth to move. Retainers or adjunctive procedures aim to maintain the teeth in the position they are in at the end of orthodontic treatment. Retainers fit over or around the teeth to prevent them moving away from that position. Adjunctive procedures may work by improving the contacts between teeth (interproximal reduction) or by cutting fibres that connect the teeth to the gum around the neck of the teeth to stop them pulling the teeth back towards their original position (pericision).

As well as ensuring teeth remain stable in their new position, retainers must keep the teeth in position without doing any harm to the teeth, gums or mouth, and should be comfortable and acceptable to wear. Retainers have the potential to cause damage to the teeth by collecting plaque and calculus, and by making it difficult for people to keep their teeth clean. This build-up of plaque may cause decay (caries) or damage to the surrounding gum. In turn, damage to the gum can cause gingivitis (inflammation of the gums) or periodontal disease (loss of attachment of the tooth to its surrounding gum and bony socket).

Why it is important to do this review

Unless we can maintain the teeth in position after orthodontic treatment, there is little if any benefit of that treatment. There are currently many different types of removable and fixed retainers, and it is unclear which retainers are the best, and for how long they should be used. There is also little known about the possible benefits and risks of adjunctive procedures aimed at reducing relapse.

It is important to summarise current knowledge about retention procedures in order to provide orthodontists and patients with an up-to-date, reliable evidence base for clinical recommendations.

This review investigated the effectiveness of different retention strategies used to stabilise tooth position after treatment with orthodontic braces. It did not attempt to identify the causes of relapse. This review looked at the effects of retainers while in place, not the long-term effects after they are no longer in use.

This is the third update of a review originally published in 2004 (Littlewood 2004; Littlewood 2006; Littlewood 2016).

OBJECTIVES

To evaluate the effects of different retainers and retention strategies used to stabilise tooth position after orthodontic braces.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs)

Types of participants

Children and adults who have had retainers fitted or adjunctive procedures undertaken following treatment with orthodontic braces. There was no restriction for the presenting malocclusion or type of active orthodontic treatment undertaken. The participants

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

had to be followed up at least three months after completing their orthodontic treatment.

We excluded:

- people who had surgical correction of the jaws;
- people with a cleft lip or palate, or both, or other craniofacial syndrome;
- people who had orthodontic treatment based on extractions alone or the fitting of a passive space maintainer, or both.

Types of interventions

Retainers or adjunctive procedures, or both, after treatment with orthodontic braces. We included only studies where the full course of definitive orthodontic treatment was completed - we excluded data on retention strategies at the end of an initial phase of removable or functional appliance treatment. We excluded studies comparing techniques (not materials) for placing the bonded retainers.

We did not include studies on aligners in this review.

Types of outcome measures

We considered four types of outcomes: stability (straightness) outcomes, failure of retainers, adverse effects of retainers, and patient-related outcome measures (PROMs) focused on patient satisfaction.

Primary outcomes

Stability

Stability could be assessed by an index of tooth irregularity, for example, Little's Irregularity Index (Little 1981), which measures how crooked anterior teeth are, or by crowding (see glossary of terms in Appendix 1). It can also be assessed by a change in the shape or size of each arch: this can be measured by intercanine width, intermolar width or arch length. The way the teeth meet together (occlusion) can also be assessed using measurements such as overjet, overbite, and assessing the quality of the final results using an index such as the Peer Assessment Rating (PAR index; Richmond 1992) or the American Board of Orthodontics objective grading system (ABO-OGS; Casko 1998). See glossary of terms in Appendix 1.

This review gave priority to reporting the Little's Irregularity Index. There is no agreed Core Outcome Measures in Effectiveness Trials (COMET) for stability studies in orthodontics at the present time, so the review authors decided to use Little's Irregularity Index for the preferred outcome measure in this review.

Settling of the occlusion was excluded as an outcome in this update since it is not directly related to stability. The assessment of stability had to be made at least three months after the fitting of the retainer or after the adjunctive procedure was carried out, or after both.

Secondary outcomes

Failure of the retainers

This assessed how long retainers lasted without breaking (in months) and how many times they needed to be replaced or repaired. If retainers were lost, we reported this as a failure, since the retainers could not fulfil their role.

Patient satisfaction

This often takes the form of a questionnaire.

Adverse effects

These were on the health of the:

- teeth (in terms of decay), assessed using indices of demineralisation or identifying the presence of caries;
- surrounding structures (gums and other supporting structures), assessed using periodontal indices or other markers of periodontal disease.

Search methods for identification of studies

The review authors, in collaboration with Cochrane Oral Health's Information Specialist, developed detailed search strategies to search for RCTs for this review.

Electronic searches

Cochrane Oral Health's Information Specialist conducted systematic searches in the following databases for relevant studies. There were no language, publication year or publication status restrictions:

- Cochrane Oral Health Trials Register (to 27 April 2022; see Appendix 2);
- Cochrane Central Register of Controlled Trials (CENTRAL; 2022, Issue 3) in the Cochrane Library (searched 27 April 2022; see Appendix 3);
- MEDLINE Ovid (1946 to 27 April 2022; see Appendix 4);
- Embase Ovid (1980 to 27 April 2022; see Appendix 5);
- OpenGrey (opengrey.eu; see Appendix 6).

Subject strategies were modelled on the search strategy designed for MEDLINE Ovid. Where appropriate, they were combined with subject strategy adaptations of the highly sensitive search strategies designed by Cochrane for identifying RCTs and controlled clinical trials (as described in the *Cochrane Handbook for Systematic Reviews of Interventions*, Version 6.3 (Lefebvre 2022)).

Searching other resources

Ongoing studies

Cochrane Oral Health's Information Specialist searched the following trials registries for ongoing studies:

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov; searched 27 April 2022; Appendix 7);
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched 27 April 2022; Appendix 8).

Checking reference lists

We checked the bibliographies of included studies and review articles that we identified for further eligible studies.

Handsearching

We handsearched conference proceedings and abstracts from the British Orthodontic Conference, the European Orthodontic Conference and the International Association for Dental Research

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

(IADR) from 2016 to 2022. We handsearched the British (BJO) and European (EJO) Journals of Orthodontics, American Journal of Orthodontics and Dentofacial Orthopedics (AJODO), the Orthodontics and Craniofacial Research; The Angle Orthodontist and Progress in Orthodontics from 2016 to 2022.

Personal communication

We contacted the authors of ongoing studies regarding the status of their studies and their willingness to share unpublished results for inclusion in this review. We contacted authors of identified RCTs where we required additional information relevant to the review that was not apparent in the published work. We asked all study authors contacted if they knew of any other published or unpublished studies relevant to the review that we had not identified. In addition, the authors of this review act as peer reviewers for many orthodontic publications so would likely be aware of any other studies. If any study submitted for review appeared to fulfil the criteria, we contacted the study authors for further details.

We did not perform a separate search for adverse effects of interventions used; we considered adverse effects described in included studies only.

Data collection and analysis

Selection of studies

For this update, two review authors (CM and AL) independently and in duplicate carried out the selection of papers and made decisions about eligibility using Covidence software. We solved any disagreement by discussion with a third review author (SL). We checked that none of the included studies in this review were retracted due to error or fraud.

When studies had multiple publications, we collated the reports of the same study so that each study, rather than each report (publication), was the unit of interest for the review, and such studies had a single identifier (based on the main publication) with multiple references. We did not impose any language or publication restrictions (published, unpublished, and in progress). If additional information was required, we contacted the author directly. We contacted authors of ongoing studies to confirm status of the study. We documented the selection process with a PRISMA flow chart in accordance with the PRISMA 2020 statement guidelines (Page 2021a).

Data extraction and management

Six review authors working in pairs (from SL, BD, DM, DB, CM, AL) performed data extraction in duplicate using a customised data collection form that included the following information: citation details of publication, if applicable; summary of the study design; participants (sample size, age, inclusion and exclusion criteria, setting of study, costs involved to participants); interventions - type of retainer, daily wearing time, type of adjunctive procedure; prescribed and actual duration of retention; quality assessment of paper (see Assessment of risk of bias in included studies); outcomes, outcome measures, and time points for measurements; primary and secondary outcome data; any additional information that could affect the assessment of the study.

Two authors (CM and AL) entered data into Review Manager Web (RevMan Web 2022). We cross-checked that the data had been entered correctly by comparing the data presented in the systematic review with those in the standardised data extraction form, which we had pre-piloted for the purpose of the review. Authors of the previous version of this review used Review Manager 5 (Review Manager 2020) to enter the data.

Assessment of risk of bias in included studies

Six review authors (SL, BD, DM, DB, CM, AL), working in pairs, used the Cochrane risk of bias tool (RoB 1) to assess the risk of bias in the studies in duplicate (Higgins 2011). We assessed the following domains: random sequence generation; allocation concealment; blinding of outcome assessors; incomplete outcome data reporting; selective outcome reporting; any other sources of bias.

Blinding of participants and clinicians is usually not possible in retention research due to the nature of the intervention. This is because it is not possible to hide either the retainer type or the adjunctive procedure. Therefore, although this puts all studies at high risk of bias for this domain, we did not include blinding of participants and personnel, that is, performance bias, in the risk of bias assessment for each study.

We assessed a study as low risk of bias overall if all the domains we assessed were judged as low risk, unclear if at least one domain was unclear, and high if at least one domain was high risk.

Measures of treatment effect

We analysed the outcomes based on intention-to-treat (ITT) analyses, where possible. For studies considered eligible for this review, we used the following analyses in line with Cochrane guidance.

For dichotomous outcomes, we expressed the estimate of effect of an intervention as risk ratios (RR), or, in case of outcomes with zero events, risk differences (RD), with corresponding 95% CIs.

For time-to-event outcomes (i.e. time to the first episode of failure), we used hazard ratios (HRs) with 95% CIs.

For continuous outcomes, we used mean differences (MD) and standard deviations (SD) to summarise the data for each group with corresponding 95% CI. Where studies reported continuous data as nonparametric data using median and interquartile range (IQR) and data passed the test of skewness, we converted median to mean and estimated the SD as IQR/1.35 as per guidance from Chapter 6 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2022). We used change-from-baseline outcomes, considering baseline as the end of active orthodontic treatment, to final position. If change data were not available, the change was calculated, or final irregularity scores were used.

Unit of analysis issues

The unit of analysis was the participant.

Dealing with missing data

We contacted study authors to obtain missing data via email and gave four weeks to reply to the request. If the corresponding authors did not respond, we attempted to contact them a second time. If we did not obtain a response, we did not include the study data in the relevant meta-analysis; we did not impute data.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Assessment of heterogeneity

We assessed the significance of any discrepancies in the estimates of the treatment effects from the different studies by means of Cochrane's test for heterogeneity and considered heterogeneity significant if the P value was more than 0.10. We used the I^2 statistic (Higgins 2003), which describes the percentage of total variation across studies that is due to heterogeneity rather than chance, with I^2 over 50% being considered moderate to high heterogeneity (Deeks 2022).

Assessment of reporting biases

Only a proportion of research projects are ultimately published in an indexed journal and become easily identifiable for inclusion in a systematic review. We investigated and attempted to minimise potential reporting biases, including publication bias or language bias, by avoiding language restrictions.

If there had been a sufficient number of studies (more than 10) in any meta-analysis, we would have assessed publication bias according to the recommendations on testing for funnel plot asymmetry described in Page 2022. If we had identified asymmetry, we would have examined possible causes.

Data synthesis

We used RevMan Web to perform statistical analysis (RevMan Web 2022). When there were studies of similar comparisons reporting the same outcome measures, we combined them in meta-analyses. We used the Mantel-Haenszel method for dichotomous outcomes to combine RRs and RDs, and the generic inverse-variance method for time-to-event outcomes and continuous outcomes to combine MDs. We used the fixed-effect model of meta-analysis.

For the patient satisfaction outcome, we grouped patient feedback from questionnaires into core groups (e.g. efficiency, speech, comfort) to ease data analysis. If we could not combine data due to differences between interventions or time points, we presented the findings in a narrative manner.

The longest follow-up time was considered in the statistical analysis and in the results when studies reported different follow-up times, unless two or more studies reported shorter follow-up times, in which case pooled analyses were performed.

Subgroup analysis and investigation of heterogeneity

We assessed clinical heterogeneity by examining the types of participants and interventions for all outcomes in each study. We did not formulate any hypotheses to be investigated by subgroup analyses.

Sensitivity analysis

We planned to undertake sensitivity analyses to examine the effect of risk of bias on the overall estimates of effect, by removing studies that were at unclear or high risk of bias from the analysis. We also planned to examine the effect of including unpublished literature on the review's findings. However, there were too few studies measuring the same comparison to undertake these analyses.

Summary of findings and assessment of the certainty of the evidence

We produced a summary of findings table for each comparison using GRADEpro GDT software. We assessed the certainty of the body of evidence with reference to the overall risk of bias in the included studies, directness of the evidence, consistency of results, precision of estimates, and risk of publication bias. We categorised the certainty of the body of the evidence for each outcome as high, moderate, low or very low. We produced summary of findings tables for the main comparisons and outcomes in this review.

RESULTS

Description of studies

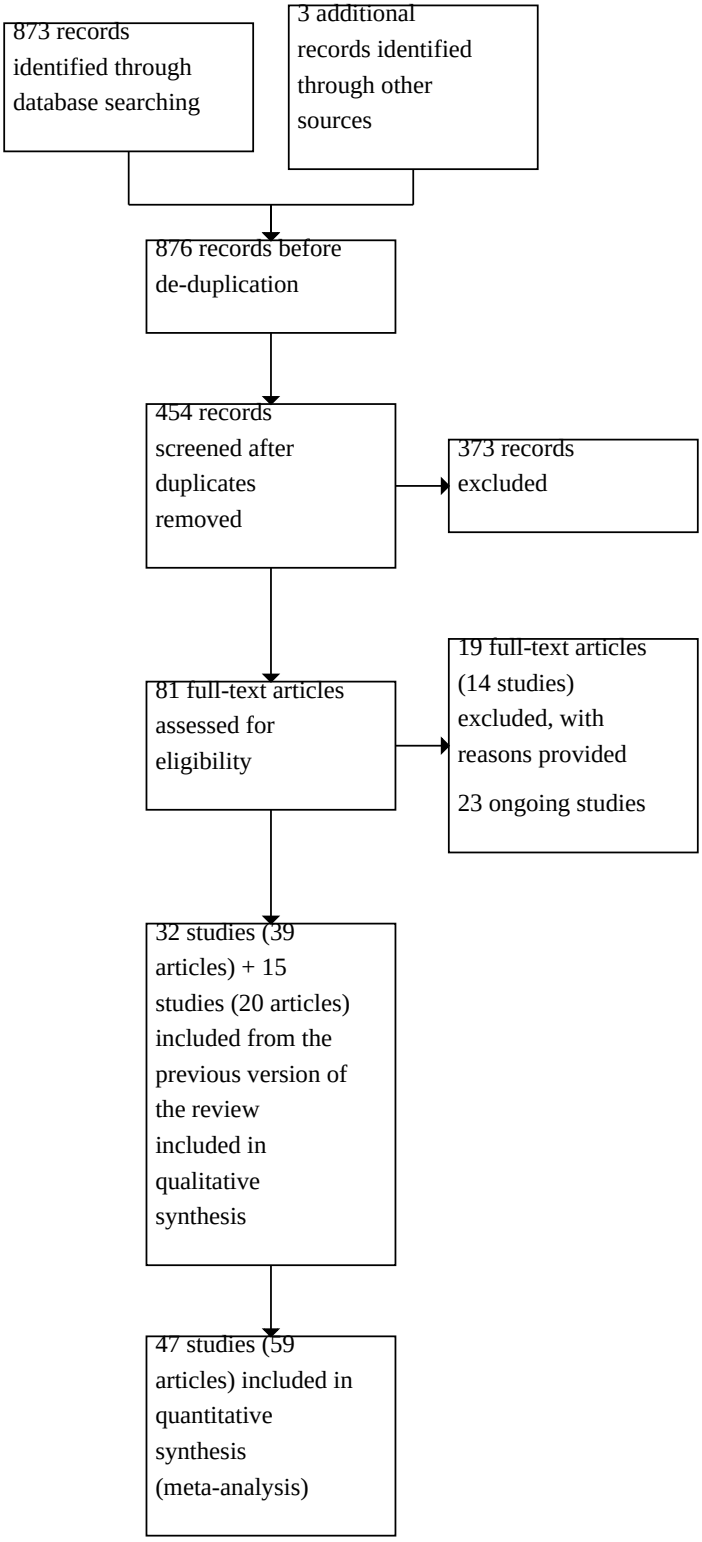
See table of Characteristics of included studies, Characteristics of excluded studies, and Characteristics of ongoing studies.

Results of the search

For this update, we re-ran searches up to 27 April 2022, and identified 873 new records, of which 422 were duplicates. We also found three additional studies through examining reference lists of included studies, review articles, conference proceedings of the above-mentioned organisations and orthodontic-focused journals. We rejected 373 irrelevant records during title and abstract screening, and excluded 14 studies (19 records) after full-text assessment, documenting our reasons in the Characteristics of excluded studies. We found 23 new ongoing studies (see Characteristics of ongoing studies), and 32 new completed studies (390 records). We found one article relating to a study (O'Rourke 2016), that had been included in the previous version of the review (Littlewood 2016). See Figure 1.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 1. Study flow diagram for studies identified for this update



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 1. (Continued)

We retained 14 included studies (20 reports) from the previous published version of the review (Littlewood 2016), and decided to include one study (Pandis 2013), that had been excluded in the last version as it assessed two types of bonding systems, but we considered the adhesives to be part of the retention, so we included it in this update. Conversely, Aslan 2013 had been included in the previous version of the review but we excluded it in this update as it used the outcome "settling of the occlusion" as a measure of stability. Of the four studies categorised as ongoing in the previous version of the review, we excluded two (ChiCTR-TRC-07000055; ChiCTR-TRC-07000454); one (Forde 2018) had been published and was included in this update; and one is still ongoing (IRCT2016022926836N1). Of the four studies categorised as 'awaiting classification' in the previous version of the review, we included one (Torkan 2014) and excluded three (Choi 2010; Liu 2010; Xu 2011).

Therefore, in total, this updated review has 47 included studies (59 reports), 49 excluded studies (54 reports) and 24 ongoing studies.

Included studies

Characteristics of study design

We included 47 RCTs (4377 participants) in this review (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Arash 2020; Årtun 1987; Ashari 2021; Bolla 2012; Edman Tynelius 2010; Eroglu 2019; Ferreira 2019; Forde 2018; Gelin 2020; Gera 2022; Gill 2007; Gunay 2018; Jowett 2022; Kartal 2021; Kaya 2019; Krämer 2019; Kumar 2011; Millett 2007; Moslemzadeh 2018; Nagani 2020; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Pandis 2013; Ramazanzadeh 2018; Rohaya 2006; Rose 2002; Rowland 2007; Saleh 2017; Salehi 2013; Scribante 2011; Scribante 2020; Sfrondini 2014; Sfrondini 2021; Shawesh 2010; Shim 2022; Sinha 2021; Soboutia 2016; Sun 2011; Thickett 2010; Torkan 2014; Wan 2017; Węgrodzka 2021; Zhu 2017). Forty-seven studies were parallel-design studies and one was a cross-over clinical study (Ferreira 2019). Thirty-eight studies were two-arm studies, seven were three-arm studies (Edman Tynelius 2010; Eroglu 2019; Moslemzadeh 2018; Naraghi 2021b; Ramazanzadeh 2018; Shim 2022; Soboutia 2016), and three were four-arm studies (Adanur-Atmaca 2021; Alrawas 2020; Årtun 1987).

Ten studies were reported in more than one publication. Eight studies each provided data in two articles that published different outcomes or follow-ups (Al-Moghrabi 2020; Årtun 1987; Forde 2018; Krämer 2019; Nagani 2020; Naraghi 2021b; O'Rourke 2016; Rowland 2007). Data from Millett 2007 were gathered from three published research abstracts, and data from Edman Tynelius 2010 were collected from three articles reporting outcomes at different time intervals. We also received data via personal email communication for 15 of the studies (Årtun 1987; Ashari 2021; Bolla 2012; Edman Tynelius 2010; Gera 2022; Gill 2007; Jowett 2022; Krämer 2019; Millett 2007; O'Rourke 2016; Pandis 2013; Rohaya 2006; Rose 2002; Saleh 2017; Thickett 2010).

Characteristics of the study settings and investigators

Thirty-four studies were undertaken in a hospital or university setting (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Ashari 2021; Eroglu 2019; Ferreira 2019; Forde 2018; Gelin 2020; Gera 2022; Gill 2007; Gunay 2018; Kartal 2021; Kaya 2019; Krämer 2019; Kumar 2011; Millett 2007; Nagani 2020; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Ramazanzadeh 2018; Rohaya 2006; Rose 2002; Scribante 2011; Scribante 2020; Sfrondini 2014; Sfrondini 2021; Shim 2022; Sinha 2021; Sun 2011; Thickett 2010; Torkan 2014; Wan 2017; Zhu 2017), 11 in a specialist practice (Arash 2020; Årtun 1987; Bolla 2012; Moslemzadeh 2018; Pandis 2013; Rowland 2007; Saleh 2017; Salehi 2013; Shawesh 2010; Soboutia 2016; Węgrodzka 2021), one in a NHS (UK National Health Service) clinic (Edman Tynelius 2010), and one in a multicentre setting combining specialist practice with hospital and university settings (Jowett 2022).

There were eight studies from the UK (Al-Moghrabi 2020; Forde 2018; Gill 2007; Jowett 2022; O'Rourke 2016; Rowland 2007; Shawesh 2010; Thickett 2010); six from Turkey (Adanur-Atmaca 2021; Alrawas 2020; Eroglu 2019; Gunay 2018; Kartal 2021; Kaya 2019), and six from Iran (Arash 2020; Moslemzadeh 2018; Ramazanzadeh 2018; Salehi 2013; Soboutia 2016; Torkan 2014); five from Italy (Bolla 2012; Scribante 2011; Scribante 2020; Sfrondini 2014; Sfrondini 2021); four from Sweden (Edman Tynelius 2010; Krämer 2019; Naraghi 2021a; Naraghi 2021b); three from China (Sun 2011; Wan 2017; Zhu 2017), and two from Malaysia (Ashari 2021; Rohaya 2006); two each from India (Kumar 2011; Sinha 2021) and USA (Årtun 1987; Shim 2022); and one each from Belgium (Gelin 2020), Brazil (Ferreira 2019), Germany (Rose 2002), Pakistan (Nagani 2020), Poland (Węgrodzka 2021), the Republic of Ireland (Millett 2007), Saudi Arabia (Saleh 2017), and Switzerland (Pandis 2013). The final study was a multicentre one from Denmark and Norway (Gera 2022).

Orthodontists provided the care for all the participants in the studies. Nineteen studies stated there was one operator (Ashari 2021; Bolla 2012; Edman Tynelius 2010; Ferreira 2019; Gera 2022; Gunay 2018; Kartal 2021; Krämer 2019; Millett 2007; Nagani 2020; Naraghi 2021a; Pandis 2013; Ramazanzadeh 2018; Rowland 2007; Scribante 2020; Sfrondini 2014; Sfrondini 2021; Shim 2022; Węgrodzka 2021); three studies stated they had two operators (Årtun 1987; Forde 2018; Gelin 2020); two studies stated they had three operators (Moslemzadeh 2018; Naraghi 2021b); one study stated they had 10 calibrated operators (Jowett 2022); and the remainder did not disclose the number of operators. Twelve studies confirmed that the retainers were fabricated by the same qualified orthodontic technician (Alrawas 2020; Arash 2020; Ferreira 2019; Gera 2022; Gunay 2018; Kartal 2021; Krämer 2019; O'Rourke 2016; Ramazanzadeh 2018; Tahir 2019; Wan 2017; Zhu 2017).

To cover the economic costs involved in studies, eight studies declared external funding sources (Al-Moghrabi 2020; Edman Tynelius 2010; Gera 2022; Gill 2007; Jowett 2022; Sfrondini 2014; Sun 2011; Wan 2017); eight studies declared internal funding sources (Alrawas 2020; Ashari 2021; Eroglu 2019; Kartal 2021; Krämer 2019; Ramazanzadeh 2018; Saleh 2017; Soboutia 2016); two studies declared internal and external funding sources (Naraghi 2021a; Naraghi 2021b); six studies declared that there was no source of funding (Arash 2020; Kaya 2019; Moslemzadeh 2018;

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nagani 2020; Scribante 2020; Sfrondini 2021); and the remainder did not disclose any source of funding.

Characteristics of the participants

There were between 15 (Ferreira 2019), and 397 (Rowland 2007), participants in the 47 studies. The studies were undertaken on both children and adults. The age range of participants across the studies was between 11 and 55 years. In 24 studies, the mean age of participants was under 18 years (Adanur-Atmaca 2021; Al-Moghrabi 2020; Edman Tynelius 2010; Eroglu 2019; Forde 2018; Gelin 2020; Gera 2022; Gill 2007; Gunay 2018; Kaya 2019; Krämer 2019; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Pandis 2013; Ramazanzadeh 2018; Rowland 2007; Scribante 2011; Shawesh 2010; Shim 2022; Sinha 2021; Sun 2011; Thickett 2010; Torkan 2014); in 17 studies the mean age was over 18 years (Alrawas 2020; Arash 2020; Ashari 2021; Bolla 2012; Ferreira 2019; Moslemzadeh 2018; Nagani 2020; Rose 2002; Saleh 2017; Salehi 2013; Scribante 2020; Sfrondini 2014; Sfrondini 2021; Soboutia 2016; Wan 2017; Węgrodzka 2021; Zhu 2017); and six studies did not state the mean age (Årtun 1987; Jowett 2022; Kartal 2021; Kumar 2011; Millett 2007; Rohaya 2006).

Characteristics of the interventions

Four studies compared more than one intervention (Alrawas 2020; Årtun 1987; Eroglu 2019; Naraghi 2021b). Therefore, the total number of studies according to the type of intervention (n = 51) is larger than the total number of included studies (n = 47).

Eight studies compared removable retainers with fixed retainers: four studies had two arms (Forde 2018; Krämer 2019; Millett 2007; O'Rourke 2016), and four studies had at least three arms and included other types of comparisons (Alrawas 2020; Årtun 1987; Eroglu 2019; Naraghi 2021b).

Twenty-two studies compared different types of fixed retainers: 17 with two arms (Arash 2020; Bolla 2012; Ferreira 2019; Gelin 2020; Gera 2022; Gunay 2018; Jowett 2022; Kartal 2021; Nagani 2020; Naraghi 2021b; Rose 2002; Salehi 2013; Scribante 2011; Sfrondini 2014; Sinha 2021; Torkan 2014; Węgrodzka 2021), four with three arms (Alrawas 2020; Årtun 1987; Shim 2022; Soboutia 2016), and one with four arms (Adanur-Atmaca 2021).

Three studies compared the same type of retainers but different bonding materials (Pandis 2013; Scribante 2020; Sfrondini 2021).

Sixteen studies compared one type of removable retainer with another (Al-Moghrabi 2020; Ashari 2021; Eroglu 2019; Gill 2007; Kaya 2019; Kumar 2011; Moslemzadeh 2018; Ramazanzadeh 2018; Rohaya 2006; Rowland 2007; Saleh 2017; Shawesh 2010; Sun 2011; Thickett 2010; Wan 2017; Zhu 2017).

One study with three arms compared a bonded mandibular canine-to-canine retainer versus mandibular anterior stripping versus positioner (Edman Tynelius 2010).

One study compared the use of retainers in cases of impacted canines versus no retention procedure (Naraghi 2021a).

Removable versus fixed retainers

Studies that compared removable versus fixed retainers made the following comparisons.

- Clear plastic versus multistrand (Alrawas 2020; Eroglu 2019; Forde 2018; Millett 2007; Naraghi 2021b; O'Rourke 2016)
- Clear plastic versus single-strand (Alrawas 2020; Krämer 2019)
- Hawley versus multistrand (Eroglu 2019)
- Removable (specific type not defined) versus thick plane stainless steel, thick spiral stainless steel and thin flexible stainless steel (Årtun 1987)

Fixed retainers

Studies that compared fixed retainers made the following comparisons.

- Fibre-reinforced composite versus multistrand stainless steel (Bolla 2012; Nagani 2020; Rose 2002; Salehi 2013; Scribante 2011; Torkan 2014);
- Fibre-reinforced composite versus spiral wire (Sfrondini 2014)
- Fibre-reinforced composite versus spiral wire versus twisted wire (Soboutia 2016)
- Multistrand stainless steel versus rigid retainer bonded only to canines (Ferreira 2019)
- Multistrand stainless steel versus single-strand (Alrawas 2020)
- Multistrand stainless steel versus thick plain titanium wire (Arash 2020)
- Multistrand stainless steel versus CAD/CAM (computer-aided design/computer-aided manufacturing) nitinol (Adanur-Atmaca 2021; Alrawas 2020; Gelin 2020; Gera 2022; Kartal 2021)
- CAD/CAM stainless steel multistrand versus lab stainless steel multistrand versus chairside rectangular chain (Shim 2022)
- Single-strand stainless steel versus CAD/CAM nitinol (Alrawas 2020)
- Rectangular chain stainless steel versus CAD/CAM nitinol (Jowett 2022)
- Rectangular braided multistrand versus round stainless-steel multistrand versus connected bonding pad retainer versus CAD/CAM nitinol (Adanur-Atmaca 2021)
- Multistrand stainless steel versus dead-soft coaxial retainers (Gunay 2018)
- Thick plane stainless steel versus thick spiral stainless steel versus thin spiral stainless steel (Årtun 1987)
- Ceramic interlocking bonded versus coaxial wire (Sinha 2021)
- Round twisted multistrand versus rectangular braided multistrand (Węgrodzka 2021)
- Different procedures or materials associated with multistrand stainless steel retainers: upper from 2 to 2 versus 3 to 3 (Naraghi 2021b)
- Different types of adhesives used to place bonded retainers: two types of composites (Scribante 2020), two types of primers (Sfrondini 2021)

Removable retainers

Studies that compared removable retainers made the following comparisons.

- Part-time versus full-time wear of clear plastic retainers (Gill 2007; Thickett 2010)
- Different thicknesses of clear plastic retainers (Zhu 2017)
- Clear plastic retainers with and without use of a mobile app (Al-Moghrabi 2020)

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

- Begg versus clear plastic retainers, with bonded retainer in lower arch in both groups (Kumar 2011)
- Hawley versus clear plastic retainers (Ashari 2021; Eroglu 2019; Kaya 2019; Moslemzadeh 2018; Ramazanzadeh 2018; Rohaya 2006; Rowland 2007; Saleh 2017; Sun 2011; Wan 2017)
- Part-time versus full-time wear of Hawley retainers (Shawesh 2010)

Characteristics of the outcomes

With the data available, it was possible to analyse the following outcomes.

- Tooth stability (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Ashari 2021; Edman Tynelius 2010; Forde 2018; Gelin 2020; Gera 2022; Gill 2007; Gunay 2018; Jowett 2022; Kaya 2019; Krämer 2019; Kumar 2011; Millett 2007; Moslemzadeh 2018; Nagani 2020; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Ramazanzadeh 2018; Rohaya 2006; Rowland 2007; Shawesh 2010; Shim 2022; Sinha 2021;; Thickett 2010)
- Survival or failure of retainers (Arash 2020; Årtun 1987; Bolla 2012; Forde 2018; Gelin 2020; Gera 2022; Gunay 2018; Jowett 2022; Kartal 2021; Krämer 2019; Millett 2007; Nagani 2020; O'Rourke 2016; Pandis 2013; Rose 2002; Rowland 2007; Salehi 2013; Scribante 2011; Scribante 2020; Sfrondini 2014; Sfrondini 2021; Shim 2022; Sinha 2021; Soboutia 2016; Sun 2011; Węgrodzka 2021; Zhu 2017)
- Participant satisfaction (Al-Moghrabi 2020; Ferreira 2019; Forde 2018; Gelin 2020; Gera 2022; Jowett 2022; Krämer 2019; Kumar 2011; Millett 2007; Rowland 2007; Saleh 2017; Scribante 2011; Wan 2017; Zhu 2017)
- Adverse effects on oral health (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Årtun 1987; Eroglu 2019; Ferreira 2019; Forde 2018; Gelin 2020; Kartal 2021; Millett 2007; Moslemzadeh 2018; O'Rourke 2016; Sinha 2021; Torkan 2014; Węgrodzka 2021)

Studies assessed tooth stability using different measurements.

- Little's Irregularity Index (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Edman Tynelius 2010; Forde 2018; Gill 2007; Gunay 2018; Jowett 2022; Kaya 2019; Krämer 2019; Kumar 2011; Millett 2007; Nagani 2020; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Ramazanzadeh 2018; Rowland 2007; Shawesh 2010; Shim 2022; Sinha 2021; Thickett 2010)
- Intercanine, interpremolar, intermolar width (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Edman Tynelius 2010; Forde 2018; Gill 2007; Gunay 2018; Jowett 2022; Kaya 2019; Krämer 2019; Kumar 2011; Millett 2007; Nagani 2020; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Ramazanzadeh 2018; Rowland 2007; Shawesh 2010; Shim 2022; Sinha 2021; Thickett 2010)
- Anterior or total arch length (Adanur-Atmaca 2021; Alrawas 2020; Edman Tynelius 2010; Forde 2018; Gelin 2020; Gera 2022; Kaya 2019; Krämer 2019; Naraghi 2021a; Naraghi 2021b; O'Rourke 2016; Ramazanzadeh 2018; Thickett 2010)
- Overjet and overbite (Edman Tynelius 2010; Forde 2018; Gill 2007; Kaya 2019; Krämer 2019; Naraghi 2021b; Thickett 2010)
- Maximum rotation (Naraghi 2021a; Naraghi 2021b; Rohaya 2006)
- Peer Assessment Rating (PAR) (Kumar 2011; Thickett 2010)
- American Board of Orthodontics (ABO) grading (Moslemzadeh 2018)

There is no agreed COMET for stability studies in orthodontics at the present time, so the review authors decided to use Little's Irregularity Index as the preferred outcome measure in this review in terms of the summary of findings tables and conclusions.

Studies assessed failure and survival of retainers by the following.

- Failure rate, including:
  - bonded retainer was partially or completely detached from the teeth, or fractured (Arash 2020; Årtun 1987; Bolla 2012; Gelin 2020; Gera 2022; Gunay 2018; Jowett 2022; Kartal 2021; Millett 2007; Nagani 2020; Pandis 2013; Rose 2002; Salehi 2013; Scribante 2011; Scribante 2020; Sfrondini 2014; Sfrondini 2021; Shim 2022; Sinha 2021; Soboutia 2016; Węgrodzka 2021);
  - removable retainer was lost, fractured, no longer fitted, or showed excessive wear (Årtun 1987; Forde 2018; Krämer 2019; Millett 2007; Rowland 2007; Sun 2011; Zhu 2017).
- Mean survival time (months/days) (Arash 2020; Soboutia 2016; Zhu 2017)
- Failure risk (hazard ratio, Kaplan Meier survival plots) (Gera 2022; Jowett 2022; Pandis 2013; Scribante 2011; Sfrondini 2014; Soboutia 2016; Sun 2011; Węgrodzka 2021)

Studies assessed patient satisfaction using questionnaires. We gathered similar questions into the following core categories.

- **Acceptance** for: "how acceptable was the retainer to wear?" (Millett 2007); “have you gotten used to retainers?” (Krämer 2019); "are you embarrassed to wear your retainer?" (Forde 2018); “how have you found wearing your retainer in comparison to your fixed appliances?" (Forde 2018); "appearance” (Saleh 2017); "self-confidence" (Saleh 2017); "embarrassed to wear retainer?" (Rowland 2007); and "worse or much worse than wearing fixed appliances?" (Rowland 2007)
- **Compliance** for: "do you wear your retainers as instructed?" (Forde 2018); “able to wear retainer as instructed?" (Rowland 2007); and "able to wear retainers away from home?” (Rowland 2007)
- **Durability** for: “durability” (Saleh 2017)
- **Efficiency** for: “are the retainers working well?" (Krämer 2019); "noticed any change in tooth position?" (Krämer 2019); "did your retainer keep your teeth straight?" (Jowett 2022)
- **Cleaning or hygiene** for: "how easy was it to keep the retainer clean?" (Millett 2007); "how have you found easy cleaning your retainers?” (Forde 2018); "was your retainer easy to look after?" (Jowett 2022); “which retainer allowed better cleaning?" (Ferreira 2019); "hygiene” (Saleh 2017); and "gingival irritation" (Saleh 2017)
- **Comfort, discomfort or pain** for: "discomfort/pain” (Krämer 2019); "was your retainer easy to look after/ comfortable?” (Jowett 2022); “which retainer was more comfortable in use?” (Ferreira 2019); "comfort" (Saleh 2017); "teeth biting: closing teeth with retainers, not chewing food” (Saleh 2017); "fitting" (Saleh 2017); "swallowing" (Saleh 2017); “soreness” (Krämer 2019); and “tension in your teeth” (Krämer 2019)

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

- **Speech** for: "has your retainer affected your speech?" (Forde 2018); "did your retainer affect speech?" (Jowett 2022); “speech” (Saleh 2017); and "number of sound distortions" (Wan 2017).
- **Appointment needed** for: "did your retainer cause a problem that meant you needed to see your orthodontist?" (Jowett 2022).
- **Aesthetics** for: “esthetic result while talking and smiling” (Scribante 2011).
- **Preference for type of retainer** for: “preference for type of retainer" (Ferreira 2019).

Studies reported outcomes at the following intervals.

- Three months (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Ashari 2021; Eroglu 2019; Forde 2018; Gunay 2018; Kartal 2021; Kumar 2011; Nagani 2020; Pandis 2013; Saleh 2017; Shim 2022; Sinha 2021; Wan 2017; Węgrodzka 2021; Zhu 2017).
- Four months (Ramazanzadeh 2018)
- Six months (Adanur-Atmaca 2021; Al-Moghrabi 2020; Alrawas 2020; Ashari 2021; Ferreira 2019; Forde 2018; Gelin 2020; Gera 2022; Gill 2007; Gunay 2018; Jowett 2022; Kartal 2021; Krämer 2019; Kumar 2011; Moslemzadeh 2018; Nagani 2020; O'Rourke 2016; Pandis 2013; Rowland 2007; Saleh 2017; Shim 2022; Sinha 2021; Torkan 2014; Węgrodzka 2021; Zhu 2017)
- Eight months (Ramazanzadeh 2018)
- Nine months (Adanur-Atmaca 2021; Al-Moghrabi 2020; Gunay 2018; Nagani 2020; Rohaya 2006)
- One year (Adanur-Atmaca 2021; Al-Moghrabi 2020; Ashari 2021; Edman Tynelius 2010; Forde 2018; Gelin 2020; Gunay 2018; Kaya 2019; Millett 2007; Nagani 2020; Naraghi 2021a; O'Rourke 2016; Pandis 2013; Scribante 2011; Sfrondini 2014; Shawesh 2010; Sun 2011; Thickett 2010; Węgrodzka 2021; Zhu 2017). One study (Scribante 2011) reported survival data every month over a 12-month period.
- Eighteen months (Krämer 2019; O'Rourke 2016; Salehi 2013; Węgrodzka 2021).
- Two years (Edman Tynelius 2010; Naraghi 2021b; Pandis 2013; Rose 2002; Scribante 2020; Sfrondini 2021; Soboutia 2016; Węgrodzka 2021). One study (Arash 2020) reported survival data every three months during the 24-month follow-up.
- Three years (Årtun 1987)
- Four years (O'Rourke 2016)
- Five years (Edman Tynelius 2010), two years with retention and three years without retention.
- Six years (Bolla 2012)

Some of the results were provided without standard deviations or other variability measures, or with data that were not amenable to being meta-analysed.

Excluded studies

See Characteristics of excluded studies. Of the 49 excluded studies:

- 19 were not RCTs (Cunning 2022; Al-Nimri 2009; Arora 2014; Atack 2007; Axelsson 1993; Barlin 2011; Basciftici 2007; Bauer 2010; Bock 2008; Brignardello-Petersen 2018; Edwards 1988; Jäderberg 2012; Knaup 2019; Lindauer 1998; Pandis 2007; Sauget 1997; Tacken 2010; Taner 2000; Vecere 1983);
- five had less than three months' follow-up (Ahrens 1981; Carvalho 2006; Choi 2010; Horton 2009; Pandian 2019);
- three investigated methods of placing or supervising retainers rather than comparing different retainers (Bazargani 2012; Bovali 2014; Conway 2011);
- three studies assessed settling, which is not considered as an outcome in this version of our review (Alkan 2020; Aslan 2013; Kara 2020);
- three had insufficient data to analyse (Haydar 1992; Larsson 1983; Stormann 2002);
- two compared fixed retainer placement methods, which is outside the scope of this review (Alaa ElDeen 2021; Bovali 2014);
- two focused on outcomes outside the scope of this review (Sohrabi 2014; Vagdouti 2019).
- one was plagiarised (Jain 2021);
- one used orthodontic treatment with aligners, which is outside the scope of this review (Sangalli 2022);
- one lacked correspondence data to request the full text published in Chinese (Liu 2010).
- one compared two similar clear plastic retainers not placed at the same time after debonding (Tahir 2019).

Eight of the excluded studies were protocols:

- three recruited insufficient participants to complete the study (ACTRN12612000670875; ISRCTN22535947; ISRCTN26364810);
- two had been registered more than 15 earlier, with no response from the study authors and no evidence of study completion or publication (ChiCTR-TRC-07000055; ChiCTR-TRC-07000454);
- one investigated methods of placing or supervising retainers rather than comparing different retainers (ISRCTN56295329);
- one was not an RCT (Cunning 2022);
- one included periodontal-compromised patients (NCT03765151).

Risk of bias in included studies

We judged 11 studies to have overall low risk of bias (Al-Moghrabi 2020; Ashari 2021; Edman Tynelius 2010; Gill 2007; Jowett 2022; Naraghi 2021b; O'Rourke 2016; Scribante 2020; Sfrondini 2021; Thickett 2010; Zhu 2017).

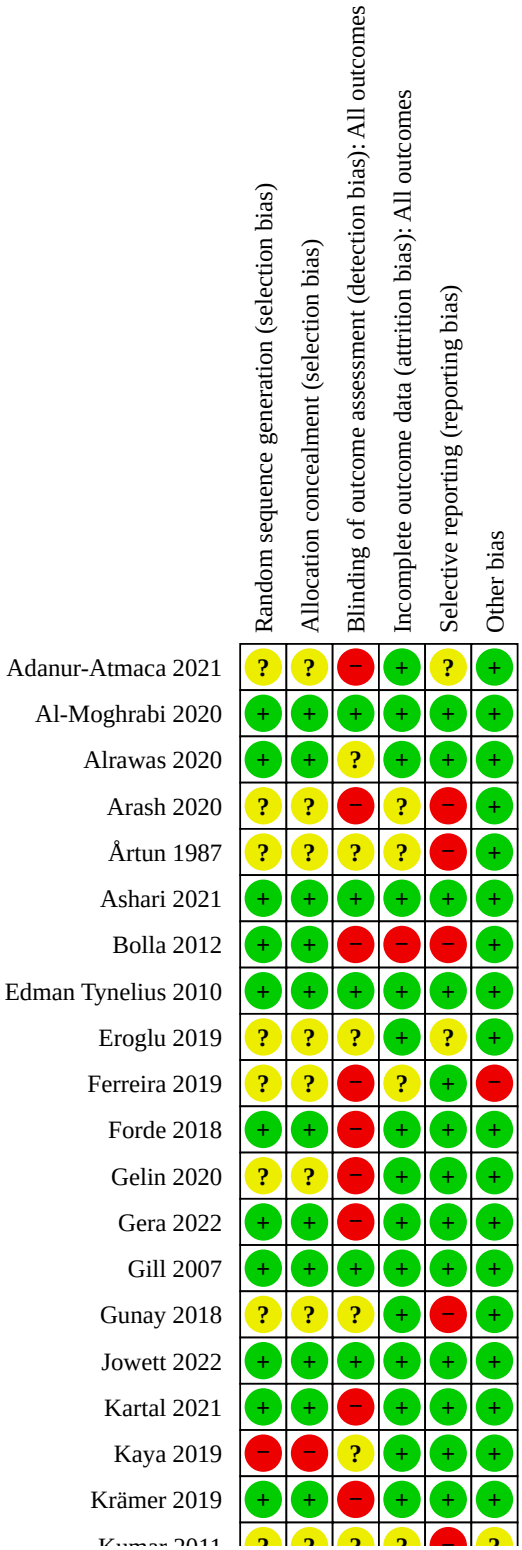
We judged 28 studies to have a high risk of bias (Adanur-Atmaca 2021; Arash 2020; Årtun 1987; Bolla 2012; Ferreira 2019; Forde 2018; Gelin 2020; Gera 2022; Gunay 2018; Kartal 2021; Kaya 2019; Krämer 2019; Kumar 2011; Millett 2007; Nagani 2020; Naraghi 2021a; Rohaya 2006; Rose 2002; Rowland 2007; Saleh 2017; Sfrondini 2014; Shawesh 2010; Shim 2022; Sinha 2021; Sun 2011; Torkan 2014; Wan 2017; Węgrodzka 2021).

In eight studies, the risk of bias was unclear (Alrawas 2020; Eroglu 2019; Moslemzadeh 2018; Pandis 2013; Ramazanzadeh 2018; Salehi 2013; Scribante 2011; Soboutia 2016).

See table of Characteristics of included studies and Figure 2and Figure 3.

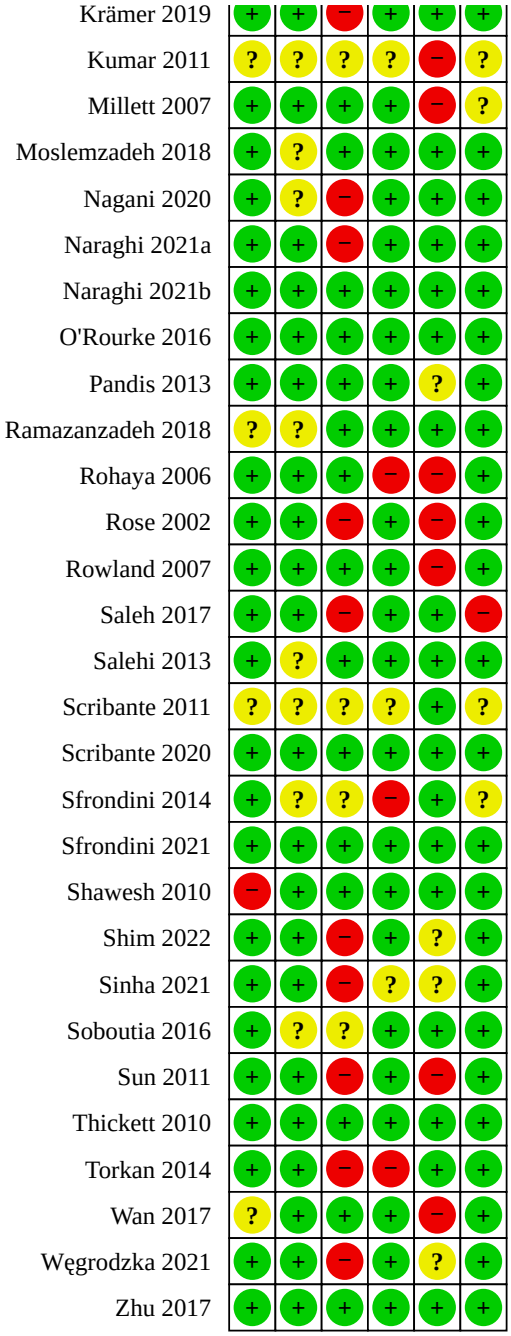
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study



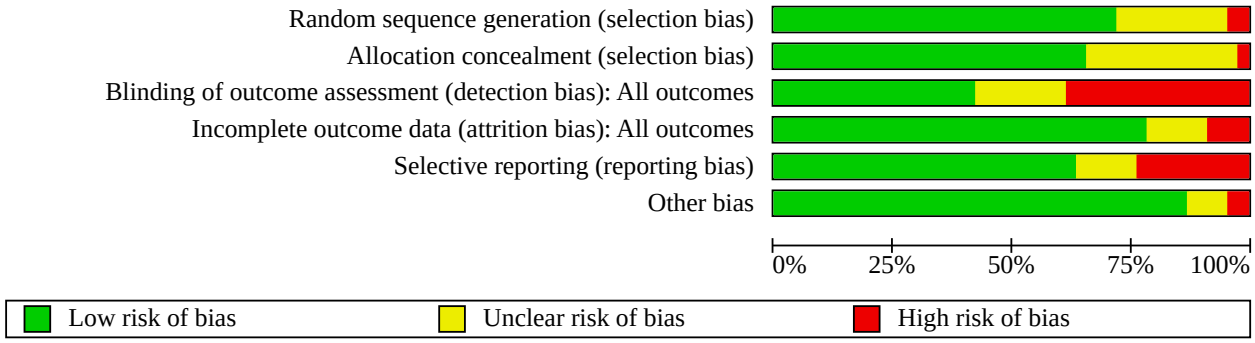
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 2. (Continued)



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Figure 3. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies



Allocation

Sequence generation

One study used an approach with a high risk of selection bias, as the technician, who had orthodontic knowledge, made the allocation without being blinded (Kaya 2019), and we judged another to be at high risk of bias as, despite adequate randomisation, there was no pre-treatment equivalence between the two intervention groups (Shawesh 2010). It was unclear how 11 studies performed the sequence generation (Adanur-Atmaca 2021; Arash 2020; Årtun 1987; Eroglu 2019; Ferreira 2019; Gelin 2020; Gunay 2018; Kumar 2011; Ramazanzadeh 2018; Scribante 2011; Wan 2017). All other studies had adequate sequence generation.

Allocation concealment

One study was at high risk of bias (Kaya 2019), and allocation concealment was unclear in 15 studies (Adanur-Atmaca 2021; Arash 2020; Årtun 1987; Eroglu 2019; Ferreira 2019; Gelin 2020; Gunay 2018; Kumar 2011; Moslemzadeh 2018; Nagani 2020; Ramazanzadeh 2018; Salehi 2013; Scribante 2011; Sfrondini 2014; Soboutia 2016). Allocation concealment was adequate in the other 31 studies.

Selection bias

Combining both aspects, two studies were at high risk of selection bias (Kaya 2019; Shawesh 2010), and 16 had an unclear risk of selection bias (Adanur-Atmaca 2021; Arash 2020; Årtun 1987; Eroglu 2019; Ferreira 2019; Gelin 2020; Gunay 2018; Kumar 2011; Moslemzadeh 2018; Nagani 2020; Ramazanzadeh 2018; Salehi 2013; Scribante 2011; Sfrondini 2014; Soboutia 2016; Wan 2017). The other 29 studies had a low risk of selection bias.

Blinding

The lack of ability to blind the participants and clinicians could potentially bias this type of research and should be noted when interpreting the findings.

Blinding of the assessors is possible, particularly when assessing stability on study models using some type of coverage to hide the retainer (Jowett 2022). Blinding of the assessors is also possible when assessing the results of participant satisfaction questionnaires. Eighteen studies had a high risk of detection bias as the outcome assessor was not blinded (Adanur-Atmaca 2021; Arash 2020; Bolla 2012; Ferreira 2019; Forde 2018; Gelin 2020; Gera 2022; Kartal 2021; Krämer 2019; Nagani 2020; Naraghi 2021a; Rose 2002; Saleh 2017; Shim 2022; Sinha 2021; Sun 2011; Torkan 2014; Węgrodzka 2021). In nine of the studies, it was unclear if assessors were blinded (Alrawas 2020; Årtun 1987; Eroglu 2019; Gunay 2018; Kaya 2019; Kumar 2011; Scribante 2011; Sfrondini 2014; Soboutia 2016). In the other 20 studies, there was low risk of detection bias as they used blinding of outcome assessors where possible.

Incomplete outcome data

Four studies had a high risk of attrition bias as they did not fully report dropouts (Bolla 2012; Rohaya 2006; Sfrondini 2014; Torkan 2014); and risk of bias in six studies was unclear (Arash 2020; Årtun 1987; Ferreira 2019; Kumar 2011; Scribante 2011; Sinha 2021). The remaining 37 studies were at low risk of attrition bias as they reported and explained dropouts and withdrawals, which were justified and similarly distributed.

Selective reporting

Twelve studies had a high risk of reporting bias.

- Eight did not report all outcomes prespecified in the methodology or protocols:
  - stability: overjet and overbite (Rowland 2007), arch length (Gunay 2018), or PAR index (Millett 2007);
  - adverse effects (Årtun 1987; Arash 2020);
  - stated more outcomes are to be reported (Rohaya 2006);
  - outcomes mentioned on a clinical trials register have not yet been reported (Sun 2011);
  - started collecting data for periodontal health, but abandoned this part way through the study due to time constraints (confirmed by contact with study author; Bolla 2012).
- Two added outcomes not previously specified in the methodology or protocols: participant satisfaction (Rowland 2007; Rose 2002)
- Two reported outcomes in part only: lower irregularity was reported, but not upper (Kumar 2011); did not report all mean values for patient satisfaction, which was measured by speech sounds through acoustic analysis (Wan 2017)

Six studies had unclear risk of reporting bias (Adanur-Atmaca 2021; Eroglu 2019; Pandis 2013; Shim 2022; Sinha 2021; Węgrodzka 2021).

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

The remaining 29 studies were at low risk of reporting bias as they showed no evidence of selective reporting.

Other potential sources of bias

We judged two studies to have a high risk of other bias. One study did not report in accordance with CONSORT guidelines, not describing important aspects of RCT design such as random sequence generation, allocation concealment, blinding, nor pre-registering the protocol on an international registry platform (Ferreira 2019). One study compared Hawley with clear plastic retainers, but all participants were also fitted with bonded lower retainers from canine to canine, as well as their removable lower retainer (Saleh 2017).

Four studies had unclear risk of other bias: there was inadequate information in two studies (Kumar 2011; Millett 2007), and two studies were not reported according to CONSORT guidelines for RCTs (Scribante 2011; Sfrondini 2014).

The other 41 studies appeared to be free of other bias.

Effects of interventions

See: **Summary of findings 1** Removable retainer (part-time) versus fixed retainer to stabilise tooth position; **Summary of findings 2** Removable retainer (full time) versus fixed retainer to stabilise tooth position; **Summary of findings 3** Fixed retainer (CAD/CAM nitinol) versus fixed retainer (conventional/analogue multistrand) to stabilise tooth position ; **Summary of findings 4** Fixed retainer (fibre-reinforced composite) versus fixed retainer (multistrand/spiral wire) to stabilise tooth position; **Summary of findings 5** Removable retainer (clear plastic) versus removable retainer (Hawley) to stabilise tooth position; **Summary of findings 6** Removable (full-time) versus removable (part-time) retainers to stabilise tooth position

Comparison 1: removable retainers versus fixed retainers

Eight studies investigated removable retainers versus fixed retainers (see Included studies above).

In two studies, the comparisons were in the upper and lower arches.

- Eroglu 2019 compared full-time upper and lower Hawley retainers and full-time upper and lower clear plastic retainers with upper and lower fixed lingual 5-stranded wire bonded to incisors and canines.
- Forde 2018 compared part-time upper and lower clear plastic retainers with upper and lower 3-stranded twistflex stainless steel wire bonded to incisors and canines. Review authors converted nonparametric data to parametric data as explained in the Methods section.

In one study, the comparison was in the upper arch only.

- Naraghi 2021b compared part-time upper clear plastic retainer with upper multistrand wire bonded only to the four incisors and bonded to the incisors and canines.

In five studies, the comparison was in the lower arch only.

- Årtun 1987 compared a removable acrylic and wire retainer with different types of fixed retainers (amount of wear not reported).
- Alrawas 2020 compared full-time clear plastic retainers with three types of fixed bonded retainers to upper and lower incisors and canines: CAD/CAM nitinol, multistrand and single strand wires.
- Krämer 2019 compared part-time partial clear plastic retainers (not covering molars) with a 0.8 mm stainless steel hard single-strand retainer bonded from canine to canine but only bonded on canines.
- Millett 2007 compared full-time lower clear plastic retainer with a multistrand stainless steel wire bonded to the canines and incisors.
- O'Rourke 2016 compared a lower clear plastic retainer with a multistrand stainless steel wire bonded to the canines and incisors. Removable retainers were used full-time for six months, then part-time. Review authors converted nonparametric data to parametric data.

Stability

Six studies reported stability using Little's Irregularity Index (Analysis 1.1).

Little's Irregularity Index

Upper arch

- Forde 2018 did not find a difference between removable clear plastic retainers used part-time and fixed multistrand wire retainers, at 12 months (MD −0.34 mm, 95% CI −0.95 to 0.27; P = 0.27; 56 participants; Analysis 1.1.1).
- Nor did Naraghi 2021b find a difference at 24 months when the fixed retainer was bonded to the four incisors (MD 0.40 mm, 95% CI −0.21 to 1.01; P = 0.20; 60 participants; Analysis 1.1.3).
- When bonded to the six anterior teeth, there was less relapse with the fixed retainer compared to the part-time clear plastic retainer (MD 0.7 mm, 95% CI 0.09 to 1.31; P = 0.02; 60 participants; Analysis 1.1.2).

Lower arch

- We were able to undertake a meta-analysis of data at six months from two studies (Alrawas 2020; O'Rourke 2016), which showed no difference between the removable clear plastic retainer worn full-time and multistrand or coaxial fixed retainers (MD 0.05 mm, 95% CI −0.00 to 0.10; P = 0.06, I^2 = 0%; 103 participants; Analysis 1.1.4).
- Alrawas 2020 also compared the removable clear plastic retainers worn full time with two other types of fixed retainers at six months and showed no difference between fixed and removable retainers, versus single-strand titanium (MD −0.10 mm, 95% CI −0.70 to 0.50; P = 0.74; 29 participants; Analysis 1.1.5); versus CAD/CAM nitinol (MD −0.06 mm, 95% CI −0.62 to 0.50; P = 0.83; 30 participants; Analysis 1.1.6).
- At 12 months, two studies comparing full-time (Millett 2007), and part-time (Forde 2018), clear plastic retainers versus fixed multistrand retainers showed differences of 0.60 mm (95% CI 0.17 to 1.03; P = 0.007; 84 participants; Analysis 1.1.7) and 0.92 mm (95% CI 0.23 to 1.61; P < 0.009; 56 participants; Analysis 1.1.8), respectively, with slightly more relapse with the lower removable.
- At 18 months, Krämer 2019 did not find a difference between a removable partial part-time clear retainer and a fixed thick plain

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

retainer bonded to canines (MD 0.07 mm, −0.20 to 0.34; P = 0.61; 95 participants; Analysis 1.1.9).

- A second publication of the O'Rourke 2016 study, four years after debonding, reported a larger mean relapse of 1.52 mm (95% CI 0.75 to 2.29; P < 0.001; 42 participants; Analysis 1.1.10) in the removable part-time clear plastic retainer compared to a coaxial fixed retainer.

Five of the six above-mentioned studies also analysed intercanine and intermolar widths, and arch length changes (Analysis 1.2; Analysis 1.3; Analysis 1.4).

Intercanine width

Upper arch

- In the upper arch, the intercanine width was similar between clear plastic part-time removable and fixed multistrand retainers at 12 months (MD −0.06 mm, 95% CI −0.48 to 0.36; P = 0.78; 56 participants; Forde 2018; Analysis 1.2.1).
- It was slightly higher at 24 months with the part-time clear plastic compared to fixed multistrand retainer bonded on the six upper anterior teeth (MD 0.50 mm, 95% CI 0.23 to 0.77; P < 0.001; 60 participants; Naraghi 2021b; Analysis 1.2.2), but similar when the fixed retainer was bonded to the four incisors (MD 0.20 mm, 95% CI −0.23 to 0.63; P = 0.36; 60 participants; Naraghi 2021b; Analysis 1.2.3).
- The removable group showed a small increase in intercanine width and the fixed group showed a small decrease (Analysis 1.2.2).

Lower arch

In the lower arch, we combined data from two studies (Alrawas 2020; O'Rourke 2016), and showed a small difference in the intercanine width at six months between removable clear plastic retainers worn full time and multistrand or coaxial fixed retainers (MD 0.11 mm, 95% CI 0.05 to 0.17; P < 0.001, I^2 = 0%; 103 participants; Analysis 1.2.4).

Intermolar width

We found no significant differences between removable and different types of fixed retainers except for the intermolar width (Analysis 1.3).

Arch length

We found no significant changes in arch length other than for the comparison between lower removable clear plastic part-time retainers versus fixed multistrand at 12 months (Forde 2018), with a small decrease in the part-time removable group and a small increase in the fixed group (MD −0.80 mm, 95% CI −1.28 to −0.32; P < 0.001; 56 participants; Analysis 1.4.7).

Overjet and overbite

Three studies (Forde 2018; Krämer 2019; Naraghi 2021b), reported overjet and overbite changes at 12, 18 and 24 months, showing no significant changes between part-time removable and fixed retainers (Analysis 1.5; Analysis 1.6).

Maximum rotation: upper arch

Naraghi 2021b investigated maximum rotation (defined as the worst single tooth contact point per participant) in the upper arch two years after debonding. We found no significant differences when comparing part-time clear plastic retainer and multistrand wire retainer bonded to the four incisors, but we found a significantly higher rotation in the removable group when the fixed retainer was bonded to the six upper anterior teeth (MD 3.70°, 95% CI 0.87 to 6.53; P = 0.01; 60 participants; Analysis 1.7.1).

Failure of retainers

Five studies reported failure of retainers (Årtun 1987; Forde 2018; Krämer 2019; Millett 2007; O'Rourke 2016), which they measured in different ways (Analysis 1.8; Analysis 1.9).

- Rate of failure in %
  - Lost removable retainers and debonded fixed retainers (Årtun 1987)
  - Lost/fractured/not fitting/ removable retainers and detachment of fixed retainers (Krämer 2019)
  - Lost/ill fitting/fractured/occlusal wear removable retainers and debonded/fractured fixed retainers (Forde 2018)
  - Fractured/lost (but not worn) removable retainers and debonded fixed retainers (Millett 2007)
  - Noncompliance levels for removable and partial detachments or history of repairs for the fixed retainers (O'Rourke 2016)
- Time to first episode of failure in days (mean survival time) (Forde 2018).

Upper arch

There was no difference in the failure rate between part-time removable and fixed multistrand wire retainers at 12 months (RR 0.81, 95% CI 0.38 to 1.71; P = 0.58; 57 participants; Forde 2018; Analysis 1.8.1).

Lower arch

Two studies compared lower clear plastic retainers worn part-time and full-time, respectively, versus fixed multistrand wire retainers at 12 months, finding different results.

- Forde 2018 found more failures in the fixed retainer group than in the part-time removable group (RR 0.44, 95% CI 0.20 to 0.98; P = 0.04; 57 participants; Analysis 1.8.2).
- In Millett 2007 there were more failures in the full-time removable retainer group than the bonded retainer group (RR 3.42, 95% CI 1.38 to 8.47; P = 0.008; 77 participants; Analysis 1.8.3).
- No differences were found between removable and thick plain wire fixed retainers at 18 months (Krämer 2019; Analysis 1.8.4),
- No differences were found between the three types of fixed retainers and the removable retainer used in Årtun 1987 at three years (Analysis 1.8.5; Analysis 1.8.6; Analysis 1.8.7).
- In O'Rourke 2016 there was a higher risk of noncompliance levels for the removable part-time retainers compared to the percentage of failures of fixed retainers at 48 months (RR 3.0, 95% CI 1.33 to 6.75; P = 0.008; 42 participants; Analysis 1.8.8).

Only one study reported the time to first episode of failure (Forde 2018), and the mean difference was not significant (Analysis 1.9).

Patient satisfaction

Three studies investigated patient satisfaction when comparing clear plastic retainers worn part-time and full-time with bonded

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

retainers (Forde 2018 (see Table 1); Krämer 2019; Millett 2007), through the following core categories.

Acceptance

Three studies reported that fixed retainers were more acceptable than removable retainers.

- Two studies reported at 12 months of follow-up, with continuous and dichotomous outcomes and with different wearing protocols, which prevented statistical pooling (Millett 2007: MD −12.84 (95% CI −18.10 to −7.09; P < 0.001; 81 participants; Analysis 1.10.1; Forde 2018; RD 0.15, 95% CI 0.00 to 0.29; P = 0.04; 56 participants; Analysis 1.11.2).
- One study reported at 18-month follow-up (Krämer 2019; MD −7.50, 95% CI −11.54 to −3.46; P < 0.001; 95 participants; Analysis 1.10.2).
- Forde 2018 also reported that there was no difference in the embarrassment felt by participants wearing either type of retainer (RD 0.07, 95% CI −0.04 to 0.19; P = 0.21; 57 participants; Analysis 1.11.1).

Compliance

One study reported that compliance with retainer instructions was lower with the part-time removable clear plastic retainers than the fixed retainers (RD 0.22, 95% CI 0.06 to 0.38; P = 0.007; 57 participants; Forde 2018; Analysis 1.12.1).

Efficiency

One study (Krämer 2019), assessed efficiency of the retainers as perceived by participants through two questions and found no difference between the two types of retainers.

- “Are the retainers working well?”: MD 2.00 (95% CI −2.75 to 6.75; P = 0.41; 95 participants; Analysis 1.13.1)
- “Noticed any change in tooth position?”: MD 2.00 (95% CI −2.66 to 6.66; P = 0.40; 95 participants; Analysis 1.13.2)

Cleaning/hygiene

Two studies reported cleaning or hygiene of fixed retainers compared to removable clear plastic retainers, with opposite findings.

- In Millett 2007,participants who wore fixed retainers found them easier to clean than those who wore removable retainers part-time (MD −10.31, 95% CI −20.05 to −0.58; P = 0.04; 81 participants; Analysis 1.14).
- In Forde 2018, however, participants who wore the fixed retainers were less likely to report finding them easy to clean than those wearing the removable retainers full-time (RR 1.39, 95% CI 1.06 to 1.83; P = 0.02; 57 participants; Analysis 1.15.1).

Comfort, discomfort or pain

Two studies assessed this outcome (Forde 2018; Krämer 2019).

- Krämer 2019 differentiated between three aspects:
  - “pain or discomfort” without significant differences (MD 5.00, 95% CI −2.21 to 12.21; P = 0.17; 95 participants; Analysis 1.16.1);
  - “soreness” with differences favouring fixed retainers (MD 2.00, 95% CI 0.72 to 3.28; P = 0.002; 95 participants; Analysis 1.16.2); and
  - “tension in your teeth” with differences favouring fixed retainers (MD 8.50, 95% CI 2.53 to 14.47; P = 0.005; 95 participants; Analysis 1.16.3).
- Forde 2018_reported differences favouring fixed retainers (RR 12.22, 95% CI 1.69 to 88.52; P = 0.01; 57 participants; Analysis 1.17.1).

We could not pool these results because Krämer 2019 reported continuous data and Forde 2018 reported dichotomous data.

Speech

One study (Forde 2018), evaluated speech and found more favourable results with fixed retainers than removable clear plastic retainers (RR 3.89, 95% CI 1.46 to 10.38; P = 0.007; 57 participants; Analysis 1.18.1).

Adverse effects on health

Five studies reported different outcomes related to oral health, when comparing removable with bonded retainers (Alrawas 2020; Eroglu 2019; Forde 2018; Millett 2007; O'Rourke 2016):

- evidence of caries (Millett 2007; Analysis 1.19);
- Plaque Index (Alrawas 2020; Eroglu 2019; Forde 2018; O'Rourke 2016; Analysis 1.20);
- Gingival Index (Alrawas 2020; Eroglu 2019; Forde 2018; Analysis 1.21) and gingival bleeding (dichotomous) (Millett 2007; (Analysis 1.22);
- Calculus Index (Forde 2018; Analysis 1.23);
- probing depth (Alrawas 2020; Eroglu 2019; Analysis 1.24) and periodontal pocketing (dichotomous) (Millett 2007; Analysis 1.25);
- bleeding on probing (Alrawas 2020; Eroglu 2019; Analysis 1.26).

It was not possible to pool results from these outcomes due to comparison of markedly different types of retainers.

Evidence of caries

There was no difference in evidence of caries when comparing lower removable clear plastic full-time versus fixed multistrand at 12 months (Millett 2007; Analysis 1.19.1).

Plaque Index: upper and lower arches

- Only one study (Forde 2018), showed differences in Plaque Index, favouring part-time clear plastic retainers versus multistrand wire bonded retainers at 12 months in the upper arch (MD −0.17, 95% CI −0.29 to −0.05; P = 0.005; 59 participants; Analysis 1.20).
- However, no differences were found in the lower arch (Analysis 1.20.2).
- Two studies (Alrawas 2020; O'Rourke 2016), compared lower full-time and part-time clear plastic retainers with different lower fixed retainers and found no differences at various time points (Analysis 1.20.3 to Analysis 1.20.6).
- One study showed pooled data for the upper and lower arches, and compared full-time Hawley (Analysis 1.20.7) and clear plastic retainers (Analysis 1.20.8) versus multistrand wire bonded retainers at three months (Eroglu 2019). No differences were found.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gingival Index: upper and lower arches

- Forde 2018 found lower Gingival Index in the part-time clear plastic retainer group in the lower arch (MD −0.34, 95% CI −0.66 to −0.02; P = 0.04; 59 participants; Analysis 1.21.2) but not in the upper arch (Analysis 1.21.1) at 12 months.
- The same studies described above for Plaque Index reported non-significant differences in Gingival Index between full-time and part-time removable retainers and fixed retainers at different time-points in the lower arch only (Analysis 1.21.3 to Analysis 1.21.6) or in both arches (Analysis 1.21.7; Analysis 1.21.8).

Calculus Index

Only Forde 2018 found higher values in the lower fixed multistrand group at 12 months (MD −0.15, 95% CI −0.29 to −0.01; P = 0.03; 59 participants; Analysis 1.22.2) compared to the part-time clear plastic retainers.

Probing depth

In Millett 2007, periodontal pocketing (probing depth), recorded as a dichotomous variable, was higher in the bonded lower retainer group (RR 0.32, 95% CI 0.12 to 0.87; P = 0.03; 84 participants; Analysis 1.24) compared to the lower removable clear plastic full-time group. The remaining studies did not show differences in probing depth when comparing removable and fixed retainers.

Bleeding on probing

- Bleeding on probing was higher in one arm of Alrawas 2020, with slightly higher values in the lower arch for clear plastic group used full-time compared to the fixed CAD/CAM nitinol bonded retainer at six months (MD 0.12, 95% CI 0.00 to 0.24; P = 0.05; 30 participants; Analysis 1.25.3). However, no differences were found for the other two types of fixed retainers used in this study (multistrand and single-strand wire retainers).
- One study (Eroglu 2019) reported differences (although clinically insignificant) in bleeding on probing, favouring the fixed multistrand retainer over the Hawley retainer (MD 0.03, 95% CI 0.00 to 0.06; P = 0.03; 30 participants; Analysis 1.25.4), but found no difference when compared to the clear plastic retainer (Analysis 1.25.5).
- Millett 2007 recorded gingival bleeding as a dichotomous variable. It was higher in the fixed retainer group (RR 0.53, 95% CI 0.31 to 0.88; P = 0.01; 84 participants; Analysis 1.26.1).

Comparison 2: fixed retainers versus fixed retainers

Twenty-five studies compared different types of fixed retainers.

Seven studies reported four comparisons in the upper and lower arches.

- Bolla 2012, Salehi 2013 and Torkan 2014 compared upper and lower fibre-reinforced composite fixed retainer versus multistrand wire fixed retainers.
- Gera 2022 and Jowett 2022 compared upper and lower CAD/CAM nitinol versus multistrand wire fixed retainers.
- Sfrondini 2021 compared upper and lower multistrand wire retainers bonded using two types of composite: highly filled orthodontic resin versus flowable nanocomposite.
- Scribante 2020 compared upper and lower multistrand wire retainers bonded using two types of primers: universal adhesive in total-etch mode versus traditional orthodontic primer.

In one study, the comparison was in the upper arch only.

- Naraghi 2021b compared upper multistrand wire bonded only to the four incisors or bonded to the six anterior teeth.

In 17 studies, the comparison was in the lower arch only.

- Adanur-Atmaca 2021 compared four types of lower fixed retainers: rectangular braided multistrand (dead-soft) wire, round multistrand stainless-steel wire, CAD/CAM nitinol retainer and 0.012 inch connected tooth-shaped bonding pad retainer on the lingual surface of each tooth.
- Alrawas 2020 compared three types of lower fixed bonded retainers: CAD/CAM nitinol, multistrand and single-strand wires.
- Arash 2020 compared lower multistrand versus soft thick plain titanium wire fixed retainers.
- Årtun 1987 compared three types of lower fixed retainers: thick (0.032 inch) plain stainless steel wire bonded only to the canines, thick (0.032 inch) spiral stainless steel wire bonded only to the canines and a thin (0.0205 inch) spiral wire bonded to the canines and incisors.
- Ferreira 2019 compared a rigid round steel wire retainer bonded only to the lower canines versus rectangular chain retainer bonded to lower incisors and canines.
- Gelin 2020 and Kartal 2021 compared lower CAD/CAM nitinol versus multistrand wire retainers.
- Gunay 2018 compared a lower indirect multistrand wire versus a direct dead-soft coaxial wire retainer.
- Nagani 2020; Rose 2002; Scribante 2011 and Sfrondini 2014 compared lower fibre-reinforced composite retainer versus multistrand wire fixed retainers.
- Pandis 2013 compared lower multistrand wire retainers bonded using two types of chemically and light-cured adhesives and composites.
- Shim 2022 compared three types of lower fixed retainers: CAD/CAM multistrand stainless steel wire, lab-based retainer (same type of wire as the CAD/CAM group, but manually bent), and rectangular chain stainless steel retainer.
- Sinha 2021 compared ceramic interlocking retainer versus flexible round coaxial wire (0.016 inch) retainer.
- Soboutia 2016 compared three types of fixed retainers: fibre-reinforced composite versus twisted ligature wire versus spiral wire (multistrand wire).
- Węgrodzka 2021 compared three-strand round twisted versus eight-strand rectangular braided retainers.

Stability

Little's Irregularity Index

Nine studies reported stability using Little's Irregularity Index (Analysis 2.1).

Upper arch

- Jowett 2022 did not find significant differences between a CAD/CAM nitinol fixed retainer and a rectangular chain fixed retainer

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

when the irregularity index was assessed at six months (Analysis 2.1.1).

- Naraghi 2021b compared a multistrand retainer bonded to the four upper incisors with the same retainer bonded to incisors and canines at 24 months, and found less relapse in the second group, but it may not be clinically significant (MD 0.3 mm, 95% CI 0.03 to 0.57; P = 0.03; 60 participants; Analysis 2.1.2).

Lower arch

Eight studies reported results at six months, and three of these studies also reported results at 12 months.

- At six months, three studies comparing CAD/CAM nitinol versus multistrand wire fixed retainers (Adanur-Atmaca 2021; Alrawas 2020; Gera 2022), did not show significant differences in the irregularity index (MD −0.03, 95% CI −0.12 to 0.06; P = 0.52; I^2= 22%; 258 participants; Analysis 2.1.3).
- When CAD/CAM nitinol retainers were compared to rectangular chain retainers (Jowett 2022; Analysis 2.1.4) and single-strand titanium retainers (Alrawas 2020; Analysis 2.1.5), no significant differences were found in the amount of relapse at six months. Non-significant differences were also found at six months in Alrawas 2020, Nagani 2020, Shim 2022 and Sinha 2021, among different types of lower fixed retainers (Analysis 2.1.6 to Analysis 2.1.11).
- Gunay 2018, however, found significantly more relapse in the direct dead-soft coaxial than in the indirect multistrand fixed retainer group at six months (MD 0.96, 95% CI 0.57 to 1.35, 120 participants; Analysis 2.1.12) and 12 months (MD 1.15 mm, 95% CI 0.74 to 1.56; P < 0.001; 120 participants; Analysis 2.1.13).
- At 12 months, Nagani 2020 also found significant differences, with less relapse in the fibre-reinforced composite group compared to multistrand retainer group (MD −0.7 mm, 95% CI −1.17 to −0.23; P = 0.003; 52 participants; Analysis 2.1.14).
- Adanur-Atmaca 2021 showed more relapse in the rectangular braided retainers (MD 0.3 mm, 95% CI 0.10 to 0.63; P = 0.007; 66 participants; Analysis 2.1.15) and in the rigid connected bonding pad retainer (MD 0.38 mm, 95% CI 0.17 to 0.58; P < 0.001; 66 participants; Analysis 2.1.17) than in the round multistrand retainers, and also less relapse in the lower CAD/CAM nitinol retainer compared to the rectangular braided (MD −0.46 mm, 95% CI −0.72 to −0.21; P < 0.001; 66 participants; Analysis 2.1.18) and rigid connected bonding pad retainer (MD −0.47 mm, 95% CI −0.66 to −0.29; P < 0.001; 66 participants; Analysis 2.1.20).

Intercanine width

Nine studies assessed intercanine width (Analysis 2.2).

Upper arch

- Gera 2022 did not find any significant differences in intercanine width changes between multistrand wire retainers and CAD/CAM nitinol retainers at six months (Analysis 2.2.1).
- Naraghi 2021b compared a multistrand retainer bonded to the four upper incisors with a retainer bonded to the six anterior teeth at 24 months and did not find any significant differences (Analysis 2.2.2).
- Jowett 2022, however, found greater decrease in intercanine width in the rectangular chain group compared to the CAD/CAM nitinol retainer group at six months (MD 1.12 mm, 95% CI 0.80 to 1.45; P < 0.001; 55 participants; Analysis 2.2.3).

Lower arch

Seven studies reported results at six months, and three of these studies also reported results at 12 months.

- At six months, four studies comparing CAD/CAM nitinol versus multistrand wire fixed retainers (Adanur-Atmaca 2021; Alrawas 2020; Gelin 2020; Gera 2022), did not show significant differences in the intercanine width changes (MD 0.04 mm, 95% CI −0.15 to 0.22; P = 0.70; I^2 = 0%; 309 participants; Analysis 2.2.4).
- Similar non-significant changes were reported by different studies comparing different types of fixed retainers or bonding procedures: two more arms of Alrawas 2020 (Analysis 2.2.5; Analysis 2.2.6), Gunay 2018 (Analysis 2.2.7), and one arm of Shim 2022 (Analysis 2.2.9).
- In contrast, Jowett 2022 found significantly greater decrease in the intercanine width in the rectangular chain group compared to the CAD/CAM nitinol retainer group (MD 0.47 mm, 95% CI 0.12 to 0.82; P = 0.008; 55 participants; Analysis 2.2.8)
- Shim 2022 found greater decrease in the rectangular chain group compared to the CAD/CAM stainless steel (MD 1.23 mm; 95% CI 0.82 to 1.64; P < 0.001; 30 participants; Analysis 2.2.10) and laboratory-constructed multistrand retainers (MD 0.92 mm, 95% CI 0.49 to 1.35; P < 0.001; 30 participants; Analysis 2.2.11).
- At 12 months, there were no significant differences between the different comparisons of fixed retainers: Gunay 2018 (Analysis 2.2.12); and different arms of Adanur-Atmaca 2021 (Analysis 2.2.13 to Analysis 2.2.17).
- We were able to combine two studies (Gelin 2020 and one arm of Adanur-Atmaca 2021; Analysis 2.2.18), but did not find significant differences in the intercanine width changes.

Interpremolar width: upper and lower arches

One study (Gera 2022), assessed interpremolar width but did not find significant differences between CAD/CAM nitinol versus multistrand wire fixed retainers in the upper and lower arches at six months (Analysis 2.3.1; Analysis 2.3.2).

Intermolar width

Four studies assessed intermolar width (Analysis 2.4).

Upper arch

- One study (Gera 2022), reported nonsignificant changes at six months (Analysis 2.4.1) and another study (Naraghi 2021b), at 24 months (Analysis 2.4.2).

Lower arch

- We found similar, non-significant results when we combined three studies (Alrawas 2020; Gelin 2020; Gera 2022), that compared CAD/CAM nitinol versus multistrand stainless steel retainers at six months (MD 0.28 mm, 95% CI −0.38 to 0.93; P = 0.41; I^2 = 0%; 243 participants; Analysis 2.4.3).
- Other results at 6 and 12 months showed the same non-significant differences (Analysis 2.4.4; Analysis 2.4.5; Analysis 2.4.6).

Anterior arch length: lower arch

Two studies (Alrawas 2020; Gelin 2020), assessed anterior arch length; Alrawas 2020 had three arms (Analysis 2.5). No significant differences were found when comparing different types of lower

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

fixed retainers at six (Analysis 2.5.1; Analysis 2.5.2; Analysis 2.5.3), and 12 months (Analysis 2.5.4).

Total arch length

Four studies assessed total arch length (Analysis 2.6).

Upper arch

- Two studies reported non-significant changes at six months (Gera 2022 – Analysis 2.6.1) and 24 months (Naraghi 2021b – Analysis 2.6.2).

Lower arch

The same non-significant results were found in the lower arch.

- We were able to combine three studies (Adanur-Atmaca 2021; Gelin 2020; Gera 2022), that compared CAD/CAM nitinol versus multistrand stainless steel retainers at six months (MD −0.02 mm, 95% CI −0.58 to 0.53; P = 0.93; I^2 = 0%; 280 participants; Analysis 2.6.3).
- At 12 months, results showed the same non-significant differences (Analysis 2.6.4. to Analysis 2.6.8).
- We were able to combine results at 12 months for two of the three studies mentioned above (Adanur-Atmaca 2021; Gelin 2020), showing non-significant differences (Analysis 2.6.9).

Overjet and overbite

Only one study (Naraghi 2021b), reported changes to overjet (Analysis 2.7) and overbite (Analysis 2.8) for the upper arch at 24 months, comparing upper fixed retainers bonded to the four anterior teeth and to the six anterior teeth. Changes in these outcomes were not significantly different between groups.

Maximum rotation

The same study (Naraghi 2021b), investigated maximum rotation, measured as the worst single tooth contact point per participant. Significantly higher rotation was found when the retainer was bonded to the four upper anterior teeth compared to bonding to the six anterior teeth (MD 2.30°, 95% CI 0.80 to 3.80; P = 0.003; 60 participants; Analysis 2.9).

Failure of retainers

Twenty studies reported failure of retainers: six of them in the upper and lower arches (Bolla 2012; Gera 2022; Jowett 2022; Salehi 2013; Scribante 2020; Sfrondini 2021), and the rest of them only in the lower arch (Arash 2020; Årtun 1987; Gelin 2020; Gunay 2018; Kartal 2021; Nagani 2020; Pandis 2013; Rose 2002; Scribante 2011; Sfrondini 2014; Shim 2022; Sinha 2021; Soboutia 2016; Węgrodzka 2021). Studies reported failure/survival of retainers in different ways:

- rate of failure in %, as number of debonded/fractured retainers (Analysis 2.10);
- mean survival time: time to first episode of failure in months (Analysis 2.10);
- failure risk (hazard ratio) (Analysis 2.12).

Failure rate

Upper arch

- At six months, Gera 2022 concluded that the failure rate was not significantly different between CAD/CAM nitinol and multistrand wire retainers (RR 0.99 mm, 95% CI 0.57 to 1.70; P = 0.96; 163 participants; Analysis 2.10.1).
- Jowett 2022, however, found a significantly higher failure rate in CAD/CAM nitinol than in rectangular chain retainers (RR 2.90, 95% CI 1.20 to 7.03; P = 0.02; 55 participants; Analysis 2.10.2).
- Two studies (Bolla 2012; Salehi 2013), compared failure rates of fibre-reinforced composite retainers with multistrand retainers in the maxilla, and it was possible to undertake a meta-analysis of these data. The pooled estimate showed an RR of 1.25 (95% CI 0.87 to 1.78; P = 0.22), indicating no difference in the failure rates. There was low heterogeneity between the studies (I^2 = 22%), with very similar methodology and materials used in both studies (Analysis 2.10.3).
- Two studies compared the failure rate associated with different bonding procedures of multistrand wire retainers at 24 months. Scribante 2020 found a significantly lower failure rate when a conventional orthodontic composite was used compared to a flowable composite (RR 0.50, 95% CI 0.28 to 0.89; P = 0.02; 600 participants; Analysis 2.10.4), and Sfrondini 2021 found significantly more bonding failures when a conventional orthodontic primer was used compared to a universal primer in total-etch mode (RR 2.23, 95% CI 1.38 to 3.59; P = 0.001; 600 participants; Analysis 2.10.5).

Lower arch

- Two studies (Gera 2022; Kartal 2021), that compared CAD/CAM nitinol versus multistrand wire retainers did not find significant differences at six months (RR 1.00, 95% CI 0.67 to 1.50; P = 0.99; I^2 = 0%; 215 participants; Analysis 2.10.6).
- Jowett 2022 and Sinha 2021 did not find significant differences when comparing failure rates of CAD/CAM nitinol versus rectangular chain retainers (Analysis 2.10.7) and ceramic interlocking bonded retainer and coaxial wire retainer (Analysis 2.10.8), respectively, at six months.
- No differences were found in the three arms of Shim 2022 either (Analysis 2.10.9. to Analysis 2.10.11) between CAD/CAM stainless steel, laboratory-fabricated stainless steel multistrand and chairside formed rectangular chain retainers.
- At 12 months (or more), different comparisons of fixed retainers did not find significant differences either for directly bonded coaxial wire versus indirectly bonded multistrand retainer (Gunay 2018; Analysis 2.10.12); CAD/CAM nitinol versus multistrand wire retainers (Gelin 2020; Analysis 2.10.13); fibre-reinforced composite retainer versus multistrand fixed retainer (RR 1.01, 95% CI 0.84 to 1.21; P = 0.95; I^2 = 60%; 7 studies, 1337 participants; Analysis 2.10.14).
- At 24 months, Soboutia 2016 found similar failure rates when comparing fibre-reinforced composite versus twisted ligature wire (Analysis 2.10.15) and spiral wire versus twisted ligature wire (Analysis 2.10.16). Węgrodzka 2021 did not find significantly different failure rates between round twisted multistrand and rectangular braided multistrand retainers at 24 months either (Analysis 2.10.17).
- Similar to their findings for the upper arch, at 24 months, Scribante 2020 found a significantly lower failure

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

rate when a conventional orthodontic composite was used compared to a flowable composite (RR 0.53, 95% CI 0.32 to 0.87; P = 0.01; 600 participants; Analysis 2.10.18), and Sfrondini 2021 found significantly more bonding failures when a conventional orthodontic primer was used compared to a universal primer in total-etch mode (RR 5.11, 95% CI 3.17 to 8.25; P < 0.001; 600 participants; Analysis 2.10.19).

- However, Pandis 2013 did not find significantly different failure rates of mandibular lingual retainers between chemically and light-cured adhesives at 24 months (Analysis 2.10.20).
- One study (Arash 2020) compared thick plain wire (canine to canine) versus thin spiral wire (incisors and canines). Significantly lower failure rates were found in the thick plain wire group (RR 0.49, 95% CI 0.25 to 0.98; P = 0.04; 240 participants; Analysis 2.10.21).
- There was no difference between the failure rates of the other three comparisons of fixed retainers in Årtun 1987 (Analysis 2.10.22; Analysis 2.10.23; Analysis 2.10.24).

Mean survival time: lower arch

Three studies reported mean survival time as the time to first episode of failure in months in the lower arch.

- Arash 2020 did not find significant differences at 24 months between a thick plain titanium wire and a twisted wire retainer (Analysis 2.11.1).
- Soboutia 2016 found significant differences at 24 months among the three types of fixed retainers analysed. The multistrand spiral wire and the fibre-reinforced composite retainers showed shorter survival time than the twisted ligature wire retainer (MD −0.99 months, 95% CI −1.33 to −0.65; P value < 0.00001; 86 participants; Analysis 2.11.2); (MD −2.40 months, 95% CI −2.77 to −2.04; P < 0.001; 87 participants; Analysis 2.11.3), respectively.
- Two studies (Rose 2002; Soboutia 2016), showed shorter survival time for fibre-reinforced composite retainer than multistrand wire retainer (MD −1.48 months, 95% CI −1.88 to −1.08; P < 0.001; I^2 = 94%; 103 participants; Analysis 2.11.4).

Failure risk

We assessed failure risk by means of hazard ratios.

Upper arch

- Jowett 2022 found significantly higher failure risk at six months for the CAD/CAM nitinol versus the multistrand wire fixed retainer (HR 2.82, 95% CI 1.00 to 7.97; P = 0.05; 55 participants; Analysis 2.12.1).
- Gera 2022 did not find significant differences in the failure risk at six months when comparing these two types of retainers (Analysis 2.12.2).

Lower arch

- The same studies (Gera 2022; Jowett 2022), also found no differences in the failure risk of CAD/CAM nitinol versus the multistrand wire fixed retainers (Analysis 2.12.3; Analysis 2.12.4).
- Sfrondini 2014 found a significantly higher failure risk for the multistrand spiral wire versus fibre-reinforced composite at 12 months, (HR 1.64, 95% CI 1.04 to 2.59; P = 0.03; 87 participants; Analysis 2.12.5).
- We found non-significant hazard ratios, indicating similar failure risks, in different studies comparing different lower fixed retainers: fibre-reinforced composite versus multistrand fixed retainers at 12 months (Scribante 2011 – Analysis 2.12.6); spiral wire versus twisted ligature wire retainers and fibre-reinforced composite versus twisted ligature wire retainers at 24 months (Soboutia 2016 – Analysis 2.12.7; Analysis 2.12.8); round twisted multistrand versus rectangular braided multistrand retainers (Węgrodzka 2021 – Analysis 2.12.9); and multistrand retainers bonded using chemically and light-cured adhesives at 24 months (Pandis 2013 – Analysis 2.12.10).

Upper and lower arches

Combined results for upper and lower arches showed significantly lower failure risk when a conventional orthodontic composite was used compared to a flowable composite (Scribante 2020: HR 0.50, 95% CI 0.34 to 0.73; P < 0.001; 100 participants; Analysis 2.12.11), and a significantly higher failure risk when a conventional orthodontic primer was used compared to a universal primer in total-etch mode (Sfrondini 2021: HR 3.77, 95% CI 2.78 to 5.11; P < 0.001; 100 participants; Analysis 2.12.12).

Patient satisfaction

Three studies (Ferreira 2019; Jowett 2022; Scribante 2011), investigated patient satisfaction when comparing different types of fixed retainers using the following core categories.

Efficiency

One study (Jowett 2022), that compared CAD/CAM nitinol fixed retainer and a rectangular chain bonded retainer of the retainers, assessed efficiency as perceived by participants. They used the question: “Did your retainer keep your teeth straight?” No significant differences were found between the groups in the upper or lower jaw (RR 1.00, 95% CI 0.91 to 1.10; P = 1.00; 62 participants; Analysis 2.13.1; Analysis 2.13.2).

Cleaning and hygiene

Two studies (Ferreira 2019; Jowett 2022) assessed cleaning and hygiene.

- Jowett 2022 compared CAD/CAM nitinol fixed retainer versus rectangular chain bonded retainer through the following question: “Was your retainer easy to look after?” No significant differences were found between the groups in the upper jaw (RD 0.00, 95% CI −0.06 to 0.06; P = 1.00; 62 participants; Analysis 2.14.1) and in the lower jaw (RD 0.03, 95% CI −0.05 to 0.12; P = 0.46; 62 participants; Analysis 2.14.2).
- Ferreira 2019 reported significant results. Participants found it easier to clean around the rigid retainer bonded only on the canines compared to the rectangular chain bonded to all the anterior teeth in the lower jaw (RD 1.00, 95% CI 0.88 to 1.12; P < 0.001; 30 participants; Analysis 2.14.3).

Comfort

Two studies (Ferreira 2019; Jowett 2022), assessed comfort.

- Jowett 2022 compared CAD/CAM nitinol fixed retainer versus rectangular chain bonded retainer through the following question: “Was your retainer comfortable?” and found no significant differences between the groups in the upper jaw (RD −0.06, 95% CI −0.17 to 0.04; P = 0.22; 62 participants; Analysis 2.15.1), or the lower jaw (RD 0.00, 95% CI −0.06 to 0.06; P = 1.00; 62 participants; Analysis 2.15.2).

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

- Ferreira 2019 reported no better comfort with the rigid retainer bonded only on the canines compared to the rectangular chain bonded retainer in the lower jaw (RD 0.20, 95% CI −0.15 to 0.55; P = 0.26; 30 participants; Analysis 2.15.3).

Speech

One study (Jowett 2022), that compared CAD/CAM nitinol fixed retainer and a rectangular chain bonded retainer, assessed speech through the following question: "Did your retainer affect speech?" with no significant differences found between the groups in the upper jaw (RD −0.03, 95% CI −0.12 to 0.05; P = 0.46; 62 participants; Analysis 2.16.1) and in the lower jaw (RD 0.00, 95% CI −0.06 to 0.06; P = 1.00; 62 participants; Analysis 2.16.2).

Additional appointments needed

One study (Jowett 2022), that compared CAD/CAM nitinol fixed retainer and a rectangular chain bonded retainer, assessed the need for additional appointments through the following question: "Did your retainer cause a problem that meant you needed to see your orthodontist?". No significant differences were found between the groups in the upper jaw (RR 2.67, 95% CI 0.78 to 9.12; P = 0.12; 62 participants; Analysis 2.17.1) and in the lower jaw (RR 1.00, 95% CI 0.27 to 3.65; P = 1.00; 62 participants; Analysis 2.17.2).

Aesthetics

One study (Scribante 2011), which compared fibre-reinforced composite versus multistrand fixed retainer, assessed aesthetics (e.g. "visibility of retainer while talking and smiling") using a visual analogue scale where 0 was a poor aesthetic result and 10 was an excellent result. It found a significant difference in favour of the fibre-reinforced composite retainer (MD 1.49, 95% CI 0.76 to 2.22; P < 0.001; 32 participants; Analysis 2.18.1).

Preference for type of retainer

One study (Ferreira 2019), reported that all participants preferred the rigid retainer bonded only on the canines compared to the rectangular chain bonded to multiple teeth in the lower jaw (RD 1.00, 95% CI 0.88 to 1.12; P < 0.001; 30 participants; Analysis 2.19.1).

Adverse effects on health

Eight studies reported different outcomes related to oral health, when comparing different types of fixed retainers (Adanur-Atmaca 2021; Alrawas 2020; Ferreira 2019; Gelin 2020; Kartal 2021; Sinha 2021; Torkan 2014; Węgrodzka 2021). All studies reported data only for the lower arch except Torkan 2014, which included both upper and lower arches. Studies reported the following outcomes.

- Plaque Index (Adanur-Atmaca 2021; Alrawas 2020; Ferreira 2019; Gelin 2020; Kartal 2021; Sinha 2021; Torkan 2014; Węgrodzka 2021; Analysis 2.20);
- Gingival Index (Adanur-Atmaca 2021; Alrawas 2020; Ferreira 2019; Gelin 2020; Kartal 2021; Sinha 2021; Torkan 2014; Węgrodzka 2021; Analysis 2.21);
- Calculus Index (Adanur-Atmaca 2021; Ferreira 2019; Gelin 2020; Torkan 2014; Analysis 1.22);
- Probing depth (Alrawas 2020; Kartal 2021; Węgrodzka 2021; Analysis 2.23);
- Bleeding on probing (Alrawas 2020; Kartal 2021; Torkan 2014; Węgrodzka 2021; Analysis 2.24);
- Marginal recession (Gelin 2020; Kartal 2021; Analysis 2.25).

Plaque Index

- Only one study (Torkan 2014), found higher Plaque Index scores at six months in the fibre-reinforced composite group compared to the spiral wire retainer group both in the upper (MD 1.00, 95% CI 0.54 to 1.46; P < 0.001; 40 participants; Analysis 2.20.1) and lower arches (MD 1.09, 95% CI 0.70 to 1.48; P < 0.001; 40 participants; Analysis 2.20.2).
- We pooled results from four studies (Adanur-Atmaca 2021; Alrawas 2020; Gelin 2020; Kartal 2021), that compared CAD/CAM nitinol versus multistrand stainless steel retainers at six months. Results showed similar Plaque Index in both groups (MD −0.09, 95% CI −0.18 to 0.01; P = 0.08; I^2 = 0%; 198 participants; Analysis 2.20.3).
- Similar non-significant differences were found between CAD/CAM nitinol, single-strand titanium and multistrand stainless steel retainers at six months (Alrawas 2020 – Analysis 2.20.4; Analysis 2.20.5).
- However, Sinha 2021 found lower Plaque Index score in the ceramic interlocking bonded group compared to the coaxial wire retainer (MD −0.40, 95% CI −0.56 to −0.24; P < 0.001; 40 participants; Analysis 2.20.6).
- The combined results from two studies (Adanur-Atmaca 2021; Gelin 2020), that compared CAD/CAM nitinol versus multistrand stainless steel retainers showed no differences at 12 months (MD −0.01, 95% CI −0.11 to 0.10; P = 0.91, I^2= 0%; 107 participants; Analysis 2.20.7).
- No differences were found for other types of lower fixed retainers neither (other arms of Adanur-Atmaca 2021; Węgrodzka 2021 – Analysis 2.20.8 to Analysis 2.20.13).

GIngival Index

Gingival Index scores were reported by the same group of studies as the Plaque Index scores.

- Torkan 2014 did not find a significant difference in the upper arch comparison (Analysis 2.21.1), but significantly higher scores were found in the fibre-reinforced composite group in the lower arch arches (MD 0.59, 95% CI 0.13 to 1.05; P = 0.47; 40 participants; Analysis 2.21.2).
- The pooled results for four studies (Adanur-Atmaca 2021; Alrawas 2020; Gelin 2020; Kartal 2021), showed lower Gingival Index scores in the CAD/CAM nitinol group compared to the multistrand retainer at six months (MD −0.14, 95% CI −0.27 to −0.01; P = 0.04; I^2 = 0%; 198 participants; Analysis 2.21.3). However, at 12 months, pooled comparison of two studies (Adanur-Atmaca 2021; Gelin 2020), showed no significant differences (MD 0, 95% CI −0.16 to 0.16; P = 0.97; I^2= 80%; 107 participants; Analysis 2.21.7).
- No significant differences were found for the other comparisons at six months (Alrawas 2020 – Analysis 2.21.4; Analysis 2.21.5; Sinha 2021 – Analysis 2.21.6), at 12 months (other arms of Adanur-Atmaca 2021 – Analysis 2.21.8 to Analysis 2.21.12) and 24 months (Węgrodzka 2021 – Analysis 2.21.13).

Calculus Index

Three studies reported the Calculus Index score in a similar way (Adanur-Atmaca 2021; Gelin 2020; Torkan 2014). No significant differences were found for all the comparisons that reported this outcome in the upper and lower arches.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

- Fibre-reinforced composite group versus spiral wire retainer (Torkan 2014 – Analysis 2.22.1; Analysis 2.22.2)
- CAD/CAM nitinol versus multistrand stainless steel retainers (Adanur-Atmaca 2021; Gelin 2020: MD −0.04, 95% CI −0.11 to 0.03; P = 0.25; I^2 = 61%; 107 participants; Analysis 2.22.3)
- Other arms of Adanur-Atmaca 2021 (Analysis 2.22.4 to Analysis 2.22.8).

Probing depth

Three studies (Alrawas 2020; Kartal 2021; Węgrodzka 2021), reported probing depth. No differences were found for the comparisons between CAD/CAM nitinol, single-strand titanium and multistrand stainless steel retainers in the lower arch (Analysis 2.23.1 to Analysis 2.23.3) at six months, and between round twisted multistrand and rectangular braided multistrand at 24 months (Analysis 2.23.4).

Bleeding on probing

Four studies reported comparisons for this outcome (Alrawas 2020; Kartal 2021; Torkan 2014; Węgrodzka 2021). No differences were found between the different types of retainers analysed: fibre-reinforced composite group versus spiral wire retainer (Torkan 2014 – Analysis 2.24.1; Analysis 2.24.2); CAD/CAM nitinol, single-strand titanium and multistrand stainless steel retainers (Alrawas 2020; Kartal 2021 – Analysis 2.24.3 to Analysis 2.24.5), and round twisted multistrand and rectangular braided multistrand (Węgrodzka 2021 – Analysis 2.24.6).

Marginal recession

We were able to combine results from two studies (Gelin 2020; Kartal 2021). The comparison between lower CAD/CAM nitinol versus multistrand stainless steel retainers at six months was not significantly different (MD −0.05, 95% CI −0.11 to 0.01; P = 0.11; 103 participants; Analysis 2.25.1).

We did not enter results from Ferreira 2019 into RevMan Web 2022 because they gave results as frequency and percentage of each index score (0, 1, 2, 3; Plaque, Gingival and Calculus) for each tooth surface. The main findings indicated that the scores for Plaque and Gingival indices were significantly higher for the rectangular chain retainer compared to the rigid retainer bonded only to 33 and 43 (P < 0.05) on the lingual and proximal surfaces. The same occurred with the Calculus Index on the lingual surfaces (P < 0.05). The Calculus Index along wire was higher for the rectangular chain wire too (P < 0.05).

Comparison 3: removable retainers versus removable retainers

Sixteen studies compared different types of removable retainers. In 13 studies, the comparisons were in the upper and lower arches (Al-Moghrabi 2020; Eroglu 2019; Gill 2007; Kaya 2019; Kumar 2011; Ramazanzadeh 2018; Rowland 2007; Saleh 2017; Shawesh 2010; Sun 2011; Thickett 2010; Wan 2017; Zhu 2017), and in three studies, the comparisons were only in the upper arch (Ashari 2021; Moslemzadeh 2018; Rohaya 2006). Studies included three types of removable retainers in these comparisons: Hawley retainer (11 studies), clear plastic retainers (15 studies) and Begg retainer (1 study).

Ten studies compared Hawley retainers with clear plastic retainers, but they were worn for different lengths of time per day (Ashari 2021; Eroglu 2019; Kaya 2019; Moslemzadeh 2018; Ramazanzadeh 2018; Rohaya 2006; Rowland 2007; Saleh 2017; Sun 2011; Wan 2017).

- Upper Hawley retainers and upper modified clear plastic retainers (0.040 inch, 1 mm) with palatal coverage for maintaining transverse expansion worn full-time for the first six months, followed by niight-time wear for the next six months (Ashari 2021)
- Upper and lower Hawley and clear plastic retainers (thickness not mentioned) worn full-time for 13 weeks (Eroglu 2019)
- Upper and lower Hawley and clear plastic retainers (0.040 inch, 1 mm), worn full-time for the first six months, followed by niight-time wear for the next six months (Kaya 2019)
- Upper Hawley and two types of upper clear plastic retainers of different thickness (1 mm and 1.5 mm, respectively) worn full-time for the six months (Moslemzadeh 2018)
- Upper and lower Hawley and clear plastic retainers (1 mm), worn full-time for the first four months, followed by niight-time wear for the next four months (Ramazanzadeh 2018)
- Upper Hawley retainers full-time for three months then six months nights only versus upper clear plastic (1 mm) retainers worn full-time for one week then part-time (3 months evenings and nights and 6 months nights only, total 9 months; Rohaya 2006)
- Upper and lower Hawley retainers worn full-time for three months then nights only (12 hours per day) versus upper and lower clear plastic retainers (1.5 mm) worn full-time for one week then part-time (12 hours per day; Rowland 2007)
- Upper and lower Hawley and clear plastic retainers (1 mm), worn full-time for the first six months, followed by niight-time wear for the next six months (Saleh 2017)
- Upper and lower Hawley retainers versus upper and lower clear plastic (0.75 mm) retainers worn full-time for 12 months (Sun 2011)
- Upper and lower Hawley retainers versus upper and lower clear plastic (0.8 mm) retainers worn full-time for three months (Wan 2017)
- One study compared part-time wear of upper and lower Hawley retainers (nights only) with full-time wear of Hawley retainers (Shawesh 2010)

One study compared upper and lower clear plastic (thickness not mentioned) retainers with Begg retainers (Kumar 2011). The participants were instructed to wear their retainers full-time for six months then part-time for six months (12 hours per day).

Two studies compared part-time wear of upper and lower clear plastic retainers with full-time wear of clear retainers (thickness not mentioned) (Gill 2007; Thickett 2010). Gill 2007 defined part-time as eight hours per day, and Thickett 2010 defined part-time as 10 hours per day.

One study compared upper and lower clear plastic retainers of different thicknesses (0.75 mm versus 1.00 mm) worn full-time for 12 months (Zhu 2017).

One study compared upper and lower clear plastic retainers (1.00 mm) worn full-time with and without access to “My retainers” mobile application reporting different outcomes at three and 12 months (Al-Moghrabi 2020).

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Stability

Studies assessed stability in a variety of ways.

- Little's Irregularity Index for each arch (Al-Moghrabi 2020; Gill 2007; Kaya 2019; Ramazanzadeh 2018; Rowland 2007; Shawesh 2010; Thickett 2010), for both arches (Kumar 2011)
- Changes in intercanine width (Ashari 2021; Gill 2007; Kaya 2019; Ramazanzadeh 2018; Rowland 2007)
- Changes in interpremolar width (Ashari 2021)
- Changes in intermolar width (Ashari 2021; Gill 2007; Kaya 2019; Ramazanzadeh 2018; Rowland 2007)
- Changes in arch length (Kaya 2019; Ramazanzadeh 2018)
- Overjet and overbite (Gill 2007; Kaya 2019)
- Stability of corrected rotations (Rohaya 2006)
- Outcome of treatment measured by PAR index (Kumar 2011)
- Outcome of treatment measured by ABO grading (Moslemzadeh 2018)

Although 13 studies that compared removable retainers with other removable retainers evaluated stability, only two studies compared similar interventions and similar outcomes (Gill 2007; Thickett 2010). However, it was not possible to pool the results in a meta-analysis as the data in the Thickett 2010 study were presented as medians, without interquartile ranges, and the study author did not respond to our request for these data.

Little's Irregularity index

Eight studies reported this index when comparing removable retainers (Analysis 3.1).

Hawley retainers versus clear plastic retainer

- Ramazanzadeh 2018 compared Hawley retainers worn full-time versus clear retainers worn full-time and part-time (3-arm study).
  - Upper arch: at four months, they found no differences between Hawley and clear retainers (Analysis 3.1.1; Analysis 3.1.3).
  - Lower arch: the clear plastic group showed significantly less crowding than Hawley group when the retainer was worn full-time (MD 0.10 mm, 95% CI 0.02 to 0.18; P = 0.02; 55 participants; Analysis 3.1.2), but not when it was worn part-time (Analysis 3.1.4). After four months, retainers were worn part-time and stability outcomes were assessed again after eight months. The upper Irregularity Index change was significantly lower in the two clear plastic groups compared to the Hawley group: four months full-time clear plastic (MD 0.38 mm, 95% CI 0.22 to 0.54; P value < 0.0001; 55 participants; Analysis 3.1.7), and four months part-time clear plastic (MD 0.45 mm, 95% CI 0.29 to 0.61; P < 0.001; 55 participants; Analysis 3.1.9). Similar significant differences were found in the lower arch at eight months: after four months full-time wear of a clear plastic retainer (MD 0.17 mm, 95% CI 0.03 to 0.31; P = 0.02; 55 participants; Analysis 3.1.8), and four months part-time wear of a clear plastic retainer (MD 0.21 mm, 95% CI 0.07 to 0.35; P = 0.003; 55 participants; Analysis 3.1.10).
- Rowland 2007 compared Hawley retainers worn full-time for three months and part-time for the next three months with clear plastic retainers worn part-time at six months.
  - Upper arch: they found a difference in the Irregularity Index in the upper arch, with 0.25 mm (95% CI 0.08 to 0.42; P = 0.004) more relapse in the Hawley retainer group (Analysis 3.1.13),
  - Lower arch: the difference in the Irregularity Index in the lower arch was 0.42 mm (95% CI 0.23 to 0.61; P < 0.001; Analysis 3.1.14).
- Kaya 2019 compared Hawley retainers versus clear plastic retainers, both of them worn full-time for six months and part-time for the next six months and found no significant differences in the upper (LII MD -0.47 mm, 95% CI −1.06 to 0.12; P=0.12; 30 participants; Analysis 3.1.17) and the lower arches (LII MD 0.01 mm, 95% CI −0.65 to 0.67; P = 0.98; 30 participants; Analysis 3.1.18) at 12 months.

Begg retainers versus clear plastic retainers

Kumar 2011 compared upper and lower Begg retainers with upper and lower clear plastic retainers (both groups had a fixed retainer in addition in the lower arch) at six months. There was a significant difference of 0.25 mm (95% CI 0.19 to 0.31; P < 0.001; Analysis 3.1.23) with more irregularity in the group wearing the Begg retainers.

Full-time versus part-time wear of clear plastic retainers

Three studies (Gill 2007; Ramazanzadeh 2018 (two arms); Thickett 2010) compared full-time wear of clear plastic retainers versus part-time wear of clear plastic retainers.

- Gill 2007 showed no statistical difference in terms of irregularity Index, in both upper and lower arches at six months (Analysis 3.1.15; Analysis 3.1.16).
- Ramazanzadeh 2018 showed the same non-significant differences in the upper (Analysis 3.1.5) and lower arches (Analysis 3.1.6) at four months and at eight months after both groups wore the clear retainers part-time for the second four-month period (Analysis 3.1.11; Analysis 3.1.12).
- Table 2 presents the results from Thickett 2010. They found no statistical difference in the irregularity index at 12 months.

While we were unable to combine the results in a meta-analysis, the finding of the three studies was that there was no evidence of a difference in stability when clear plastic retainers were worn full-time or part-time.

- Al-Moghrabi 2020 did not find significant differences at 12 months in the Irregularity Index of the upper (Analysis 3.1.19) and lower arches (Analysis 3.1.20) when clear plastic retainers were worn full-time with or without access to “My retainers” mobile application.

Full-time versus part-time wear of Hawley retainers

Shawesh 2010 compared full-time wear of Hawley retainers (6 months worn full-time and 6 months worn part-time) with part-time wear of Hawley retainers at 12 months. They showed no statistical difference in Irregularity Index in both arches between the two wear regimens (upper: MD 0.00 mm, 95% CI −0.41 to 0.41; P = 1.00; 1 study, 52 participants; Analysis 3.1.21) (lower: MD 0.20 mm, 95% CI −0.28 to 0.68; P = 0.41; 1 study, 52 participants; Analysis 3.1.22).

Intercanine, interpremolar and intermolar width changes

Six studies reported intercanine width changes (Analysis 3.2), one study reported interpremolar width changes (Analysis 3.3), and six

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

reported intermolar width changes (Analysis 3.4), when comparing removable retainers.

Hawley retainers versus clear plastic retainers

- Ramazanzadeh 2018 found no significant differences in intercanine and intermolar widths between groups when comparing Hawley retainers worn full-time versus clear retainers worn full-time and part-time (3-arm study) at four months (Analysis 3.2.1 to Analysis 3.2.4; Analysis 3.4.1 to Analysis 3.4.4); and after eight months, when retainers were used part-time after the initial four-month period (Analysis 3.2.7 to Analysis 3.2.10; Analysis 3.4.7 to Analysis 3.4.10).
- Ashari 2021 compared upper Hawley retainers and upper modified clear plastic retainers with palatal coverage. They found no differences in intercanine (Analysis 3.2.13), interpremolar (Analysis 3.3.1) and intermolar (Analysis 3.4.13; Analysis 3.4.14) width changes when worn full-time for the first six months, nor when this period was followed by night-time wear for the next six months (Analysis 3.2.14; Analysis 3.3.2; Analysis 3.4.15; Analysis 3.4.16).
- Rowland 2007 compared full-time Hawley retainers with part-time clear plastic retainers and found no difference in intercanine (Analysis 3.2.15; Analysis 3.2.16) and intermolar widths (Analysis 3.4.17; Analysis 3.4.18) in both arches at six months.
- Kaya 2019 compared Hawley retainers versus clear plastic retainers, both of them worn full-time for six months and part-time for the next six months, and found no significant differences in the upper intercanine (Analysis 3.2.19) and intermolar width (Analysis 3.4.21) changes, and in the lower arches (Analysis 3.2.20; Analysis 3.4.22) at 12 months.

Full-time versus part-time wear of clear plastic retainers

Three studies (Gill 2007; Ramazanzadeh 2018 (2 arms); Thickett 2010) compared intercanine (Analysis 3.2.17; Analysis 3.2.18; Analysis 3.2.5; Analysis 3.2.6; Analysis 3.2.11; Analysis 3.2.12; Table 2) and intermolar width changes (Analysis 3.4.19; Analysis 3.4.20; Analysis 3.4.5; Analysis 3.4.6; Analysis 3.4.11; Analysis 3.4.12; Table 2) and found no differences between groups in the upper and lower arches.

Arch length changes

Three studies (Kaya 2019; Ramazanzadeh 2018; Thickett 2010), reported arch length changes (Analysis 3.5, Table 2). All three studies found no differences between groups in the upper and lower arches when comparing Hawley and clear plastic retainers (Ramazanzadeh 2018; Thickett 2010; Analysis 3.5.1 to Analysis 3.5.12; Table 2) or clear plastic retainers worn full-time versus part-time (Kaya 2019; Analysis 3.5.13; Analysis 3.5.14).

Overjet and overbite

Three studies (Kaya 2019; Gill 2007; Thickett 2010), reported overjet and overbite changes (Analysis 3.6; Analysis 3.7, Table 2). No differences were found between groups when comparing Hawley and clear plastic retainers (Gill 2007; Thickett 2010; Analysis 3.6.1; Analysis 3.7.1; Table 2) or clear plastic retainers worn full-time versus part-time (Kaya 2019; Analysis 3.6.2, Analysis 3.7.2).

Rotations

Rohaya 2006 investigated the ability to maintain previously rotated teeth in a stable position, comparing Hawley retainers and clear plastic retainers in the upper arch. They found that clear plastic retainers were 4.88 times more likely to hold derotated teeth stable than Hawley retainers (95% CI 1.13 to 21.07; P = 0.03; Analysis 3.8). This is likely to be clinically significant.

Maintaining outcome of treatment

- Kumar 2011 evaluated this outcome using the PAR index. There was a reduction in the quality of the result of 1.71 points (95% CI 1.44 to 1.98; P < 0.001; Analysis 3.9) on the PAR index in the Begg retainer group, indicating more relapse in this group than in the clear plastic retainer group.
- Moslemzadeh 2018 evaluated this outcome using the ABO grading (Analysis 3.10). They found no differences between the upper Hawley retainer group and the groups wearing upper clear plastic retainers of different thicknesses (Analysis 3.10.1 to Analysis 3.10.3).

Failure of retainers

Three studies reported failure of retainers for the upper and lower arches together (Rowland 2007; Zhu 2017), or measured per arch (Sun 2011). Studies measured failures in different ways.

- Rate of failure in %:
  - total number of failures (Sun 2011; Zhu 2017; Analysis 3.11);
  - number of different type of failure: (Analysis 3.12 to Analysis 3.15);
    - fracture, loss, non-fitting, and abrasion (Sun 2011; Zhu 2017);
    - broken, lost, non-fitting retainers (Rowland 2007);
- Mean survival time in days (Zhu 2017; Analysis 3.16);
- Risk of failure (hazard ratio) (Sun 2011; Zhu 2017; Analysis 3.17).

Hawley retainers versus clear plastic retainers

Sun 2011 compared the failure rate of Hawley retainers and clear plastic (0.75 mm) retainers worn full-time for 12 months. They assessed the failure rate separately for upper and lower retainers. For comparisons with other studies, we combined upper and lower data. Sun 2011 showed no difference in total failure rates in the upper arch (Analysis 3.11.1). However, in the lower arch there was an increased chance of clear plastic retainers failing compared with Hawley retainers (RR 0.60, 95% CI 0.43 to 0.83; P = 0.002; 111 participants; Analysis 3.11.2), which is clinically significant. When we combined upper and lower total failures, we also found a significant result (RR 0.72, 95% CI 0.57 to 0.90; P = 0.005; 222 participants; Analysis 3.11.3).

- Reasons for failure
  - Broken retainers: rates were similar in the upper and lower arches together (Analysis 3.12.2), and in the upper arch only (Analysis 3.12.3). In the lower arch, fractures were-higher in the 0.75 mm thickness clear plastic group than in the Hawley group (RR 0.51, 95% CI 0.30 to 0.85; P = 0.01; 111 participants; Analysis 3.12.4).
  - Lost retainers: rates were similar between groups in the upper arch (Analysis 3.13.3) and for the lower arch (Analysis 3.13.4). However, when upper and lower failures were added together, significantly larger number of clear plastic retainers

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

were lost compared to Hawley retainers (RR 0.55, 95% CI 0.31 to 0.97; P = 0.04; 222 participants; Analysis 3.13.2).
    - Retainers not fitting (Analysis 3.14.2 to Analysis 3.14.4), and retainers with serious abrasions (Analysis 3.15.1 to Analysis 3.15.3): there were no significant differences between the two types of retainers.
- The risk of failure according to the hazard ratio was not significantly different for the Hawley retainer group compared to the clear plastic retainer group Analysis 3.17.1), and for maxillary retainers (Hawley and clear plastic) compared to the mandibular retainers (Analysis 3.17.2).
- Rowland 2007 also compared failure rates of upper and lower Hawley retainers worn full-time for three months then nights only versus upper and lower clear plastic retainers (1.5 mm) worn part-time at six months. They assessed failure rates by recording how many retainers broke, how many were lost, and how many did not fit. Hawley retainers were more likely to break, giving a RR of 3.21 (95% CI 1.55 to 6,64; P = 0.002; Analysis 3.12.1), which is clinically significant. There was no difference in the number of retainers that were lost (Analysis 3.13.1) or did not fit (Analysis 3.14.1).

Different thickness of clear plastic retainers

Zhu 2017 compared the failure rate of 1 mm-thick clear plastic and 0.75 mm-thick clear plastic (0.75 mm) retainers worn full-time for 12 months.

- Total failure rates: Zhu 2017 showed no difference in the total failure rates between the 1 mm- and 0.75 mm-thick retainers for the upper and lower arches together (Analysis 3.11.4).
- Reasons for failure
    - Rate of broken retainers: this was significantly lower in the 1 mm-thick retainer group (RR 0.23, 95% CI 0.05 to 1.01; P = 0.05; 72 participants; Analysis 3.12.5).
    - Number of lost retainers (Analysis 3.13.5): there were no significant differences between the two thickness of retainers.
    - Retainers not fitting (Analysis 3.14.5): there were no significant differences between the two thickness of retainers.
    - Retainers with serious abrasions (Analysis 3.15.4):there were no significant differences between the two thickness of retainers.
- Mean survival time in days (Analysis 3.16.1) was not significantly different between thicknesses.
- Risk of failure according to the hazard ratio (Analysis 3.17.3) was not significantly different between thicknesses.

Patient satisfaction

Three studies investigated patient satisfaction when comparing Hawley retainer versus clear plastic retainer (Rowland 2007; Saleh 2017; Wan 2017) using the following core categories.

Acceptance

Two studies reported significant results for acceptance of the retainer. Participants found it more acceptable to wear clear plastic retainers than Hawley retainers in all the analyses.

- Saleh 2017 reported this core category through the following two issues:
    - "appearance" (MD −2.77, 95% CI −3.24 to −2.30; P < 0.001; 86 participants; Analysis 3.18.1) and
    - "self-confidence" (MD −2.18, 95% CI −2.56 to −1.80; P < 0.001; 86 participants; Analysis 3.18.2).
- Rowland 2007 reported this core category through the following questions:
    - "embarrassed to wear your retainer?" (R 2.42, 95% CI 1.30 to 4.49; P = 0.005; 348 participants; Analysis 3.19.1) and
    - "worse or much worse than wearing fixed appliance?" (RR 9.37, 95% CI 3.80 to 23.10; P < 0.001; 349 participants; Analysis 3.19.2).

Although both studies reported six-month outcomes, we could not pool these results statistically because the wearing protocol was different in both studies, and one study reported continuous data (Saleh 2017), while the other reported dichotomous data (Rowland 2007).

Compliance

- One study (Rowland 2007) reported that the clear plastic retainer group were more likely to report being able to wear the retainer than the Hawley group, with an RR of 3.03 (95% CI 1.46 to 6.30; P = 0.003; 347 participants; Analysis 3.20.1).
- There was no difference in the ability to wear the retainers away from home (Analysis 3.20.2).

Cleaning and hygiene

- One study (Saleh 2017), assessed cleaning and hygiene through the following two issues:
    - "hygiene", with no differences (MD −0.42, 95% CI −0.89 to 0.05; P = 0.08; 86 participants; Analysis 3.21.1) and
    - "gingival irritation", with the clear plastic retainer less irritating to the gingiva than the Hawley retainer (MD −1.28, 95% CI −1.73 to −0.83; P < 0.001; 86 participants; Analysis 3.21.2).

Comfort

Two studies (Rowland 2007; Saleh 2017), assessed comfort.

- Saleh 2017 reported this core category through the following four issues:
    - "comfort" with differences favouring clear plastic retainer (MD −1.86, 95% CI −2.19 to −1.53; P < 0.001; 86 participants; Analysis 3.22.1);
    - "biting with retainers" with differences favouring Hawley retainer (MD 0.91, 95% CI 0.44 to 1.38; P < 0.001; 86 participants; Analysis 3.22.2);
    - "fitting" with no differences between the two types of retainers (Analysis 3.22.3) and
    - "swallowing" with differences favouring clear plastic retainer (n = 86; MD −1.28, 95% CI −1.68 to −0.88; P < 0.001; Analysis 3.22.4).
- Rowland 2007 reported no difference between the two retainers in the amount of discomfort experienced (RR 1.43, 95% CI 0.77 to 2.65; P = 0.25; 349 participants; Analysis 3.23.1).

Although both studies reported six-month comfort-related outcomes, we could not pool these results statistically because one study reported continuous data (Saleh 2017), and the other

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

dichotomous data (Rowland 2007). Furthermore, the wearing protocol (full-time or part-time) differed between studies.

Speech

Two studies (Saleh 2017; Wan 2017), assessed speech.

- Saleh 2017 reported that participants with the clear plastic retainer felt that they were less likely to cause a speech problem compared to the Hawley retainer (MD −1.25, 95% CI −1.69 to −0.81; P < 0.001; 86 participants; Analysis 3.24.1).
- Wan 2017 reported no difference in the number of sound distortions (RR 3.00, 95% CI 0.12 to 72.68; P = 0.50; 180 participants; Analysis 3.25).

Durability

One study (Saleh 2017), assessed durability, with differences favouring the Hawley retainer (MD 1.83, 95% CI 1.07 to 2.59; P < 0.001; 86 participants; Analysis 3.26.1).

Adverse effects on health

Three studies reported different outcomes related to oral health (Al-Moghrabi 2020; Eroglu 2019; Moslemzadeh 2018), when comparing different types of removable retainers. Studies reported the following outcomes.

- Plaque Index (Al-Moghrabi 2020; Eroglu 2019; Analysis 3.27)
- Gingival Index (Eroglu 2019; Moslemzadeh 2018; Analysis 3.28)
- Probing depth (Al-Moghrabi 2020; Eroglu 2019; Analysis 3.29)
- Bleeding on probing (Al-Moghrabi 2020, Eroglu 2019; Analysis 3.30)

Hawley retainers versus clear plastic retainers

- Eroglu 2019 found no differences in Plaque Index scores at three months when upper and lower full-time Hawley retainers and full-time clear plastic retainers were compared (Analysis 3.27.1). Gingival Index (Analysis 3.28.4), probing pocket depth (Analysis 3.29.1), and bleeding on probing (Analysis 3.30.1) were similar for the upper and lower arches when full-time Hawley retainers and full-time clear plastic retainers were compared at 3 months.
- Moslemzadeh 2018 compared three types of removable retainers in the upper arch worn full-time for six months: Hawley retainer, 1.5 mm clear plastic retainer and 1 mm clear plastic retainer in the upper arch. No significant different Gingival Index scores were found between groups (Analysis 3.28.1 to Analysis 3.28.3).

Clear plastic retainers with and without the use of mobile app

Al-Moghrabi 2020 found no significant differences in Plaque Index scores (Analysis 3.27.2; Analysis 3.27.3), probing pocket depth (Analysis 3.29.2; Analysis 3.29.3), and bleeding on probing (Analysis 3.30.2; Analysis 3.30.3), for the upper and lower arches between both retainer groups at 12 months, whether using the mobile app or not.

Comparison 4: Upper clear plastic (UCP) + lower fixed canine to canine versus UCP + lower interproximal reduction (stripping) versus positioner

Upper removable and lower fixed retainer versus upper removable retainer and lower adjunctive procedure

Edman Tynelius 2010 compared an upper clear plastic retainer worn part-time and lower rigid stainless steel retainer from lower canine to canine with an upper clear plastic retainer worn part-time and the adjunctive procedure of interproximal reduction in the lower arch, with no lower retainer. They measured stability at one and two years of retention. Change data were reported for Little's Irregularity Index in both arches, upper and lower intercanine and intermolar width, arch length, overjet and overbite.

- After one year:
  - there were differences in the lower intercanine width (MD 1.00 mm, 95% CI 0.78 to 1.22; P < 0.001), which may be clinically significant (Analysis 4.6.1); in the lower arch length (MD 0.70 mm, 95% CI 0.2 to 1.20; P = 0.006; Analysis 4.8.1) and in overbite (MD −0.60, 95% CI −1.17 to -0.03; P = 0.04; Analysis 4.10.1), which are unlikely to be clinically significant.
  - There was no difference in the Little's Irregularity Index in either arch, intercanine width in the upper arch, intermolar width in either arch, arch length in the upper arch, and overjet (Analysis 4.1.1; Analysis 4.2.1; Analysis 4.3.1; Analysis 4.4.1; Analysis 4.5.1; Analysis 4.7.1; Analysis 4.9.1).
- After two years:
  - there were differences in the lower intercanine width (MD 1.20 mm, 95% CI 0.74 to 1.66; P value < 0.0001), which may be clinically significant (Analysis 4.6.2); and in overjet (MD −0.8 mm, 95% CI −1.43 to −0.17; P = 0.01; Analysis 4.9.2) and overbite (MD −0.80 mm, 95% CI −1.37 to −0.23; P = 0.006; Analysis 4.10.2), which are unlikely to be clinically significant.
  - There was no difference in the Little's Irregularity Index in either arch, intercanine width in the upper arch, intermolar width in either arch, arch length in either arch, overjet or overbite (Analysis 4.1.2; Analysis 4.2.2; Analysis 4.3.2; Analysis 4.4.2; Analysis 4.5.2; Analysis 4.7.2; Analysis 4.8.2).

Upper removable and lower fixed retainer versus positioner

Edman Tynelius 2010 compared an upper clear plastic retainer worn part-time and lower rigid stainless steel retainer from lower canine to canine with a prefabricated positioner covering the teeth in the maxilla and mandible worn part-time. They measured stability at one and two years of retention. Change data were reported for Little's Irregularity Index in both arches, upper and lower intercanine and intermolar width, arch length, overjet and overbite.

- After one year:
  - there were differences in the lower Little's Irregularity Index (MD −0.50 mm, 95% CI −1.01 to 0.01; P = 0.05; Analysis 4.5.3), lower intercanine width (MD 0.70 mm, 95% CI 0.26 to 1.14; P = 0.002; Analysis 4.6.3) and lower intermolar width (MD 0.80 mm, 95% CI 0.06 to 1.54; P = 0.03; Analysis 4.7.3), which are unlikely to be clinically significant.
  - There was no difference in the upper Little's Irregularity Index, intercanine width in the upper arch, intermolar width in the upper, arch length in either arch, overjet or

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

overbite (Analysis 4.1.3; Analysis 4.2.3; Analysis 4.3.3; Analysis 4.4.3; Analysis 4.8.3; Analysis 4.9.3; Analysis 4.10.3).

- After two years:
  - there was a difference in the upper intercanine width (MD 0.80 mm, 95% CI 0.10 to 1.50; P = 0.03; Analysis 4.2.4), but this is unlikely to be clinically significant; lower Little's Irregularity Index (MD −1.00 mm, 95% CI −1.65 to −0.35; P = 0.003; Analysis 4.5.4), which may be clinically significant; lower intercanine width (MD 1.30 mm, 95% CI 0.76 to 1.84; P < 0.001; Analysis 4.6.4), which may be clinically significant; lower intermolar width (MD 1.00 mm, 95% CI 0.07 to 1.93; P = 0.03; (Analysis 4.7.4), which may be clinically significant; and overjet (MD −0.70 mm, 95% CI −1.31 to −0.09; P = 0.02; Analysis 4.9.4), but this is unlikely to be clinically significant.
  - There was no difference in the Little's Irregularity Index in the upper arch, upper intermolar width, arch length in either arch, and overbite (Analysis 4.1.4; Analysis 4.3.4; Analysis 4.4.4; Analysis 4.8.4 and Analysis 4.10.4).

Upper removable retainer and lower adjunctive procedure versus positioner

Edman Tynelius 2010 compared an upper clear plastic retainer worn part-time and lower adjunctive procedure of interproximal stripping with a positioner worn part-time. They measured stability at one and two years of retention. Change data were reported for Little's Irregularity Index in both arches, upper and lower intercanine and intermolar width, arch length, overjet and overbite.

- After one year:
  - there was a difference in upper intercanine width (MD 1.00 mm, 95% CI −0.01 to 2.01; P value = 0.05; Analysis 4.2.5), which is clinically significant, and in lower intercanine width (MD 0.70 mm, 95% CI 0.26 to 1.14; P = 0.002; Analysis 4.6.5) and lower arch length (MD −0.50 mm, 95% CI −0.95 to −0.05; P = 0.03; Analysis 4.8.5), but these are unlikely to be clinically significant.
  - There was no difference in the Little's Irregularity Index in both arches, intermolar widths in both arches, upper arch length, overjet and overbite (Analysis 4.1.5; Analysis 4.3.5; Analysis 4.4.5; Analysis 4.5.5; Analysis 4.7.5; Analysis 4.9.5; and Analysis 4.10.5).
- After two years:
  - there was a difference in upper intercanine width (MD 0.90 mm, 95% CI 0.17 to 1.63; P = 0.02; Analysis 4.2.6); and in lower irregularity index (MD −0.70 mm, 95% CI −1.37 to −0.03; P = 0.04; Analysis 4.5.6), which are not clinically significant.
  - There was no difference in Little's Irregularity Index in the upper arch, intercanine width in the lower arch, intermolar width in both arches, arch length in both arches, overjet and overbite (Analysis 4.1.6; Analysis 4.3.6; Analysis 4.4.6; Analysis 4.6.6; Analysis 4.7.6; Analysis 4.8.6; Analysis 4.9.6; and Analysis 4.10.6).

Same authors reported at five years follow-up (Edman Tynelius 2010). However, these results are not included in this review because after two years, all retainers were stopped and bonded retainers removed.

Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines

Naraghi 2021a compared an upper retainer versus no retention after alignment of maxillary impacted canines. They measured stability in the maxilla at one year, using the following change-from-baseline-outcomes: Little's irregularity index, arch length, intercanine width, intermolar width and maximum rotation.

- After one year, the non-retention group showed significantly higher irregularity index (MD −0.90 mm, 95% CI −1.33 to −0.47; P < 0.001; 63 participants; Analysis 5.1), and decreased arch length (MD 0.30 mm, 95% CI 0.03 to 0.57; P = 0.03; 63 participants; Analysis 5.4) compared to the retention group. These results were not clinically significant. The most rotated teeth were maintained better with retainers (MD −1.70°, 95% CI −3.20 to −0.20; P = 0.03; 63 participants; Analysis 5.5). There was no difference in intercanine width (Analysis 5.2) and intermolar width (Analysis 5.3) between the two groups.

DISCUSSION

Summary of main results

Please note that the evidence summarised here is only low to very low certainty, and therefore ongoing and future research may very well change these findings.

Comparison 1: Removable retainers versus fixed retainers

Stability

We found limited evidence to suggest that multistrand fixed retainers were slightly better at maintaining straight teeth as measured by Little's Irregularity Index (LII) than removable clear plastic retainers in the upper arch (1 study, 60 participants, part-time use (Naraghi 2021b)), and in the lower arch (1 study full-time use (Millett 2007); 2 studies, part-time use (Forde 2018; O'Rourke 2016)). These three studies of 182 participants in total showed significant differences favouring multistrand/coaxial fixed retainers when the follow-up was 12 months or longer (regardless of the regimen of use).

Participants wearing removable retainers showed a slightly greater increase in the intercanine width compared to those with the fixed retainers, especially in the upper arch (1 study, part-time use (Naraghi 2021b); 2 studies, full-time use (Alrawas 2020; O'Rourke 2016)), but there was no difference in maintaining intermolar width, overjet and overbite. Arch length was not significantly different between removable and fixed retainers in one study of part-time use (Forde 2018). Relapse of rotated teeth was higher in the removable compared to the fixed retainer group, when bonded to the incisors and canines (1 study, part-time use (Naraghi 2021b)).

There was no difference in stability between clear retainers and CAD/CAM nickel-titanium wires (1 study, Alrawas 2020), or clear plastic retainers and rigid retainers bonded only to the canine teeth (1 study, Krämer 2019).

Failure

There was conflicting evidence about failure rates of removable clear plastic retainers versus multistrand fixed retainers (1 study, full-time use (Millett 2007); 1 study, part-time use (Forde 2018)). When the removable retainers were used full-time, more failures

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

were reported with the removable retainers than the fixed retainers. However, the opposite result was found when the removable retainer was used part-time. An interesting finding was that there appeared to be a significantly higher risk of patients no longer wearing the removable part-time retainers compared to the percentage of failures of fixed retainers after four years (O'Rourke 2016).

Satisfaction

There was some evidence to show that patient satisfaction was higher with fixed retainers than clear plastic retainers. Participants reported that the fixed retainers were more acceptable and comfortable to wear, speech was easier, and they found it easier to comply with regimens of use (3 studies, Forde 2018; Krämer 2019; Millett 2007). Both retainers were perceived by study participants as equally efficient (1 study, Krämer 2019). It remains unclear which retainers are easier to keep clean, as two studies report opposite results (part-time use, Forde 2018; full-time use, Millett 2007). Participants who used removable retainers part-time found them easier to clean (1 study, Forde 2018), but those who wore the removable retainer full-time found it more difficult to clean than participants who had fixed multistrand retainers (1 study, Millett 2007).

Oral health

There was limited evidence to suggest that bonded retainers have the potential to lead to worse oral health. Evidence of caries was not significantly different when comparing fixed retainers and removable clear plastic retainers (1 study, Millett 2007). Plaque Index scores were significantly lower in the part-time clear plastic retainer group versus multistrand fixed retainer group in the upper arch, but no different in the lower arch (1 study, Forde 2018). Higher Gingival Index and Calculus Index scores were found in the lower arch when fixed multistrand retainers were used compared to part-time use of clear plastic retainers (Forde 2018). The rest of included studies comparing different types of removable and fixed retainers did not show differences in periodontal measurements. Only one study (Forde 2018, 60 participants), reported significant differences; three studies (130 participants with 3, 6 and 48 months of follow-up (Alrawas 2020; Eroglu 2019; O'Rourke 2016)) found no differences, which could be explained by the fact that participants performed better oral care due to their participation in a research study (Hawthorne effect; Braunholtz 2001).

In only one study, periodontal pocketing and bleeding on probing, recorded as a dichotomous variable, was significantly higher in the lower fixed retainer group compared to those wearing lower removable clear plastic full-time (Millett 2007, follow-up 12 months). However, this periodontal pocketing result should be viewed with caution as the two groups were not equal at debond, with four times as many participants having pocketing in the fixed retainer group. Two RCTs (73 participants) with three and six months of follow-up, did not find differences between clear plastic retainers and multistrand retainers (Alrawas 2020; Eroglu 2019). On the other hand, bleeding on probing was slightly higher in the lower arch for clear plastic group used full-time compared to the fixed CAD/CAM nitinol fixed retainer (Alrawas 2020). No significant differences in oral health were found between full-time Hawley removable retainers and multistrand fixed (bonded) retainer (1 study, Eroglu 2019), except for a clinically insignificant difference in bleeding on probing, favouring the fixed multi-strand retainer over the Hawley retainer.

Comparison 2: Fixed retainers versus fixed retainers

Stability

Most of the studies that compared fixed retainers found similar Irregularity Index results regardless of the type of retainer used. Less relapse was found when the upper multistrand retainer was bonded to the six anterior teeth rather than only the four incisors (Naraghi 2021b). Indirectly bonded multistrand retainers showed less relapse than the directly bonded coaxial retainers (Gunay 2018). However, more relapse was found in the round multistrand retainer group compared to the fibre-reinforced composite group when the follow-up reached 12 months (Nagani 2020). Based on one study, lower CAD/CAM nitinol and round stainless steel multistrand retainers had less relapse than rectangular multistrand and connected bonding pad retainer on incisors and canines at 12 months (Adanur-Atmaca 2021). Intercanine width changes were similar in most studies except for one that found significantly greater decrease in the rectangular chain group compared to the CAD/CAM nitinol retainer group (Jowett 2022). Similar results were found when comparing the rectangular chain boned retainer group with CAD/CAM multistrand stainless steel and laboratory-made multistrand retainers (Shim 2022). Interpremolar and intermolar widths, anterior and total arch lengths, and overbite and overjet changes were similar among all comparisons. Relapse of rotation was higher in the retainer bonded to four incisors compared to the retainer bonded to the six anterior upper teeth (Naraghi 2021b).

Failure

We were able to carry out four meta-analyses for this outcome: upper and lower fibre-reinforced composite versus multistrand fixed retainer showed no significant differences in failure rates. Nor did the comparisons between lower CAD/CAM nitinol versus multistrand retainers. Lower thick plain wire (canine to canine, bonded only on the canines) showed significantly fewer failures than thin spiral wire bonded to all incisors and canines at 24 months (Arash 2020), but not at 36 months (Årtun 1987). CAD/CAM nitinol fixed retainers showed higher failure rates than rectangular chain fixed retainers in the upper arch but not in the lower arch (Jowett 2022). Regarding the bonding materials, a significantly lower failure rate was found when a conventional orthodontic composite was used compared to a flowable composite (Scribante 2020), and significantly more bonding failures were reported when a conventional orthodontic primer was used compared to a universal primer in total-etch mode (Sfrondini 2021). No differences were found when using chemically or light-cured adhesives (Pandis 2013). Many of the included studies found no difference in failure rates of different types of fixed retainers. Regarding mean survival time, two studies reported significant differences (Rose 2002; Soboutia 2016). The fibre-reinforced composite fixed retainer showed the shortest survival time, followed by the multistrand spiral wire and the twisted ligature wire retainers. Sfrondini 2014 found opposite results, with a significantly higher failure risk for the multistrand spiral wire versus fibre-reinforced composite. Jowett 2022 found significantly higher failure risk for the CAD/CAM nitinol versus the multistrand wire fixed retainer in the upper arch, but not in the lower arch, which is in agreement with the results reported for failure rate.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Satisfaction

There was some evidence to show similar patient satisfaction between the rigid retainer (bonded only on the canines), the CAD/CAM nitinol and the rectangular chain retainer in terms of efficacy, comfort, speech and number of appointments required to resolve any retainer issues. The rigid retainer seems to be preferred by patients compared to the rectangular chain, highlighting its ease of cleaning (Ferreira 2019). No differences were found in the cleaning capacity of rectangular chain retainers and CAD/CAM nitinol retainers (Jowett 2022). Fibre-reinforced composite retainer was found to be aesthetically more favourable while smiling or talking compared to the multistrand retainer (Scribante 2011).

Oral health

One RCT (40 participants) suggested that oral health was worse with fibre-reinforced composite retainers compared to spiral wire retainers. Plaque Index score was higher at six months in the ribbon retainer group, both in the upper and lower arches, and Gingival Index scores were significantly higher in the lower arch too. We were able to meta-analyse four studies that compared CAD/CAM nitinol versus multistrand stainless steel retainers at six months, with non-significant results (Adanur-Atmaca 2021; Alrawas 2020; Gelin 2020; Kartal 2021). There was no heterogeneity, and they showed similar Plaque Index scores in both groups. One study reported lower Plaque Index scores in a ceramic interlocking bonded group compared to a coaxial wire retainer group (Sinha 2021). Only this study included this type of ceramic interlocking bonded retainer. At 12 months, no significant difference was found among different types of lower fixed retainers.

Meta-analysis of four RCTs showed significantly lower Gingival Index scores in the CAD/CAM nitinol group compared to the multistrand retainer at six months (Adanur-Atmaca 2021; Alrawas 2020; Gelin 2020; Kartal 2021). However, at 12 months, pooled comparison of two studies showed no difference between the four types of fixed retainers (Adanur-Atmaca 2021; Gelin 2020). Further evidence is necessary to obtain conclusive results for these comparisons.

Other comparisons of different types of fixed retainers did not show differences regarding Plaque and Gingival indices. No significant differences were found for the comparisons that reported Calculus Index, probing depth and bleeding on probing changes, and marginal recessions using different types of fixed retainers. In one study, which reported data that could not be entered into RevMan, all oral indices (Plaque, Gingival and Calculus) showed significantly higher values (frequency and percentage) with rectangular chain retainers compared to rigid retainers bonded only canine to canine, suggesting better oral health with the retainers bonded only to the canine teeth (Ferreira 2019).

Comparison 3: Removable retainers versus removable retainers

Stability

We could not perform meta-analysis for any stability outcome due to the different regimens of use of removable retainers and different follow-ups.

Part-time wear versus full-time wear of removable retainers

Two studies failed to find a difference in stability between part-time and full-time wear of clear plastic retainers (moderate-certainty evidence) (Gill 2007; Thickett 2010). There was limited evidence from one study that part-time wear of Hawley retainers was as effective for preventing relapse of irregular and crowded teeth as full-time wear (Shawesh 2010).

Hawley retainers versus clear plastic retainers

There was some evidence to suggest that clear plastic retainers offer better retention in the upper and lower arches than Hawley retainers. One study showed a small increase in Irregularity Index in both arches with Hawley retainers even though clear plastic retainers were worn part-time from the beginning of the retention period (Rowland 2007). Unless the Irregularity Index was restricted to one tooth contact, particularly in the upper arch, this small increase in the irregularity index was unlikely to be significant. Similar results were seen in some arms of Ramazanzadeh 2018, who combined different regimens of use. No differences between the different types of removable retainers were reported in intercanine, interpremolar, and intermolar width changes, arch length, overjet, or overbite changes. Both retainer types maintained the outcome of treatment results in a similar way. One study showed that clear plastic retainers prevented relapse of previously rotated teeth in the upper arch better than Hawley retainers (Rohaya 2006).

Begg retainers versus clear plastic retainers

There was limited evidence from one study that clear plastic retainers may retain the outcome of orthodontic treatment better than Begg retainers (Kumar 2011).

Monitoring via mobile application

One study compared clear plastic retainers with and without adjunctive use of a mobile app and found no differences in Irregularity Index (Al-Moghrabi 2020).

Failure

Hawley retainers versus clear plastic retainers

We found limited evidence from two studies that could not be combined. In one study (Rowland 2007), when Hawley retainers worn three months full-time then nine months part-time, were compared to clear plastic retainers worn part-time, there were more breakages with Hawley retainers. In another study (Sun 2011), when both sets of retainers were worn full-time, there were more failures and more losses in the clear plastic retainer group. However, hazard ratios (failure risks) were similar for both types of full-time removable retainers, and for the comparison between upper and lower arches.

Different thickness of clear plastic retainers

Failure rates and failure risk (hazard ratio) of clear plastic retainers worn full-time did not differ between the 0.75 mm thickness group and the 0.1 mm thickness group, but, breakages were higher in the thinner group. Mean survival time was similar for both thicknesses of retainers (Zhu 2017).

Satisfaction

Patient satisfaction and self-confidence was higher with clear plastic retainers compared to Hawley retainers (Saleh 2017;

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rowland 2007, Wan 2017). Participants reported that clear plastic retainers were more acceptable and comfortable to wear, easier to clean and to speak with than Hawley retainers, and that they were more likely to comply with regimens of use. The Hawley retainer was favoured for its durability and better settling of the teeth in occlusion when wearing the retainer (Saleh 2017). No differences were found between the two retainers in pain or discomfort (Rowland 2007).

Oral health

Plaque and Gingival indices, and probing depth and bleeding on probing were similar among the types of removable retainers compared (Hawley versus clear plastic retainers and clear retainers with and without adjunctive use of a mobile app) (Al-Moghrabi 2020).

Comparison 4: Upper removable clear plastic and lower fixed retainer (canine to canine) versus upper removable clear plastic and lower adjunctive procedure (interproximal reduction) versus positioner

Based on the results of one study (Edman Tynelius 2010), all three approaches maintained the teeth in a reasonably stable position for the first year. However, after two years, there was greater irregularity in the lower arch with the positioner. The fact that adequate stability was maintained in the group with an adjunctive procedure, but no lower retainer is interesting, because in recent decades it has been presumed that all participants had to wear some sort of retainer to maintain stability. It is important to emphasise that the participants initially presented with normal jaw relationships in all three dimensions (anteroposterior, vertical and transverse proportions), and so were essentially Class I crowding cases. Typically, teeth do not need to be moved as far in these cases. When trying to straighten teeth and obtain a good bite when the jaws do not meet together well, teeth need to be moved further to compensate for this. As a result, it may not be possible to generalise these results to most people undergoing orthodontic treatment.

Comparison 5: Upper fixed retainer versus no retention after treatment of maxillary impacted canines

One study (Naraghi 2021a) found that while upper arch length increased in the retention group and decreased in the non-retention group, and corrected rotations were better maintained in the retention group, there was no difference between groups in stability (Little's Irregularity Index) or intercanine and intermolar width.

Overall completeness and applicability of evidence

The included studies reported a considerable variety of different types of fixed and removable retainers, different follow-up periods, and different regimens of use for removable retainers (full- and part-time wear over different periods of time). Some studies involved people who had completed growth, while others still had some residual growth remaining. Therefore, pooled comparisons were almost never possible.

Overall, in this update we included 48 studies, (33 more than in the previous version that included 15 studies) investigating different approaches to retention and multiple outcomes. We included only RCTs in this review, which we recognise is only a fraction of the total available body of evidence.

Most of the included studies evaluated one or more of three main comparisons: removable retainers versus fixed retainers; different types of fixed retainers or bonding materials; and different types of removable retainers. One study compared a combination of upper clear plastic and lower bonded retainers versus upper clear plastic retainers with lower adjunctive procedures versus positioner (Edman Tynelius 2010). It should be noted that there were very strict inclusion criteria for this study and, while the study was well conducted with a low risk of bias, the findings may only be applicable to people with similar malocclusions (i.e. cases of tooth crowding with normal positions of the jaws in all three dimensions). One study compared the use of upper fixed retention versus no retention at all after treatment of maxillary impacted canines (Naraghi 2021a).

There are currently no agreed Core Outcome Measures in Effectiveness Trials (COMET) for relapse studies in orthodontics. The review authors, drawing on their clinical experience, chose to prioritise Little's Irregularity Index as the preferred measure of the primary outcome of this review. Only three of the 48 studies reported outcomes beyond two years; nine reported results at two years. This is important as retainers may be needed indefinitely to reduce the risk of relapse, but the findings of this review predominantly inform clinicians and patients of the short-term effects of different approaches to retention.

When interpreting the results of this review, it is important to remember that the results of the studies may be affected by the age of the participants and whether they are still growing, type and severity of initial malocclusions, treatment procedures, such as extractions, and amount and type of tooth movement. The impact of these factors is still not fully understood. We identified 24 ongoing studies that may be included in the next update of this review (see Characteristics of studies awaiting classification and Characteristics of ongoing studies). Hopefully, the inclusion of these studies in future updates of this review will allow for the consideration of these confounding factors.

Quality of the evidence

We report the overall certainty of the evidence in the summary of findings tables (Summary of findings 1 to Summary of findings 6). It is mostly very low for the primary and secondary outcomes, downgraded mainly due to risk of bias, imprecision, and small numbers of RCTs contributing to the main findings. Whilst we identified many different treatment options, generally there was a shortage of good-quality RCTs available for each comparison; in many instances, there is only one study per comparison. When there was high heterogeneity, we downgraded for inconsistency. Most of the studies were at high risk of detection bias because it is very difficult to blind assessors in these types of interventions.

Future research is very likely to have an important impact on our confidence in the estimate of the effect and likely to change the estimate. Although many studies have been added since the last update of this review, we still have very low confidence in the findings and so any conclusions that could be drawn are tentative. Moreover, recent developments in techniques and materials led to an increased number of RCTs, which normally tend to overestimate the effect of the new intervention (novelty bias; Seehra 2021), instead of replicating existing preliminary conclusions.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Potential biases in the review process

This review update adheres to the standards for the planning, conduct and reporting of updates of Cochrane Intervention Reviews as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2022), and in the MECIR (Methodological Expectations of Cochrane Intervention Reviews) standards (Higgins 2021a).

We used a broad, sensitive search strategy of multiple databases, and also searched for unpublished studies and data, without language restrictions. Therefore, we hope that potential publication bias in the review process has been minimised. We made an effort to contact the main authors of ongoing studies (registered protocols) and were able to obtain some preliminary unpublished data. We also tried to contact authors of included studies for any missing data.

We could not add data from one of the studies to RevMan (RevMan Web 2022), as they were reported as medians without interquartile ranges (Thickett 2010). We presented these results separately in Table 2.

Agreements and disagreements with other studies or reviews

Removable versus fixed retainers

Our review found no moderate- or high-certainty evidence on which to base recommendations, but, to date, part-time removable retainers have shown slightly less stability than fixed retainers, although, when a full-time protocol was used, the difference was minor or even clinically irrelevant. People wearing clear plastic removable retainers had slightly better periodontal health after one year compared to people wearing fixed retainers.

Al-Moghrabi 2016 conducted a systematic review, including RCTs and prospective cohort studies, to determine the effects of fixed and removable orthodontic retainers on periodontal health, survival and failure rates of retainers, cost-effectiveness, and the impact of orthodontic retainers on patient-reported outcomes. Similar to our review, Al-Moghrabi 2016 found a lack of evidence to endorse the use of one type of orthodontic retainer, due to a lack of high-quality, relevant research, and concluded that further, well-designed prospective studies are needed to elucidate the benefits and potential harms associated with orthodontic retainers.

Bellini-Pereira 2022 conducted another systematic review comparing treatment stability, periodontal effects, and failure rates of fixed versus clear plastic retainers. The authors concluded that both retainers had the same efficacy in terms of maintaining stability in the long term. Fixed retainers were associated with greater plaque and calculus accumulation than clear plastic ones after 12 months. Fixed retainers presented greater failure rates in the lower arch. The Bellini-Pereira 2022 authors mixed different types of bonded fixed retainers, and due to the substantial clinical and methodological heterogeneity between studies, quantitative analysis was not feasible.

Fixed versus fixed retainers

This review found no moderate- or high-certainty evidence on which to base recommendations, but, to date, in comparison to multistrand fixed retainers, participants were more satisfied with fibre-reinforced composite fixed retainers. The fibre-reinforced composite fixed retainers demonstrated slightly more stability but were more likely to fail earlier and showed higher gingival inflammation than multistrand fixed retainers.

Iliadi 2015 conducted a systematic review that evaluated the risk of failure of fixed orthodontic retention protocols. RCTs, prospective and retrospective studies were included. The inclusion of retrospective studies is likely to have biased the findings. Moreover, only two studies that compared fibre-reinforced composite versus multistrand wire retainers were considered in the meta-analysis, while our review included seven studies (Analysis 2.10.14). Authors concluded that the quality of the available evidence was low. No conclusive evidence was found in order to guide orthodontists in the selection of the best protocol.

Buzatta 2017 conducted a systematic review that compared the gingival parameters associated with fixed retainer designs that allow unobstructed interproximal flossing, with the ones that do not (plain versus waved retainers). Four studies (one cross-sectional and three non-RCTs) were included, and showed equivocal results. Authors concluded that there was not enough scientific evidence either to support or to refute an association between the retainer design and gingival health, flossing frequency, or patient comfort.

Arn 2020 conducted a systematic review that evaluated the effects of fixed orthodontic retainers on periodontal health. It included RCTs, prospective cohort studies, and cross-sectional studies. Results were heterogeneous. The authors concluded that orthodontic fixed retainers seem to be a retention strategy compatible with periodontal health, or at least, not related to severe detrimental effects on the periodontium.

Jedliński 2021 conducted a systematic review that evaluated failures rates of fixed retainers. It included RCTs, prospective cohort studies and retrospective studies. Jedliński 2021 concluded that no wire or fibre splint presented superior characteristics in terms of failure rate.

Removable versus removable retainers

According to our review, there is no moderate- or high-certainty evidence from which to draw firm conclusions, but, to date, Hawley removable retainers when worn part-time resulted in lower stability and poorer patient satisfaction than clear plastic retainers, although they did survive longer. They showed similar stability when worn full-time. Regarding the wearing protocol of removable retainers, there was no difference in stability for part-time or full-time wear of either clear plastic retainers or Hawley retainers, but their effects on oral health or patient satisfaction have not yet been evaluated.

Mai 2014 conducted a systematic review that compared Hawley retainers to clear plastic retainers. They included RCTs, quasi-randomised and controlled clinical studies. They reported that there were no differences with respect to changes in intercanine and intermolar widths between both retainers, with similar results to our conclusions.

Al Rahma 2018 conducted a systematic review to investigate the performance of the Hawley-type retainers compared to other types of removable retainers and wearing schedules. Although 10 RCTs were included, authors concluded that, due to the multitude of parameters, which may have affected the results of the included

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

studies, good practice would suggest further research is required to increase both the quantity and quality of information available.

Outhaisavanh 2020 conducted a systematic review that compared the use of Hawley and clear plastic retainers for maintaining tooth position following fixed orthodontic appliance. The conclusion of their qualitative synthesis was similar to our findings based on two different studies with different wear regimes: clear plastic retainers provided slightly better relapse prevention of incisor irregularity than Hawley retainers in both arches.

Li 2021 conducted a systematic review that compared periodontal health with Hawley versus clear plastic retainers. The authors used endpoint comparisons rather than change-from-baseline outcomes. This may affect the results as it does not consider the between-person variability. Li 2021 concluded from the quantitative synthesis that people using the Hawley retainer had slightly better gingival health and a lower Plaque Index than those using clear plastic retainers, but the clinical relevance of these findings is questionable (the mean difference ranged from −0.01 mm to 0.15 mm).

AUTHORS' CONCLUSIONS

Implications for practice

The evidence is only low or very low certainty for retention procedures.

Part-time removable retainers showed slightly less stability than fixed retainers, but not of a clinically significant amount, and when a full-time wear protocol was used for the removable retainers, the difference was clinically irrelevant. People wearing clear plastic removable retainers had slightly better periodontal health after one year compared to those wearing fixed retainers.

Regarding new design and manufacturing technologies for fixed retainers, computer-aided design/computer-aided manufacturing (CAD/CAM) nitinol fixed retainers showed no clinically meaningful difference in stability and periodontal health at one year, and no difference in survival compared to analogue multistrand fixed retainers.

Regarding the use of aesthetic fixed retainers (made of fibre-reinforced composite), in comparison to multistrand fixed retainers, participants were more satisfied with fibre-reinforced composite fixed retainers. The fibre-reinforced composite fixed retainers demonstrated slightly more stability, though not of a clinically significant amount, and they were more likely to fail earlier and showed higher levels of gingival inflammation than multistrand fixed retainers.

Regarding different types of removable retainers, Hawley removable retainers worn part-time resulted in poorer stability and patient satisfaction than clear plastic retainers, but they did survive longer. Both types showed similar stability when worn full-time.

Regarding the wearing protocol of removable retainers, there was no difference in stability between part-time and full-time wear of clear plastic retainers or Hawley retainers.

Overall, there is no moderate- or high-certainty evidence from which to draw conclusions on retention procedures for stabilising tooth position after treatment with orthodontic braces. High-quality RCTs with well-planned protocols are needed that follow up participants for at least two years. These studies should measure and report the outcomes tooth stability, retainer survival, adverse effects (caries and periodontal disease) and patient satisfaction, and compare different retainers and different wear protocols.

Implications for research

Retention studies are not easy to undertake, but several randomised controlled clinical studies have now been completed showing that this research is feasible. Relapse is a long-term problem and long-term follow-up of participants is practically difficult and financially demanding. However, given that the vast majority of people requiring orthodontic treatment undergo a phase of retention, this vital area of orthodontic research should continue to be given priority. To minimise the risk of bias, future studies in this field should have the following features:

- adequate allocation concealment and appropriate generation of randomisation;
- blinding of outcome assessors;
- full and accurate reporting and analysis of withdrawals and dropouts;
- reporting of all data collected;
- a priori sample size calculations;
- clear inclusion and exclusion criteria;
- follow-up for at least two years, given the long-term nature of the problem of relapse.

Although some randomised controlled trials (RCTs) have addressed aspects of retention given below, further high-quality studies in these areas are needed to increase the evidence base:

- to compare different types of retainers (fixed and removable);
- to investigate the effects of different adjunctive techniques, such as pericision and interproximal reduction;
- to further investigate whether adjunctive techniques, for example interproximal reduction, may be sufficient without retainers to provide adequate retention;
- to investigate levels of patient satisfaction with different retention regimens, including no retention.

Appropriate outcomes to investigate include:

- stability;
- survival of retainers;
- adverse effects on oral health;
- patient satisfaction.

In particular, it would be useful to assess these outcomes over the long term.

There are studies currently ongoing, some of which fulfil these criteria. If completed, these will be included in the next update of this review.

ACKNOWLEDGEMENTS

For the 2023 update, we would like to thank Laura MacDonald (Managing Editor, Cochrane Oral Health) for assistance in the development of the review and managing the editorial process, Anne Littlewood (Information Specialist, Cochrane Oral Health) for

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

developing the search strategy and running the searches. We thank Anne-Marie Glenny (Co-ordinating Editor, Cochrane Oral Health), Jayne Harrison (Editor, Cochrane Oral Health) and the rest of Cochrane Oral Health for their kind assistance in completing this update. We thank John Fricker and Ross Hobson for peer review feedback, and Denise Mitchell for copy editing.

We would also like to express our sincere thanks to Akshaya Pandian, Asma Ashari, Dalya Al Moghrabi, Despina Koletsi, Farhan Bazargani, Hari Christian Petsos, Marie A Cornelis, Mohammad Y Hajeer, Norhidayah Mohd Tahir, Paul Schneider, Piotr S Fudalej, Raghda Alaa, Rungtiwa Jirawiwatsaree, Silvio Bellini-Pereira, Vinay Kumar Chugh, Wan Nurazreena Wan Hassan, Yanqi Yang; as the corresponding authors of ongoing and published studies who have shared, reported or clarified relevant data on their research for the completion of this review.

For previous versions, we would like to thank Sue Furness, Anne Littlewood, Philip Riley, Sylvia Bickley, Emma Tavender, Luisa Fernandez, Laura MacDonald, Anne-Marie Glenny and the rest of the Cochrane Oral Health Group, and referees Susan Cunningham, Hend Elsayed and John Fricker.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

REFERENCES

References to studies included in this review

Adanur-Atmaca 2021 *{published data only}*

Adanur-Atmaca R, Çokakoğlu S, Öztürk F. Effects of different lingual retainers on periodontal health and stability. *Angle Orthodontist* 2021;**91**(4):468-76. [DOI: 10.2319/110220-904.1.]

Al-Moghrabi 2020 *{published data only}***10.1093/ejo/cjz088**

* Al-Moghrabi D, Pandis N, McLaughlin K, Johal A, Donos N, Fleming PS. Evaluation of the effectiveness of a tailored mobile application in increasing the duration of wear of thermoplastic retainers: a randomized controlled trial. *European Journal of Orthodontics* 2020;**42**(5):571-9. [DOI: 10.1093/ejo/cjz088]

Zhang X, Al-Moghrabi D, Pandis N, Shah S, Fleming PS. The effectiveness of a bespoke mobile application in improving adherence with removable orthodontic retention over 12 months: a randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2022;**161**(3):327-37. [DOI: 10.1016/j.ajodo.2021.09.010]

Alrawas 2020 *{published and unpublished data}*

Alrawas MB, Kashoura Y, Tosun Ö, Öz U. Comparing the effects of CAD/CAM nickel-titanium lingual retainers on teeth stability and periodontal health with conventional fixed and removable retainers: a randomized clinical trial. *Orthodontics & Craniofacial Research* 2020;**24**(2):241-50. [DOI: 10.1111/ocr.12425]

Arash 2020 *{published data only}***10.1186/s40510-020-00315-7**

Arash V, Teimoorian M, Farajzadeh Jalali Y, Sheikhzadeh S. Clinical comparison between multi-stranded wires and single strand ribbon wires used for lingual fixed retainers. *Progress in Orthodontics* 2020;**21**(1):22. [DOI: 10.1186/s40510-020-00315-7]

Årtun 1987 *{published and unpublished data}*

Artun 1987 [pers comm]. Cochrane systematic review on orthodontic retention [personal communication]. Letter to Dr Artun (lead author of paper) 24 March 2002.

* Årtun J, Spadafora AT, Shapiro PA, McNeill RW, Chapko MK. Hygiene status associated with different types of bonded, orthodontic canine-to-canine retainers. A clinical trial. *Journal of Clinical Periodontology* 1987;**14**(2):89-94.

Årtun J, Spadafora AT, Shapiro PA. A 3-year follow-up study of various types of orthodontic canine-to-canine retainers. *European Journal of Orthodontics* 1997;**19**(5):501-9.

Ashari 2021 *{unpublished data only}*

Ashari 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Asma Ashari 8 July 2021, 26 August 2022.

* Ashari A, Xian L, Syed Mohamed AMF, Wahab RMA, Kit YC, Tata MD, et al. One-year comparative assessment of retention of arch width increases between modified vacuum-formed and Hawley retainers: a multicenter randomized clinical trial. *Angle Orthodontics* 2022;**92**(2):197-203. [DOI: 10.2319/050921-363.1]

Bolla 2012 *{published and unpublished data}*

Bolla 2012 [pers comm]. Thanks for the article…and 4 very quick questions about the excellent research! [personal communication]. Email to: Mauro Cozzani 6 March 2013.

* Bolla E, Cozzani M, Doldo T, Fontana M. Failure evaluation after a 6-year retention period: a comparison between glass fiber-reinforced (GFR) and multistranded bonded retainers. *International Orthodontics* 2012;**10**:16-28.

Edman Tynelius 2010 *{published data only (unpublished sought but not used)}*

Edman Tynelius 2010 [pers comm]. Cochrane systematic review on orthodontic retention [personal communication]. Email. Details unavailable.

* Edman Tynelius G, Bondemark L, Lilja-Karlander E. Evaluation of orthodontic retention capacity after one year of retention - a randomized controlled trial. *European Journal of Orthodontics* 2010;**32**(5):542-7.

Edman Tynelius G, Bondemark L. A randomized controlled trial of three orthodontic retention methods in Class I four premolar extraction cases - stability after 2 years in retention. *Orthodontics and Craniofacial Research* 2013;**16**:105-15.

Edman Tynelius G, Petrén S, Bondemark L, Lilja-Karlander E. Five-year postretention outcomes of three retention methods - a randomized controlled trial. *European Journal of Orthodontics* 2015;**37**(4):345-53. [DOI: 10.1093/ejo/cju063]

Eroglu 2019 *{published data only}***10.1016/j.ajodo.2018.08.022**

Eroglu AK, Baka ZM, Arslan U. Comparative evaluation of salivary microbial levels and periodontal status of patients wearing fixed and removable orthodontic retainers. *American Journal of Orthodontics and Dentofacial Orthopedics* 2019;**156**(2):186-92. [DOI: 10.1016/j.ajodo.2018.08.022]

Ferreira 2019 *{published data only}***10.1590/2177-6709.24.3.064-070.oar**

Ferreira LA, Sapata DM, Provenzano MGA, Hayacibara RM, Ramos AL. Periodontal parameters of two types of 3 x 3 orthodontic retainer: a longitudinal study. *Dental Press Journal of Orthodontics* 2019;**24**(3):64-70. [DOI: 10.1590/2177-6709.24.3.064-070.oar]

Forde 2018 *{published and unpublished data}*

* Forde K, Storey M, Littlewood SJ, Scott P, Luther F, Kang J. Bonded versus vacuum-formed retainers: a randomized controlled trial. Part 1: stability, retainer survival, and patient satisfaction outcomes after 12 months. *European Journal of Orthodontics* 2018;**40**(4):387-98.

Storey M, Forde K, Littlewood SJ, Scott P, Luther F, Kang J. Bonded versus vacuum-formed retainers: a randomized controlled trial. Part 2: periodontal health outcomes after 12 months. *European Journal of Orthodontics* 2018;**40**(4):399-408.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gelin 2020 *{published and unpublished data}*10.4041/kjod.2020.50.6.373

Gelin E, Seidel L, Bruwier A, Albert A, Charavet C. Innovative customized CAD/CAM nickel-titanium lingual retainer versus standard stainless-steel lingual retainer: a randomized controlled trial. *Korean Journal of Orthodontics* 2020;**50**(6):373-82. [DOI: 10.4041/kjod.2020.50.6.373]

Gera 2022 *{unpublished data only}*

Gera 2022 [pers comm]. Monitoring an ongoing study's progress [personal communication]. Email to: Marie Cornelis 14 June 2022.

Gera A, Pullisaar H, Cattaneo PM, Gera S, Vandevska V, Cornelis MA. CAD/CAM fixed retainers vs. conventional multistranded fixed retainers in orthodontic patients. Comparison of posttreatment stability, retainer failures, and patient satisfaction. A 6-month follow-up of a two-center randomized controlled clinical trial. *European Journal of Orthodontics* 2023;**45**(1):58-67.

Gill 2007 *{published and unpublished data}*

Gill 2007 [pers comm]. Retention WFO paper for Cochrane review - a few quick questions! [personal communication]. Email to: Daljit S Gill 10 January 2010.

Gill DS, Naini FB, Jones A, Tredwin CJ. Part-time versus full-time retainer wear following fixed appliance therapy: a randomized prospective controlled trial. *World Journal of Orthodontics* 2007;**8**:300-6.

Gunay 2018 *{published and unpublished data}*

Gunay F, Oz AA. Clinical effectiveness of 2 orthodontic retainer wires on mandibular arch retention. *American Journal of Orthodontics and Dentofacial Orthopedics* 2018;**153**(2):232-8.

Jowett 2022 *{published and unpublished data}*

Jowett A [pers comm]. Data from Jowett 2021 [personal communication]. Email to: C Martin 12 October 2021.

Jowett A, Littlewood S, Hodge T, Dhaliwal H, Wu J. CAD/CAM nitinol bonded retainer versus a chairside rectangular chain bonded retainer: a multicentre randomised controlled trial. *Journal of Orthodontics* 2022;**0**(0):14653125221118935. [DOI: 10.1177/14653125221118935]

Kartal 2021 *{published data only}*10.1007/s00056-020-00243-5

Kartal Y, Kaya B, Polat-Özsoy Ö. Comparative evaluation of periodontal effects and survival rates of Memotain and five-stranded bonded retainers. *Journal of Orofacial Orthopedics* 2021;**82**(1):32-41. [DOI: 10.1007/s00056-020-00243-5]

Kaya 2019 *{published data only}*10.5152/TurkJOrthod.2019.18045

Kaya Y, Tunca M, Keskin S. Comparison of two retention appliances with respect to clinical effectiveness. *Turkish Journal of Orthodontics* 2019;**32**(2):72-8. [DOI: 10.5152/TurkJOrthod.2019.18045]

Krämer 2019 *{published data only}*10.1093/ejo/cjz072

Krämer 2019 [pers comm]. Urgent data request [personal communication]. Email to: Anke Krämer 12 Jan 2022.

* Krämer A, Sjöström M, Hallman M, Feldmann I. Vacuum-formed retainer versus bonded retainer for dental stabilization in the mandible—a randomized controlled trial. Part I: retentive capacity 6 and 18 months after orthodontic treatment. *European Journal of Orthodontics* 2019;**42**(5):551-8. [DOI: 10.1093/ejo/cjz072]

Krämer A, Sjöström M, Hallman M, Feldmann I. Vacuum-formed retainers and bonded retainers for dental stabilization—a randomized controlled trial. Part II: patients' perceptions 6 and 18 months after orthodontic treatment. *European Journal of Orthodontics* 2021;**43**(2):136-43. [DOI: 10.1093/ejo/cjaa039]

Kumar 2011 *{published and unpublished data}*

Kumar AG, Bansal A. Effectiveness and acceptability of Essix and Begg retainers: a prospective study. *Australian Orthodontic Journal* 2011;**27**:52-6.

Millett 2007 *{published and unpublished data}*

McDermott P, Field D, Erfida I, Millett DT. Operator and patient experiences with fixed or vacuum formed retainers. In: Irish Division IADR Conference Abstract 0017. Cork, 2007.

McDermott P, Millett DT, Field D, Van den Heuvel A, Erfid I. Lower incisor retention with fixed or vacuum formed retainers. In: IADR Conference Abstract 0642. Toronto, 2008.

Millett 2007 [pers comm]. Queries about Cork retention studies [personal communication]. Email to: D Millett 10 January 2010.

* Millett DT, McDermott P, Field D, Erfida I, Doubleday B, Vandenheuvel A, et al. Dental and periodontal health with bonded or vacuum-formed retainer. In: IADR Conference Abstract 3168. Toronto, 2008.

Moslemzadeh 2018 *{published and unpublished data}*

Moslemzadeh SH, Sohrabi A, Rafighi A, Farshidnia S. Comparison of stability of the results of orthodontic treatment and gingival health between Hawley and vacuum-formed retainers. *Journal of Contemporary Dental Practice* 2018;**19**(4):443-9.

Nagani 2020 *{published data only}*10.1186/s12903-020-01167-7

* Nagani NI, Ahmed I, Tanveer F, Khursheed HM, Farooqui WA. Clinical comparison of bond failure rate between two types of mandibular canine-canine bonded orthodontic retainers- a randomized clinical trial. *BMC Oral Health* 2020;**20**(1):180. [DOI: 10.1186/s12903-020-01167-7.]

Nagani NI, Ahmed I. Effectiveness of two types of fixed lingual retainers in preventing mandibular incisor relapse. *Journal of College of Physicians and Surgeons Pakistan* 2020;**30**(3):282-6. [DOI: 10.29271/jcpsp.2020.03.282]

Naraghi 2021a *{published data only}*10.1093/ejo/cjaa010

Naraghi S, Ganzer N, Bondemark L, Sonesson M. Comparison of post-treatment changes with and without retention in adolescents treated for maxillary impacted canines—a randomized controlled trial. *European Journal of Orthodontics* 2021;**43**(2):121-7. [DOI: 10.1093/ejo/cjaa010]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Naraghi 2021b *{published data only}*10.1093/ejo/cjaa077

* Naraghi S, Ganzer N, Bondemark L, Sonesson M. Stability of maxillary anterior teeth after 2 years of retention in adolescents: a randomized controlled trial comparing two bonded and a vacuum-formed retainer. *European Journal of Orthodontics* 2021;**43**(2):152-8. [DOI: 10.1093/ejo/cjaa077]

Sonesson M, Naraghi S, Bondemark L. Cost analysis of two types of fixed maxillary retainers and a removable vacuum-formed maxillary retainer: a randomized controlled trial. *European Journal of Orthodontics* 2022;**44**(2):197-202. [DOI: 10.1093/ejo/cjab080]

O'Rourke 2016 *{published and unpublished data}*

Al-Moghrabi D, Johal A, O'Rourke N, Donos N, Pandis N, Gonzales-Marin C, et al. Effects of fixed vs removable orthodontic retainers on stability and periodontal health: 4-year follow-up of a randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2018;**154**(2):167-74.e161.

O'Rourke 2016 [pers comm]. Possible duplicated publication [personal communication]. Email to: Dalya Al-Moghrabi 2 May 2022.

* O'Rourke N, Albeedh H, Sharma P, Johal A. Effectiveness of bonded and vacuum-formed retainers: a prospective randomized controlled clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2016;**150**:406-15.

Pandis 2013 *{published and unpublished data}*

Pandis 2013 [pers comm]. Data request for a systematic review [personal communication]. Email to: Piotr S. Fudalej and Nikolaos Pandis 30 April 2022.

Pandis N, Fleming PS, Kloukos D, Polychronopoulou A, Katsaros C, Eliades T. Survival of bonded lingual retainers with chemical or photo polymerization over a 2-year period: a single-center, randomized controlled clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2013;**144**:169-75.

Ramazanzadeh 2018 *{published and unpublished data}*

Ramazanzadeh B, Ahrari F, Hosseini ZS. The retention characteristics of Hawley and vacuum-formed retainers with different retention protocols. *Journal of Clinical & Experimental Dentistry* 2018;**10**(3):e224-31.

Rohaya 2006 *{published and unpublished data}*

Rohaya 2012 [pers comm]. Quick clarification point from your retention study with Rohaya [personal communication]. Email to: jangus.scott@btinternet.com 13 June 2012.

Rohaya MA, Shahrul Hisham ZA, Doubleday B. Randomised clinical trial: comparing the efficacy of vacuum-formed and Hawley retainers in retaining corrected tooth rotations. *Malaysian Dental Journal* 2006;**27**(1):38-44.

Rose 2002 *{published and unpublished data}*

Rose 2002 [pers comm]. Cochrane systematic review on orthodontic retention [personal communication]. Email details unavailable.

Rose E, Frucht S, Jonas IE. Clinical comparison of a multistranded wire and a direct-bonded polyethylene ribbon-reinforced resin composite used for lingual retention. *Quintessence International* 2002;**33**(8):579-83.

Rowland 2007 *{published and unpublished data}*

Hichens L, Rowland H, Williams A, Hollinghurst S, Ewings P, Clark S, et al. Cost-effectiveness and patient satisfaction: Hawley and vacuum-formed retainers. *European Journal of Orthodontics* 2007;**29**:372-8.

* Rowland H, Hichens L, Williams A, Hills D, Killingback N, Ewings P, et al. The effectiveness of Hawley and vacuum-formed retainers: a single-center randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2007;**132**:730-7.

Saleh 2017 *{published and unpublished data}*10.1093/ejo/cjx024

Saleh 2017 [pers comm]. Data request [personal communication]. Email to: Mohammad Y Hajeer 12 January 2022.

Saleh M, Hajeer MY, Muessig D. Acceptability comparison between Hawley retainers and vacuum-formed retainers in orthodontic adult patients: a single-centre, randomized controlled trial. *European Journal of Orthodontics* 2017;**3**(4):453-61. [DOI: 10.1093/ejo/cjx024]

Salehi 2013 *{published data only}*

Salehi P, Najafi HZ, Roeinpeikar SM. Comparison of survival time between two types of orthodontic fixed retainer: a prospective randomized clinical trial. *Progress in Orthodontics* 2013;**14**:25.

Scribante 2011 *{published data only}*

Scribante A, Sfondrini MF, Broggini S, D'Allocco M, Gandini P. Efficacy of esthetic retainers: clinical comparison between multistranded wires and direct-bond glass fiber-reinforced composite splints. *International Journal of Dentistry* 2011;**2011**:548356. [DOI: 10.1155/2011/548356]

Scribante 2020 *{published data only}*10.3390/polym12040963

Scribante A, Gallo S, Turcato B, Trovati F, Gandini P, Sfondrini MF. Fear of the relapse: effect of composite type on adhesion efficacy of upper and lower orthodontic fixed retainers: in vitro investigation and randomized clinical trial. *Polymers (Basel)* 2020;**12**(4):963. [DOI: 10.3390/polym12040963]

Sfrondini 2014 *{published data only}*

Sfondrini MF, Fraticelli D, Castellazzi L, Scribante A, Gandini P. Clinical evaluation of bond failures and survival between mandibular canine- to-canine retainers made of flexible spiral wire and fiber-reinforced composite. *Journal of Clinical and Experimental Dentistry* 2014;**6**(2):e145-9.

Sfrondini 2021 *{published data only}*10.3390/ma14061341

Sfondrini MF, Gallo S, Turcato B, Montasser MA, Albelasy NF, Vallittu PK, et al. Universal adhesive for fixed retainer bonding: in vitro evaluation and randomized clinical trial. *Materials (Basel)* 2021;**14**(6):1341. [DOI: 10.3390/ma14061341]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shawesh 2010 *{published and unpublished data}*

Shawesh M, Bhatti B, Usmani T, Mandall N. Hawley retainers full or part time? A randomized clinical trial. *European Journal of Orthodontics* 2010;**32**:165-70.

Shim 2022 *{published data only}*

Shim H, Foley P, Bankhead B, Kim KB. Comparative assessment of relapse and failure between CAD/CAM stainless steel and standard stainless steel fixed retainers in orthodontic retention patients: a randomized controlled trial. *Angle Orthodontist* 2022;**92**(1):87-94. [DOI: 10.2319/121720-1015.1.]

Sinha 2021 *{published data only}*

Sinha A, Sonar S, Batra P, Raghavan S. A prospective randomized controlled trial on the comparative clinical efficiency and hygiene of a ceramic inter locking retainer and a flexible spiral wire bonded retainer. *Indian Journal of Dental Research* 2021;**32**(2):174-80. [DOI: 10.4103/ijdr.IJDR_1050_20]

Soboutia 2016 *{published data only}*

Soboutia F, Rakhshan V, Saravid MG, Zamaniand A, Shariatie M. Two-year survival analysis of twisted wire fixed retainer versus spiral wire and fiber-reinforced composite retainers: a preliminary explorative single-blind randomized clinical trial. *Korean Journal of Othodontics* 2016;**46**(2):104-10.

Sun 2011 *{published and unpublished data}*

Sun J, Yu MY, Liu L, Chen L, Li HW, Zhang L, et al. Survival time comparison between Hawley and clear overlay retainers: a randomized trial. *Journal of Dental Research* 2011;**90**:1197-201.

Thickett 2010 *{published and unpublished data}*

Thickett 2010 [pers comm]. Cochrane systematic review on orthodontic retention [personal communication]. Email details unavailable.

Thickett E, Power S. A randomized clinical trial of thermoplastic retainer wear. *European Journal of Orthodontics* 2010;**32**:1-5. [DOI: 10.1093/ejo/cjp061]

Torkan 2014 *{published and unpublished data}***10.1186/s40510-014-0047-8**

Torkan S, Oshagh M, Khojastepour L, Shahidi S, Heidari S. Clinical and radiographic comparison of the effects of two types of fixed retainers on periodontium - a randomized clinical trial. *Progress in Orthodontics* 2014;**15**(1):47. [DOI: 10.1186/s40510-014-0047-8]

Wan 2017 *{published and unpublished data}***10.2319/012716-76.1.**

Wan J, Wang T, Pei X, Wan Q, Feng W, Chen J. Speech effects of Hawley and vacuum-formed retainers by acoustic analysis: a single-center randomized controlled trial. *Angle Orthodontics* 2017;**87**(2):286-92. [DOI: 10.2319/012716-76.1.]

Węgrodzka 2021 *{published data only}*

Węgrodzka E, Kornatowska K, Pandis N, Fudalej PS. A comparative assessment of failures and periodontal health between 2 mandibular lingual retainers in orthodontic patients. A 2-year follow-up, single practice-based randomized trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2021;**160**(4):494-502.e1. [DOI: 10.1016/j.ajodo.2021.02.015]

Zhu 2017 *{published data only}***10.1016/j.ajodo.2016.10.019**

Zhu Y, Lin J, Long H, Ye N, Huang R, Yang X, et al. Comparison of survival time and comfort between 2 clear overlay retainers with different thicknesses: a pilot randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2017;**151**(3):433-9. [DOI: 10.1016/j.ajodo.2016.10.019]

References to studies excluded from this review

ACTRN12612000670875 *{published data only}*

ACTRN12612000670875. A prospective study of orthodontic retention outcomes with different retention strategies. www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=362581 (first received 22 June 2012).

Ahrens 1981 *{published and unpublished data}*

Ahrens DG, Shapira Y, Kuftinec MM. An approach to rotational relapse. *American Journal of Orthodontics* 1981;**80**(1):83-91.

Alaa ElDeen 2021 *{published and unpublished data}*

Alaa ElDeen 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Raghda Alaa 8 Jul 2021, 30 Nov 2021.

Alaa ElDeen R, El Dakroury A, El Sharaby F, Abou el Fotouh M, El Beialy A. Bond failure of 3D printed positioning tray for fixed retainers. *Acta Scientific Dental Sciences* 2021;**5.7**:03-09.

Alkan 2020 *{published data only}***10.1111/joor.12970**

Alkan Ö, Kaya Y, Keskin S. Computerized occlusal analysis of Essix and Hawley retainers used during the retention phase: a controlled clinical trial. *Journal of Orofacial Orthopedics / Fortschritte der Kieferorthopädie* 2020;**81**(5):371-81.

Alkan Ö, Kaya Y. Changes in occlusal surface area and occlusal force distribution following the wear of vacuum-formed, Hawley and bonded retainers: a controlled clinical trial. *Journal of Oral Rehabilitation* 2020;**47**(6):766-74. [DOI: 10.1111/joor.12970]

Al-Nimri 2009 *{published data only}*

Al-Nimri K, Al Habashneh R, Obeidat M. Gingival health and relapse tendency: a prospective study of two types of lower fixed retainers. *Australian Orthodontic Journal* 2009;**25**(2):142-6.

Arora 2014 *{published and unpublished data}*

Arora S, Naik RD. Comparing relapse after occlusion using Begg and biocryl retainers. *Orthodontic Update* 2014;**7**:59-66.

Aslan 2013 *{published and unpublished data}*

Aslan BI, Dincer M, Salmanli I, Qasem MA. Comparison of the effects of modified and full-coverage thermoplastic retainers on occlusal contacts. *Orthodontics: the Art and Practice of Dentofacial Enhancement* 2013;**14**:e198-208.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Atack 2007 *{published data only}*

Atack N, Harradine N, Sandy JR, Ireland AJ. Which way forward? Fixed or removable lower retainers. *Angle Orthodontist* 2007;**77**(6):954-9.

Axelsson 1993 *{published and unpublished data}*

Axelsson S, Zachrisson BU. Clinical experience with direct-bonded labial retainers. *European Journal of Orthodontics* 1993;**4**(15):333.

Barlin 2011 *{published and unpublished data}*

Barlin S, Smith R, Reed R, Sandy J, Ireland AJ. A retrospective randomized double-blind comparison of the effectiveness of Hawley vs vacuum-formed retainers. *Angle Orthodontist* 2011;**81**(3):404-9.

Basciftici 2007 *{published data only}*

Basciftci FA, Uysal T, Sari Z, Inan O. Occlusal contacts with different retention procedures in 1-year follow-up period. *American Journal of Orthodontics and Dentofacial Orthopedics* 2007;**131**(3):357-62.

Bauer 2010 *{published and unpublished data}*

Bauer EM, Behrents R, Oliver DR, Buschang PH. Posterior occlusion changes with a Hawley vs perfector and Hawley retainer. *Angle Orthodontist* 2010;**80**:853-60.

Bazargani 2012 *{published data only}*

Bazargani F, Jacobsona S, Lennartssonb B. A comparative evaluation of lingual retainer failure bonded with or without liquid resin. *Angle Orthodontist* 2012;**82**:84-7.

Bock 2008 *{published data only}*

Bock N, Ruf S. Post-treatment occlusal changes in Class II division 2 subjects treated with the Herbst appliance. *European Journal of Orthodontics* 2008;**30**(6):606-13.

Bovali 2014 *{published and unpublished data}*

Bovali 2014 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Marie A Cornelis 8 Jul 2021, 14 Jun 2022.

* Bovali E, Kiliaridis S, Cornelis MA. Indirect vs direct bonding of mandibular fixed retainers in orthodontic patients: a single-center randomized controlled trial comparing placement time and failure over a 6-month period. *American Journal of Orthodontics and Dentofacial Orthopedics* 2014;**146**(6):701-8.

Cornelis MA, Egli F, Bovali E, Kiliaridis S, Cattaneo PM. Indirect vs direct bonding of mandibular fixed retainers in orthodontic patients: Comparison of retainer failures and posttreatment stability. A 5-year follow-up of a single-center randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2022;**162**(2):152-61.

Egli F, Bovali E, Kiliaridis S, Cornelis MA. Indirect vs direct bonding of mandibular fixed retainers in orthodontic patients: comparison of retainer failures and posttreatment stability. A 2-year follow-up of a single-center randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics* 2017;**151**(1):15-27.

Brignardello-Petersen 2018 *{published data only}***10.1016/j.adaj.2017.12.008**

Brignardello-Petersen R. No evidence of differences important to patients between bonded and vacuum-formed retainers after 1 year of use. *Journal of American Dental Association* 2018;**149**(4):e71. [DOI: 10.1016/j.adaj.2017.12.008] [PMID: 29395006]

Carvalho 2006 *{published data only}*

Carvalho CV, Bauer FP, Romito GA, Pannuti CM, De Micheli G. Orthodontic extrusion with or without circumferential supracrestal fiberotomy and root planing. *International Journal of Periodontics & Restorative Dentistry* 2006;**26**(1):87-93.

ChiCTR-TRC-07000055 *{published data only}*

ChiCTR-TRC-07000055. The impact of the Hawley's removable retainer and clear overlay removable retainer on orthodontic patients' oral hygiene. www.chictr.org.cn/showprojen.aspx?proj=9470 (first received 8 May 2007).

ChiCTR-TRC-07000454 *{published data only}*

ChiCTR-TRC-07000454. Hawley's retainer and transparent retainer for the oral cavity hygiene in orthodontic patients. www.chictr.org.cn/showprojen.aspx?proj=9080 (first received 26 April 2007).

Choi 2010 *{unpublished data only}*

Choi J-H, Moon C-H. Discomfort caused by the circumferential comfortable retainer (CCR) as a removable maxillary retainer. *Korean Journal of Orthodontics* 2010;**40**(5):325-33.

Conway 2011 *{published and unpublished data}*

Conway M, Littlewood SJ, Hodge T, Nelson-Moon Z. Bonded retainers placed under rubber dam: an exploratory RCT. *Journal of Orthodontics* 2011;**38**:299.

Cunning 2022 *{published data only}*

ACTRN12618002022246. Comparative assessment of survival, stability and occlusal settling between full-time and part-time wear regime of two types of thermoplastic retainers: a randomised controlled clinical trial. www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=376518 (first received 17 December 2018).

* Cunning B, Sahhar A, Lawrence S, Manton D, Schneider P. Comparative assessment of the survival, stability and occlusal settling between two types of thermoplastic retainers: a prospective clinical trial. *Australasian Orthodontic Journal* 2022;**38**(1):74-87.

Schneider 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Paul Schneider 9 July 2021, 3 November 2021.

Edwards 1988 *{published and unpublished data}*

Edwards JG. A long-term prospective evaluation of the circumferential supracrestal fiberotomy in alleviating orthodontic relapse. *American Journal of Orthodontics and Dentofacial Orthopedics* 1988;**93**(5):380-7.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Edwards JG. A surgical procedure to eliminate rotational relapse. *American Journal of Orthodontics* 1970;**57**(1):35-46.

Haydar 1992 *{published and unpublished data}*

Haydar B, Ciger S, Saatçi P. Occlusal contact changes after the active phase of orthodontic treatment. *American Journal of Orthodontics and Dentofacial Orthopedics* 1992;**102**(1):22-8.

Horton 2009 *{published data only}*

Horton JK, Bushang PH, Oliver DR, Behrents RG. Comparison of the effects of Hawley and perfector/spring aligner retainers on postorthodontic occlusion. *American Journal of Orthodontics and Dentofacial Orthopedics* 2009;**135(6)**:729-36.

ISRCTN22535947 *{published data only}*

ISRCTN22535947. The benefit of prolonged orthodontic retention - a randomised clinical trial. www.isrctn.com/ISRCTN22535947 (first received 30 September 2005).

ISRCTN26364810 *{published data only}*

ISRCTN56295329. Orthodontic treatment with a Twin Block appliance randomised to fixed or removable orthodontic retention. www.isrctn.com/ISRCTN56295329 (first received 19 January 2001).

ISRCTN56295329 *{published data only}*

ISRCTN56295329. How should patients be supervised during orthodontic retention? A randomised controlled trial. www.isrctn.com/ISRCTN56295329 (first received 28 September 2007).

Jäderberg 2012 *{published and unpublished data}*

Jäderberg S, Feldmann I, Engström C. Removable thermoplastic appliances as orthodontic retainers. *European Journal of Orthodontics* 2012;**34**:475-9.

Jahanbin 2014 *{published data only}*

Jahanbin A, Ramazanzadeh B, Ahrari F, Forouzanfar A, Beidokhti M. Effectiveness of Er:YAG laser-aided fiberotomy and low-level laser therapy in alleviating relapse of rotated incisors. *American Journal of Orthodontics and Dentofacial Orthopedics* 2014;**146**(5):565-72. [DOI: 10.1016/j.ajodo.2014.07.006]

Jain 2021 *{published data only}*

Jain 2021 [pers comm]. Data request for a systematic review [personal communication]. Email to: Shalu Jain 29 April 2022.

Jain S, Kauser A, Tiwari HD, Ashna A, Ahmad A, Yadav N, et al. Efficacy of fixed versus removal retainer post orthodontic treatment: a comparative study. *Turkish Journal of Physiotherapy and Rehabilitation* 2021;**32**(3):9479-85.

Kara 2020 *{published data only}***10.1016/j.ajodo.2019.05.020**

Kara B, Yilmaz B. Occlusal contact area changes with different retention protocols: 1-year follow-up. *American Journal of Orthodontics and Dentofacial Orthopedics* 2020;**157**(4):533-41. [DOI: 10.1016/j.ajodo.2019.05.020] [PMID: 32241360]

Knaup 2019 *{published data only}***10.1007/s00056-019-00169-7**

Knaup I, Wagner Y, Wego J, Fritz U, Jäger A, Wolf M. Potential impact of lingual retainers on oral health: comparison between conventional twistflex retainers and CAD/CAM fabricated nitinol retainers: a clinical in vitro and in vivo investigation. *Journal of Orofacial Orthopedics* 2019;**80**(2):88-96. [DOI: 10.1007/s00056-019-00169-7] [PMID: 30778609]

Larsson 1983 *{published and unpublished data}*

Larsson E, Schmidt G. The effect of the supra-alveolar soft tissue on the relapse of orthodontic treatment. *British Journal of Orthodontics* 1983;**10**(1):50-2.

Lindauer 1998 *{published data only}*

Lindauer SJ, Shoff RC. Comparison of Essix and Hawley retainers. *Journal of Clinical Orthodontics* 1998;**32**:95-7.

Liu 2010 *{published data only}*

Yam Liu. Application of fiber-reinforced composite as fixed lingual retainer. *Hua Xi Kou Qiang Yi Xue Za Zhi* 2010;**28**(3):290-3.

NCT03765151 *{published data only}*

NCT03765151. The effectiveness of low-level laser therapy in orthodontic retention for periodontal-compromised patients. clinicaltrials.gov/ct2/show/NCT03765151 (first received 5 December 2018).

Yang 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Yanqi Yang 8 Jul 2021.

Pandian 2019 *{unpublished data only}*

Pandian 2019a [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Akshaya Pandian 9 Jul 2021.

Pandian 2019b [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Akshaya Pandian 9 July 2021.

Pandis 2007 *{published data only}*

Pandis N, Vlahopoulos K, Madianos P, Eliades T. Long-term periodontal status of patients with mandibular lingual fixed retention. *European Journal of Orthodontics* 2007;**29**(5):471-6.

Sangalli 2022 *{published data only}*

Sangallia L, Savoldic F, Dalessandria D, Viscontia L, Massettia F, Bonettia S. Remote digital monitoring during the retention phase of orthodontic treatment: a prospective feasibility study. *Korean Journal of Orthodontics* 2022;**52**:123-30.

Sauget 1997 *{published and unpublished data}*

Sauget E, Covell DA, Boero RP, Lieber WS. Comparison of occlusal contacts with use of Hawley and clear overlay retainers. *Angle Orthodontist* 1997;**67**(3):223-30.

Sohrabi 2014 *{published data only}*

Sohrabi A, Rafighi A, Hossein Moslemzadeh S, Foughi Moghaddam S, Rousoli F. Comparation of two different retention techniques on re-opening of extraction space at different intervals during the retention phase in fixed

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

orthodontic patients . *Iran Journal of Orthodontics* 2014;**9**(3):e3747. [DOI: 10.17795/ijo-3747]

Stormann 2002 *{published data only}*

Stormann I, Ehmer U. A prospective randomized study of different retainer types. *Journal of Orofacial Orthopedics* 2002;**63**(1):42-50.

Tacken 2010 *{published data only}*

Tacken MP, Cosyn J, De Wilde P, Aerts J, Govaerts E, Vannet BV. Glass fibre reinforced versus multistranded bonded orthodontic retainers: a 2 year prospective multi-centre study. *European Journal of Orthodontics* 2010;**32**(2):117-23.

Tahir 2019 *{published and unpublished data}*

Mohd Tahir N, Wan Hassan WN, Saub R. Comparing retainers constructed on conventional stone models and on 3D printed models: a randomized crossover clinical study. *European Journal of Orthodontics* 2019;**41**(4):370-80.

Taner 2000 *{published and unpublished data}*

Taner T, Haydar B, Kavuklu I, Korkmaz A. Short-term effects of fiberotomy on relapse of anterior crowding. *American Journal of Orthodontics and Dentofacial Orthopedics* 2000;**118**(6):617-23.

Vagdouti 2019 *{published data only}*

Vagdouti G, Karvouni E, Bitsanis E, Koletsia D . Objective evaluation of compliance after orthodontic treatment using Hawley or vacuum-formed retainers: a 2-center randomized controlled trial over a 3-month period. *American Journal of Orthodontics and Dentofacial Orthopedics* 2019;**156**(6):717-26.

Vecere 1983 *{published and unpublished data}*

Vecere JW. Extraction space closure stability following canine retraction and periodontal surgery. *American Journal of Orthodontics* 1983;**84**(1):89-90.

Xu 2011 *{published data only}*

Xu X-C, Li R-M, Tang G-H. Clinical evaluation of lingual fixed retainer combined with Hawley retainer and vacuum-formed retainer. *Shanghai Kou Qiang Yi Xue* 2011;**20**(6):623-6.

References to ongoing studies

CTRI/2018/10/016151 *{published and unpublished data}*

* CTRI/2018/10/016151. Evaluation of success of two orthodontic retainer wires on maxillary and mandibular arch retention over a 12-month period: a randomized controlled trial. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=28736&EncHid=&userName=CTRI/2018/10/016151 (first received 25 October 2018).

Kumar 2021. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Vinay Kumar Chugh 8 Jul 2021.

CTRI/2019/05/019136 *{published and unpublished data}*

CTRI/2019/05/019136. To compare lingual bonded retainer post orthodontically. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=30363&EncHid=&userName=CTRI/2019/05/019136 (first received 15 May 2019).

CTRI/2020/04/024461 *{published data only}*

CTRI/2020/04/024461. Comparison of effects of Hawley's and Van der Linden retainers on final teeth position. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=38236&EncHid=&userName=CTRI/2020/04/024461 (first received 3 April 2020).

CTRI/2020/04/024486 *{published data only}*

CTRI/2020/04/024486. Comaprison of effect of Hawley and vacuum formed retainer on speech. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=38717&EncHid=&userName=CTRI/2020/04/024486 (first received 7 April 2020).

CTRI/2021/02/031387 *{published data only}*

CTRI/2021/02/031387. Straight or wave retainer; the better one. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=51805&EncHid=&userName=CTRI/2021/02/031387 (first received 12 February 2021).

CTRI/2021/02/031498 *{published data only}*

CTRI/2021/02/031498. Evaluation of periodontal health and stability following fixed orthodontic treatment by the use of two types of retainer. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=50001&EncHid=&userName=CTRI/2021/02/031498 (first received 7 November 2021).

CTRI/2021/03/032110 *{published data only}*

CTRI/2021/03/032110. Comparison of effects of Essix and modified Essix retainer on final teeth positioning. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=53187&EncHid=&userName=CTRI/2021/03/032110 (first received 18 March 2021).

CTRI/2021/11/037851 *{published data only}*

CTRI/2021/11/037851. Changes in corrected position of teeth following braces treatment. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=60020&EncHid=&userName=CTRI/2021/11/037851 (first received 8 November 2021).

CTRI/2021/11/038102 *{published data only}*

CTRI/2021/11/038102. A comparison of clinical efficiency of nanocomposite and conventional composite used for bonding lower fixed retainers. ctri.nic.in/Clinicaltrials/pmaindet2.php?trialid=53674&EncHid=&userName=CTRI/2021/11/038102 (first received 17 November 2021).

DRKS00016710 *{published and unpublished data}*

* DRKS00016710. Influence of fixed and removable orthodontic retainers on the gingival health of mandibular anterior teeth - a randomized, clinical controlled trial. drks.de/search/en/trial/DRKS00016710 (first received 5 September 2019).

Petsos 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Hari Christian Petsos 19 July 2021, 21 August 2022.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

IRCT2016022926836N1 *{published data only}*

IRCT2016022926836N1. Comparison of survival time and satisfaction of Hawley and vacuum-formed retainers in orthodontic patients – a randomized clinical trial. en.irct.ir/trial/22137?revision=22137 (first received 29 April 2016).

IRCT2017021722951N3 *{published data only}*

IRCT2017021722951N3. Comparison of orthodontic treatment results stability and gingival health between Hawley and vacuum-formed retainers: randomized clinical trial. en.irct.ir/trial/19683 (first received 28 April 2017).

IRCT20210221050435N2 *{published data only}*

IRCT20210221050435N2. Level of Metalloproteinase Matrix 8 in gingival crevicular fluid of mandibular anterior teeth with two types of fixed and removable orthodontic retainers. en.irct.ir/trial/56547 (first received 7 October 2021).

NCT02742493 *{published and unpublished data}*

Koletsi 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Despina Koletsi 9 July 2021, 19 August 2022.

* NCT02742493. Effectiveness of 3 orthodontic retention schemes on post-treatment stability and gingival recession. clinicaltrials.gov/ct2/show/NCT02742493 (first received 19 April 2016).

NCT03844425 *{published data only}*

NCT03844425 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Wan Nurazreena Wan Hassan 8 July 2021.

* NCT03844425. Microbial colonisation on vacuum-formed retainers constructed on conventional models and three dimensional (3D) reconstructed models. clinicaltrials.gov/ct2/show/NCT03844425 (first received 18 February 2019).

NCT03954145 *{published data only}*

NCT03954145. Comparison of bonding failure rates of fixed mandibular lingual retainers in a French sample aged 11 or older comparing prior enamel sandblasting with conventional pumice polishing. clinicaltrials.gov/ct2/show/NCT03954145 (first received 17 May 2019).

NCT04578704 *{published and unpublished data}*

* NCT04578704. Stability of bimaxillary proclination cases following fixed appliances treatment: a randomized control trial. clinicaltrials.gov/ct2/show/NCT04578704 (first received 8 October 2022).

Tahir 2022 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Norhidayah Mohd Tahir 8 Jul 2021, 19 Aug 2022.

NCT04828096 *{published and unpublished data}*

Bazargani 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Farhan Bazargani 8 Jul 2021, 18 Aug 2022.

* NCT04828096. Retention with three different bonded retainers a multicenter, randomized controlled trial with 5-year follow-up. clinicaltrials.gov/ct2/show/NCT04828096 (first received 1 April 2021).

NCT04847323 *{published and unpublished data}*

Bellini-Pereira 2022 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Silvio Bellini-Pereira 8 July 2021, 26 August 2022.

* NCT04847323. Clinical effectiveness of bonded versus vacuum-formed retainers after 12 months: a randomized controlled trial. clinicaltrials.gov/ct2/show/NCT04847323 (first received 19 April 2021).

NCT05006339 *{published data only}*

NCT05006339. The use of remote monitoring for orthodontic retention review . clinicaltrials.gov/ct2/show/NCT05006339 (first received 16 August 2021).

NCT05014308 *{published data only}*

NCT05014308. Changes in dentition and hygiene in the retention phase of orthodontic treatment. clinicaltrials.gov/ct2/show/NCT05014308 (first received 20 August 2021).

NCT05121220 *{published data only}*

NCT05121220. Changes in dentition and hygiene in the retention phase of orthodontic treatment. clinicaltrials.gov/ct2/show/NCT05121220 (first received 16 November 2021).

NCT05250765 *{published data only}*

NCT05250765. Comparison of efficiency and effectiveness of two types of bonded orthodontic retainers: an RCT. clinicaltrials.gov/ct2/show/ (first received 22 February 2022).

TCTR20190416001 *{published and unpublished data}*

Jirawiwatsaree 2021 [pers comm]. Cochrane Review update: Retention procedures for stabilising tooth position after treatment with orthodontic braces [personal communication]. Email to: Rungtiwa Jirawiwatsaree 8 July 2021.

* TCTR20190416001. An evaluation of surface electromyography and masticatory performance in patients wearing thermoplastic oral appliance: a randomized clinical trial. www.thaiclinicaltrials.org/show/TCTR20190416001.

Additional references

Aasen 2005

Aasen T, Espeland L. An approach to maintain orthodontic alignment of lower incisors without the use of retainers. *European Journal of Orthodontics* 2005;**27**:209-14.

Al Rahma 2018

Al Rahma WJ, Kaklamanos EG, Athanasiou AE. Performance of Hawley-type retainers: a systematic review of

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

randomized clinical trials. *European Journal of Orthodontics* 2018;**40**(2):115-25. [DOI: 10.1093/ejo/cjx036]

Al-Moghrabi 2016

Al-Moghrabi D, Pandis N, Fleming PS. The effects of fixed and removable orthodontic retainers: a systematic review. *Progress in Orthodontics* 2016;**17**(1):24. [DOI: 10.1186/s40510-016-0137-x]

Arn 2020

Arn ML, Dritsas K, Pandis N, Kloukos D. The effects of fixed orthodontic retainers on periodontal health: a systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics* 2020;**157**(2):156-64. [DOI: 10.1016/j.ajodo.2019.10.010]

Bellini-Pereira 2022

Bellini-Pereira SA, Aliaga-Del Castillo A, Dos Santos CC, Henriques JF, Janson G, Normando D. Treatment stability with bonded versus vacuum-formed retainers: a systematic review of randomized clinical trials. *European Journal of Orthodontics* 2022;**44**(22):187-96. [DOI: 10.1093/ejo/cjab073]

Braunholtz 2001

Braunholtz DA, Edwards SJ, Lilford RJ. Are randomized clinical trials good for us (in the short term)? Evidence for a "trial effect". *Journal of Clinical Epidemiology* 2001;**54**:217-24.

Buzatta 2017

Buzatta LN, Shimizu RH, Shimizu IA, Pachêco-Pereira C, Flores-Mir C, Taba M Jr, et al. Gingival condition associated with two types of orthodontic fixed retainers: a meta-analysis. *European Journal of Orthodontics* 2017;**39**(4):446-52. [DOI: 10.1093/ejo/cjw057]

Casko 1998

Casko JS, Vaden JL, Kokich VG, Damone J, James RD, Cangialosi TJ, et al. Objective grading system for dental casts and panoramic radiographs. American Board of Orthodontics. *American Journal of Orthodontics and Dentofacial Orthopedics* 1998;**114**(5):589-9.

Covidence [Computer program]

Covidence. Version accessed 1 April 2022. Melbourne, Australia: Veritas health Innovation. Available at covidence.org.

Deeks 2022

Deeks JJ, Higgins JP, Altman DG editor(s). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.

GRADEpro GDT [Computer program]

GRADEpro GDT. Version accessed 1 May 2022. Hamilton (ON): McMaster University (developed by Evidence Prime). Available at gradepro.org.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557-60.

Higgins 2011

Higgins JP, Altman DG, Sterne JA editor(s). Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. training.cochrane.org/handbook/archive/v5.1/.

Higgins 2021a

Higgins JP, Lasserson T, Chandler J, Tovey D, Thomas, J, Flemyng E, et al. Methodological Expectations of Cochrane Intervention Reviews (MECIR). Available from community.cochrane.org/mecir-manual February 2021.

Higgins 2022

Higgins JP, Li T, Deeks JJ editor(s). Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.

Iliadi 2015

Iliadi A, Kloukos D, Gkantidis N, Katsaros C, Pandis N. Failure of fixed orthodontic retainers: a systematic review. *Journal of Dentistry* 2015;**43**(8):876-96. [DOI: 10.1016/j.jdent.2015.05.002]

Jedliński 2021

Jedliński M, Grocholewicz K, Mazur M, Janiszewska-Olszowska J. What causes failure of fixed orthodontic retention? - systematic review and meta-analysis of clinical studies. *Head & Face Medicine* 2021;**17**(1):32. [DOI: 10.1186/s13005-021-00281-3]

Lefebvre 2022

Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Marshall C, et al. Technical Supplement to Chapter 4: Searching for and selecting studies. In: Higgins JP, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.

Li 2021

Li B, Xu Y, Lu C, Wei Z, Li Y, Zhang J. Assessment of the effect of vacuum-formed retainers and Hawley retainers on periodontal health: a systematic review and meta-analysis. *PLoS One* 2021;**16**(7):e0253968. [DOI: 10.1371/journal.pone.0253968]

Little 1981

Little RM, Wallen TR, Riedel RA. Stability and relapse of mandibular anterior alignment - first premolar extraction cases treated by traditional edgewise orthodontics. *American Journal of Orthodontics* 1981;**80**(4):349-65.

Little 1988

Little RM, Riedel RA, Årtun J. An evaluation of changes in mandibular anterior alignment from 10 to 20 years

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

postretention. *American Journal of Orthodontics and Dentofacial Orthopedics* 1988;**93**(5):423-8.

Mai 2014

Mai W, He J, Meng H, Jiang Y, Huang C, Li M, et al. Comparison of vacuum-formed and Hawley retainers: a systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics* 2014;**145**(6):720-7.

Melrose 1998

Melrose C, Millett DT. Toward a perspective on orthodontic retention? *American Journal of Orthodontics and Dentofacial Orthopedics* 1998;**113**(5):507-14.

Outhaisavanh 2020

Outhaisavanh S, Liu Y, Song J. The origin and evolution of the Hawley retainer for the effectiveness to maintain tooth position after fixed orthodontic treatment compare to vacuum-formed retainer: a systematic review of RCTs . *International Orthodontics / College Europeen d'Orthodontie* 2020;**18**(2):225-36. [DOI: 10.1016/j.ortho.2020.02.008]

Page 2021a

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**29**(372):n71. [DOI: 10.1136/bmj.n71]

Page 2022

Page MJ, Higgins JP, Sterne JA. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook.

Review Manager 2020 [Computer program]

Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: The Cochrane Collaboration, 2020.

RevMan Web 2022 [Computer program]

Review Manager Web (RevMan Web). Version 4.11.0. The Cochrane Collaboration, 2022. Available from revman.cochrane.org.

Richmond 1992

Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Jones R, Stephens CD, et al. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *European Journal of Orthodontics* 1992;**14**(2):125-39.

Seehra 2021

Seehra J, Stonehouse-Smith D, Pandis N. Assessment of early exaggerated treatment effects in orthodontic interventions using cumulative meta-analysis. *European Journal of Orthodontics* 2021;**43**(5):601-5. [DOI: 10.1093/ejo/cjab042]

References to other published versions of this review

Littlewood 2004

Littlewood SJ, Millett DT, Doubleday B, Bearn DR, Worthington HV. Retention procedures for stabilising tooth position after treatment with orthodontic braces. *Cochrane Database of Systematic Reviews* 2004, Issue 1. Art. No: CD002283. [DOI: 10.1002/14651858.CD002283.pub2]

Littlewood 2006

Littlewood SJ, Millett DT, Doubleday B, Bearn DR, Worthington HV. Retention procedures for stabilising tooth position after treatment with orthodontic braces. *Cochrane Database of Systematic Reviews* 2006, Issue 1. Art. No: CD002283. [DOI: 10.1002/14651858.CD002283.pub3]

Littlewood 2016

Littlewood SJ, Millett DT, Doubleday B, Bearn DR, Worthington HV. Retention procedures for stabilising tooth position after treatment with orthodontic braces. *Cochrane Database of Systematic Reviews* 2016, Issue 1. Art. No: CD002283. [DOI: 10.1002/14651858.CD002283.pub4]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Adanur-Atmaca 2021

Study characteristics	
Methods	4-arm parallel RCT
Participants	**Number recruited**: 132.33 per group **Age (years):** 16 (SD 3.8) **Sex (M:F):** 40:92 **Inclusion criteria**: nonextraction treatment in the mandible; moderate irregularity before treatment according to LII; good oral hygiene (absence of visible plaque and redness in the gingiva); no caries

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Adanur-Atmaca 2021 *(Continued)*

	Exclusion criteria: not mentioned **Setting**: orthodontic department of Pamukkale University, Turkey
Interventions	**4 types of lower lingual fixed retainers** • Group 1: rectangular braided multistrand (dead-soft) wire • Group 2: round multistrand stainless-steel wire • Group 3: CAD/CAM nitinol retainer (Memotain) • Group 4: ⌀ 0.012 inch connected bonding pad retainer
Outcomes	Stability: LII, intercanine width, arch length Adverse effects on periodontal health: Plaque, Gingival and Calculus Indices **Time points:** 3, 6, 9 and 12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information on random sequence generation method
Allocation concealment (selection bias)	Unclear risk	Not reported in the published paper; insufficient information to permit judgement
Blinding of outcome assessment (detection bias) All outcomes	High risk	Blinding of clinicians was not possible in this study because the outcome assessor and clinician were the same person. Likely to be influenced by lack of blinding
Incomplete outcome data (attrition bias) All outcomes	Low risk	All randomised participants were analysed by ITT analysis. No missing outcome data
Selective reporting (reporting bias)	Unclear risk	The study protocol is not available; insufficient information to permit judgement
Other bias	Low risk	The study appears to be free of other sources of bias.

Al-Moghrabi 2020

Study characteristics

Methods	2-arm parallel RCT
Participants	**Number recruited**: 84 **Age (years):** control group 17.20 (SD 1.89), intervention group 17.24 (SD 2.00) **Sex (M:F):** 1:1 in both groups **Inclusion criteria**: age 12-21 years; planned for removable retention with clear plastic retainers; no medication that would affect gingival health; permanent dentition

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Al-Moghrabi 2020 *(Continued)*

	Exclusion criteria: inability to access or use smartphone; cleft lip/palate or other craniofacial anomaly; history of periodontal disease **Setting**: hospital (Institute of Dentistry, Barts and The London School of Medicine and Dentistry), UK Treatment provided for free on NHS
Interventions	**Use of mobile application to increase duration of use of clear plastic retainers:** • Group 1 (intervention): access to "My retainers" mobile application • Group 2 (control): no access to this application "Maxillary and mandibular clear plastic retainers (Essix ACE® Plastic 1 mm in thickness (DENTSPLY)) were fitted 7–10 days following debond. All participants were instructed to wear TPRs on a full-time basis (22 hours) for 6 months, followed by part-time wear (8 hours) for a further 6 months. A TheraMon® micro-electronic sensor (MC Technology GmbH, Hargelsberg, Austria) was embedded in the maxillary CP retainer in all participants following a standardized laboratory technique."
Outcomes	Stability: LII, and transverse dimension changes (intercanine, interpremolar, and intermolar widths; only in the second publication that reports changes at 12 months) Patient satisfaction: questionnaire on participant experience and knowledge related to retention (only in the first publication, at 3 months) Adverse effects on periodontal health: plaque scores, bleeding on probing and probing depths for each arch **Time points:** 3, 6, 9 and 12 months
Notes	Contacted author for relevant information

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomization was based on computer-generated random numbers", stratified by sex
Allocation concealment (selection bias)	Low risk	Concealed in opaque, sealed envelopes
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"The use of coded study models and data ensured that the researcher was kept blind to the treatment group when undertaking measurements and during data analysis. The statistician was also kept blind to group allocation."
Incomplete outcome data (attrition bias) All outcomes	Low risk	"Imputation of missing data was undertaken to account for losses and to compensate for uncertainty surrounding missing values"
Selective reporting (reporting bias)	Low risk	No suggestion that there were other outcomes that were not reported
Other bias	Low risk	No other area of bias detected

Alrawas 2020

Study characteristics

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alrawas 2020 *(Continued)*

Methods	4-arm parallel RCT
Participants	**Number recruited**: 60, 15 per group **Age (years):** 19.5 (SD 7) **Sex (M:F):** 17:43 **Inclusion criteria:** irregularity of lower anteriors; completed fixed appliances; no retreatment; no circumferential supracrestal fibreotomy; perfect alignment of lower anterior teeth at end of treatment **Exclusion criteria:** poor oral hygiene; history of periodontal disease; antibiotics before gingival/periodontal exam; medication affecting periodontal health; pregnancy; diabetes; smoking; orthognathic surgery **Setting:** university hospital orthodontic clinic, Turkey
Interventions	**4 types of lower retainers** • Group 1: ⌀ 0.012 x 0.018 inch CAD/CAM nitinol bonded retainer • Group 2: ⌀ 0.017 inch twisted multi-strand stainless steel bonded retainer • Group 3: ⌀ 0.027 x 0.011 inch single-strand titanium wire • Group 4: clear plastic (CP) – full-time wear for 6 months
Outcomes	Stability • Little's irregularity index (mm) • Anterior dental arch length (mm) • Intercanine width (mm) • Intermolar width (mm) Adverse effects on periodontal health (only net changes between groups reported) • Plaque Index • Gingival Index • Bleeding on probing • Pocket depth **Time points:** 3 and 6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Electronic randomisation online software was used to generate random numbers for all participants based on inclusion and exclusion criteria and type of interventions (www.random.org)"
Allocation concealment (selection bias)	Low risk	"Sealed in an opaque envelope that was prepared in advance by a complete different practitioner to ensure that both the participants and the authors were masked in order to avoid selection bias in our study"
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	“Lingual surfaces were blinded from the observer during the assessment of teeth stability in order to minimize measurement bias. However, blinding was not an option during the clinical periodontal examination."

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Alrawas 2020 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Few dropouts (only 1 in each of 2 of the groups) - unwilling to attend appointment
Selective reporting (reporting bias)	Low risk	No suggestion that other outcomes were measured but not reported
Other bias	Low risk	No other area of bias detected

Arash 2020

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Numbers recruited**: 260 randomised, 250 included in statistical analysis (10 dropouts due to non-attendance at follow-up appointments). Group 1 (twisted): 138 participants, Group 2 (ribbon): 112 participants **Age (years):** 20 (SD 4.35) **Sex (M:F):** 44:68 **Inclusion criteria:** class I occlusion with good oral hygiene **Exclusion criteria:** deep overbite or any sign of periodontal problems like bleeding on probing or pocket depth > 3 mm **Setting:** private practice at an orthodontic office, Iran
Interventions	**Two types of lower fixed lingual retainers** • Group 1: fixed lingual mandibular retainers from canine-canine using ⌀ 0.0175 inch stainless steel twisted wire (G&H Orthodontics, USA) • Group 2: fixed lingual mandibular retainers from canine-canine using single-strand thick plain titanium wire retainer (Retainium, Reliance Orthodontics, USA)
Outcomes	Survival: failure rate in terms of detachments, survival/failure time (months) **Time points:** every 3 months during 24-month follow-up
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Insufficient information about the sequence generation process to permit judgement: "The patients were randomly divided into two groups"
Allocation concealment (selection bias)	Unclear risk	Method of concealment not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessment is mentioned. It is likely that the outcome measurement would be influenced by lack of blinding

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Arash 2020 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No reasons for missing data provided
Selective reporting (reporting bias)	High risk	Not all of the study's prespecified primary outcomes have been reported, e.g. side effects
Other bias	Low risk	No other area of bias detected

Ashari 2021

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 35 • Hawley retainer (HR) = 17 (F: 12; M: 5) • Modified clear plastic retainer (mCP) = 18 (F: 11; M: 7) **Age (years)**: 21.97 (SD 5.03) **Sex (M:F)**: 12:23 **Inclusion criteria**: aged ≥ 13 years at the time of debond, had existing pretreatment dental cast and undergone > 3 mm of maxillary dentoalveolar expansion **Exclusion criteria**: not reported **Setting**: Orthodontic University Clinic Kebangsaan, Malaysia
Interventions	**Two types of removable upper retainers** • Group 1: Hawley retainer (HR) • Group 2: modified clear plastic retainer (mCP) with palatal coverage for maintaining transverse expansion Participants were instructed to wear the retainers full-time for the first 6 months, followed by nighttime wear for the next 6 months. Text reminders were given once a month.
Outcomes	Stability: effectiveness in maintaining transverse expansion measured (in mm) as: • intercanine width (ICW); • interpremolar width (IPMW); • interfirst molar mesiobuccal cusp width (IFMW1); • interfirst molar distobuccal cusp width (IFMW2). **Time points**: debond, 3, 6 and 12 months
Notes	

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ashari 2021 *(Continued)*

Random sequence generation (selection bias)	Low risk	Computer-generated randomisation sequence
Allocation concealment (selection bias)	Low risk	Contacted the centre by phone after patients agreed to participate as to prevent selection bias and protect the assignment sequence until allocation
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"Each dental cast with an identity document was measured by a calibrated researcher who was blinded to the retention regime provided to each participant."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropout numbers similar in both groups and similar reasons given
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	No other area of bias detected

Bolla 2012

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 85 **Age (years)**: 21.9 (SD not provided) **Sex (M:F)**: 29:56 **Inclusion criteria:** people who had been treated with fixed appliances, with: good oral hygiene; mild crowding in upper labial segments; treated non-extraction; post-treatment Class I relationship with an overjet and overbite of 1-3 mm; correct anterior Bolton's index **Exclusion criteria:** dental agenesis or anomalies; enamel decalcification or white spots on lingual surface of anterior teeth; periodontal disease; considerable dental abrasion; parafunctional habits such as nocturnal bruxism **Setting**: university clinic in Italy and participants had to pay for treatment (information provided by study authors)
Interventions	**Two types of fixed retainers (upper and lower arches)** • Group 1: fibre-reinforced composite retainers were silanised, 14 µm, unidirectional, glass fibre devices embedded in a resin matrix • Group 2: multistrand retainers were ⌀ 0.0175 inch dead-soft passive multistrand wire Both types were bonded lateral incisor to lateral incisor in upper, and canine to canine in the lower arch. All were placed under rubber dam using a standardised bonding process.
Outcomes	Survival: number of detachments or bond failures **Time point:** 6 years
Notes	No measures of patient satisfaction, effect on health or stability were recorded.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Bolla 2012 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Contacted study authors who confirmed, "the operator tossed a coin for every patient" after agreeing to be in the study
Allocation concealment (selection bias)	Low risk	Contacted study authors who confirmed "the operator tossed a coin for every patient", which is considered a valid random component according to RoB 2 tool guidelines.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not feasible as the outcome assessor could not be blinded to the presence or absence of the wire
Incomplete outcome data (attrition bias) All outcomes	High risk	Paper recorded no dropouts, but on contacting the study authors, they reported that, "we described only the patients that completed the trial, we estimate, now, we had a drop out of 15-20 per group"
Selective reporting (reporting bias)	High risk	Contacted authors who reported, "we started to check the periodontium, but it was making the visit too long for the administration therefore we had to terminate that evaluation" They also started "evaluating the bacterial flora taking a sample every 3 months, unfortunately we ran out of funds"
Other bias	Low risk	No other areas of risk of bias identified

Edman Tynelius 2010

Study characteristics	
Methods	3-arm parallel RCT
Participants	**Numbers recruited**: 75 participants, 25 per group **Sex (M:F)**: 30:45 **Age (years)**: group 1 mean 14.1 (SD 1.3), group 2 mean 14.7 (SD 1.8), group 3 mean 14.3 (SD 1.5) **Inclusion criteria**: people who had completed upper and lower fixed appliances, with: no other previous experience of orthodontic treatment; permanent dentition; crowding in both arches; normal skeletal and dentoalveolar sagittal, vertical and transverse relationships; Class I molar relationship or up to 3 mm anterior or posterior deviation; treatment plan involving extraction of 4 premolars **Exclusion criteria**: no additional criteria specified **Setting**: Swedish National Health Service clinic
Interventions	**Two types of removable retainer and a positioner** • Group 1: upper clear plastic retainer and a ⌀ 0.7 mm stainless steel-bonded retainer bonded to the lower canines only. Clear plastic retainers were worn 22-24 hours for the first 2 days and then nights only after this. The bonded retainer was 0.7 mm stainless steel and bonded to the lingual of the canines only • Group 2: upper clear plastic retainer with interproximal stripping of the lower anterior teeth, but no mandibular retainer. Clear plastic retainers were worn 22-24 hours for the first 2 days and then nights only after this. Interproximal stripping was undertaken 5-6 weeks before debond and at debond.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Edman Tynelius 2010 *(Continued)*

	• Group 3: prefabricated positioner covering all the maxillary and mandibular teeth. Positioner was worn 30 min during the day (and participant was encouraged to bite into it) and also during sleep. It was a soft plastic device covering all teeth In the second year of retention, participants wore their removable retainers every other night. After 2 years, all retainers were stopped and bonded retainers removed.
Outcomes	Stability: based on LII in the upper and lower labial segments, intercanine width, intermolar width, arch length, overjet and overbite. Stability data collected after 1 year **Time points:** 1, 2 and 5 years post debond. Change data available for years 1 and 2. Final LII data only in both arches were available for 5 years post-debond
Notes	No measurement of survival of retainers, adverse effects on health or patient satisfaction It was reported that 2 participants did not comply with the positioner. On contact with the study author, the failure rates for the bonded retainers after 1 year were: 18 participants had no failures, 5 had 1 failure, 1 had 2 failures and 1 had 4 failures. Cost-effectiveness also reported, but not an outcome in this review.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"The generation of randomisation sequence was performed in blocks of five to ensure that equal numbers of patients were allocated to each of the three retention groups". The study author confirmed appropriate random number generation.
Allocation concealment (selection bias)	Low risk	Direct communication with the study author confirmed that they used adequate concealment, and neither the participant nor clinician could have known of the intervention to be used in advance. This was done by placing ballots in a hat by an independent observer, so that the clinician could not know what intervention was going to be drawn out.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The paper reported that the outcome assessor was blinded to the retention protocol.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropouts were clearly reported and not included in the final analysis.
Selective reporting (reporting bias)	Low risk	All the data were reported over the 5-year study period.
Other bias	Low risk	No other areas of bias identified

Eroglu 2019

Study characteristics

Methods	3-arm parallel RCT
Participants	**Numbers recruited**: 45 participants (15 per group)

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Eroglu 2019 *(Continued)*

	Age (years): 15.2 (SD 2.1) (range, 11.0-20.9) **Sex(M:F):** 11:34 **Inclusion criteria**: no caries, no antibiotic use in past 3 months, no smoking, no periodontal or systemic disorder, no prosthesis in mouth, about to complete orthodontic treatment **Exclusion criteria**: not mentioned **Setting**: hospital (Department of Orthodontics, Faculty of Dentistry, Selcuk University, Konya, Turkey)
Interventions	**3 types of upper and lower retainers:** • Group 1: fixed lingual (6 anterior teeth, ⌀ 0.0215 inch 5-stranded wire) retainer • Group 2: clear plastic retainer • Group 3: Hawley retainer The participants were instructed to wear their removable retainers all day, except during meals, and to brush them after toothbrushing.
Outcomes	Oral health, reported using periodontal indices: • probing depth • Gingival Index • bleeding on probing • Plaque Index **Time points:** debond, 1, 5 and 13 weeks (to report changes at 13 weeks only for this review)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"One of the retainers routinely used in the retention phase was applied randomly" States participants randomly divided into the 3 groups but no further details given
Allocation concealment (selection bias)	Unclear risk	Not indicated
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not indicated whether trained clinician was blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data complete
Selective reporting (reporting bias)	Unclear risk	It is unclear if other outcomes were recorded but not reported in this paper.
Other bias	Low risk	No other biases identified

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ferreira 2019

Study characteristics	
Methods	2-arm cross-over RCT
Participants	**Numbers recruited**: 15 **Age (years)**: 18-25 (SD not provided) **Sex (M:F):** not reported **Inclusion criteria**: good alignment of mandibular anterior teeth, previous orthodontic treatment **Exclusion criteria**: under orthodontic treatment, having severe crowding of mandibular anterior teeth **Setting**: hospital (Universidade Estadual de Maringá, Departamento de Odontologia (Maringá, Brazil)
Interventions	**2 types of lower fixed retainers**; participants wore 1 retainer for 6 months, then a 15-day wash-out period, then 6 months with the other retainer • Group 1: Ortho-Flextech (rectangular chain) bonded all 6 anterior mandibular teeth (⌀ 0.2 x 0.7 mm chain stainless steel) • Group 2: canine and canine rigid retainer bonded only to the canines but not the incisors (⌀ 0.8 mm diameter round stainless steel wire)
Outcomes	Oral health (periodontal indices): Plaque Index; Gingival Index; Calculus Index described as frequency and percentage for each tooth surface (mesial, distal, lingual and vestibular – by vestibular they mean buccal) Patient satisfaction: questionnaire on satisfaction (4 questions: comfort in use, better cleaning, need to use floss, preference for type of retainer – expressed as percentage) **Time points:** periodontal indices measured at T0 (baseline), T1 (after 6 months of retainer wear), T2 (wash-out – removal of the first retainer, removal of residual resin, dental polishing, and waiting period of 15 days), T3 (after second retainer used for 6 months). Questionnaire evaluated at T1 and T3 (following 6 months' wear of each retainer)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) All outcomes	High risk	No mention of blinding for outcomes
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not reported
Selective reporting (reporting bias)	Low risk	All outcomes reported

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ferreira 2019 *(Continued)*

Other bias	High risk	The only suggestion of randomisation is in the abstract saying this is "randomised longitudinal trial". Reporting does not follow CONSORT guidelines. No mention of other important issues pertinent to RCTs

Forde 2018

Study characteristics	
Methods	Multicentre, 2-arm parallel RCT
Participants	**Numbers recruited**: 60 (30 in each group) **Sex (M:F)**: 27:33 **Age (years) (median (IQR))**: bonded retainer group: 16 (2); clear plastic retainer group: 17 (4) **Inclusion criteria:** completion of fixed appliance therapy (involving both dental arches) with a satisfactory correction of the presenting malocclusion and dental alignment; good general health; demonstration of a good standard of oral hygiene (determined through questioning and clinical examination); full and normal complement of teeth in the upper and lower labial segments; teeth within the labial segments to be a regular size and shape; willing to consent to the study and comply with the study regime; may have presented with any malocclusion prior to orthodontic treatment; may have been managed on an extraction (premolar or molar) or a non-extraction basis; may have been treated with removable or functional appliances in conjunction with their orthodontic treatment; may have had adjunctive surgery **Exclusion criteria:** any general medical health problems that may influence gingival health, such as those necessitating antibiotic cover as bacteraemia prophylaxis, diabetes mellitus, epilepsy, or physical or mental disability; poor periodontal health, including the presence of supragingival calculus or periodontal pocketing > 3 mm; periodontal problems with either a pending referral to the periodontal department or a history of periodontal management during the orthodontic treatment; gross or uncontrolled caries; absent or diminutive lateral incisors; a starting malocclusion requiring extreme transverse **Setting:** 3 hospital orthodontic departments in the UK
Interventions	**A removable retainer versus a fixed retainer** • Group 1: bonded retainers (⌀ 0.0195 inch, 0.45 mm, 3-stranded twistflex stainless steel wire) upper and lower, canine-canine. Bonded to each tooth with composite resin • Group 2: clear plastic upper and lower retainers, to cover all fully erupted teeth and extend half way across the occlusal surface of the most distal molar. Used at night (part-time)
Outcomes	Stability (upper and lower arches) • LII mm • Intercanine width (ICW) mm • Intermolar width (IMW) mm • Arch length (AL) mm • Overjet (OJ) mm • Overbite (OB) mm Survival: time to first episode of failure; failure of retainers as % of lost/ill fitting/fractured/occlusal wear removable retainers and debonded/fractured fixed retainers Patient satisfaction questionnaire Adverse effects on periodontal health assessed using Calculus Index, Gingival Index and Plaque Index

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Forde 2018 *(Continued)*

	Time points: 3, 6 and 12 months post de-bond
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Participants were allocated randomly to receive either upper and lower BRs [bonded retainers] or upper and lower VFRs"
Allocation concealment (selection bias)	Low risk	"Using sequentially numbered, opaque, sealed envelopes, prepared in advance of the trial by a colleague independent of the study, using a computerized randomisation programme."
Blinding of outcome assessment (detection bias) All outcomes	High risk	High for stability measurements and periodontal outcomes Low for questionnaire Blinding of the outcome assessor looking at stability on the digital models was not possible, but blinding of the patient satisfaction assessor was undertaken by removing patient identifying information from the completed questionnaires
Incomplete outcome data (attrition bias) All outcomes	Low risk	Low dropout
Selective reporting (reporting bias)	Low risk	ITT analysis included
Other bias	Low risk	No other areas of bias identified

Gelin 2020

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 62 randomised, 61 (T0) to 41 (T3) included in the statistical analysis (reason for dropouts: participants discontinued their participation) **Age (years)**: 17 (SD 4.4) (control group: 16 (SD 3.18); test group: 18 (SD 5.14)) **Sex (M:F):** 18:43 **Inclusion criteria:** undergoing fixed orthodontic appliance at the mandible; end of treatment defined as Class I, i.e. complete correction of upper and lower maxillary and mandibular crowding, optimal overjet and overbite (from 1-2 mm), coincident upper and lower midlines, occlusion according to the 6 keys of Andrews; adequate oral hygiene; no missing teeth **Exclusion criteria:** craniofacial or severe asymmetry syndromes; history of periodontal disease; dental restorations or missing teeth on the anterior mandibular area; previous anterior fixed retainer placement; nickel allergy **Setting:** Department of Orthodontics, University Hospital of Liège, Belgium
Interventions	**2 types of fixed retainers**

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gelin 2020 *(Continued)*

- Group 1: fixed lingual mandibular retainers from canine to canine using a rectangular ⌀ 0.014 × 0.014 inch memory shape customised CAD/CAM nitinol retainer (Memotain; CA Digital GmbH, Mettmann, Germany)
- Group 2: fixed lingual mandibular retainers from canine to canine using a round ⌀ 0.0175 inch 6-strand twisted stainless-steel wire retainer (Supra-Flex; RMO Europe, Illkirch-Graffenstaden, France)

Outcomes	Stability: LII (primary outcome), intercanine width, interpremolar width, anterior arch length, total arch length Survival: retainer and bonding integrity Patient satisfaction: overall satisfaction (VAS), discomfort of the tongue (VAS; SD not provided, only means) Adverse effects on periodontal health: Gingival Index, Plaque Index, recession level, Calculus Index **Time points:** T0: bonding of the lingual fixed retainer, T1: 1 month after T0, T2: 6 months after T0, T3: 12 months after T0
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Insufficient information about the sequence generation process to permit judgement: "Patients were randomized sequentially by blocks of 10 (5 in each group) until completion of the required sample size"
Allocation concealment (selection bias)	Unclear risk	Method of concealment is not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessment is mentioned. However, with the retainers in place, it is possible that the outcome measurement is likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Similar number of dropouts per group. Reasons do not seem to be related to the outcome
Selective reporting (reporting bias)	Low risk	All outcomes reported; no changes to the protocol after study commencement
Other bias	Low risk	No other bias identified

Gera 2022

Study characteristics

Methods	Multicentre, 2-arm parallel RCT
Participants	**Number recruited**: 181 randomised participants, 163 participants included in the statistical analysis at 6 months of follow-up (reasons for dropouts: 13 participants did not attend follow-up, 4 digital models were lost and 1 participant needed re-treatment) **Age (years)**: control group: 16.2 (SD 3); test group: 16.2 (SD 2.8) **Sex (M:F):** 85:96

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gera 2022 *(Continued)*

	Inclusion criteria: good general health, age range 12-25 years, presence of all maxillary and mandibular anterior teeth with normal shape and size, completion of a treatment with full fixed appliance **Exclusion criteria:** cleft lip and/or palate or any other craniofacial syndrome, compromised oral hygiene or periodontal disease, having had orthognathic surgery, lingual appliance treatments, enamel hypoplasia, fluorosis, active caries, restorations or fractures in the anterior teeth, having had separate arch debondings with a difference of > 2 months in between jaws, re-treatments, having been assigned to another retention protocol due to initial malocclusion **Setting:** university setting at 2 postgraduate orthodontic clinics (Aarhus University, Denmark and University of Oslo, Norway)
Interventions	**2 types of fixed retainers** • Group 1: CAD/CAM fixed retainers: ⌀ 0.014 x 0.014 inch rectangular nitinol retainers (Memotain, CA Digital, Hilden, Germany) • Group 2: conventional multistrand fixed retainers: ⌀ 0.0215 inch 6-stranded stainless steel ((SS) G&H, Franklin, Indiana, USA)
Outcomes	Stability • LII (primary outcome) • Intercanine width • Interpremolar width • Intermolar length • Arch length Survival: retainer and bonding integrity Patient satisfaction: overall satisfaction (VAS) **Time point**: 6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer random number generator - quote: “Randomized in 1:1 ratio into blocks of four following an online computer-generated sequence (randomization.com)”
Allocation concealment (selection bias)	Low risk	Sequentially numbered, opaque, sealed envelopes - quote: “Allocation concealment was ensured by using sequentially numbered, opaque, sealed envelopes prepared in advance of the trial by a colleague not involved in the study.”
Blinding of outcome assessment (detection bias) All outcomes	High risk	Incomplete blinding of outcome assessment - quote: “Blinding of the outcome assessor regarding stability (based on digital models) and failures (based on clinical assessment) was not possible due to the evident differences between the retainers, but blinding the assessor for patient satisfaction was ensured by removing patient identifying information from the completed forms.”
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data are balanced in numbers across intervention groups, with similar reasons for missing data across groups.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gera 2022 *(Continued)*

Selective reporting (reporting bias)	Low risk	The study protocol is available and all of the study's prespecified (primary and secondary) outcomes that are of interest in the review have been reported in the prespecified way.
Other bias	Low risk	No other bias detected

Gill 2007

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 60 participants **Sex (M:F)**: 25:32 **Age (years)**: group 1: 13.72 (SD 2.5), group 2: 13.37 (SD 1.42) **Inclusion criteria**: patients treated with upper and lower pre-adjusted edgewise appliances (both extraction and non-extraction cases) **Exclusion criteria**: hypodontia where a Hawley-type retainer was required for space maintenance; cases where rapid maxillary expansion was used or surgically assisted expansion; non-orthognathic cases where significant movement of the lower incisors required a bonded retainer; people with reduced periodontal support **Setting**: hospital department in the UK; participants treated free of charge
Interventions	**Two types of wearing protocols for clear plastic retainers** • Group 1: part-time wear of clear plastic retainers (sleeping hours with minimum of 8 hours of wear) • Group 2: full-time wear of clear plastic retainers. The clear plastic retainers were full-arch coverage and made from 1 mm thick polycarbonate sheets.
Outcomes	Stability assessed using LII, intercanine width and intermolar width in both arches, and overjet and overbite. PAR index reported in a graph, but data not sufficiently clear for further analysis **Time point:** after 6 months
Notes	Study authors reported the type of initial malocclusion, whether the participants were treated by extraction or non-extraction and the time in treatment. Despite randomisation, there were more Class II participants in the part-time group.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The study authors confirmed that randomisation was generated by rolling a die and allocating the group according to odd or even numbers.
Allocation concealment (selection bias)	Low risk	"Identical opaque sealed envelopes containing a slip for group allocations were used for the purpose of random allocation"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome assessor was blind to the treatment group

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gill 2007 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Loss of follow-up data was reported in CONSORT diagram. The effect of this was not analysed, but as dropout was low, it is unlikely to have had a significant effect.
Selective reporting (reporting bias)	Low risk	Study authors confirmed that they did not collect any other outcome data and there was no selective reporting.
Other bias	Low risk	No other potential sources of bias identified

Gunay 2018

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 120 in total, 60 per group **Sex (M:F)**: 37:83 **Age (years)**: group 1: 15.7, group 2: 16.2 (SD not provided) **Inclusion criteria:** patients had undergone fixed appliance therapy, 60 with premolar extractions 60 without extractions; no missing mandibular incisors; no restorations on the mandibular incisors that might affect retainer bonding on lingual enamel surfaces; no morphologic crown anomalies **Exclusion criteria:** patients with rotations **Setting**: university hospital, Samsun, Turkey
Interventions	**Two different techniques for bonding two types of lingual wire retainers** • Group 1: ⌀ 0.0175 inch 6-strand stainless steel wire (Ortho Technology, Lutz, USA) was used, the lingual retainers were fabricated on plaster models, and a silicon transfer key was used. • Group 2: ⌀ 0.0195 inch dead-soft coaxial wire (Respond; Ormco, Orange, USA) was used, and the lingual retainers were made in the participant's mandibular arch without a study model.
Outcomes	Stability assessed using LII, intercanine width and arch length (mm) (data not available) Survival: failure rate as a percentage **Time points**: 3, 6, 9 and 12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The study authors state “the patients were divided into 2 groups randomly with equal numbers of extraction and nonextraction patients”.
Allocation concealment (selection bias)	Unclear risk	No comment was made in the paper.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No comment was made in the paper.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Gunay 2018 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	No comment was made in the paper, but it appears from the failure data that there were no dropouts.
Selective reporting (reporting bias)	High risk	They measured arch length but did not report the data.
Other bias	Low risk	No areas of risk of bias were identified.

Jowett 2022

Study characteristics	
Methods	Multicentre 2-arm parallel RCT
Participants	**Number recruited**: 62 in total, 31 per group **Mean age (SD)**: not provided **Sex (M:F)**: not provided **Inclusion criteria:** undergone upper and lower fixed appliance orthodontic treatment of any malocclusion (with or without extractions) with satisfactory correction; normal size, shape and number of teeth; proper oral hygiene; good health; able to comply with the trial regime; had given written informed consent **Exclusion criteria:** cleft palate and/or other severe facial deformities; nickel allergy; poor periodontal health; gross or uncontrolled caries; prosthodontic requirements; extreme transverse correction required **Setting**: 3 study centres: university teaching hospital, district general hospital and specialist orthodontic practice, National Institute of Health (NIH) UK
Interventions	**Two types of fixed lower retainers** • Group 1: lower CAD/CAM nitinol bonded retainers (Memotain). Custom retainer with silicone transfer jig • Group 2: lower chairside rectangular chain bonded retainers (Ortho-FlexTech) fitted directly at the chairside Common standard operating procedure for placement of Memotain and Ortho-FlexTech bonded retainers
Outcomes	Stability was measured by changes in LII and intercanine width. Results provided as medians and IQR. Survival was recorded as the time to the first episode of failure. Reported as hazard ratios Patient satisfaction was determined by a questionnaire. **Time points:** debond and 6 months
Notes	Contacted author for relevant information
Risk of bias	

Bias	**Authors' judgement**	**Support for judgement**

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Jowett 2022 *(Continued)*

Random sequence generation (selection bias)	Low risk	www.sealedenvelope.com was used to allocate participants.
Allocation concealment (selection bias)	Low risk	"simple randomisation by list was used with randomly allocated block sizes of 2, 4 and 6 to allocate participants to treatment groups. Once set-up, the list was immediately permanently deleted."
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"It was not possible to blind the operator and patient to the type of retainer used, but in order to blind the assessor a new technique was used to mask the retainer type using a Blu-Mousse coverage."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropout numbers similar in both groups and similar reasons given; however, study was discontinued before reaching the full sample size due to the high number of failures in the upper Memotain retainer group.
Selective reporting (reporting bias)	Low risk	All outcomes were reported.
Other bias	Low risk	No other areas of bias were identified.

Kartal 2021

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 52 randomised, 52 analysed (26 per group, no dropouts were observed) **Age (years (SD)):** group 1 (Memotain): 15.65 (2.17), group 2 (5-stranded): 18.42 (5.17) **Sex (M:F)**: 20:32 **Inclusion criteria:** presence of 6 mandibular anterior teeth involving 4 incisors and 2 canines; no caries, restorations, fractures, periodontal disease of mandibular anterior teeth; adequate oral hygiene and treatment compliance **Exclusion criteria**: not mentioned **Setting:** university setting at the orthodontic clinic of Baskent University, Ankara, Turkey
Interventions	**Two types of fixed lower retainers** • Group 1: fixed lingual mandibular retainers from canine to canine using CAD/CAM Memotain retainers (CA-Digital, Mettmann, Germany) made of ⌀ 0.014 × 0.014 inch rectangular nickel–titanium wires • Group 2: fixed lingual mandibular retainers from canine to canine using handmade ⌀ 0.0215 inch 5-stranded stainless steel coaxial wire retainers (GC Orthodontics Inc., Alsip, USA)
Outcomes	Survival: failure rate Adverse effects on periodontal health: • Plaque Index • Gingival Index • Probing depth • Marginal recession • Bleeding on probing

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kartal 2021 *(Continued)*

	Time points: 1 week post-debonding, 1, 3 and 6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	“Picking one of the opaque, sealed envelopes, which were prepared before the trial and each contained a card indicating the allocated group number”
Allocation concealment (selection bias)	Low risk	Participants and investigators enrolling participants could not foresee assignment: "sealed envelopes”
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessment is mentioned; however, with the retainers in place, it is possible that the outcome measurement is likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing data, no dropouts: “No dropouts were observed during the 6-month follow-up period in this study”
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	No other bias detected

Kaya 2019

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 30 **Age (years):** 17.03 (SD 3.16), range 11.16-23.5 **Sex (M:F):** not reported **Inclusion criteria:** fixed orthodontic treatment with the straight-wire technique using 0.018-inch slot Roth brackets; Class I skeletal pattern; no previous orthodontic treatment; achievement of optimum occlusion; treatment compatible with the use of a retainer and long-term follow-up; good oral hygiene **Exclusion criteria:** necessity of using a bonded retainer; placement of a tooth in the retainer due to congenital tooth deficiency; cleft lip and palate; orthognathic surgery **Setting:** university hospital orthodontic department, Turkey
Interventions	**Two types of removable retainers** • Group 1: upper-lower clear plastic retainers (Essix, 0.040 inch sheets (Dentsply Raintree Essix, New Orleans, Louisiana, USA)) • Group 2: upper-lower Hawley retainer Full-time for 6 months then night only for 6 months
Outcomes	Stability

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kaya 2019 *(Continued)*

- Overjet and overbite (mm)
- Intercanine, intermolar and arch length in both arches (mm)
- LII in each arch (mm)
- Changes in lateral cephalometric measurements. These are not relevant for this review.

Time point: 1 year

Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"The retainer type for each patient was randomly allocated by the technician." There is significant risk of bias in this, as the technician, who has orthodontic knowledge, is making this decision.
Allocation concealment (selection bias)	High risk	Participant and operator were not aware, but technician made the decision with the models available.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	This could have been done, but it was not stated that it was done.
Incomplete outcome data (attrition bias) All outcomes	Low risk	No flow chart diagram but no reports of dropouts
Selective reporting (reporting bias)	Low risk	No suggestion that other outcomes were not reported
Other bias	Low risk	No other bias detected

Krämer 2019

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 104 (52 per group) **Age (years):** 17.1 (SD not provided) **Sex (M:F):** 1:1 **Inclusion criteria:** fixed appliances in both jaws (standard straightwire, 0.022 slot size, MBT braces prescription) **Exclusion criteria:** rapid maxillary expansion; orthognathic surgery; segmented appliance; agenesis or extraction of mandibular incisors; orofacial cleft; other craniofacial syndromes **Setting:** public hospital, Sweden
Interventions	**A removable mandibular retainer compared to a fixed mandibular retainer**

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Krämer 2019 *(Continued)*

	• Group 1: removable clear plastic (vacuum-formed) retainer: 1 mm Essix C+ partial coverage - covering teeth from premolar to premolar but not molars (full-time wear for 1 week, then nights only up to 12 months, then alternate nights) • Group 2: fixed canine to canine retainer: ⌀ 0.8 mm stainless steel hard single-strand (Remanium wire Dentaurum, Ispingen, Germany) wire bonded only to canines (not incisors) All participants received a clear plastic retainer (1.0 mm Essix C+ Plastic; Dentsply, Sarasota, USA) in the maxilla covering all erupted teeth.
Outcomes	Stability: LII, arch length (right and left and total), intercanine width, intermolar width, overjet and overbite (all in mm) Survival: expressed as % failure (lost/not fitting/detachment) Patient satisfaction: 100 mm VAS for each of 21 questions; scales (ordinal responses) used for 3 questions **Time points**: 6 and 18 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"The allocation sequence for both groups was computer generated by a statistician"
Allocation concealment (selection bias)	Low risk	"...concealed in sequentially numbered, opaque, and sealed envelopes until randomization. The envelopes were prepared by a colleague who was independent of the trial"
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not possible. "Blinding of the examiner for the assessment of the anonymized digital models was only feasible for measurements at T0 and T1, as the bonded retainers were visible on the dental casts at T2 and T3"
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts, shown in CONSORT diagram. ITT analysis was performed for missing outcome data. "If there were missing values for LII at T2 and T3, we assumed a worst-case scenario by imputation of the 75th percentile value for respective group at current time point"
Selective reporting (reporting bias)	Low risk	Other outcomes in separate paper (second reference of the study)
Other bias	Low risk	No other biases identified

Kumar 2011

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 224 participants **Sex (M:F)**: not reported **Age (years)**: not reported

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Kumar 2011 *(Continued)*

	Inclusion criteria: people who had pre-adjusted edgewise appliance in each arch and willing to wear upper and lower retainers **Exclusion criteria**: none reported **Setting**: university clinic in Department of Orthodontics & Dentofacial Orthopedics, College of Dental Science, Davangere, Karnataka, India
Interventions	**Two types of removable retainers** • Group 1: upper and lower Begg retainers. Acrylic baseplates with labial bow of 0.9 mm stainless steel wire. Labial bow had U-loops opposite the premolars and extended past the last erupted molars • Group 2: upper and lower clear plastic (vacuum-formed) - Essix A+, (Raintree Essix, New Orleans, USA) retainers All participants had bonded wire placed in the lower as well. Participants were instructed to wear retainers 24 h/day for 6 months (and then 12 h/day for the following 6 months - although the 1-year data were not reported). They were instructed to eat with the retainers in place and remove them after eating for cleaning.
Outcomes	Stability assessed using PAR and LII. The results for the lower irregularity were reported, but not the upper. Participant satisfaction data included the mean, but no SDs, so the data could not be analysed using Review Manager software (RevMan Web 2022). We contacted the authors to see if they could provide the raw data to allow analysis, but received no reply. **Time points**: relapse was assessed at 2 time points after debond: 3 and 6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not reported - study authors contacted about this but no reply received
Allocation concealment (selection bias)	Unclear risk	Not reported - study authors contacted about this but no reply received
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported - study authors contacted about this but no reply received
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not reported - study authors contacted about this but no reply received
Selective reporting (reporting bias)	High risk	Results of irregularity reported in the lower but not the upper arch
Other bias	Unclear risk	Insufficient data in the paper to determine this. Study authors contacted but no reply received

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Millett 2007

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 85 participants **Sex (M:F)**: 39:46 **Age (years)**: not reported **Inclusion criteria**: all participants had undergone fixed appliance treatment in at least the lower arch with satisfactory alignment and: were in good health; brushed teeth at least twice a day; had full complement of teeth in lower labial segment with these teeth being of normal size and shape; were willing to comply with study protocol; included extraction and non-extraction cases **Exclusion criteria:** general medical health problems such as diabetes mellitus, epilepsy, physical or mental disability; poor periodontal health; uncontrolled caries; absent or unusually shaped lower incisors; cleft palate or other severe facial deformities **Setting**: hospital department in Cork, Republic of Ireland
Interventions	**A removable retainer compared to a fixed retainer** • Group 1: bonded retainer was ⌀ 0.018 inch stainless steel multistrand wire bonded to lingual of all lower incisors and canines • Group 2: clear plastic retainer was full occlusal coverage worn full-time from the day it was fitted In the upper arch, the clinician chose the upper retainer (Hawley or clear plastic).
Outcomes	Stability assessed by LII. The initial protocol discussed the use of PAR index, but these data were not collected. Survival of retainers - clear plastic retainers were assessed as failed if they fractured or were lost (but not worn); bonded retainers were assessed as failed if debonded (clarified this with the study authors). Adverse effects on health - caries (evidence of decalcification) and periodontal health (gingival bleeding and periodontal probing > 3 mm) was assessed for lower anterior teeth. Participant satisfaction - participant questionnaire assessing retainer wear and acceptability recorded. Data for both retainers available for the following questions: • "How acceptable did you find the retainer brace to wear?" • "How easy was it to keep your retainer brace clean?" • "How happy are you with the way your teeth look now?" (at the end of the retention period) **Time point**: 1 year after debond
Notes	Additional outcome was operator perception of each retainer type; this will not be reported here as it is not an outcome in this review.

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Study authors confirmed computer-generated random sequence
Allocation concealment (selection bias)	Low risk	Sealed consecutive opaque envelopes in a separate co-ordinating centre

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Millett 2007 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Contact with the study author confirmed that assessor was blinded to intervention type.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropouts were reported. Study authors confirmed the data were analysed on an ITT basis.
Selective reporting (reporting bias)	High risk	Not all outcomes were reported in the abstracts that had been described in the protocol (PAR index for occlusal change was not collected). Further data will be reported after 2 years of wear.
Other bias	Unclear risk	As this study was only available as research abstracts, it is unclear as to whether other bias exists.

Moslemzadeh 2018

Study characteristics	
Methods	3-arm parallel RCT
Participants	**Number recruited**: 66 (22 per group) **Age (years)**: range 11-35, mean (SD) per group: Hawley 17.97 (3.87), 1.5 mm clear plastic 19.61 (5.22), 1 mm clear plastic 20.23 (5.55) **Sex (M:F):** not presented in the paper, just that groups were matched in terms of numbers of male and female participants **Inclusion criteria**: completion of orthodontic treatment with an optimal occlusion and an indication for using a removable retainer in the maxilla **Exclusion criteria**: loss of several teeth and the need for their prosthetic replacement, bruxism, syndromic conditions (cleft lip and palate, hemifacial microsomia), allergic to acrylic resin, temporomandibular junction disorders, oral habits leading to malocclusion, undergoing maxillary expansion treatment, early debonding, generalised spacing, any specific systemic condition, periodontal diseases, use of tobacco and use of any specific medication **Setting**: private orthodontic clinic, Tabriz, Iran
Interventions	**3 types of removable retainers** • Group 1: Hawley retainer • Group 2: clear plastic (vacuum-formed) retainer with a thickness of 1.5 mm • Group 3: clear plastic (vacuum-formed) retainer with a thickness of 1 mm
Outcomes	ABO-OGS index was used to evaluate the stability of the results of orthodontic treatment. (A special ABO measuring gauge was used to evaluate the 7 parameters of occlusion, including alignment, buccolingual inclination, occlusal contacts, overjet, interproximal contacts, marginal ridges, and occlusal relationship.) Gingival Index (GI) was used to evaluate gingival health (score 0: normal gingiva with no inflammation; score 1: mild inflammation; slight changes in colour and slight oedema, with no bleeding on probing; score 2: moderate inflammation; oedema, with bleeding on probing; Score 3: severe inflammation; mild oedema, ulceration and tendency for spontaneous bleeding). **Time points:** baseline and 6 months after delivery of retainers

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Moslemzadeh 2018 *(Continued)*

Notes

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"The Randlist software program was used to randomly assign the subjects to three groups."
Allocation concealment (selection bias)	Unclear risk	Not mentioned in the paper
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Low for ABO grading, Unclear for Gingival Index "The ABO-OGS measurements were made by one operator blinded to the type of the retainers used by the patients" No mention in the text as to whether the operator scoring the gingival health was aware of the type of retainer used.
Incomplete outcome data (attrition bias) All outcomes	Low risk	All participants completed the study and none was excluded from the study.
Selective reporting (reporting bias)	Low risk	All data appears to have been reported and there were no dropouts.
Other bias	Low risk	No areas of risk of bias were identified.

Nagani 2020

Study characteristics

Methods	Multicentre, 2-arm parallel RCT
Participants	**Number recruited**: 54 randomised, 52 included in statistical analysis (2 dropouts due to loss to follow-up) **Age (years)**: mean 21.5 (SD 3.6) **Sex (M:F):** 8:44 **Inclusion criteria:** non extraction cases < 45 years, with moderate crowding and normal facial pattern who agreed to visit after every 3 months for a follow-up of 1 year and were treated with MBT fixed appliance **Exclusion criteria:** congenitally missing anterior teeth and poor oral hygiene **Setting:** university setting at the Department of Orthodontics, Dr Ishrat Ul Ebad Khan Institute Of Oral Health Sciences (DIKIOHS) and Dow University of Health Sciences (DUHS), Pakistan
Interventions	**2 types of fixed lingual lower retainers** • Group 1: fibre-reinforced composite retainers from canine to canine using fibre-reinforced composite retainers (INOD, U.P. Fiber Splint, 2 mm) • Group 2 (control group): wire retainers from canine to canine using multistranded stainless steel wire retainers (All Star Orthodontics, ⌀ 0.0175 inch).

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Nagani 2020 *(Continued)*

Outcomes	Stability: LII (contact point displacement) Survival of retainers • Failure rate (bond failure per tooth) • Failure pattern: Adhesive Remnant Index (ARI) **Time points:** assessed at 3, 6, 9 and 12 months after insertion. Data reported as medians and IQR
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	“randomly divided through computer generated software (Microsoft Excel) into two groups”
Allocation concealment (selection bias)	Unclear risk	Method of concealment is not described
Blinding of outcome assessment (detection bias) All outcomes	High risk	Authors stated: "although blinding was not possible for the interventions given as they were visible on the surfaces of the teeth, blinding was done for the outcome assessment which minimizes the observation and detection biases. However, with the retainer in place, it is impossible to blind the outcome assessor."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups
Selective reporting (reporting bias)	Low risk	The study protocol is available and all of the study’s prespecified (primary and secondary) outcome were reported.
Other bias	Low risk	No other bias identified

Naraghi 2021a

Study characteristics

Methods	2-arm parallel RCT
Participants	**Number recruited**: 63 randomised participants, 63 analysed participants (Group 1 = 32 vs Group 2 = 31) - 2 participants in the non-retention group (G2) were considered treatment failures in the final analysis as the participant received a retention device to prevent further relapses. **Age (years)**: mean 12.9 (SD 1.5) **Sex (M:F):** 24:39 **Inclusion criteria**: having at least 1 impacted or unerupted maxillary canine; moderate irregularity of the maxillary 6 anterior teeth of 4–6 mm according to Little’s index (LI) **Exclusion criteria**: agenesis or extracted maxillary anterior teeth; rotations exceeding 45° of one of the maxillary incisors; in need of orthognathic surgery; with syndromes

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Naraghi 2021a *(Continued)*

	Setting: university setting at the Orthodontic Clinic in Växjö, Public Dental Service, Kronoberg County, Sweden
Interventions	**Retention vs non-retention** • Group 1: retention group: participants in the retention group were fitted with the upper removable clear plastic (vacuum-formed) retainer on the day of debond and were instructed to wear the retainer 22-24 h/day until the first visit after 4 weeks. Following that, participants were instructed to wear the clear plastic retainer for 10–12 h/day. In case of pre-treatment spacing, the clear plastic retainer was replaced by a bonded retainer to all 6 teeth from canine to canine (Penta-One 0.0195 inch, Masel, Carlsbad, California, USA). • Group 2: non-retention group
Outcomes	Stability: LII, arch length, intercanine width, intermolar width, maximum rotation **Time points:** assessed at the end of active treatment and after 1-year follow-up
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Block randomisation was performed taking random note from sealed opaque envelopes, quote: "took a note from the first envelope"
Allocation concealment (selection bias)	Low risk	"Opaque envelopes contained 20 sealed notes each…"
Blinding of outcome assessment (detection bias) All outcomes	High risk	Because of the type of treatment, neither participant nor operator was blinded to participant group.
Incomplete outcome data (attrition bias) All outcomes	Low risk	ITT analysis. Low dropout number and reasons given
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	No other bias detected

Naraghi 2021b

Study characteristics

Methods	3-arm parallel RCT
Participants	**Number recruited**: 90 randomised participants, 90 analysed participants (Group 1 = 30 vs Group 2 = 30 vs Group 3 = 30) **Age** (years): 13.9 (SD 1.7) **Sex (M:F):** 36:54 **Inclusion criteria**: adolescents treated with fixed appliance in the maxilla or the maxilla and mandible

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Naraghi 2021b *(Continued)*

	Exclusion criteria: having clefts or syndromes, agenesis or extracted maxillary anterior teeth, or having undergone orthognathic surgery **Setting:** university setting located at the Orthodontic Clinic Växjö, Public Dental Service, Region Kronoberg, Sweden
Interventions	**Two types of fixed retainers and a removable retainer** • Group 1: fixed maxillary multistrand retainers from 13–23, Penta-One ⌀ 0.0195 inch (Masel, Carlsbad, California, USA) • Group 2: fixed maxillary multistrand retainers from bonded retainer 12–22, Penta-One ⌀ 0.0195 inch (Masel, Carlsbad, California, USA) • Group 3: removable clear plastic retainer covering the maxillary teeth including the second molars. Essix, Erkodur, 1.5 mm 120ø, (Erkodent Erich Kopp GmbH, Pfalzgrafenweiler, Germany). The participants followed a standard protocol for wear: 22–24 hours/day during the first 4 weeks, then every night. One year after debonding, the wearing time was reduced to every other night.
Outcomes	Stability: LII, arch length, intercanine width, intermolar width, overjet, overbite, maximum rotation **Time points:** before treatment (T0), after removal of the fixed orthodontic appliances/start of retention (T1), after 2 years of retention (T2)
Notes	Second reference reports additional outcome (primary) mentioned: cost analysis was made regarding chair time costs based on the costs per hour for the specialist in orthodontics, and material costs plus any eventual costs for repairs of the appliance; this is not reported here as it was not one of our outcomes.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"The randomization notes were delivered in a sealed opaque envelope each. Every new participant picked an envelope and revealed the group assignment by opening the envelope"
Allocation concealment (selection bias)	Low risk	"The randomization notes were delivered in a sealed opaque envelope each. Every new participant picked an envelope and revealed the group assignment by opening the envelope"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Study casts were anonymised before the assessment, and the assessor was blinded at T2 for the allocation groups. Before taking impression and evaluations all retainers were removed, there were no traces of any retainer at the time of analysis.
Incomplete outcome data (attrition bias) All outcomes	Low risk	ITT analysis. Low dropout number and reasons given
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	No other bias detected

O'Rourke 2016

Study characteristics

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

O'Rourke 2016 *(Continued)*

Methods	2-arm parallel RCT
Participants	**Number recruited**: 82 (59 female and 23 male); at 4-year follow-up: 42 (10 men, 32 women) **Sex (M:F):** 23:59; at 4-year follow-up: 10:32 **Mean age (years):** 17.73 (SD 3.52); at 4-year follow-up: 21.15 (SD 2.41) **Inclusion criteria**: people who had completed treatment with 0.022 x 0.028 inch pre-adjusted Edge-wise fixed appliances with: pre-treatment records and study models available to confirm the presence of pre-treatment labial segment crowding or spacing; clinically acceptable alignment at the end of treatment **Exclusion criteria**: completed treatment early or had repeated breakages during treatment; poor oral hygiene during treatment; prosthodontic needed in the lower arch at end of treatment; history of peri-odontal disease; learning difficulties **Setting**: NHS hospital orthodontic department in London, UK
Interventions	**A removable retainer compared to a fixed retainer** • Group 1: clear plastic retainer (Essix ACE 0.030 inch) in lower arch fitted within 7 days of debond. To wear full-time wear for the first 6 months followed by 6 months of nighttime wear and a further 6 months of alternate night wear • Group 2: bonded retainer canine to canine ⌀ 0.0175 inch co-axial stainless steel bonded to every tooth passively with Transbond LR
Outcomes	Stability (reported in both publications): assessed at 6, 12, 18 months (O'Rourke 2016) and at 48 months (Al-Moghrabi 2018: secondary ref under O'Rourke 2016) by measuring: LII, intercanine width, intermolar width, arch length, extraction site space opening Failure of retainers (Al-Moghrabi 2018) at 48 months Oral health: assessed as gingival inflammation – modified Gingival Index; calculus – part of Oral Hy-giene index; plaque levels – modified Quigley–Hein Plaque Index; clinical attachment level – mm; bleeding on probing – present/absent
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence genera-tion (selection bias)	Low risk	"Used electronic randomisation programme generating a random allocation sequence"
Allocation concealment (selection bias)	Low risk	"Allocation was concealed in blackout envelopes that were identical, tamper evident and prepared in advance"
Blinding of outcome as-sessment (detection bias) All outcomes	Low risk	Models measured in random order and blinded for data analysis
Incomplete outcome data (attrition bias) All outcomes	Low risk	O'Rourke 2016**:** operators used an ITT analysis. The study authors reported that any participants who failed to attend their appointments were analysed in their original group. Imputation values were used in case of missing data, calculated as the mean of the outcome being measured. The authors acknowl-edged that due to a higher number of dropouts at 18 months, the use of impu-tation may give increased uncertainty to the results.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

O'Rourke 2016 *(Continued)*

		Al-Moghrabi 2018 (see O'Rourke 2016**):** dropouts were significant over the 4-year period, although the final sample of 42 was just 6 fewer than that obtained 2.5 years previously. Dropout rates were balanced between the groups, and the main reason for failure to attend was logistical, ensuring that missing data occurred at random, and therefore the risk of attrition bias was minimised.
Selective reporting (reporting bias)	Low risk	No registration on clinical trials registers, but study authors confirmed there were no other data to report (apart from cost-effectiveness data, which is not relevant to this review)
Other bias	Low risk	No other potential areas of bias identified

Pandis 2013

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 220 participants (110 in each group) **Median age (IQR)(years):** 16.0 (2) **Sex (M:F):** 60:160 **Inclusion criteria**: no active caries, restorations, fractures on the mandibular anterior teeth or periodontal disease; adequate oral hygiene **Exclusion criteria**: unwilling to be assigned to any of the approaches or had any abnormal oral or medical condition contraindicating fixed retention **Setting**: private practice at Bern, Switzerland
Interventions	**2 types of adhesives and composites** • Group 1: chemical group; maximum Cure 2-part liquid adhesive (Reliance Orthodontic Products, Itasca, Ill) and Excel 2-part paste (Reliance Orthodontic Products) • Group 2: photo polymerisation group; light-cured liquid (Assure; Reliance Orthodontic Products) and paste in 2 layers (Flow-Tain; Reliance Orthodontic Products) All participants received a soft-bonded lingual retainer of 0.022 inch (Tru-Chrome multi-stranded wire; Rocky Mountain Orthodontics, Denver, Colo) that was fabricated intraorally.
Outcomes	Survival: failure of the lingual retainer, failure risk (hazard ratio), adhesive remnant index (ARI) scores **Time points:** 3, 6, 12, 18, 24 months
Notes	

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Well-reported randomisation process, using a computer random number generator (Stata software, Corp, College Station, Tex) in random permuted blocks of 20 participants, ensuring equal distribution in the 2 groups

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Pandis 2013 *(Continued)*

Allocation concealment (selection bias)	Low risk	Central allocation by the practice manager responsible for opening the next envelope in sequence
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The outcome assessment was blinded. "Blinding was applicable for outcome assessment only."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Trial authors imputed missing data using appropiate methods. Trial authors carried out imputations under the missing at random assumption.
Selective reporting (reporting bias)	Unclear risk	The study protocol is not available. Insufficient information to permit judgement of 'low risk' or 'high risk'
Other bias	Low risk	The study appears to be free of other sources of bias.

Ramazanzadeh 2018

Study characteristics	
Methods	3-arm parallel RCT
Participants	**Number recruited**: 90 participants (28, 27, 27 in each group) **Age**: mean (SD) per group: Group 1: Hawley 16.2 (2.4) Group 2: Essix 4 months 16.5 (2.4) Group 3: Essix 1 week 16.4 (2.8) **Sex (M:F):** 30:52 **Inclusion criteria**: age range between 14-30 years; treated with fixed orthodontic appliances in both jaws; removed 4 premolar teeth for orthodontic reasons; presented 4-7 mm of crowding at the start of orthodontic treatment; intended to wear maxillary and mandibular retainers **Exclusion criteria:** single-arch or sectional fixed orthodontic treatment; hypodontia requiring tooth replacement in the retainer; need for placing bonded retainer; poor periodontal condition; presence of cleft lip or palate; orthodontic treatment combined with orthognathic surgery; need for additional procedures such as interdental stripping or aesthetic restorative treatment **Setting**: Dental Research Center, Mashhad, Iran
Interventions	**Two types of removeable retainers, one with two different protocols:** • Group 1: Hawley retainers 4 months full time and then night only • Group 2: clear plastic retainer (Essix) 4 months full-time and then night only • Group 3: clear plastic retainer (Essix) 1 week full-time and then night only
Outcomes	Stability: intercanine width, intermolar width, arch length and the LII, presumably all measured in mm. This is not explicit in the text but the study authors do state they used digital dial callipers with accuracy of 0.1 mm. **Time points:** debond, 4 and 8 months later
Notes	

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Ramazanzadeh 2018 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not discussed in paper
Allocation concealment (selection bias)	Unclear risk	Not discussed in paper
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcomes measured on randomised study models by single operator who was blinded to the groups to which the participants had been allocated
Incomplete outcome data (attrition bias) All outcomes	Low risk	Low dropout evenly spread amongst the groups
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	No other biases identified

Rohaya 2006

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 218 participants (this figure differed from the published paper and was provided by the study author) **Age**: not reported **Sex (M:F):** not reported **Inclusion criteria**: participants received upper fixed appliance treatment (and sometimes lower fixed in addition). They may have also had an initial stage of functional appliance and possibly extractions. In addition, participants were prepared to wear either a Hawley or a clear plastic (vacuum-formed) retainer **Exclusion criteria**: previous orthognathic surgery; hypodontia treated by re-opening for prosthetic teeth; marked reduction in periodontal bone support; upper arches that had been substantially expanded with either a rapid maxillary expansion device or a quadhelix; cleft lip and palate or other craniofacial syndrome **Setting**: district general hospital in Malaysia, treatment free of charge
Interventions	**Two types of removable retainers** • Group 1: upper Hawley retainer full-time for 3 months then 6 months of nights only (9 months in total) • Group 2: upper clear plastic retainer (Essix C, Raintree Essix) full-time for 1 week then evenings and nights for 3 months, then nights only for 6 months (9 months in total)
Outcomes	Stability: assessed only in relation to previously rotated upper teeth (upper incisors, canines and premolars). Only participants with ≥ 1 rotated tooth included in analysis

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rohaya 2006 *(Continued)*

	Time point: 9 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	We contacted study authors, who confirmed that this was undertaken appropriately. Computer program used to produce randomisation and then placed in numbered sealed envelopes prepared in advance by someone else
Allocation concealment (selection bias)	Low risk	We contacted study authors, who confirmed that this was undertaken appropriately. Computer program used to produce randomisation and then placed in numbered sealed envelopes prepared in advance by someone else
Blinding of outcome assessment (detection bias) All outcomes	Low risk	We contacted study authors, who reported about outcome assessor blinding: "A separate person assessed the models without knowing which retainer had been worn. The models were randomised and only identifiable by a number, the code was broken only after all models had been measured and the results were then linked to the retainer groups for statistical analysis"
Incomplete outcome data (attrition bias) All outcomes	High risk	The number of dropouts was very high (32% in Hawley group and 39% in clear plastic retainer group) and these were not analysed or taken into account.
Selective reporting (reporting bias)	High risk	Other outcomes collected but not reported. We contacted study authors to request this information - they reported that they are planning to publish this in the future.
Other bias	Low risk	No other biases identified

Rose 2002

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 20 **Sex (M:F)**: 12:8 **Mean age (years)**: 22.4 (SD 9.7) **Inclusion criteria**: previously identified as needing a lower bonded retainer by the orthodontist in charge of the clinic, healthy periodontal condition, good oral hygiene, no previous bonded retainer, demonstrated consistent compliance during fixed appliances **Exclusion criteria**: none reported **Setting**: university clinic, Department of Orthodontics, Freiburg University, Germany
Interventions	**2 types of lower fixed retainers:** • Group 1: fibre-reinforced composite retainer ('Ribbond') was cut to the correct length, pre-treated with adhesive bis-GMA sealant and the bonded with 'Heliosite' orthodontic composite resin • Group 2: multistrand wire retainer was ⌀ 0.0175 inch passive multistrand wire

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rose 2002 *(Continued)*

	Both types were bonded to all teeth, canine to canine in the lower arch. The multistrand retainers were bonded under rubber dam.
Outcomes	Survival: failure rates **Time points:** recorded at 2 years
Notes	Paper describes patient satisfaction with the retainer types but no data presented on this

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Appropriate randomisation (confirmed with study author - used random number generation. NB this information is from a previous version of the review and we have no details about the email received).
Allocation concealment (selection bias)	Low risk	Confirmed as appropriate after contact with study author (randomisation was prepared independent to clinician so could not be aware of allocation)
Blinding of outcome assessment (detection bias) All outcomes	High risk	Impossible to blind the assessor to outcomes in this study
Incomplete outcome data (attrition bias) All outcomes	Low risk	There were no dropouts (confirmed with study author; NB this information is from a previous version of the review and we have no details about the email received).
Selective reporting (reporting bias)	High risk	Single outcome reported. Description of patient satisfaction with retainer type, but no data reported
Other bias	Low risk	No other areas of bias identified

Rowland 2007

Study characteristics

Methods	2-arm parallel RCT
Participants	**Number recruited**: 397 **Sex (M:F):** 156:241 **Mean age**: group 1: 14.8 years (SD 1 year 8 months), group 2: 15 years (SD 1 year 5 months) **Inclusion criteria**: all participants had preadjusted edgewise fixed appliances, with pretreatment records, treatment plan and study models available, and were willing to wear maxillary and mandibular retainers **Exclusion criteria**: single arch or sectional fixed appliance treatment; hypodontia requiring tooth replacement on the retainer; poor periodontal status; early debonding; transfer patients; learning difficulties; cleft lip or palate **Setting**: NHS specialist practice in Bristol, UK, treatment free of charge by 1 operator
Interventions	**Two types of removable retainers** • Group 1: upper and lower Hawley retainers, worn 3 months full-time then 12 h/day for 3 months

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Rowland 2007 *(Continued)*

	• Group 2: upper and lower clear plastic retainers - worn full-time for 1 week then 12 h/day. Full coverage design made from 1.5 mm 'Erkodur' blank Data collected after 6 months of retention
Outcomes	Stability: measured using: LII, intercanine width, intermolar width. Authors reported that tooth rotations, overjet and overbite analysis were measured, but no data were provided in the paper. Patient satisfaction assessed by questionnaire. Results of the most relevant questions provided in the paper Survival of retainers - number of broken retainers recorded for each type of retainer Cost-effectiveness was also reported, but was not one of the outcomes for our review. **Time points:** measured at 6-month period of retention
Notes	Assessment of stability data provided as medians and IQRs. This could not be analysed using Review Manager 5 statistical software (Review Manager 2020), so we requested the raw data, which the study authors provided. Helen Worthington (author of this review) then converted these into data that could be put into Review Manager 5. Patient satisfaction data were dichotomised for full analysis as follows: • responses related to whether retainers were worn away from home dichotomised by grouping "always" and "nearly always" together • amount of discomfort data were dichotomised by grouping "never" and "on occasion" together • comparison to fixed appliances was dichotomised by grouping together "much better" and "better", and grouping together "worse" and "much worse"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Used block randomisation (taken from "www.randomisation.com")
Allocation concealment (selection bias)	Low risk	Concealment done by ensuring randomisation was performed after consent was obtained, using the computer program www.randomisation.com, so the clinician and participant could not know what intervention would be generated
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome assessor blinded to the type of retainer used
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number of dropouts clearly reported and described
Selective reporting (reporting bias)	High risk	Tooth rotations, overjet and overbite data measured but not reported. The authors also only reported the "most relevant" questions on the patient satisfaction questionnaire.
Other bias	Low risk	Pretreatment equivalence shown for both groups and no other apparent bias present

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saleh 2017

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 94; 86 completed the study **Sex (M:F):** 37:49 **Mean age (years)**: 26.3 (SD 6) **Inclusion criteria**: treatment only with fixed appliances, no lateral expansion treatment, no hypodontia, no cleft lip and palate, no surgical corrections, no extraction-based plans, ≥ 18 years, willingness to wear maxillary and mandibular removable retainers **Exclusion criteria**: cleft lip and palate, hypodontia, treated with orthognathic surgery, only one arch treatment **Setting**: Saudi Swiss Consultant Dental Centre, Saudi Arabia
Interventions	**Two types of removable retainers** • Group 1: Hawley retainer fitted with upper and lower Hawley retainers at debond • Group 2: clear plastic retainer fitted with upper and lower at debond Participants were instructed to wear both retainers 24 h/day for the first 6 months and 12 h/day (nighttime wear) for the next 6 months. Fixed lower lingual retainers extending from canine to canine, made from multistranded stainless steel twisted wire were bonded in all participants.
Outcomes	Patient satisfaction: questionnaire using 10 cm VAS (10 = most favourable) with 10 questions on teeth: 1) biting (closing teeth with retainers in, not chewing food); 2) how well the appliance fitted; 3) speech; 4) appearance; 5) oral hygiene; 6) durability; 7) gingival irritation; 8) swallowing; 9) self-confidence; 10) comfort **Time points:** data collected after 1 week, 3 months, 6 months of retention
Notes	As all participants were also fitted with lower fixed lingual retainers, extending from canine to canine, this does not give a fair representation of the efficacy of the lower Hawley or lower VFR alone.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Manual allocation by staff not involved in the research, "participants picked a concealed opaque envelope from a black plastic box. This box contained 94 envelopes with 47 containing the letter H (i.e. HR) and the other 47 containing the letter V (i.e. VFR)".
Allocation concealment (selection bias)	Low risk	Allocation was concealed using opaque envelopes
Blinding of outcome assessment (detection bias) All outcomes	High risk	High for recording of responses on the questionnaires: "Questionnaires were filled in front of the treating doctor at their normal follow-up visits"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number of dropouts clearly reported and described

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Saleh 2017 *(Continued)*

Selective reporting (reporting bias)	Low risk	All planned outcomes were assessed.
Other bias	High risk	The paper claims to be comparing Hawley retainers with clear plastic retainers, but all participants were fitted with lower bonded retainers canine to canine too.

Salehi 2013

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 142 **Sex (M:F):** 59:83 **Mean age (years)**: group 1: 18.1 (SD 5.23), group 2: 18.2 (SD 4.81) **Inclusion criteria**: people treated with standard edgewise fixed appliances, good oral hygiene, healthy periodontium, no previous bonded retainer **Exclusion criteria**: deep overbite, traumatic parafunctional habits such as bruxism and clenching **Setting**: clinic of 1 operator in Iran; participants paid for treatment
Interventions	**Two types of upper and lower fixed retainers** • Group 1: fibre-reinforced composite retainer ('Ribbond') was cut to the correct length, pre-treated with adhesive bis-GMA sealant and bonded with 'Heliosite' orthodontic composite resin • Group 2: multistrand wire retainer was ⌀ 0.0175 inch passive multistrand wire Both types were bonded to all teeth, canine to canine Both retainers bonded under rubber dam with Heliosite composite resin using Fluoro Bond adhesive bis-GMA sealant
Outcomes	Survival: retainer assessed as failed if it debonded from tooth or fractured **Time points:** at 18 months
Notes	No assessment of stability, adverse effects on health or patient satisfaction

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random table number table used
Allocation concealment (selection bias)	Unclear risk	No concealment described. Study authors contacted to clarify this but no reply received
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Study authors noted that the statistician analysing the data was blinded

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Salehi 2013 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Low risk	Number of dropouts clearly reported and described
Selective reporting (reporting bias)	Low risk	No evidence or suggestion of selective reporting
Other bias	Low risk	No other suggestions of bias noted

Scribante 2011

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 34 **Sex (M:F):** 9:25 **Mean age (years)**: 14.3 (SD not provided) **Inclusion criteria**: correct dental alignment, need for permanent orthodontic retention in the lower anterior segments, free of occlusal interferences in order to eliminate the influence of trauma **Exclusion criteria**: not mentioned **Setting**: Unit of Orthodontics and Paediatric Dentistry, University of Pavia, Italy
Interventions	**Two types of lower lingual fixed retainers:** • Group 1: ⌀ 0.0175 mm multistranded wire • Group 2: fibre-reinforced composite retainer preimpregnated with Transbond XT Primer (3M Unitek, Monrovia, California, USA)
Outcomes	Survival: numbers (percentage) of total detachment at the end of follow-up, Kaplan-Meier survival plots and hazard ratio Patient satisfaction: opinion about the aesthetic result while talking and smiling by means of VAS **Time points:** every month for 12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Just mentioned "randomly assigned"
Allocation concealment (selection bias)	Unclear risk	Not mentioned
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not mentioned

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Scribante 2011 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	The number of participants is not clear: "34 patients were enroled: 17 with a multistrand wire retainer, while 15 with a polyethylene fiber-reinforced resin composite"
Selective reporting (reporting bias)	Low risk	No dropouts in the 12-month follow-up
Other bias	Unclear risk	Not reported using CONSORT standards

Scribante 2020

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 100 **Sex (M:F):** 54:46 **Mean age (years):** 25 (SD not provided) **Inclusion criteria**: being at the end of an orthodontic treatment requiring the subsequent application of a lingual and palatal fixed retainer **Exclusion criteria**: reported facial trauma, onychophagia and habit of biting pencils or pens **Setting**: Unit of Orthodontics and Paediatric Dentistry, University of Pavia, Italy
Interventions	**Two types of composite used for adhesion of fixed upper and lower retainers** (multistranded wire; Ortosmail Krugg): • Group 1 (control): highly filled orthodontic resin Transbond XT (XT) • Group 2 (intervention): flowable nanocomposite Filtek Supreme XTE (FL)
Outcomes	Survival of retainers: numbers and failure rates of fixed retainers for the overall, upper and lower teeth in the two groups tested (all teeth; distinguishing canines and incisors) **Time points:** 24 months of follow-up, monthly visits
Notes	Two parts reported: in vitro and in vivo (RCT)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	"...block randomization table, the data analyst provided a randomization sequence..."
Allocation concealment (selection bias)	Low risk	"...sequentially numbered and sealed envelopes with the allocation cards previously prepared..."
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Participants, data assessor and data analyst were always blinded during the study; none of them knew the treatment administered to each participant
Incomplete outcome data (attrition bias)	Low risk	No dropouts, no missing outcome data, no variations to the outcome occurred after the study commencement

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Scribante 2020 *(Continued)*

All outcomes		
Selective reporting (reporting bias)	Low risk	All planned outcomes reported
Other bias	Low risk	No other biases identified

Sfrondini 2014

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 100; 87 analysed and 13 dropped out due to non-attendance at follow-up appointments **Sex (M:F):** 35:52 **Mean age (years)**: 24 (SD not provided) **Inclusion criteria**: previously undergoing fixed orthodontic treatment (average duration of 22 months) **Exclusion criteria**: extraction cases, short duration of treatment, and orthognathic surgery patients **Setting**: university setting at the Department of Orthodontics, University of Pavia, Italy
Interventions	**Two types of lower fixed lingual retainers from right to left canine:** • Group 1: fibre-reinforced composite retainer (Everstick Ortho, Stick Tech Ltd, Turku, Finland; Ø 0.5 mm) • Group 2: flexible multi-stranded stainless steel spiral wire (Ortosmail, Krugg spa, Milan, Italy; Ø 0.0175 inch) retainer
Outcomes	Primary: failure rate in terms of detachment Secondary: failure risk (in months) **Time point**: 12 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Retainer type assigned to participants using randomisation tables
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No information provided

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sfrondini 2014 *(Continued)*

Incomplete outcome data (attrition bias) All outcomes	High risk	13 participants dropped out of the study for missing at least one follow-up visit; 10 of them were in the same group (multistranded stainless steel wire retainer)
Selective reporting (reporting bias)	Low risk	No suggestion that there were other outcomes that were not reported
Other bias	Unclear risk	Not reported according to CONSORT guidelines

Sfrondini 2021

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 100 **Sex (M:F):** 48:52 **Mean age**: (overall age and SD not provided) Control group: male: 27 years, 2 months; female: 26 years, 11 months Intervention group: male: 24 years, 2 months; female: 23 years, 10 months **Inclusion criteria**: being at the end of an orthodontic treatment requiring the subsequent application of a lingual and palatal fixed retainer **Exclusion criteria**: reported facial trauma, onychophagia and habit of biting pencils or pens **Setting**: Unit of Orthodontics and Paediatric Dentistry, University of Pavia, Italy
Interventions	**Two primers used with the same composite to bond fixed upper and lower retainers:** • Group 1 (control): traditional orthodontic primer (Transbond XT Primer, 3M ESPE) (XT Primer) • Group 2 (intervention): universal adhesive (Scotchbond Universal, 3M ESPE) (SB) in total-etch mode
Outcomes	Survival of retainers Numbers and failure rates of fixed retainers for the overall, upper, and lower teeth in the two groups tested (all teeth; distinguishing canines and incisors) **Time points:** 24 months of follow-up, monthly visits
Notes	Two parts reported: In vitro and in vivo (RCT)

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	"Block randomization table, the data analyst provided a randomization sequence"
Allocation concealment (selection bias)	Low risk	"Sequentially numbered and sealed envelopes with the allocation cards previously prepared"
Blinding of outcome assessment (detection bias)	Low risk	Participants, data assessor and data analyst were blinded during the study

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sfrondini 2021 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	Low risk	No dropouts, no missing outcome data, no variations to the outcome occurred after the study commencement
Selective reporting (reporting bias)	Low risk	All outcomes reported
Other bias	Low risk	No other biases identified

Shawesh 2010

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 67 **Sex (M:F):** 22:45 **Mean age (years)**: group 1: 15.6 (SD 1.6), group 2: 15.8 (SD 1.2) **Inclusion criteria**: aged 10-16 years; labial segment crowding or tooth contact displacement at start of treatment; clinically acceptable labial segment alignment at end of treatment; good oral hygiene **Exclusion criteria**: lack of consent; severe rotations or mid-line diastema suggesting need for bonded retainer; restorative need in the labial segment, e.g. implant, bridges or missing teeth **Setting**: NHS specialist practice in Manchester, UK; treatment was free of charge
Interventions	**Two types of removable retainers:** • Group 1: part-time wear of Hawley retainers for 1 year (retainers worn only at night) • Group 2: full-time wear of Hawley retainers for 6 months then nights only for 6 months
Outcomes	Stability assessed using LII and crowding assessment **Time point:** at 1 year
Notes	No patient satisfaction assessment, survival of retainers or assessment of effect on dental health

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Very clear description of appropriate randomisation process: "The subjects were randomly allocated to one of the two retention regimen groups using a restricted randomisation technique, in blocks of 12, to ensure equal numbers were allocated to each group. The allocation was decided by throwing an unweighted di where throws of 1, 2 or 3 = group 1 and 4, 5, or 6 = group 2. From this random list, the retention regime was recorded along side a patient identification number" Study authors pointed out that despite the appropriate use of randomisation, there was not complete equivalence between groups at the start of treatment. The group with full-time wear for the first 6 months had 1 mm more crowd-

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shawesh 2010 *(Continued)*

		ing and nearly 2 mm more irregularity at the start of treatment. The authors warned that the results should therefore be viewed with caution.
Allocation concealment (selection bias)	Low risk	After randomisation the allocation was "sealed in numbered opaque envelopes and held in a central place. Thus, neither the clinician nor the patient knew their group allocation prior to consenting to the study"
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Outcome assessor was blinded to the retention regimen used
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number of dropouts clearly reported and described
Selective reporting (reporting bias)	Low risk	No evidence or suggestion of selective reporting
Other bias	Low risk	No other bias detected

Shim 2022

Study characteristics	
Methods	3-arm parallel RCT
Participants	**Number recruited**: 75 (group 1 30; group 2 20; group 3 25); 46 attended 3-month visit (group 1 16; group 2 16; group 3 14) and those participants were analysed **Sex (M:F):** 18:29 (group 1: 9M, 7F; group 2: 4M, 12F; group 3: 5M, 9F) **Median age (IQR)**: group 1: 16.5 (15.3); group 2: 15.8 (13.6); group 3: 15.2 (13.6) **Inclusion criteria**: completion of comprehensive orthodontic treatment; Class I molar and canine relationship based on Andrew's criteria for finishing; need for a fixed retainer on the mandibular anterior teeth assessed before debonding **Exclusion criteria**: refusal to participate in the study and conditions contraindicating fixed retainers, including poor oral hygiene and pathology **Setting**: orthodontic clinic of Saint Louis University, Saint Louis, Missouri, USA
Interventions	**3 types of lower fixed retainers:** • Group 1: CAD/CAM multistrand stainless steel wire (Dentaflex wires, Dentaurum GmbH & Co., Ispringen, Germany) • Group 2: lab-based (lab) retainer, same type of wire as the CAD/CAM group, but manually bent • Group 3: chairside (traditional) rectangular chain stainless steel (Ortho-FlexTech, Reliance, Itasca, Ill)
Outcomes	Stability: LII and intercanine width changes Failure rate**:** complete or partial detachment were considered failures **Time points**: 3 and 6 months
Notes	

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Shim 2022 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Well-informed randomisation process generation. A simple software-generated randomisation was performed by a statistician not involved in the study.
Allocation concealment (selection bias)	Low risk	The sequence generator was contacted by phone for group assignment after determining patient eligibility for enrolment.
Blinding of outcome assessment (detection bias) All outcomes	High risk	"The principal investigator and the clinicians bonding the fixed retainers could not be blinded regarding the assignment of patients to groups in the clinical setting. Upon data analysis, patients' records were encrypted to minimize observer and measurement bias." It is likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups
Selective reporting (reporting bias)	Unclear risk	The present study was not registered. Insufficient information to permit judgement. The published report includes the main primary outcomes.
Other bias	Low risk	The study appears to be free of other sources of bias.

Sinha 2021

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 61 fulfilled inclusion criteria, 40 participated in the study **Sex (M:F):** 20:20 **Mean age (years)**: 15.61 (SD 1.13) years for flexible spiral wire group and 15.83 (SD 1.07) years for ceramic interlocking bonded group **Inclusion criteria**: healthy gingival status with probing depth < 3 mm with no radiographic evidence of bone loss, LII ≤ 2 mm and patients treated with optimum occlusion with the all-conventional orthodontic treatment objectives satisfied. The patients had undergone orthodontic treatment without extractions. **Exclusion criteria**: impacted teeth, anomalies in the lower anterior teeth **Setting**: Department of Orthodontics and Dentofacial Orthopedics, Institute of Dental Studies and Technologies (IDST), Ghaziabad, Uttar Pradesh, India
Interventions	**Two types of fixed lower lingual retainers:** • Group 1: ceramic interlocking retainer canine to canine (Ling Lock 3 M Unitek, Monrovia, USA) • Group 2: flexible round coaxial wire (⌀ 0.016 inch) retainer canine to canine
Outcomes	Stability: LII and 3-D model deviations (in the X, Y and Z axis i.e. the displacement in the mesiodistal, labiolingual and vertical direction, respectively) Failure rate: bond failures Oral health: Plaque Index and Gingival Index

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sinha 2021 *(Continued)*

	Time points: 3 and 6 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A computer-generated random allocation sequence was reported.
Allocation concealment (selection bias)	Low risk	The authors reported the sequences were concealed.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Primary investigator and data analyser were impossible to blind due to the nature of the study.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Insufficient reporting of attrition/exclusions to permit judgement
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement
Other bias	Low risk	The study appears to be free of other sources of bias.

Soboutia 2016

Study characteristics	
Methods	3-arm parallel RCT
Participants	**Number recruited**: 150 randomised, 128 analysed in statistical analysis (22 participants dropped out due to failure to regularly attend the follow-up sessions) **Sex (M:F):** 60:68 **Mean age (years)**: 18.0 (SD 3.6) (range 13-25 years) **Inclusion criteria**: healthy people who had no history of previous dental extraction or orthognathic surgery, with an original indication for non-extraction treatment (Class I crowding, incisor mandibular plane angle [IMPA] < 92°, vertically normal or horizontal and no pattern of vertical excess, and with a crowding extent and soft tissue characteristics appropriate for non-extraction treatment), with ideally aligned dentition, a Class I relationship with an overbite/overjet between 1 and 3 mm, no issues contraindicating or interfering with retention, bruxism or clenching, advanced dental abrasion/attrition/erosion, and a need for mandibular canine-to-canine fixed retention. In addition, patients who had good oral hygiene, healthy periodontal condition and no previous history of using bonded retainers **Exclusion criteria**: patients with widespread probing depths > 3 mm and radiographic evidence of periodontal bone loss **Setting**: private office in Iran
Interventions	**Three types of lower fixed lingual retainers from right to left canine:** • Group 1: fibre-reinforced composite retainer (Everstick Ortho; Stick Tech, Turku, Finland)

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Soboutia 2016 *(Continued)*

- Group B: flexible multistranded stainless steel spiral wire (Ortosmail, Krugg spa, Milan, Italy; Ø 0.0175 inch) retainer
- Group C**:** twisting 2 ⌀ 0.009 inch dead soft wires (Ligature wire; 3M Unitek, Monrovia, CA, USA) retainer

Outcomes	Failure rate in terms of first detachment of any composite pad or retainer breakage/distortion Survival of retainers (in months). Kaplan-Meier survival plots and hazard ratio **Time point**: 24 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Using a random number table"
Allocation concealment (selection bias)	Unclear risk	Not reported
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number of dropouts clearly reported and described
Selective reporting (reporting bias)	Low risk	All outcomes were reported according to its material and method section
Other bias	Low risk	No other area of bias detected

Sun 2011

Study characteristics

Methods	2-arm parallel RCT
Participants	**Numbers recruited**: 120 participants **Sex (M:F):** 60:60 **Mean age (years)**: 14.7 (range 12-17) **Inclusion criteria**: aged < 18 years; all second molars erupted and in occlusal contact; agreed to take part and signed consent with parent or legal guardian present **Exclusion criteria**: allergy to acrylic resin; did not accept either of the 2 types of retainers; unable to comply with follow-up appointments during the 12 months with "the possibility of serious relapse" **Setting**: university clinic at West China Stomatology Hospital of Sichuan, China
Interventions	**Two types of removable retainers** - Group 1: upper and lower Hawley retainers

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Sun 2011 *(Continued)*

	• Group 2: upper and lower clear plastic retainers (0.75 mm thick Biolon, Dreve Dentamid GmbH, Unna, Germany) Participants were asked to wear both types of retainer full-time except during meals. Reviewed at 3, 6 and 12 months, and participants contacted if the retainer failed
Outcomes	**Outcome** relevant to review: survival of retainers Failures of retainers assessed by how often they failed. Retainers could fail by fracture, "no longer fitting", local serious abrasion causing penetration and loss of retainers (reported by participant or observed at appointment) **Time point:** final outcome at 1 year
Notes	Additional outcomes of effect on oral hygiene mentioned on the clinical trials record but not reported here

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random sequence. Researcher managing this did not participate in the enrolment procedure
Allocation concealment (selection bias)	Low risk	Both participant and researcher blinded to randomisation until after the participant had agreed to take part in the study. The method of concealment not described, although study authors reported that they concealed allocation to avoid selection bias and it was done by someone not involved in the enrolment procedure
Blinding of outcome assessment (detection bias) All outcomes	High risk	Impossible to blind the assessor to the type of retainer in this type of study, as the assessor had to record the type of failure/breakage of the retainer, and would therefore see the different types of removable retainers
Incomplete outcome data (attrition bias) All outcomes	Low risk	Loss to follow-up was small and they were reported and evenly distributed between each group (5 in the Hawley group and 4 in the clear overlay group)
Selective reporting (reporting bias)	High risk	The clinical trials record reported that oral hygiene was the primary outcome, but it is not reported in the paper. We contacted the study authors about this, but did not receive a reply
Other bias	Low risk	No other biases identified

Thickett 2010

Study characteristics

Methods	2-arm parallel RCT
Participants	**Number recruited**: 62 **Sex (M:F):** 26:36 **Age (years)**: group 1 mean 13.6 (SD 1.5), group 2 mean 13.8 (SD 1.5) **Inclusion criteria**: malocclusion requiring premolars extracted, no previous orthodontic treatment

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Thickett 2010 *(Continued)*

	Exclusion criteria: required fixed retention; had had functional appliance treatment, extra oral orthopaedic force, craniofacial anomalies or orthognathic surgery **Setting**: NHS hospital department in Bournemouth, UK; treatment was free of charge
Interventions	**Two types of removable retainers wearing protocols** • Group 1: full-time wear for 0-3 months after debond, then 10 h/day for 3-6 months, then alternate nights for 6-9 months after debond, then 1-2 times a week for 9-12 months after debond • Group 2: part-time wear (10 h/day) for 0-3 months after debond, then 10 h/day for 3-6 months, then alternate nights for 6-9 months after debond, then 1-2 times a week for 9-12 months after debond Both groups had clear plastic retainers made from Essix B material (GAC International, Bohemia, New York, USA).
Outcomes	**Outcome** relevant to review: stability Stability assessed using LII, intercanine width, intermolar width, arch length, overjet, overbite, PAR Data collected at 1 year
Notes	The study authors mistakenly interchange the terms irregularity and crowding, when they mean irregularity. Data were non-parametric so could not be entered into Review Manager 5 to be included in meta-analysis. Results presented in Table 2. PAR results were not presented in a format that allowed analysis. We requested these from the study authors, but they were not provided.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The paper stated that the participants were allocated by "random number generation" - we contacted the study authors and they clarified that this was done appropriately using random number generation (NB reference to author clarification is from the previous version of the review and we do not have any details about the emails sent or received in relation to this).
Allocation concealment (selection bias)	Low risk	Authors explained that an independent researcher randomised the participant to the intervention after consent was obtained.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Authors confirmed that the outcome assessor was blinded to the retainer type.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Number of dropouts reported but no explanation given. Study authors contacted and they replied that dropouts were not included in the analysis. Study authors confirmed that the dropouts were due to participants not attending review appointment.
Selective reporting (reporting bias)	Low risk	No evidence of selective reporting
Other bias	Low risk	"There was non-compliance with the retention regimen, but as this was a effectiveness rather than an efficacy study, this is acceptable."

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Torkan 2014

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Numbers recruited**: 30 participants (15 in each group) **Sex (M:F):** SW group: 27% male, 73% female; FRC: 40% male, 60% female) **Mean age (years)**: spiral wire group: 15.7 (SD 2.1); polyethylene ribbon-reinforced resin group: 16.2 (SD 1.9) **Inclusion criteria**: healthy with no systemic disease; age 13-19 years; willing to take part in study; treated with upper and lower fixed appliances **Exclusion criteria**: presence of any syndrome; missing or extraction of any anterior teeth **Setting**: hospital (Orthodontics Department, Shiraz University of Medical Sciences, Shiraz, Iran)
Interventions	**Two types of fixed upper and lower retainers** • Group 1: fibre-reinforced composite retainer (fibre strand: NSI Ltd., Hornsby, Australia, composite: Filtek Z250, 3M Unitek, USA) bonded on the lingual surfaces of the six upper and lower anterior teeth • Group 2: multi-stranded ⌀ 0.0175 inch stainless steel wire retainer (American Orthodontics, Sheboygan, WI, USA) bonded with a thin layer of flowable composite resin (Filtek Flow, 3M ESPE, St. Paul, MN, USA) to the six upper and lower anterior teeth
Outcomes	Oral health: • Plaque Index • Calculus Index • Gingival Index • Bleeding on probing (all on lingual surfaces of anterior teeth) • Change in periodontal ligament width measured on 2 upper anterior teeth and 1 lower tooth over 6 months (% of decreased width of periodontal ligament, showing improvement) **Time points:** at debond and after 6 months
Notes	

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was ensured by permutation of blocks with 10 blocks including 4 participants in each block.
Allocation concealment (selection bias)	Low risk	Allocation was concealed from the orthodontist until after the placement of the retainer.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding for periodontal indices The radiologists were blinded to the study and therefore not aware whether the radiographs were taken at the beginning of the study or after 6 months, but it would not be possible to blind them to the type of retainer (which would appear different on radiographs).
Incomplete outcome data (attrition bias) All outcomes	High risk	10 dropped out of the study

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Torkan 2014 *(Continued)*

Selective reporting (reporting bias)	Low risk	No evidence of selective reporting
Other bias	Low risk	No other risk identified

Wan 2017

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 20 **Sex (M:F):** 7:13 (HR group: 5 men, 5 women; VFR group: 2 men, 8 women) **Mean age (years)**: HR group: 24.6 (SD 2.6); VFR group: 24.1 (SD 3.1) **Inclusion criteria**: had completed active orthodontic treatment **Exclusion criteria**: cleft lip or cleft palate, surgical correction of the jaws, dialects, hearing and speech disorders, temporomandibular joint dysfunction syndrome, < 18 years, suffering from serious periodontitis **Setting**: West China Hospital of Stomatology, Sichuan University, Chengdu, China
Interventions	**2 types of removable retainer** • Group 1: upper and lower Hawley retainers • Group 2: upper and lower clear plasticretainers Participants were instructed to wear the maxillary and mandibular retainers full-time throughout the 3-month period of the study, except while eating or brushing their teeth
Outcomes	Patient satisfaction: assessment of speech sounds performed objectively using acoustic analysis: • Numbers of sound distortions • Changes in articulation Vowels were analysed in terms of 3 formant frequencies (F1, F2, and F3) Hz Voiceless fricatives were analysed in terms of upper boundary frequency (UBF) Hz **Time points**: before wearing, immediately after wearing, and at 24 hours, 1 week, 1 month and 3 months Changes in articulation not reported at T5, only number of sound distortions.
Notes	

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The study authors state they used “sortation randomization method”. Not sure what this means
Allocation concealment (selection bias)	Low risk	Allocation concealment was achieved with opaque sealed envelopes

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Wan 2017 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Operators for clinical procedures and valuators for data collection or analysis were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	No participants were lost to follow-up
Selective reporting (reporting bias)	High risk	Examples of some mean values for the acoustic analysis data were reported, but not all
Other bias	Low risk	No other biases identified

Węgrodzka 2021

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 133 (65 in round twisted (RT) group and 66 in rectangular braided (RB) group) **Sex (M:F):** 42:91 **Median age (years)**: 24.6 (IQR 17.2-32.4) **Inclusion criteria**: age 15-50 years at debonding; all mandibular permanent incisors and canines present; no active caries, no restorations, no fractures on the mandibular incisors and canines, no periodontal disease; and retention plan including only retainers bonded from 3-3 **Exclusion criteria**: inadequate hygiene, need for restorative or surgical treatment, active periodontal disease, or removable retainer as an adjunct to a bonded retainer **Setting**: orthodontic private practice, Krakow, Poland
Interventions	**2 types of fixed retainers** • Group 1: 3-strand round twisted (RT) • Group 2: 8-strand rectangular braided (RB) fixed retainers bonded to all 6 anterior teeth in the mandible
Outcomes	Survival time of retainers. failure rates (first-time failure of retainer) and failure risk (hazard ratio) Oral health: periodontal indices (periodontal index, bleeding on probing, plaque and gingival indices, pocket depth) **Time points**: baseline, at debonding and retainer placement, 3, 6, 12, 18 and 24 months
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation is very well reported: software-generated, block-randomised and equal-distributed
Allocation concealment (selection bias)	Low risk	Allocation concealment was achieved with sequentially numbered, opaque, sealed envelopes prepared before the study. The study co-ordinator (practice

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Węgrodzka 2021 *(Continued)*

		manager) was responsible for opening envelopes in sequence and implementing the randomisation process.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Blinding was not possible because the retainers could not be masked during clinical assessments. It is likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Few missing outcome data and balanced in numbers between intervention groups, with similar reasons for missing data between groups
Selective reporting (reporting bias)	Unclear risk	This study did not have a published protocol prior to the initiation of the study. Insufficient information to allow judgements, although the publication reports all the outcomes indicated in the material and methods section.
Other bias	Low risk	The study appears to be free of other sources of bias.

Zhu 2017

Study characteristics	
Methods	2-arm parallel RCT
Participants	**Number recruited**: 80 (42 in 0.75-mm group, 38 in 1-mm group); 72 completed study (37 in 0.75-mm group, 35 in 1-mm group) **Sex (M:F):** 20:60 **Age (years)**: mean 19.3 (range 12-33) **Inclusion criteria**: healthy without systemic or oral diseases; central incisors, canines, and first molars of both arches were present; the study and signed informed consent forms were fully understood, either by participants or their parents **Exclusion criteria**: oral habits such as bruxism and clenching; did not accept vacuum-formed retainers; unable to comply with follow-up during the 12-month period **Setting**: West China Hospital of Stomatology, Sichuan University, Chengdu, China
Interventions	**2 types of clear plastic removable retainer** • Group 1**:** 0.75 mm thickness upper and lower clear plastic retainers • Group 2: 1.00 mm thickness upper and lower clear plastic retainers Participants were instructed to wear the retainers 24 h/day during the 12-month observation period, except during meals and cleaning. They were told to avoid pigmented and fizzy drinks.
Outcomes	Survival time of retainers (time from fitting to failure) - days Patient satisfaction measured using a questionnaire, the SF-36 scale, at T1; patient comfort assessed with a VAS at T0 and T1 **Time points**: T0 when the fixed orthodontic appliances were removed and the retention appliance inserted, then at 1, 3, 6 and 12 months after wearing the retainers
Notes	
Risk of bias	

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Zhu 2017 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomisation was accomplished by tossing a coin.
Allocation concealment (selection bias)	Low risk	Allocations concealed in sequentially numbered, opaque, sealed envelopes.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"The participants and the researcher were both blinded after assignment to the interventions"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Overall rate of loss to follow-up 10%. 8 dropouts. Baseline characteristics similar between groups
Selective reporting (reporting bias)	Low risk	Tables and raw data provided in supplementary material
Other bias	Low risk	No other biases detected

Årtun 1987

Study characteristics	
Methods	4-arm parallel RCT
Participants	**Number recruited**: 49 **Sex (M:F)**: not reported **Age (years)**: mean not reported, range 12-55 **Inclusion criteria**: treated with edgewise light-wire fixed appliances by 2 orthodontists in 1 practice, variety of malocclusions **Exclusion criteria**: none reported **Setting**: specialist practice in Seattle, USA Stratified according to age (adult/adolescent), sex and gingival condition
Interventions	**3 types of fixed retainer and 1 removable retainer** • Group 1: thick, plain ⌀ 0.032 inch stainless steel bonded retainer bonded to lower canines only • Group 2: thick spiral ⌀ 0.032 inch stainless steel spiral wire bonded retainer bonded only to lower canines • Group 3: thin flexible ⌀ 0.0205 inch stainless steel spiral wire bonded to all lower anterior 6 teeth • Group 4: removable retainer. Exact design not described but appeared in a figure to have acrylic covering the lower anterior 6 teeth except incisal edges and supported by metal framework both buccally and lingually. Exact extent of retainer not clear. Amount of wear not reported At time of placement of retainer, all participants were taught correct oral hygiene procedures with toothbrush and floss. All retainers were in the lower arch only.
Outcomes	Stability was reported using LII, but there were insufficient data to analyse for this review.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Årtun 1987 *(Continued)*

	Survival of retainers was assessed by describing the number of bonded retainers that debonded and how many removable retainers were lost. Adverse effects on health were reported in terms of effects on periodontal health and caries, but full data were not available so statistical analysis was not possible on this particular outcome. **Time point**: after 3 years
Notes	Contact from study authors confirmed that no data were available in addition to those in the published paper. There were insufficient data in the paper to allow analysis of stability or adverse effects.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No description given of randomisation
Allocation concealment (selection bias)	Unclear risk	No description given of concealment method
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No description given of blinding method
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Not clear from the paper
Selective reporting (reporting bias)	High risk	Data missing from a number of outcomes
Other bias	Low risk	No other biases identified

ABO-OGS: American Board of Orthodontics objective grading system; **ARI:** adhesive remnant index: **CAD/CAM:** computer-aided design/computer-aided manufacturing; **F:** female **GMA:** glycidyl methacrylate; **h/day:** hours per day; **ITT:** intention-to-treat; **IQR:** interquartile range; **LII:** Little's Irregularity Index; **LR:** lingual wire; **M:** male; **MBT:** brand of braces; **MSLR:** multi-stranded stainless steel lingual wire; **NHS:** National Health Service (UK); **PAR:** peer assessment rating index; **RCT:** randomised controlled trial; **SD:** standard deviation; **SF-36:** short form health survey; **SSLR:** single-strand lingual wire; **TPR:** thermoplastic retainer; **VAS:** visual analogue scale

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
ACTRN12612000670875	Contacted author and insufficient participants were recruited, so study was abandoned
Ahrens 1981	Follow-up too short. Study only followed up changes for 30 days
Al-Nimri 2009	Allocation was not randomised (treatments were allocated alternately)
Alaa ElDeen 2021	Studies comparing different methods of bonded retainer placement were not included in this review.
Alkan 2020	Ineligible focus/outcome: settling of retainers (quality of occlusion and occlusal forces after 3 months, 6 months and 1 year ("...occlusal force distribution (OFD), individual tooth force (ITF), and occlusal surface area...") - this is not stability

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Arora 2014	Lead author confirmed by email that the groups were allocated alternately (not by randomisation) (NB this is from previous version of review and we do not have details about the email from the author).
Aslan 2013	Assesses settlement not stability: "The purpose of this prospective study was to determine the number of contacts in centric occlusion during retention with modified and full-coverage thermoplastic Essix retainers."
Atack 2007	Retrospective review paper
Axelsson 1993	Contact with authors confirmed study was retrospective. NB this was from previous version of the review and we do not have details about the email from the author.
Barlin 2011	Contact with authors confirmed study was retrospective. NB this was from previous version of the review and we do not have details about the email from the author.
Basciftici 2007	Allocation was not randomised
Bauer 2010	Not a true randomised controlled trial. Based on a sample from a previous study based on 50 participants. In this study, they claimed to allocate randomly, but there were 28 participants in 1 group and 22 in the other group. This was done deliberately as the authors reported they wanted to "oversample" as they expected problems in the Perfector group.
Bazargani 2012	Described different methods of placing fixed retainers, not different types of retainers
Bock 2008	Retrospective study
Bovali 2014	Studies comparing different methods of bonded retainer placement were not included in this review.
Brignardello-Petersen 2018	This is a brief comment on the published article by Forde 2018. This type of article does not provide primary data and is outside the scope of this review.
Carvalho 2006	Participant had individual teeth treated for 8 weeks and then the retention was only for 8 weeks.
ChiCTR-TRC-07000055	The authors were contacted several times to inquire about the status of the study, but we did not receive a response.
ChiCTR-TRC-07000454	This protocol was registered in 2007. The authors were contacted several times to inquire about the status of the study, but we did not receive a response.
Choi 2010	1 month follow-up
Conway 2011	Study compared different methods of attaching the same retainer - so did not fit the review criteria
Cunning 2022	Ineligible study design: controlled clinical trial. The authors were contacted. The comparison of retainer type (aligners: Vivera vs Duran) did not follow a random sequence generation. Quote: "The patient is given the choice of which retainer they would prefer"
Edwards 1988	Not a randomised controlled trial. Interventions were allocated alternately, so the sequence generation was not random, and there was no allocation concealment. Although the study had been an included study in a previous version of this review, we decided to remove quasi-randomised trials from the inclusion criteria, to reduce the risk of bias. In addition, the study had substantial dropout, and the effects of the dropouts were not discussed or analysed. We decided that the lack

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
	of adequate sequence generation and allocation concealment produced a large risk of bias, so we excluded this study from the review.
Haydar 1992	Insufficient data available in paper to allow statistical analysis
Horton 2009	Retention phase only followed up for 2 months (less than the required 3 months for this review)
ISRCTN22535947	Correspondence with study author confirmed insufficient participants attended for recall to allow analysis of data. NB this was from previous version of the review and we do not have details about the email from the author.
ISRCTN26364810	Personal correspondence with lead author confirmed the retention part of the study was never completed. NB this was from previous version of the review and we do not have details about the email from the author.
ISRCTN56295329	Study was a method of supervision of retainers, not comparison of different retention regimens.
Jahanbin 2014	1 month follow-up
Jain 2021	Plagiarised study, confirmed and reported by original review authors
Jäderberg 2012	We contacted the authors who explained that treatment groups were allocated "alternately" and not by randomisation. NB this was from previous version of the review and we do not have details about the email from the author.
Kara 2020	Study design unclear - study author did not mention randomisation. We have tried to contact the authors. This study also included outcome domains (e.g. settling) and outcome measures (e.g. mean occlusal contact area in mm^2) outside the scope of this review.
Knaup 2019	Study design: retrospective and in vitro study. Quote: "The first part was a retrospective controlled clinical study which was intended to function as a pilot study; the second part an in vitro investigation of the possible retainer material-dependent biofilm formation"
Larsson 1983	Insufficient data to allow analysis
Lindauer 1998	Prospective, but non-randomised study
Liu 2010	Correspondence not available to request the full text published in Chinese
NCT03765151	Study was performed on periodontal-compromised patients
Pandian 2019	Study follow-up 1 month
Pandis 2007	Retrospective study
Sangalli 2022	Active treatment with aligners is outside the scope of this review.
Sauget 1997	Not a randomised controlled trial. Interventions were allocated alternately, so the sequence generation was not random, and there was no allocation concealment. Although the study was included study in a previous version of this review, we decided to remove quasi-randomised trials from the inclusion criteria, to reduce the risk of bias. We decided that the lack of adequate sequence generation and allocation concealment produced a large risk of bias, so we excluded this study from the review
Sohrabi 2014	This study included outcome measures (e.g. re-opening of interdental spaces) outside the scope of this review.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Study	Reason for exclusion
Stormann 2002	Insufficient data in the original publication. Contacted study authors who were unwilling to release any further data. NB this was from previous version of the review and we do not have details about the email from the author
Tacken 2010	Allocation not randomised
Tahir 2019	In this crossover RCT, the removable retainers were not placed at the same time. Test retainers were placed one week after debonding while control retainers were placed on the same day. Quote: "This first non-interventional VFR (VFR-CV) was issued on the same day of debond. Because of the limitation of our facility to construct the VFR-3D within a day, the interventional VFRs were standardized to be issued a week after debond."
Taner 2000	Not a randomised controlled trial. After personal communication with the authors, they confirmed that the participants were allocated alternately, not randomly. NB this was from previous version of the review and we do not have details about the email from the author. As the interventions were allocated alternately, the sequence generation was not random, and there was no allocation concealment. Although the study had been an included study in a previous version of this review, we decided to remove quasi-randomised trials from the inclusion criteria, to reduce the risk of bias. We decided that the lack of adequate sequence generation and allocation concealment produced a large risk of bias, so we excluded this study from the review.
Vagdouti 2019	Ineligible focus/outcome: patient compliance through the use of a thermosensitive microsensor.
Vecere 1983	Reviewed study author's thesis, which confirmed that the study was retrospective
Xu 2011	Full text requested from the study authors for the first time on 27 January 2016 and for the second time on 24 November 2021. No response has been received to date.

VFR: vacuum-formed retainer

Characteristics of ongoing studies *[ordered by study ID]*

CTRI/2018/10/016151

Study name	Evaluation of success of two orthodontic retainer wires on maxillary and mandibular arch retention over a 12-month period: a randomized controlled trial
Methods	Study design: single-blinded (participants), 2-armed, parallel, randomised, active control trial Country: India Duration of follow-up: time frame: debond and start of the retention phase, 3, 6 and 12 months Inclusion criteria • Male or female from 13-40 years • Presence of the 6 anterior teeth in non-extraction case and 6 anterior teeth and 2 premolars in premolar extraction case • No active caries, restorations, fractures, or periodontal disease of these teeth • Optimum dental health with no need of restorations on the teeth that might affect retainer bonding on lingual enamel surfaces • No morphologic crown anomalies • Ability to maintain oral hygiene • Informed consent signed by the participant before bonding the fixed retainers Exclusion criteria

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2018/10/016151 *(Continued)*

	• Compromised periodontal health • Missing incisor or canine • Any systemic disease • Case with extraction of incisor or canine
Participants	The study aims to enrol and randomise 60 participants.
Interventions	• Group A (active comparator): 0.0195 inch coaxial wire (Ortho Organizers, Mount Holly, NJ) • Group B (experimental intervention): 8 Braid Flat- Dead Soft Stainless Steel 6 inch retainer wire (Ortho-Direct, MO, USA)

Outcomes	Primary outcome
	• Failure rate (wire breakage/wire debond).
	Secondary outcomes
	• Stability - LII
	• Patient satisfaction - set of questions:
	  ◦ any food lodgement?
	  ◦ any difficulty in brushing?
	  ◦ any irritation to tongue?
	• Periodontal health status - Plaque Index, Gingival Index, Calculus Index and Papillary Bleeding Index (PBI)

Starting date	25 October 2018
Contact information	Principal Investigator: Dr Vinay Kumar Chugh Institution: Department of Dentistry, Basni-2 AIIMS Jodhpur Rajasthan, India Email: drvinaychd@yahoo.com
Notes	Contacted author - study ongoing

CTRI/2019/05/019136

Study name	To evaluate the clinical efficacy of four different lingually bonded retainers
Methods	Study design: single-centre, double-blinded, double-dummy, 4-armed, parallel RCT Setting: conducted in a university settling Country: India Duration of follow-up: debond (baseline) and at 3, 6 and 9 months Inclusion criteria • Male or female from 14-25 years of age • No active periodontal and gingival conditions • Consented to be a part of the study • Had completed fixed orthodontic treatment and were debonded • Compliant enough for regular follow-up visits Exclusion criteria

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2019/05/019136 *(Continued)*

	• Active periodontal and gingival condition
Participants	The study aims to enrol and randomise 60 participants.
Interventions	• Group A: coaxial wire (0.0155 inch) • Group B: flat titanium wire • Group C: stainless steel ligature wire (0.010 inch) • Group D: titanium molybdenum alloy wire (0.018 inch)
Outcomes	Primary outcomes (periodontal parameters) • Plaque Index • Gingival Index • Probing depth Secondary outcome • Efficacy of wires
Starting date	3 June 2019
Contact information	Principal Investigator: Dr Maharabam Gyanada Institution: Orthodontics and Dentofacial Orthopaedics department, Subharti Dental College, Meerut, India Email: genumaharabam@gmail.com
Notes	Contacted study author - awaiting a response

CTRI/2020/04/024461

Study name	Comparison of occlusal settling, patient compliance and retention ability between Hawley's and Van Der Linden retainers: a randomized controlled trial
Methods	Study design: single-centre, double-blinded (participant and outcomes assessor), 2-armed, parallel, randomised, active control trial Setting: conducted in a university settling Country: India Duration of the follow-up: at time of retainer delivery (T0), 3 months of retention (T1) and 6 months of retention (T2) Inclusion criteria • Male or female from 18-40 years • Gave consent for the proposed study • Had undergone fixed orthodontic treatment • Had completed orthodontic treatment and were in retention phase of treatment • Systemically healthy Exclusion criteria • Patients with craniofacial or dentofacial syndromes • Patients with diabetes • Patients having habit of smoking

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2020/04/024461 *(Continued)*

	• Patients who are suffering from periodontal disorder • Non-compliant patients
Participants	The study aims to enrol and randomise 80 participants.
Interventions	• Group A (experimental): Van der Linden retainer • Group B (active control): Hawley retainer
Outcomes	Primary outcome • Occlusal settling Secondary outcomes • Stability - LII (LII), intercanine width (ICW) and intermolar width (IMW) • Patient compliance
Starting date	20 April 2020
Contact information	Principal Investigator: Dr Nikunj Patel Institution: Department Of Orthodontics & Dentofacial Orthopaedics, Manubhai Patel Dental College, India Email: nikunj_84@yahoo.co.in
Notes	Contacted study author, awaiting a response

CTRI/2020/04/024486

Study name	Evaluation of speech adaptation to Hawleys retainer and Essix retainer & comparison between two by analyzing the frequency of vowels and consonants: an in vivo study
Methods	Study design: single-blinded (participant), 2-armed, parallel RCT Setting: conducted in a university and hospital setting Country: India Duration of follow-up: before wearing retainers, immediately after delivery of retainer, 24 hours, 1 week, 3 weeks, 3 months and 6 months Inclusion criteria • Patients who have completed their orthodontic treatment • Patients aged between 18-35 years • Patients who give consent for the proposed study • Patients who are systemically healthy Exclusion criteria • Patients with cleft lip or cleft palate • Patients who have undergone surgical correction of the jaws • Patients who have hearing and speech disorders • Patients who are suffering from serious periodontitis
Participants	The study aims to enrol and randomise 66 participants.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2020/04/024486 *(Continued)*

Interventions	• Group A (experimental): clear plastic retainers of 2-3 mm thickness, constructed on vacuum former • Group B (active control): Hawley retainers constructed with a cold cure acrylic base plate, Adams clasps, and a labial bow
Outcomes	Primary outcome • Alteration in frequency of vowels and consonants during wear of Hawley and vacuum-formed retainer. Participants instructed to pronounce vowels such as a, e, i, o & u and consonants such as t, d, n, s, z, m, n, c, h & g. These were recorded through microphone. The recorded data were analysed later by using frequency analysing software.
Starting date	15 April 2020
Contact information	Principal Investigator: Dr Kalpesh Patel Institution: Department of orthodontics Manubhai Patel Dental College, India Email: dr_kalpeshp@yahoo.co.in
Notes	Contacted study author, awaiting a response

CTRI/2021/02/031387

Study name	Comparison of two different designs of fixed retainers for periodontal health of anterior teeth and their clinical efficacy
Methods	Study design: double-blinded (participant and outcome assessor), 2-armed, parallel, randomised, active-control trial Setting: conducted in a university and hospital settling Country: India Duration of follow-up: 1, 2, 3 and 4 months Inclusion criteria • Male or female from 10-50 years • No missing mandibular incisors • No restorations on the mandibular incisors that might affect retainer bonding on lingual enamel surfaces • No morphologic crown anomalies • Well aligned incisors after orthodontic treatment completion will be included Exclusion criteria • A professional dental cleaning within the last 4 months • A history of diabetes • A habit of smoking • Pre-existing periodontal disease • Post-orthodontic periodontal disease • Antibiotic prophylaxis prior to periodontal data collection, current use of antibiotics • Pregnancy
Participants	The study aims to enrol and randomise 30 participants.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2021/02/031387 *(Continued)*

Interventions	• Group A: wave-type retainer. Wave retainer from canine of left side in mandible to canine on right side of arch • Group B: straight twisted wire bonded lingual retainer. Straight retainer from canine of left side in mandible to canine on right side of arch
Outcomes	Primary outcomes • Stability - LII Secondary outcomes (periodontal parameters) • Plaque Index • Gingival pocket probing depths • Gingival recession • Bleeding on probing • Oral hygiene - questionnaire for frequency of brushing, flossing, ease of flossing and comfort of retainers
Starting date	18 February 2021
Contact information	Principal Investigator: Dr Udita Thakkar Institution: Department of Orthodontics, Manubhai Patel Dental College, India Email: druditamehta@gmail.com
Notes	Contacted study author, awaiting a response

CTRI/2021/02/031498

Study name	Evaluation of periodontal health and stability following fixed orthodontic treatment by the use of two types of retainer
Methods	Study design: single-centre, 2-armed, parallel, RCT Setting: conducted in a university setting. Country: India Duration of follow-up: debonding, after 1 month, 8 months and 12 months Inclusion criteria • Male or female from 16- 26 years of age • Only premolar extraction • No missing maxillary incisors or molars • Fixed appliance treatment involving both arches; preadjusted edgewise appliances • Pretreatment records, study models available Exclusion criteria • Non extraction cases • Patients having craniofacial or dentofacial syndromes • Single-arch or sectional fixed appliance treatment • Cases requiring bonded retainers • Patients having poor periodontal status

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2021/02/031498 *(Continued)*

	• Cases that have to be debonded prior to finishing and detailing stage of treatment
Participants	The study aims to enrol and randomise 30 participants.
Interventions	• Group A: clear plastic vacuum-formed retainer (Essix) • Group B: Hawley retainer
Outcomes	Primary outcomes • Stability • Periodontal health
Starting date	22 February 2021
Contact information	Principal Investigator: Dr Mona Prabhakar Institution: Faculty of dental sciences, SGT University Gurgaon, India Email: prabhakarmona@gmail.com
Notes	Contacted study author, awaiting a response

CTRI/2021/03/032110

Study name	Comparison of effects of Essix and modified Essix retainer on final teeth positioning
Methods	Study design: double-blinded (participant, outcome assessor), 2-armed, parallel, randomised, active-control trial Setting: conducted in a university and hospital setting Country: India Duration of follow-up: debond, 3 months and 6 months Inclusion criteria • Male or female from 18-40 years of age • Patients who give consent for the proposed study • Patients who have undergone fixed orthodontic treatment • Patients who have completed their orthodontic treatment and are in retention phase of treatment • Patients who are systemically healthy Exclusion criteria • Patients having craniofacial or dentofacial syndromes • Patients treated for anterior open bite • Patients suffering from diabetes • Patients having habit of smoking • Patients who are suffering from periodontal disorder • Non-compliant patients
Participants	The study aims to enrol and randomise 80 participants.
Interventions	• Group 1: clear plastic (Essix) retainer. After completion of fixed orthodontic treatment, patients will be given Essix retainer to were in maxillary arch. Participant impression and bite registration

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2021/03/032110 *(Continued)*

	will be taken at time of retainer delivery, 3 months and 6 months of retention; at 3 months and 6 months a questionnaire will also be given to participants to fill in. • Group 2: modified clear plastic (Essix) retainer. After completion of fixed orthodontic treatment, participants will be given modified Essix retainer to wear in maxillary arch. Participant impression and bite registration will be taken at time of retainer delivery, 3 months and 6 months of retention; at 3 months and 6 months a questionnaire will also be given to patients to fill in. Participants will be instructed to wear the retainers 24 h/day for 6 months, excluding while eating and brushing their teeth.
Outcomes	Primary outcomes • Change in occlusal contacts • Change in patient compliance, assessed by questionnaires
Starting date	23 March 2021
Contact information	Principal Investigator: Dr Kalpesh Patel Institution: Department of orthodontics Manubhai Patel Dental College, India Email: dr_kalpeshp@yahoo.co.in
Notes	Contacted author, awaiting a response

CTRI/2021/11/037851

Study name	Comparative evaluation of short term relapse following orthodontic space closure with & without fibreotomy intervention: a randomized controlled trial
Methods	Study design: single-blinded (outcome assessor), 2-armed, parallel, RCT Country: India Duration of follow-up: debond and start of the retention phase (T0), 3-month follow-up (T1), 6-month follow-up (T2) Inclusion criteria • The study involves completely corrected malocclusion cases that underwent extraction. • Cases with class I molar and canine relationship, correct crown angulation, crown inclination and tight contacts Exclusion criteria • Patients with spacing, rotation, teeth out of arch, increased overjet and overbite • Patients with detrimental oral habits • Patients with poor periodontal health and gingival recession
Participants	The study aims to enrol and randomise 24 participants.
Interventions	• Intervention: circumferential supracrestal fibreotomy (to remove the supracrestal fibres from 35-45 with periotome) • Comparator: no adjunctive procedure
Outcomes	Primary outcome • Relapse potential

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CTRI/2021/11/037851 *(Continued)*

Starting date	20 November 2021
Contact information	Principal investigator: Dr Pavithranand Institution: Department of Orthodontics, Indira Gandhi Institute of Dental Sciences, Pondicherry Email: pavithranandammayappan@igids.ac.in
Notes	

CTRI/2021/11/038102

Study name	A comparison of clinical efficiency of nanocomposite and conventional composite used for bonding lower fixed retainers
Methods	Study design: double-blinded (participants and outcome assessor), 2-armed, parallel RCT Country: India Duration of follow-up: debond and start of the retention phase, 3 months, 6 months, 9 months and 12 months Inclusion criteria • Adult patients visiting the department for orthodontic treatment, and have undergone debonding and indicated for fixed retention in the mandibular arch Exclusion criteria • Patients with chronic periodontal diseases • Missing mandibular anterior teeth
Participants	The study aims to enrol and randomise 40 participants.
Interventions	Intervention: nanoparticle containing composite: 1% TiO2 Comparator: Transbond XT conventional composite
Outcomes	Primary outcome • Survival rate Secondary outcome • Periodontal health
Starting date	26 November 2021
Contact information	Principal investigator: Dr Shravanthi Institution: Department of Orthodontics and Dentofacial Orthopedics, Dayananda Sagar College of Dental Sciences, India Email: gauthamkalladka@gmail.com
Notes	

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

DRKS00016710

Study name	Influence of fixed and removable orthodontic retainers on the gingival health of mandibular anterior teeth - a randomized, clinical controlled trial
Methods	Study design: single-centre, 2-armed, parallel, randomised, active-control trial Setting: conducted at a university Country: Germany Duration of follow-up: debonding, after 6 months (± 1) and 12 months (± 1) Inclusion criteria • Male or female aged ≥ 12 years, with no maximum age limit • Presence of all lower anterior teeth (33-43) • Patients with orthodontic brace who are in need of a retention after removing all brackets • Healthy patients with no general diseases • No Periodontal Screening Index Code • Written informed consent (of the parents, legal guardian) Exclusion criteria • Pregnancy
Participants	The study aims to enrol and randomise 36 participants.
Interventions	• Group A (experimental): fixed steel retainer on the lingual sides of teeth 33-43 • Group B (active control): modified retention splint (removable) incorporating only the coronal part of teeth 33-43, used for the first 3 months all day and for the following 9 months only at night
Outcomes	Primary outcome • Change of mean recession depth Secondary outcomes • Pocket probing depth • Clinical attachment level • Gingival Index • Plaque Index
Starting date	17 July 2019
Contact information	Principal Investigator: Dr Hari Petsos Institution: Johann Wolfgang Goethe-University Frankfurt/Main, Germany Email: petsos@med.uni-frankfurt.de
Notes	Contacted study author, study ongoing

IRCT2016022926836N1

Study name	Comparison of survival time and satisfaction of Hawley and vacuum-formed retainers in orthodontic patients – a randomized clinical trial

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

IRCT2016022926836N1 *(Continued)*

Methods	Study design: multi-centre, single-blinded, 3-armed, parallel, randomised, active-control trial Setting: conducted in 3 private clinics Country: Iran Duration of follow-up: 1, 3, 6 months after debonding Inclusion criteria • Male or female with no age limit • Orthodontic treatment in an optimal occlusion • Indication of use removable retainer • Patients willing to participate in the study Exclusion criteria • Prosthetic treatment of missing teeth • Bruxism • Syndromic disease (cleft lip & palate, hemifacial microsomia) • Alergy to acrylic resins • Temporomandibular disorder • Detrimental oral habits • Maxillary expansion treatment • Early debonding • Generalised spacing
Participants	The study aims to enrol and randomise 105 participants.
Interventions	• Group 1 (experimental): clear plastic retainer with 1 mm thickness in upper arch • Group 2 (experimental): clear plastic retainer with 1.5 mm thickness in upper arch • Group 3 (active control): Hawley retainer in upper arch
Outcomes	Primary outcome • Patient satisfaction Secondary outcomes • Survival time - clinical evaluation of retainers for fracture • Clinical evaluation for fracture, crack, colour change, retainer loss and fitness
Starting date	24 April 2016
Contact information	Principal Investigator: Dr Seyed Hossein Moslemzadeh Institution: Tabriz University of Medical Sciences, Iran Email: hmoslemzade@yahoo.com
Notes	Contacted study author, awaiting a response

IRCT20170217229951N3

Study name	Comparison of orthodontic treatment results stability and gingival health between Hawley and vacum-formed retainers: randomized clinical trial

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

IRCT20170217722951N3 *(Continued)*

Methods	Study design: single-centre, single-blinded, 2-armed, parallel, randomised, active-control trial Country: Iran Duration of the follow-up: time frame: before retainer delivery and 6 months after delivery Inclusion criteria • Completion of orthodontic treatment in a proper occlusion • Indication of removable retainer Exclusion criteria • The absence of several teeth and the need to replace them with prosthesis • Patients with bruxism • Syndromic disease (lip and palate clefts, hemi-facial microsomia) • Patients with allergy to acrylic resin • Patients with temporomandibular joint problems
Participants	The study aims to enrol and randomise 66 participants.
Interventions	• Group 1 (experimental): clear plastic retainer with a thickness of 1 mm retainer • Group 2 (experimental): clear plastic retainer with a thickness of 1.5 mm retainer • Group 3 (active control): Hawley retainer
Outcomes	Primary outcomes • Stability - objective grading system method • Gingival Index
Starting date	28 April 2017
Contact information	Principal Investigator: Dr Seyed Hossein Moslemzadeh Institution: Tabriz University of Medical Sciences, Iran Email: hmoslemzade@yahoo.com
Notes	Contacted study author, awaiting a response

IRCT20210221050435N2

Study name	Level of metalloproteinase matrix 8 in gingival crevicular fluid of mandibular anterior teeth with two types of fixed and removable orthodontic retainers
Methods	Study design: double-blinded (participants and research recruiters), 2-armed, parallel RCT Country: Iran Duration of follow-up: debond and start of the retention phase, 30 days, 90 days Inclusion criteria • Patients who have successfully corrected malocclusion and proper dental alignment • Patient's condition is favourable in terms of general health • Absence of craniofacial syndromes and facial clefts • Absence of severe and uncontrolled caries

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

IRCT20210221050435N2 *(Continued)*

	• Patient should have good oral hygiene and be periodontally normal • Patient should have all the anterior mandibular teeth and the teeth should be of normal shape and size Exclusion criteria • Lack of patient co-operation to refer • Observe poor periodontal status during the study • Break and not using the retainer • Specific conditions of COVID-19 infection that prevent the patient from visiting
Participants	The study aims to enrol and randomise 30 participants.
Interventions	Clear plastic retainer will be placed in the upper jaw for both groups. • Intervention A: lower stainless steel, 8-strand, dead soft fixed retainer • Intervention B: clear plastic retainer, removable (1 mm thick)
Outcomes	Primary outcome • Level of matrix metalloproteinase 8 after orthodontic debonding Secondary outcomes • Bleeding on probing • LII • Overbite • Overjet
Starting date	1 July 2021
Contact information	Principal investigator: Dr Hossein Institution: Dental school, Yazd University of Medical Sciences, Iran Email: hosseinaghily@gmail.com
Notes	

NCT02742493

Study name	Effectiveness of 3 orthodontic retention schemes on post-treatment stability and gingival recession
Methods	Study design: single-centre, open-label, 3-armed, parallel, randomised, active-control trial Country: Greece Duration of follow-up: 1 year in retention, 2 years and 5 years Inclusion criteria • Both sexes • Age between 10 and 16 years • At the finishing stage of orthodontic treatment, programmed for debonding • Class I to Class II molar relationship before the beginning of orthodontic treatment • Non-extraction treatment plan with fixed straight-wire appliances

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT02742493 *(Continued)*

	• Space deficiencies and irregularity of no more than 6 mm in the maxillary/mandibular arch (Irregularity Index) Exclusion criteria • No missing teeth, impacted teeth or other dental anomalies • No congenital anomalies/syndromes
Participants	The study aims to enrol and randomise 60 participants.
Interventions	• Group 1 (experimental intervention): lower fixed canine and canine retainer (0.028 inch stainless steel bonded to canines) and upper removable Hawley • Group 2 (experimental intervention): lower fixed canine to canine retainer (7-strand Twistflex 0.027 inch bonded to all 6 anterior teeth) and upper removable Hawley • Group 3 (active comparator): lower Hawley-type and upper Hawley removable retainer
Outcomes	Primary outcomes • Stability - LII, intercanine width, interpremolar width, intermolar width, arch length, overbite and overjet Secondary outcomes • Survival rate - bond failures of fixed retainers or breakages of removable ones • Adverse effects - gingival recessions
Starting date	19 April 2016
Contact information	Principal Investigator: Dr Despina Koletsi Institution: Department of Orthodontics, School of Dentistry, University of Athens, Greece Email: d.koletsi@gmail.com
Notes	Contacted study author, study ongoing

NCT03844425

Study name	Microbial colonisation on vacuum-formed retainers constructed on conventional models and three dimensional (3D) reconstructed models
Methods	Study design: single-centre, single-blinded (outcomes assessor), 3-armed, parallel RCT Setting: postgraduate Orthodontic Clinic, Faculty Of Dentistry, University Of Malaya Country: Malaysia Duration of follow-up: debond and 6 months Inclusion criteria • Fixed appliance on upper and lower arches • About to debond • Planned for clear plastic retainers for retention • Non-smoking patients • No systemic disease Exclusion criteria

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT03844425 *(Continued)*

	• Undergoing sectional fixed appliance/single arch treatment • Indicated for double retention regime with bonded retainers • Patient with smoking habit • Any systemic disease that may affect the salivary flow • History of use of antimicrobial mouthwash < 1 month prior to debond
Participants	The study aims to enrol and randomise 36 participants.
Interventions	• Group 1 - experimental: clear plastic retainer from stereolithography (SLA) constructed on 3D re-constructed models using SLA technique • Group 2 - experimental: clear plastic retainer from fused deposition modelling (FDM) constructed on 3D reconstructed models using FDM technique • Group 3 - active comparator: conventional clear plastic retainer constructed on conventional stone models
Outcomes	Primary outcome • Microbial colonisation on retainers Secondary outcomes • Surface roughness of retainers • Patient compliance in wearing retainers
Starting date	2 December 2019
Contact information	Principal Investigator: Dr Wan Nurazreena Wan Hassan Institution: Department Of Paediatric Dentistry & Orthodontics, Faculty Of Dentistry, University of Malaya Email: wannurazreena@um.edu.my
Notes	Contacted study author, study ongoing

NCT03954145

Study name	Comparison of bonding failure rates of fixed mandibular lingual retainers in a French sample aged 11 or older comparing prior enamel sandblasting with conventional pumice polishing
Methods	Study design: single-centre, double-blinded (investigator, outcomes assessor), 2-armed, parallel, randomised, active-control trial Setting: conducted in a hospital Country: France Duration of follow-up: debond, 3, 6, 9, 12, 15 and 18 months Inclusion criteria • Minimum of 11 years of age • At completion of fixed orthodontic therapy • Presence of all 6 mandibular canines and incisors Exclusion criteria: • Active decay

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT03954145 *(Continued)*

	• Presence of dental restorations • Tooth fracture • Periodontal disease • Structural anomalies affecting the lingual side of the mandibular incisors and canines • Cleft lip and palate • Craniofacial syndrome • Inability to obtain the patient informed consent
Participants	The study aims to enrol and randomise 88 participants.
Interventions	• Group 1 (active comparator) or pumicing group: lingual surfaces of the lower incisors and canines are pumiced before bonding the lingual retainer. Teeth are being cleaned using a brush loaded with pumice and mounted on a low speed contra-angle. Teeth are then etched and the multi-braided retainer wire is being bonded with composite onto the lingual surfaces of the lower canines and incisors. • Group 2 (experimental intervention) or sandblasting group: enamel is being initially prepared through sandblasting the lingual surfaces of the mandibular canines and incisors. Sandblasting is performed with the help of a MicroEtcher IIATM (Danville), which is projecting 50 µm Al2O3 particles onto the enamel surfaces. Teeth are subsequently etched and the retainer is then being bonded with composite onto the lingual surfaces of the mandibular incisors and canines.
Outcomes	Primary outcome: debonding failure rate
Starting date	8 October 2019
Contact information	Principal Investigator: not found Institution: University Hospital, Strasbourg, France Email: not found
Notes	Study reported in ClinicalTrials.gov as completed; however, study authors could not be located

NCT04578704

Study name	Stability of bimaxillary proclination cases following fixed appliances treatment: a randomized control trial
Methods	Study design: single-centre, double-blinded (participant, outcomes assessor), 3-armed, parallel RCT Setting: conducted in a university settling Country: Malaysia Duration of follow-up: debond, and 1 year post-debond of orthodontic fixed appliances Inclusion criteria • Pretreatment bimaxillary proclination patients with interincisal angle < 125° • Patients with fixed orthodontic appliances treatment on both arches consisting of the extraction of 4 premolars • Both arches are indicated for retainers. • Pretreatment lateral cephalometric radiograph of adequate diagnostic quality Exclusion criteria

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT04578704 *(Continued)*

	• Pretreatment spacing, e.g. median diastema • No change or proclined lower labial segment post-treatment • Early debond patient • Patient intended to relocate within study period • Hypodontia that require prosthesis, cleft lip and/or cleft palate, or orthognathic cases • Patients with learning disabilities and unable to listen, speak and read written instructions in Bahasa Melayu or English
Participants	The study aims to enrol and randomise 27 participants, 9 per group.
Interventions	• Group 1 (experimental intervention): clear plastic (CP) retainer. Participants will be instructed to wear the removable CP vacuum form retainer for 8-10 h/day and only remove the retainers during tooth brushing. The CP retainer will be constructed using thermoplastic material that will be purchased from manufacturer at 1.0 mm thickness. • Group 2 (experimental intervention): fixed bonded retainer. Bonded retainer will be done using Ormco response deadsoft retainer wire size 0.190 inch round, made in USA by ORMCO Corporation, the fixed bonded wire will be extended distal to the extraction spaces that were previously closed by fixed appliances on both the upper and lower arches. The bonded retainer will be worn by patients for 24 hours. • Group 3 (experimental intervention): double regime (CP retainer and fixed bonded retainer). Constructed with same properties as CP retainer and bonded retainer. Standard oral and written instruction will be given to participants regarding retainer's use.
Outcomes	Primary outcome • Stability - change in the incisal inclinations. Measure in degrees any changes in the upper incisor to maxillary plane angle, lower incisor to mandibular plane angle and changes in interincisal angle. Secondary outcomes: • Stability - soft tissue changes. Measure in degrees any changes in the nasolabial angle, labiomental angle. The upper and lower lips distances in relation to Rickett's Esthetic line will be measured in mm. • Stability of dentition. Any irregularity at lower labial segment will be measured using LII measured in mm. • Changes in arches parameters. Intra-arch changes will be measured in mm based on intermolar width, intercanine width, arch length, overbite and overjet. • Impact of retainers on oral health-related quality of life. Patient perception of retainer will be measured using an Oral Health Related Quality of Life (OHIP) questionnaire. Patients will be assessed on how often they experienced the problems mentioned in the questionnaire throughout the time they use the retainers - very often, quite often, sometimes, never, don't know.
Starting date	10 September 2019
Contact information	Principal Investigator: Dr Norhidayah Mohd Tahir Institution: Faculty of Dentistry University of Malaya, Malaysia Email: norhidayahmt@um.edu.my
Notes	Contacted author, study ongoing

NCT04828096

Study name	Retention with three different bonded retainers a multicenter, randomized controlled trial with 5-year follow-up

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT04828096 *(Continued)*

Methods	Study design: multi-centre (Örebro, Malmö, Barcelona), single-blinded (outcomes assessor), 3-armed, parallel, randomised, active-control trial Country: Spain and Sweden Setting: conducted in a university Duration of follow-up: 2 and 5 years Inclusion criteria • Patients requiring mandibular lingual fixed canine-to-canine retainer after undergoing orthodontic treatment with fixed appliances • Sound enamel on the lingual surface • Calculus-free lingual surface in the lower anterior teeth Exclusion criteria • Not mentioned
Participants	The study aims to enrol and randomise 324 participants.
Interventions	• Group 1: Penta-one retainer. Penta-one wire retainer is to be bonded on the lower anterior 6 teeth (canine-to-canine). Half of the retainers will be bonded with sandblasted enamel and half without. • Group 2: Ortho-FlexTech retainer. Ortho-FlexTech chain wire retainer is to be bonded on the lower anterior 6 teeth (canine-to-canine). Half of the retainers are going to be bonded with sandblasted enamel and half without. • Group 3: Memotain retainer. Memotain retainer is to be bonded on the lower anterior 6 teeth (canine-to-canine). Half of the retainers are going to be bonded with sandblasted enamel and half without.
Outcomes	Primary outcome • Failure rates Secondary outcomes • Complications. Incidence of complications (if any) such as retraction of gingival tissue • Cost-effectiveness • Effect of micro-etching. Effect of sandblasting on bonding strength and failure rate is to be evaluated
Starting date	25 March 2021
Contact information	Principal Investigator: Dr Farhan Bazargani Institution: Postgraduate Dental Education Center, Örebro, Sweden Email: farhan.bazargani@regionorebrolan.se
Notes	Contacted study author, study ongoing. Preliminary results due in 2024

NCT04847323

Study name	Clinical effectiveness of bonded versus vacuum-formed retainers after 12 months: a randomized controlled trial

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT04847323 *(Continued)*

Methods	Study design: single-centre, single-blinded (outcomes assessor), 2-armed, parallel, randomised, active-control trial Setting: conducted at a university setting Country: Brazil Duration of follow-up: debonding, after 3, 6 and 12 months Inclusion criteria • Patients with the proper completion of orthodontic treatment and correction of the initial malocclusion • Acceptable oral hygiene (assessed through clinical examination) • Presence of all teeth up to the second molars • Clinically acceptable teeth alignment Exclusion criteria • Patients with any systemic condition that may have an influence on periodontal health • Patients with facial deformities • Initial malocclusion that required extreme corrections, large transverse expansions or orthognathic surgery • History of periodontal diseases
Participants	The study aims to enrol and randomise 50 participants.
Interventions	• Group 1: bonded V-bend retainers. The V-bend retainer is bonded in the lingual surface of the anterior teeth. The retainer is constructed using 0.024 inch stainless steel wires. Differently from the conventional bonded retainers, this retainer presents V-bends in the sagittal direction, parallel to the occlusal plane in each interproximal contact point of the incisors and canines. The retainers will be bonded after adequate etching with phosphoric acid and application of adhesive with a low viscosity resin. • Group 2: clear plastic retainers. The removable clear plastic retainers will be made of acetate 1 mm thickness. They will be made at the same appointment of debonding using plaster models. The patients will be instructed to use the retainers only during nights.
Outcomes	Primary outcome • Changes in LII. Stability will be evaluated comparing the changes in the LII between time points. The average of the linear distances between the anatomical contact points of the anterior teeth will be considered. Secondary outcomes • Intercanine distance. Distance between the cusp tips of the right and left permanent canines. In the maxilla and mandible • Intermolar distance. Distance between the tips of the mesiobuccal cusps of the first right and left permanent molars. In the maxilla and mandible • Overjet. Distance between the incisal edge of the maxillary central incisor and the incisal edge of the mandibular incisor, parallel to the occlusal plane • Overbite. Distance between the incisal edge of the upper central incisor and the incisal edge of the lower incisor, perpendicular to the occlusal plane • Calculus accumulation. Periodontal status will be evaluated regarding the amount of calculus accumulation using the Simplified Oral Hygiene Index from Greene and Vermillion, 1978. The extension of calculus will be estimated visually in all erupted teeth. Each tooth will receive a score ranging from 0-3. Minor scores will indicate absence of calculus while higher scores will indicate greater calculus accumulation on the tooth surfaces. • Gingival health. Periodontal status will be evaluated regarding the patients' gingival health using the Gingival Index described by Loe, 1967. Gingival health will be estimated clinically with a peri-

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT04847323 *(Continued)*

	odontal probe in all erupted teeth. Each tooth will receive a score ranging from 0-3. Minor scores will indicate acceptable gingival health while higher scores will indicate greater inflammation and unacceptable gingival health. • Plaque accumulation. Periodontal status will be evaluated regarding the amount of plaque accumulation using the Plaque Index from Loe, 1967. The extension of plaque will be estimated clinically using a periodontal probe in all erupted teeth. Each tooth will receive a score ranging from 0-3. Minor scores will indicate absence of plaque while higher scores will indicate greater plaque accumulation on the tooth surfaces. • Retainer survival rates. The number and cause of failures for each patient will be recorded during the 12-month follow-up. From these data, the total survival rate of the retainers will be estimated as a percentage, as well as its major causes of failure. • Patient perception. Patients' perception will be assessed using a VAS containing a questionnaire with 9 questions related to the use of the retainers. Each question will receive a score ranging from 0-10. Minor scores will indicate that the patient felt extremely uncomfortable during the retention period while higher scores will indicate that the patients felt highly comfortable.
Starting date	1 February 2021
Contact information	Principal Investigator: Dr. Silvio Augusto Bellini-Pereira Institution: Department of Orthodontics, Bauru Dental School, University of São Paulo, Brazil. Email: sbellinipereira@gmail.com
Notes	Contacted study author, study ongoing

NCT05006339

Study name	The use of remote monitoring for orthodontic retention review
Methods	Study design: single-blinded (outcome assessors), 2-armed, parallel, RCT Country: Australia Duration of follow-up: 4 years Inclusion criteria • Consent to participate in the trial and comply with the retention regime and able to follow instructions • Permanent dentition • Good general health • Good oral hygiene and good periodontal health • Owning or having access to a mobile/smart phone compatible with Dental Monitoring application (DM app is available for use on all iOS and Android devices) Exclusion criteria • Congenital anomalies or craniofacial syndromes • Poor oral hygiene and poor periodontal health • Missing teeth and teeth with poor enamel quality • Patients who were treated with orthognathic surgery • Patients who are unwilling or unable to follow the instructions provided
Participants	The study aims to enrol and randomise 82 participants.
Interventions	• Intervention A: dental device monitoring

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT05006339 *(Continued)*

	• Comparator B: clinic reviews (orthodontic retention review via in-office visits as per routine care intervention)
Outcomes	Primary outcome: stability • LII • Spacing, if present • Intercanine width • Intermolar width • Overjet • Overbite Secondary outcomes • Failure (identification: yes or no) • Failure (identification: duration, in days) • Compliance • Fit of retainer (space in mm between retainer and teeth) • Cavity presence • Gingivitis • Patient satisfaction • Cost effectiveness
Starting date	1 December 2021
Contact information	Principal Investigator: Oyku Dalci Institution: Sydney Local Health District, The University of Sydney, Australia Email: oyku.dalci@sydney.edu.au
Notes	This study is estimated to be completed in September 2023.

NCT05014308

Study name	Changes in dentition and hygiene in the retention phase of orthodontic treatment
Methods	Study design: single-centre, single-blinded *(outcomes assessor)*, 3-armed, parallel RCT Country: Saudi Arabia Duration of follow-up: after 6 and 12 months of retention Inclusion criteria • Individuals who had lower anterior irregularity (crowding, rotation, buccally or lingually displaced) will be selected. • Good oral hygiene, healthy gingiva and no evidence of bone loss as seen on the radiograph • Healthy systemic conditions • Patients pass the growth spurt according to cervical vertebral maturation method as described by Baccetti et al • Availability of good-quality records at pretreatment and post-treatment Exclusion criteria • Patients that have any signs of active periodontal disease or evidence of bone loss on the radiograph

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT05014308 *(Continued)*

	• Patients with systemic disease such as osteoporosis and diabetes • Smoking patients • Study casts did not include all mandibular permanent incisors, canines and first molars
Participants	39 randomised participants
Interventions	• Control: no interventions • Intervention 1: piezocision group. Prodecure: piezoelectric alveolar decortication and fibreotomy according to Dibart et al. (no reference provided by study authors) • Intervention 2: fibreotomy group. Prodecure: piezoelectric alveolar decortication and fibreotomy according to Edward et al. (no reference provided by study authors)
Outcomes	Lower stability: LII
Starting date	15 March 2016
Contact information	Complete study according to clinicaltrials.gov. The study authors were contacted without success
Notes	

NCT05121220

Study name	Changes in dentition and hygiene in the retention phase of orthodontic treatment
Methods	Study design: single-blinded (investigator), 3-armed, parallel RCT Country: Croatia Duration of follow-up: 1, 2 and 3 years Inclusion criteria • Finished orthodontic treatment Exclusion criteria • Hypodontia • Missing teeth
Participants	The study aims to enrol and randomise 152 participants.
Interventions	• Group 1: lower fixed rectangular retainer 0.673 x 0.268 mm (8-stranded wire Bond A Braid (Reliance Orthodontic Products, USA) will be bonded in mandible canine to canine on each tooth using Filtek flowable composite (Ivoclar Vivadent, Lichtenstein) and Multilink adhesive (3M, USA)) • Group 2: lower fixed round retainer 0.40 mm (6-stranded wire (Forestadent, Germany) will be bonded on 6 mandibular frontal teeth, canine to canine on each tooth using Filtek flowable composite (Ivoclar Vivadent, Lichtenstein) and Multilink adhesive (3M, USA)) • Group 3: lower clear plastic vacuum-formed removable retainer
Outcomes	Primary outcome • LII (mm) • Failure of the retainer (debonding, breakage) • Plaque Index (Silness and Loe) Secondary outcomes:

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT05121220 *(Continued)*

	• Intercanine width change (mm) • Calculus surface index • Bleeding on probing • Satisfation with stability • Total costs • Duration of orthodontic treatment (months)
Starting date	15 November 2020
Contact information	Principal Investigator: Silvija Kanižaj Ugrin Institutions: • University of Zagreb School of Dental Medicine, Zagreb, Croatia • University of Rijeka Faculty of Dental Medicine, Rijeka, Croatia Email: kanizaj.ugrin@gmail.com
Notes	This study is estimated to be complete in November 2023.

NCT05250765

Study name	Comparison of efficiency and effectiveness of two types of bonded orthodontic retainers: an RCT
Methods	Study design: 4-armed, parallel, RCT Country: Belgium Duration of follow-up: measured before removing braces, after 6 months, 1 year and 2 years Inclusion criteria • Fixed orthodontic appliances upper and lower jaw • Patient stays for 2 years in Belgium • Parental consent • Proper oral hygiene Exclusion criteria • Orthognathic surgery • Craniofacial disorders • Cleft lip palate patients • Orthodontic treatment without fixed appliances • Extra retention other than bonded wire in the lower jaw
Participants	The study aims to enrol and randomise 100 participants.
Interventions	Experimental groups • Intervention 1: bonding a twistflex retainer with rubber dam (0.0195 inch dead-soft coaxial wire (Respond; Ormco, Orange, Calif)) • Intervention 2: bonding an orthoflex retainer with rubber dam (Reliance orthodontic products, Itasca III, USA) Active comparator groups

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

NCT05250765 *(Continued)*

	• Control 1: bonding a twistflex retainer under relative isolation (hygrophormic suction, cotton pads) • Control 2: bonding an Orthoflex retainer under relative isolation (hygrophormic suction, cotton pads)
Outcomes	Primary outcome • LII Secondary outcomes • Failure of the retainer considered as: debonding of the retainer. Date and place of reparation registered • Failure of the retainer considered as: possible activation (unwanted tooth movement possible torque differences) • Probing depth of the 6 anterior teeth (6 places per tooth) • Bleeding on probing • Plaque Index
Starting date	1 September 2020
Contact information	Principal Investigator: Prof Guy De Pauw Institution: University Hospital, Ghent, Belgium
Notes	This study is estimated to be complete in September 2023.

TCTR20190416001

Study name	An evaluation of surface electromyography and masticatory performance in patients wearing thermoplastic oral appliance: a randomized clinical trial
Methods	Study design: single-centre, 2-armed, parallel, randomised, active-control trial Setting: conducted at a university Country: Thailand Duration of follow-up: pre-treatment, after 3 months, after 6 months Inclusion criteria • All participants have a natural dentition with no missing teeth except for the third molars • All participants are dentate with a Class I occlusion, overjet and overbite < 3 mm • All participants had no edentulous area, no significant facial deformity or symptoms of craniomandibular disorders, and no temporomandibular symptoms • Participants had previous orthodontic treatment by either extraction or non-extraction, and have to wear removable retainers to maintain the occlusion. Exclusion criteria • Systemic diseases that may affect activities of the masticatory muscles • Medications, such as muscle relaxants, that may affect muscle function
Participants	The study aims to enrol and randomise 60 participants.
Interventions	• Group 1 (experimental): clear plastic vacuum-formed appliance group. All thermoform appliances are manufactured, under standardised laboratory conditions and following the manufacturer's

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TCTR20190416001 *(Continued)*

	instructions, using a dental laboratory thermoforming machine (BIOSTAR-Scheu-Dental GmbH, Germany). The clear plastic retainers are made from 0.75-mm thick Essix ACE plastic foil (ACE, Raintree Essix, Glenroe) • Group 2 (active control): wrap-around retainer group. The wrap-around retainers had the acrylate base placed orally, and the labial archwire was made from round, 0.8 mm diameter, stainless steel wire (Dentaurum, Ispringen, Germany). The labial arches surrounded all erupted teeth including the second molars and formed a U-shaped loop between the canines and first premolars.
Outcomes	Primary outcomes • Muscle ability using surface electromyographic measurement • Masticatory performance using chewing efficiency measurement Secondary outcomes • Subjective evaluation of masticatory ability using the Food Intake Ability Index • Oral health-related quality of life
Starting date	1 January 2019
Contact information	Principal Investigator: Dr Rungtiwa Jirawiwatsaree Institution: Orthodontic department, Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand Email: mue.jirawiwat@gmail.com
Notes	Contacted study author; study currently under review in a scientific journal and about to be published. The study authors were unable to share any information about their study in the meantime.

LII: Little's Irregularity Index; **RCT:** randomised controlled trial; **VAS:** visual analogue scale

DATA AND ANALYSES

Comparison 1. Removable versus fixed retainers

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Stability: Irregularity Index	6		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.1.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	-0.34 [-0.95, 0.27]
1.1.2 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.70 [0.09, 1.31]
1.1.3 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.40 [-0.21, 1.01]
1.1.4 Lower removable clear plastic full-time vs fixed multistrand/coaxial at 6 months	2	103	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.00, 0.10]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1.5 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.70, 0.50]
1.1.6 Lower removable clear plastic full-time vs CAD-CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.62, 0.50]
1.1.7 Lower removable clear plastic full-time vs fixed multistrand at 12 months	1	84	Mean Difference (IV, Fixed, 95% CI)	0.60 [0.17, 1.03]
1.1.8 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	0.92 [0.23, 1.61]
1.1.9 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months	1	95	Mean Difference (IV, Fixed, 95% CI)	0.07 [-0.20, 0.34]
1.1.10 Lower removable clear plastic part-time vs fixed coaxial at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	1.52 [0.75, 2.29]
1.2 Stability: intercanine width	5		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.2.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.48, 0.36]
1.2.2 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.50 [0.23, 0.77]
1.2.3 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.23, 0.63]
1.2.4 Lower removable clear plastic full-time vs fixed multistrand/coaxial at 6 months	2	103	Mean Difference (IV, Fixed, 95% CI)	0.11 [0.05, 0.17]
1.2.5 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.06 [-0.48, 0.60]
1.2.6 Lower removable clear plastic full-time vs fixed CAD-CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.12, 0.32]
1.2.7 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	-0.19 [-0.53, 0.15]
1.2.8 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months	1	95	Mean Difference (IV, Fixed, 95% CI)	-0.08 [-0.29, 0.13]
1.2.9 Lower removable clear plastic part-time vs fixed coaxial at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	-0.24 [-0.87, 0.39]
1.3 Stability: intermolar width	5		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.3.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	0.18 [-0.56, 0.92]
1.3.2 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.53, 0.33]
1.3.3 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.43, 0.43]
1.3.4 Lower removable clear plastic full-time vs fixed multistrand/coaxial at 6 months	2	103	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.23, 0.03]
1.3.5 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.03 [-0.84, 0.90]
1.3.6 Lower removable clear plastic full-time vs fixed CAD-CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.25 [-1.19, 0.69]
1.3.7 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	0.43 [-0.30, 1.16]
1.3.8 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months	1	95	Mean Difference (IV, Fixed, 95% CI)	0.24 [-0.15, 0.63]
1.3.9 Lower removable clear plastic part-time vs fixed coaxial at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	0.57 [-0.36, 1.50]
1.4 Stability: arch length	5		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.4.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	0.46 [-0.16, 1.08]
1.4.2 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.37, 0.17]
1.4.3 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.17, 0.37]
1.4.4 Lower removable clear plastic full-time vs fixed multistrand/coaxial at 6 months	2	103	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.11, 0.14]
1.4.5 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.31 [-0.84, 0.22]
1.4.6 Lower removable clear plastic full-time vs fixed CAD-CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.73, 0.43]
1.4.7 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	-0.80 [-1.28, -0.32]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.4.8 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months	1	95	Mean Difference (IV, Fixed, 95% CI)	0.19 [-0.12, 0.50]
1.4.9 Lower removable clear plastic part-time vs fixed multistrand at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.84, 0.54]
1.5 Stability: overjet	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.5.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	0.21 [-0.07, 0.49]
1.5.2 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months	1	95	Mean Difference (IV, Fixed, 95% CI)	0.08 [-0.19, 0.35]
1.5.3 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.53, 0.33]
1.5.4 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.64, 0.44]
1.6 Stability: overbite	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.6.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months	1	56	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.45, 0.15]
1.6.2 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months	1	95	Mean Difference (IV, Fixed, 95% CI)	0.15 [-0.09, 0.39]
1.6.3 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.33, 0.53]
1.6.4 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.43, 0.43]
1.7 Stability: maximum rotation	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.7.1 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	3.70 [0.87, 6.53]
1.7.2 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	1.40 [-1.36, 4.16]
1.8 Failure of retainers	5		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.8.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	57	Risk Ratio (M-H, Fixed, 95% CI)	0.81 [0.38, 1.71]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.8.2 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	57	Risk Ratio (M-H, Fixed, 95% CI)	0.44 [0.20, 0.98]
1.8.3 Lower removable clear plastic full-time vs fixed multistrand at 12 months	1	77	Risk Ratio (M-H, Fixed, 95% CI)	3.42 [1.38, 8.47]
1.8.4 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months	1	104	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.21, 4.73]
1.8.5 Lower removable (acrylic & wire) vs thick plain canine to canine at 36 months	1	25	Risk Ratio (M-H, Fixed, 95% CI)	1.57 [0.16, 15.16]
1.8.6 Lower removable (acrylic & wire) vs thick spiral canine to canine at 36 months	1	27	Risk Ratio (M-H, Fixed, 95% CI)	0.46 [0.10, 2.12]
1.8.7 Lower removable (acrylic & wire) vs thin spiral wire (incisors and canines) at 36 months	1	25	Risk Ratio (M-H, Fixed, 95% CI)	0.52 [0.11, 2.61]
1.8.8 Lower removable part-time clear plastic vs fixed coaxial at 48 months	1	42	Risk Ratio (M-H, Fixed, 95% CI)	3.00 [1.33, 6.75]
1.9 Mean survival time (days)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.9.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	60	Mean Difference (IV, Fixed, 95% CI)	38.80 [-15.80, 93.40]
1.9.2 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	60	Mean Difference (IV, Fixed, 95% CI)	85.60 [31.44, 139.76]
1.10 Patient satisfaction: acceptance	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.10.1 Lower removable clear plastic full-time vs fixed multistrand at 12 months (specifically, "How acceptable was the retainer to wear?")	1	81	Mean Difference (IV, Fixed, 95% CI)	-12.84 [-18.60, -7.09]
1.10.2 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "Have you gotten used to retainers?")	1	95	Mean Difference (IV, Fixed, 95% CI)	-7.50 [-11.54, -3.46]
1.11 Patient satisfaction: acceptance (dichotomous)	1		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
1.11.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "Are you embarrassed to wear your retainer?" (dichotomised))	1	57	Risk Difference (M-H, Fixed, 95% CI)	0.07 [-0.04, 0.19]
1.11.2 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "How have you found	1	56	Risk Difference (M-H, Fixed, 95% CI)	0.15 [0.00, 0.29]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
wearing your retainer in comparison to your fixed appliances?" (dichotomised))				
1.12 Patient satisfaction: compliance (dichotomous)	1		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
1.12.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "Do you wear your retainers as instructed?" (dichotomised))	1	57	Risk Difference (M-H, Fixed, 95% CI)	0.22 [0.06, 0.38]
1.13 Patient satisfaction: efficiency	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.13.1 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "Are the retainers working well?")	1	95	Mean Difference (IV, Fixed, 95% CI)	2.00 [-2.75, 6.75]
1.13.2 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "Noticed any change in tooth position?")	1	95	Mean Difference (IV, Fixed, 95% CI)	2.00 [-2.66, 6.66]
1.14 Patient satisfaction: cleaning/hygiene (continuous)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.14.1 Lower removable clear plastic full-time vs fixed multistrand at 12 months (specifically, "How easy was it to keep the retainer clean?")	1	81	Mean Difference (IV, Fixed, 95% CI)	-10.31 [-20.05, -0.58]
1.15 Patient satisfaction: cleaning/hygiene (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.15.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "How have you found easy cleaning your retainers?")	1	57	Risk Ratio (M-H, Fixed, 95% CI)	1.39 [1.06, 1.83]
1.16 Patient satisfaction: discomfort/pain	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.16.1 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "discomfort/pain")	1	95	Mean Difference (IV, Fixed, 95% CI)	5.00 [-2.21, 12.21]
1.16.2 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "soreness")	1	95	Mean Difference (IV, Fixed, 95% CI)	2.00 [0.72, 3.28]
1.16.3 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "tension in your teeth")	1	95	Mean Difference (IV, Fixed, 95% CI)	8.50 [2.53, 14.47]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.17 Patient satisfaction: discomfort/pain (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.17.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months	1	57	Risk Ratio (M-H, Fixed, 95% CI)	12.22 [1.69, 88.52]
1.18 Patient satisfaction: speech (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.18.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "Has your retainer affected your speech?")	1	57	Risk Ratio (M-H, Fixed, 95% CI)	3.89 [1.46, 10.38]
1.19 Oral health: evidence of caries (dichotomous)	1		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
1.19.1 Lower removable full-time clear plastic vs fixed multistrand at 12 months	1	84	Risk Difference (M-H, Fixed, 95% CI)	0.00 [-0.05, 0.05]
1.20 Oral health: Plaque Index	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.20.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	59	Mean Difference (IV, Fixed, 95% CI)	-0.17 [-0.29, -0.05]
1.20.2 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	59	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.32, 0.02]
1.20.3 Lower removable clear plastic full-time vs fixed multistrand at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.21 [-0.51, 0.10]
1.20.4 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.28 [-0.74, 0.18]
1.20.5 Lower removable clear plastic full-time vs CAD/CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.27, 0.35]
1.20.6 Lower removable clear plastic part-time vs fixed coaxial at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-1.21, 0.21]
1.20.7 Upper & lower removable Hawley full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.06, 0.02]
1.20.8 Upper & lower removable clear plastic full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.04, 0.14]
1.21 Oral health: Gingival Index	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.21.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	59	Mean Difference (IV, Fixed, 95% CI)	-0.33 [-0.67, 0.01]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

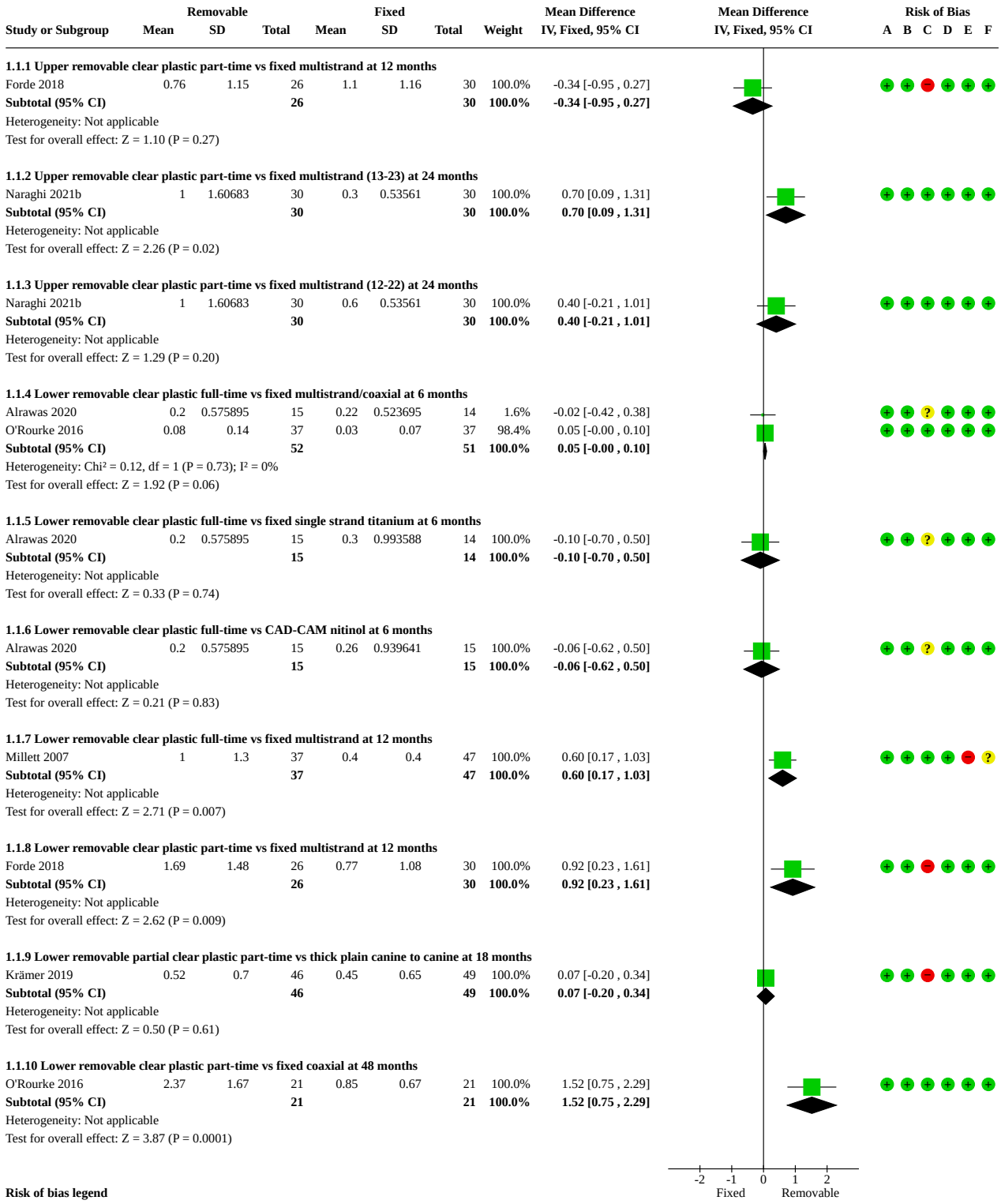
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.21.2 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	59	Mean Difference (IV, Fixed, 95% CI)	-0.34 [-0.66, -0.02]
1.21.3 Lower removable clear plastic full-time vs fixed multistrand at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.12 [-0.50, 0.26]
1.21.4 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.06 [-0.39, 0.50]
1.21.5 Lower removable clear plastic full-time vs fixed CAD/CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.12 [-0.25, 0.49]
1.21.6 Lower removable clear plastic part-time vs fixed coaxial at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	0.50 [-0.84, 1.84]
1.21.7 Upper & lower removable Hawley full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-0.15, 0.05]
1.21.8 Upper & lower removable clear plastic full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.16, 0.24]
1.22 Oral health: Calculus Index	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.22.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months	1	0	Mean Difference (IV, Fixed, 95% CI)	Not estimable
1.22.2 Lower removable clear plastic part-time vs fixed multistrand at 12 months	1	59	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.29, -0.01]
1.22.3 Lower removable clear plastic part-time vs fixed coaxial at 48 months	1	42	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.45, 0.45]
1.23 Oral health: probing depth	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.23.1 Lower removable clear plastic full-time vs fixed multistrand at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.09 [-0.32, 0.14]
1.23.2 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.28, 0.16]
1.23.3 Lower removable clear plastic full-time vs CAD/CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.32, 0.12]
1.23.4 Upper & lower removable Hawley full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.13 [-0.37, 0.11]
1.23.5 Upper & lower removable clear plastic full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.07 [-0.27, 0.13]
1.24 Oral health: periodontal pocketing (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.24.1 Lower removable clear full-time plastic vs fixed multistrand at 12 months	1	84	Risk Ratio (M-H, Fixed, 95% CI)	0.32 [0.12, 0.87]
1.25 Oral health: bleeding on probing	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
1.25.1 Lower removable clear plastic full-time vs fixed multistrand at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.13, 0.15]
1.25.2 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.18, 0.12]
1.25.3 Lower removable clear plastic full-time vs CAD/CAM nitinol at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.12 [0.00, 0.24]
1.25.4 Upper & lower removable Hawley full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.03 [0.00, 0.06]
1.25.5 Upper & lower removable clear plastic full-time vs fixed multistrand at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.02, 0.04]
1.26 Oral health: gingival bleeding (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.26.1 Lower removable clear plastic full-time vs fixed multistrand at 12 months	1	84	Risk Ratio (M-H, Fixed, 95% CI)	0.53 [0.31, 0.88]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.1. Comparison 1: Removable versus fixed retainers, Outcome 1: Stability: Irregularity Index



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

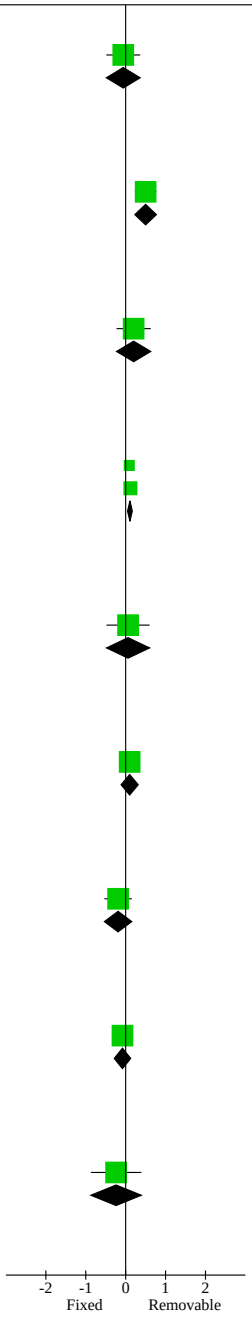
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.1. (Continued)

(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.2. Comparison 1: Removable versus fixed retainers, Outcome 2: Stability: intercanine width

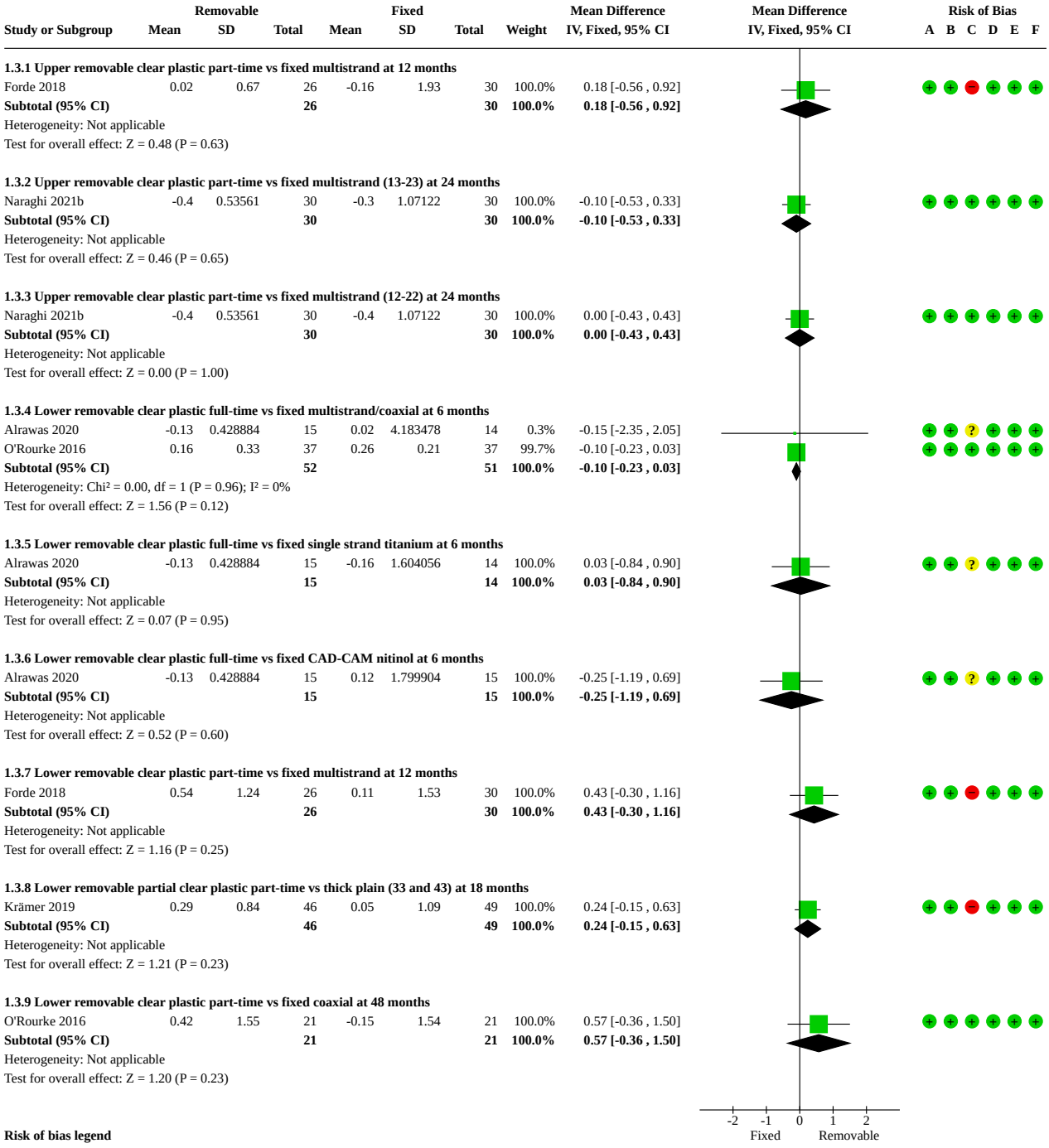
Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F
1.2.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months									
Forde 2018	0.19	0.76	26	0.25	0.86	30	100.0%	-0.06 [-0.48 , 0.36]	+ + − + + +
Subtotal (95% CI)			**26**			**30**	**100.0%**	**-0.06 [-0.48 , 0.36]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.28 (P = 0.78)									
1.2.2 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months									
Naraghi 2021b	0.2	0.53561	30	-0.3	0.53561	30	100.0%	0.50 [0.23 , 0.77]	+ + + + + +
Subtotal (95% CI)			**30**			**30**	**100.0%**	**0.50 [0.23 , 0.77]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 3.62 (P = 0.0003)									
1.2.3 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months									
Naraghi 2021b	0.2	0.53561	30	0	1.07122	30	100.0%	0.20 [-0.23 , 0.63]	+ + + + + +
Subtotal (95% CI)			**30**			**30**	**100.0%**	**0.20 [-0.23 , 0.63]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.91 (P = 0.36)									
1.2.4 Lower removable clear plastic full-time vs fixed multistrand/coaxial at 6 months									
Alrawas 2020	0.000001	0.003036	15	-0.09	0.174672	14	42.1%	0.09 [-0.00 , 0.18]	+ + ? + + +
O'Rourke 2016	0.23	0.15	37	0.11	0.19	37	57.9%	0.12 [0.04 , 0.20]	+ + + + + +
Subtotal (95% CI)			**52**			**51**	**100.0%**	**0.11 [0.05 , 0.17]**	
Heterogeneity: Chi² = 0.24, df = 1 (P = 0.62); I² = 0%									
Test for overall effect: Z = 3.55 (P = 0.0004)									
1.2.5 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months									
Alrawas 2020	0.000001	0.003036	15	-0.06	1.030968	14	100.0%	0.06 [-0.48 , 0.60]	+ + ? + + +
Subtotal (95% CI)			**15**			**14**	**100.0%**	**0.06 [-0.48 , 0.60]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.22 (P = 0.83)									
1.2.6 Lower removable clear plastic full-time vs fixed CAD-CAM nitinol at 6 months									
Alrawas 2020	0.000001	0.003036	15	-0.1	0.425397	15	100.0%	0.10 [-0.12 , 0.32]	+ + ? + + +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.10 [-0.12 , 0.32]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.91 (P = 0.36)									
1.2.7 Lower removable clear plastic part-time vs fixed multistrand at 12 months									
Forde 2018	-0.45	0.64	26	-0.26	0.67	30	100.0%	-0.19 [-0.53 , 0.15]	+ + − + + +
Subtotal (95% CI)			**26**			**30**	**100.0%**	**-0.19 [-0.53 , 0.15]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.08 (P = 0.28)									
1.2.8 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months									
Krämer 2019	-0.11	0.61	46	-0.03	0.38	49	100.0%	-0.08 [-0.29 , 0.13]	+ + − + + +
Subtotal (95% CI)			**46**			**49**	**100.0%**	**-0.08 [-0.29 , 0.13]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.76 (P = 0.45)									
1.2.9 Lower removable clear plastic part-time vs fixed coaxial at 48 months									
O'Rourke 2016	-0.52	1.19	21	-0.28	0.88	21	100.0%	-0.24 [-0.87 , 0.39]	+ + + + + +
Subtotal (95% CI)			**21**			**21**	**100.0%**	**-0.24 [-0.87 , 0.39]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.74 (P = 0.46)									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

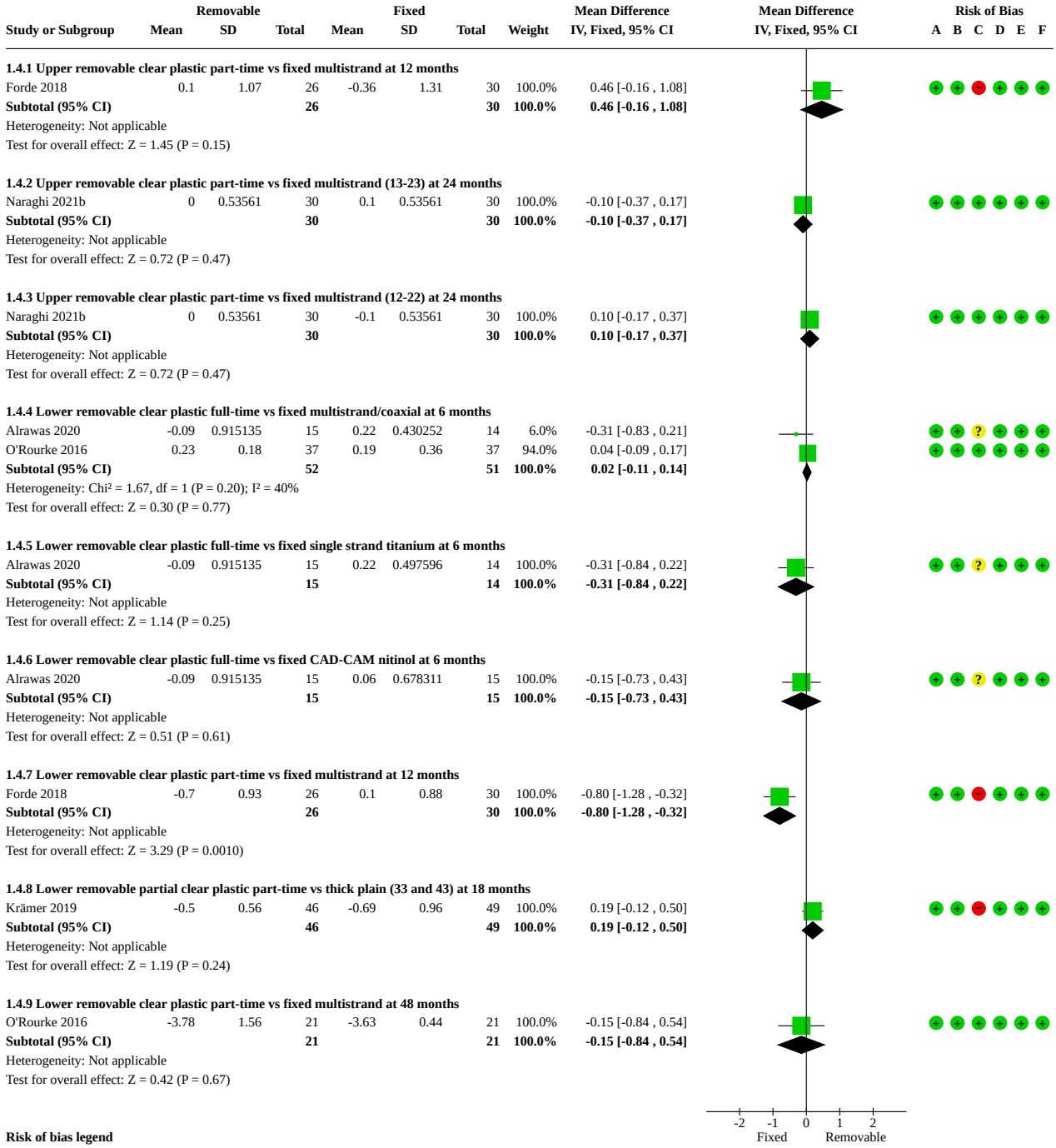
Analysis 1.3. Comparison 1: Removable versus fixed retainers, Outcome 3: Stability: intermolar width



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.4. Comparison 1: Removable versus fixed retainers, Outcome 4: Stability: arch length

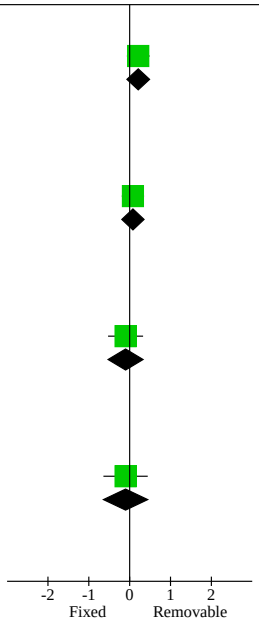


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.5. Comparison 1: Removable versus fixed retainers, Outcome 5: Stability: overjet

Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F
1.5.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months									
Forde 2018	0.12	0.51	26	-0.09	0.55	30	100.0%	0.21 [-0.07 , 0.49]	+ + - + + +
Subtotal (95% CI)			**26**			**30**	**100.0%**	**0.21 [-0.07 , 0.49]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.48 (P = 0.14)									
1.5.2 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months									
Krämer 2019	-0.02	0.75	46	-0.1	0.59	49	100.0%	0.08 [-0.19 , 0.35]	+ + - + + +
Subtotal (95% CI)			**46**			**49**	**100.0%**	**0.08 [-0.19 , 0.35]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.58 (P = 0.57)									
1.5.3 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months									
Naraghi 2021b	0	1.07122	30	0.1	0.53561	30	100.0%	-0.10 [-0.53 , 0.33]	+ + + + + +
Subtotal (95% CI)			**30**			**30**	**100.0%**	**-0.10 [-0.53 , 0.33]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.46 (P = 0.65)									
1.5.4 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months									
Naraghi 2021b	0	1.07122	30	0.1	1.07122	30	100.0%	-0.10 [-0.64 , 0.44]	+ + + + + +
Subtotal (95% CI)			**30**			**30**	**100.0%**	**-0.10 [-0.64 , 0.44]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.36 (P = 0.72)									



Risk of bias legend

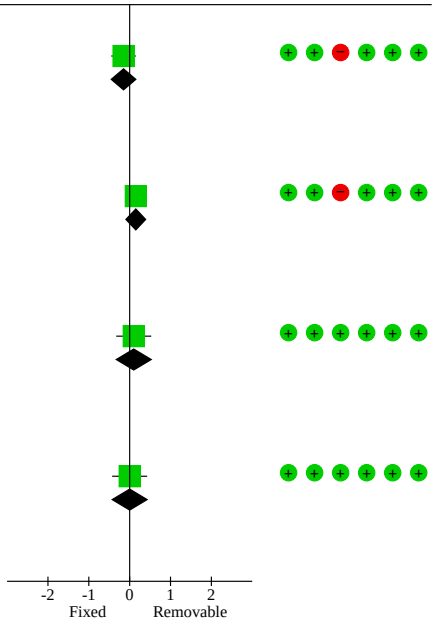
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.6. Comparison 1: Removable versus fixed retainers, Outcome 6: Stability: overbite

Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI
1.6.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months								
Forde 2018	0.15	0.61	26	0.3	0.53	30	100.0%	-0.15 [-0.45 , 0.15]
Subtotal (95% CI)			**26**			**30**	**100.0%**	**-0.15 [-0.45 , 0.15]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.97 (P = 0.33)								
1.6.2 Lower removable partial clear plastic part-time vs thick plain (33 and 43) at 18 months								
Krämer 2019	0.51	0.53	46	0.36	0.67	49	100.0%	0.15 [-0.09 , 0.39]
Subtotal (95% CI)			**46**			**49**	**100.0%**	**0.15 [-0.09 , 0.39]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.22)								
1.6.3 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months								
Naraghi 2021b	0.3	1.07122	30	0.2	0.53561	30	100.0%	0.10 [-0.33 , 0.53]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**0.10 [-0.33 , 0.53]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.46 (P = 0.65)								
1.6.4 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months								
Naraghi 2021b	0.3	1.07122	30	0.3	0.53561	30	100.0%	0.00 [-0.43 , 0.43]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**0.00 [-0.43 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								

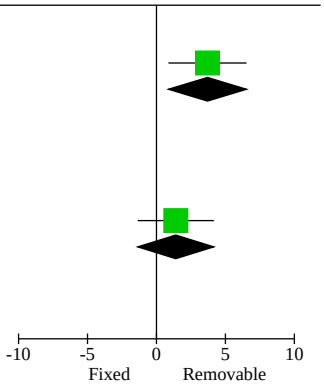
Test for subgroup differences: Chi² = 0.00, df = 3 (P < 0.00001), I² = 0%



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

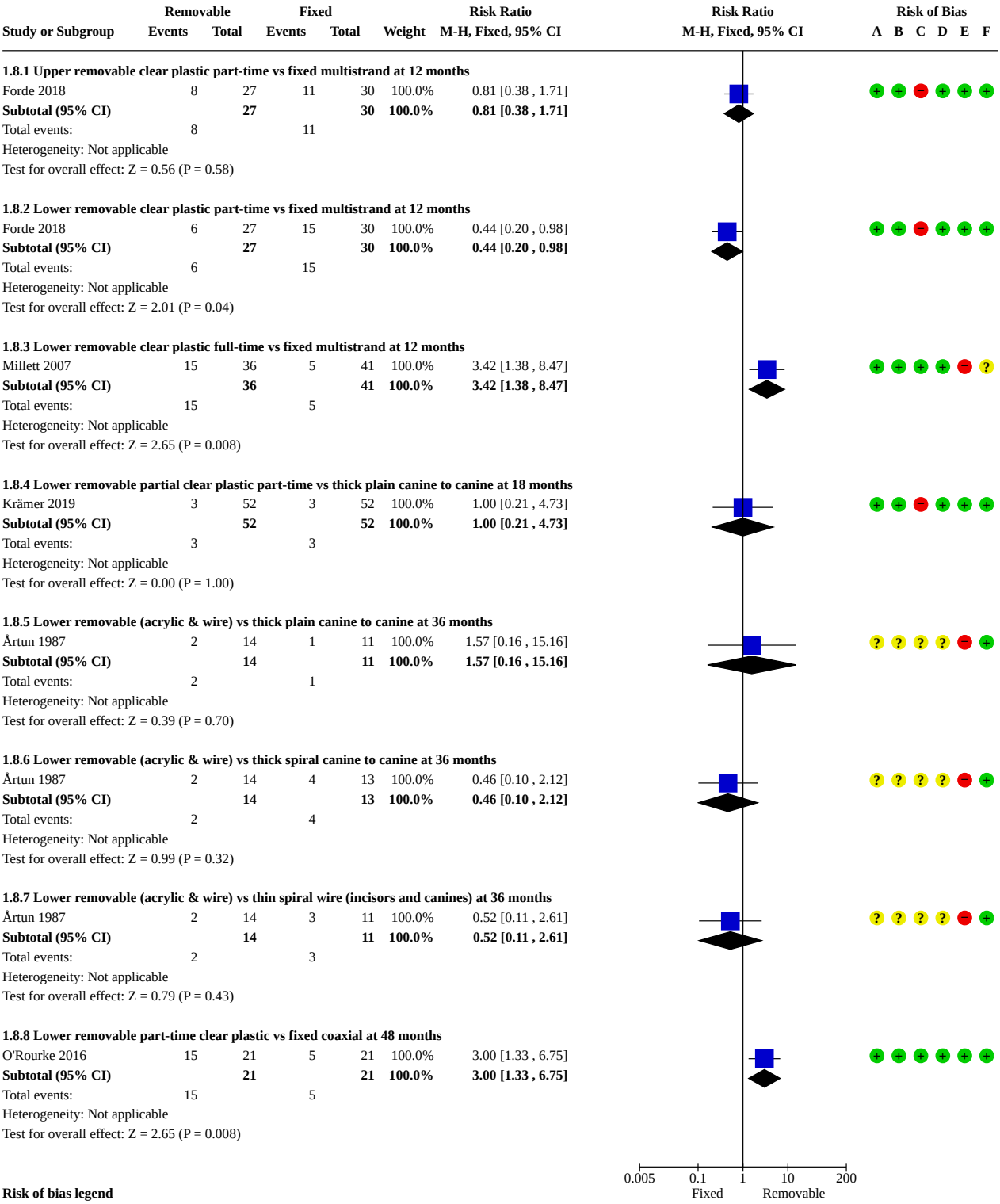
Analysis 1.7. Comparison 1: Removable versus fixed retainers, Outcome 7: Stability: maximum rotation

Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI
1.7.1 Upper removable clear plastic part-time vs fixed multistrand (13-23) at 24 months								
Naraghi 2021b	6.6	7.230733	30	2.9	3.213659	30	100.0%	3.70 [0.87 , 6.53]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**3.70 [0.87 , 6.53]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.56 (P = 0.01)								
1.7.2 Upper removable clear plastic part-time vs fixed multistrand (12-22) at 24 months								
Naraghi 2021b	6.6	7.230733	30	5.2	2.678049	30	100.0%	1.40 [-1.36 , 4.16]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**1.40 [-1.36 , 4.16]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.99 (P = 0.32)								



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.8. Comparison 1: Removable versus fixed retainers, Outcome 8: Failure of retainers



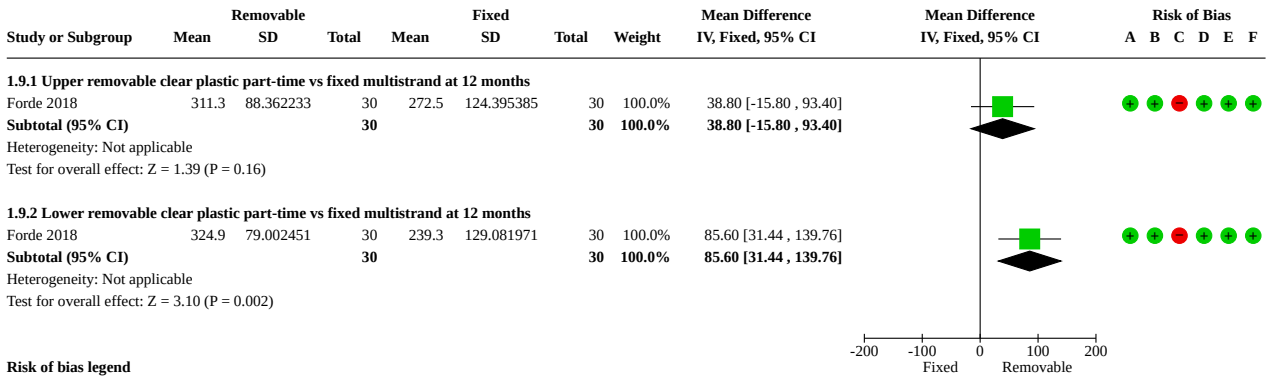
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.8. (Continued)

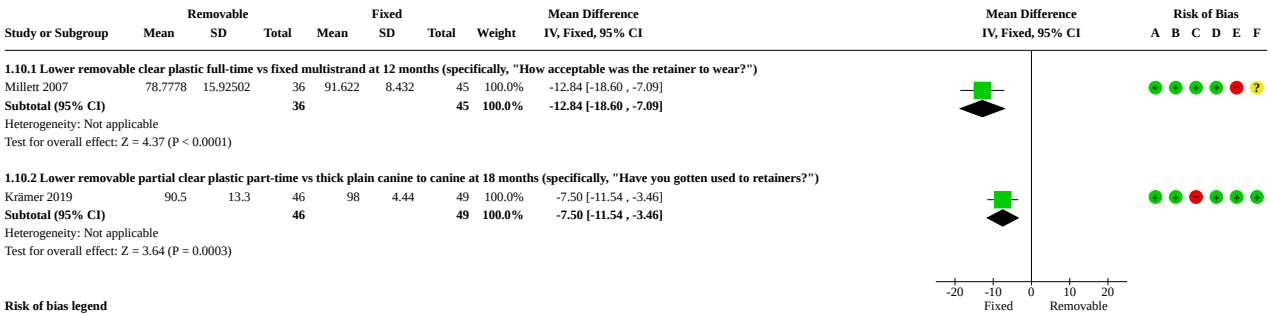
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.9. Comparison 1: Removable versus fixed retainers, Outcome 9: Mean survival time (days)



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

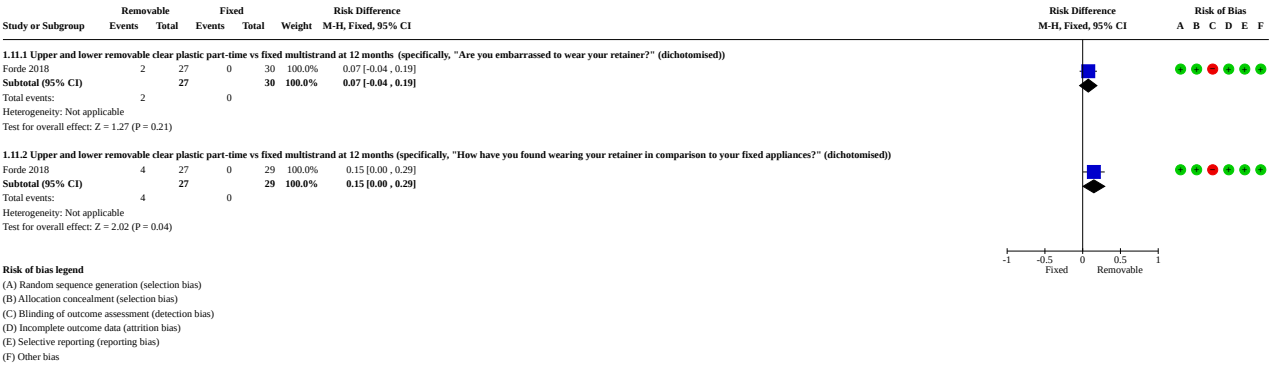
Analysis 1.10. Comparison 1: Removable versus fixed retainers, Outcome 10: Patient satisfaction: acceptance



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

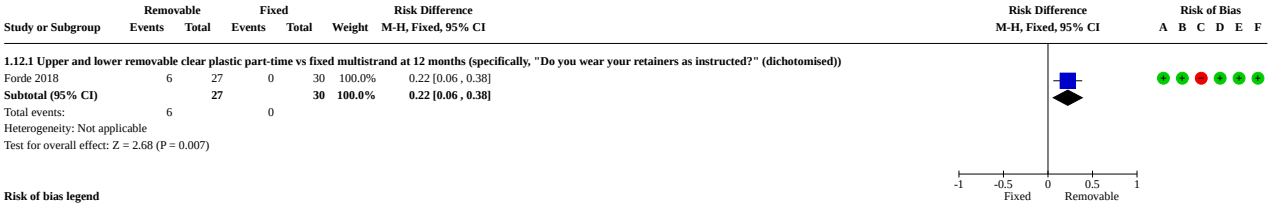
Analysis 1.11. Comparison 1: Removable versus fixed retainers, Outcome 11: Patient satisfaction: acceptance (dichotomous)



Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Difference M-H, Fixed, 95% CI
1.11.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "Are you embarrassed to wear your retainer?" (dichotomised))						
Forde 2018	2	27	0	30	100.0%	0.07 [-0.04 , 0.19]
Subtotal (95% CI)		**27**		**30**	**100.0%**	**0.07 [-0.04 , 0.19]**
Total events:	2		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.27 (P = 0.21)						
1.11.2 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "How have you found wearing your retainer in comparison to your fixed appliances?" (dichotomised))						
Forde 2018	4	27	0	29	100.0%	0.15 [0.00 , 0.29]
Subtotal (95% CI)		**27**		**29**	**100.0%**	**0.15 [0.00 , 0.29]**
Total events:	4		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.02 (P = 0.04)						

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

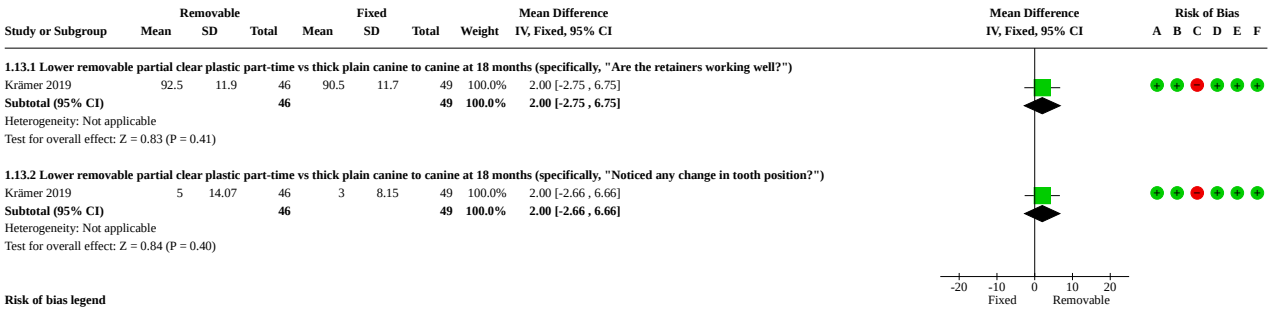
Analysis 1.12. Comparison 1: Removable versus fixed retainers, Outcome 12: Patient satisfaction: compliance (dichotomous)



Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Difference M-H, Fixed, 95% CI
1.12.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "Do you wear your retainers as instructed?" (dichotomised))						
Forde 2018	6	27	0	30	100.0%	0.22 [0.06 , 0.38]
Subtotal (95% CI)		**27**		**30**	**100.0%**	**0.22 [0.06 , 0.38]**
Total events:	6		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.68 (P = 0.007)						

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.13. Comparison 1: Removable versus fixed retainers, Outcome 13: Patient satisfaction: efficiency



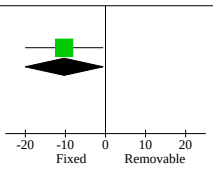
Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI
1.13.1 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "Are the retainers working well?")								
Krämer 2019	92.5	11.9	46	90.5	11.7	49	100.0%	2.00 [-2.75 , 6.75]
Subtotal (95% CI)			**46**			**49**	**100.0%**	**2.00 [-2.75 , 6.75]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.83 (P = 0.41)								
1.13.2 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "Noticed any change in tooth position?")								
Krämer 2019	5	14.07	46	3	8.15	49	100.0%	2.00 [-2.66 , 6.66]
Subtotal (95% CI)			**46**			**49**	**100.0%**	**2.00 [-2.66 , 6.66]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.84 (P = 0.40)								

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.14. Comparison 1: Removable versus fixed retainers, Outcome 14: Patient satisfaction: cleaning/hygiene (continuous)

Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI
1.14.1 Lower removable clear plastic full-time vs fixed multistrand at 12 months (specifically, "How easy was it to keep the retainer clean?")								
Millett 2007	71.44	25.33	36	81.755	17.57	45	100.0%	-10.31 [-20.05 , -0.58]
Subtotal (95% CI)			**36**			**45**	**100.0%**	**-10.31 [-20.05 , -0.58]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.08 (P = 0.04)								

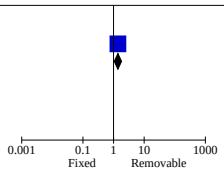




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.15. Comparison 1: Removable versus fixed retainers, Outcome 15: Patient satisfaction: cleaning/hygiene (dichotomous)

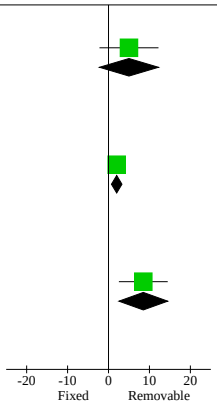
Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.15.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "How have you found easy cleaning your retainers?")						
Forde 2018	25	27	20	30	100.0%	1.39 [1.06 , 1.83]
Subtotal (95% CI)		**27**		**30**	**100.0%**	**1.39 [1.06 , 1.83]**
Total events:	25		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.34 (P = 0.02)						

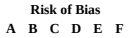


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.16. Comparison 1: Removable versus fixed retainers, Outcome 16: Patient satisfaction: discomfort/pain

Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI
1.16.1 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "discomfort/pain")								
Krämer 2019	6	24.9	46	1	1.48	49	100.0%	5.00 [-2.21 , 12.21]
Subtotal (95% CI)			**46**			**49**	**100.0%**	**5.00 [-2.21 , 12.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.36 (P = 0.17)								
1.16.2 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "soreness")								
Krämer 2019	3	3.89	46	1	2.22	49	100.0%	2.00 [0.72 , 3.28]
Subtotal (95% CI)			**46**			**49**	**100.0%**	**2.00 [0.72 , 3.28]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.05 (P = 0.002)								
1.16.3 Lower removable partial clear plastic part-time vs thick plain canine to canine at 18 months (specifically, "tension in your teeth")								
Krämer 2019	16	15.5	46	7.5	14.07	49	100.0%	8.50 [2.53 , 14.47]
Subtotal (95% CI)			**46**			**49**	**100.0%**	**8.50 [2.53 , 14.47]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.79 (P = 0.005)								



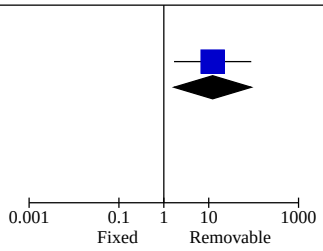


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.17. Comparison 1: Removable versus fixed retainers, Outcome 17: Patient satisfaction: discomfort/pain (dichotomous)

Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.17.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months						
Forde 2018	11	27	1	30	100.0%	12.22 [1.69 , 88.52]
Subtotal (95% CI)		**27**		**30**	**100.0%**	**12.22 [1.69 , 88.52]**
Total events:	11		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.48 (P = 0.01)						

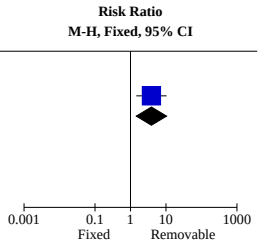




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

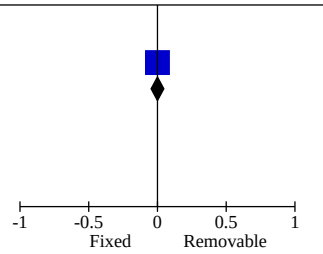
Analysis 1.18. Comparison 1: Removable versus fixed retainers, Outcome 18: Patient satisfaction: speech (dichotomous)

Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.18.1 Upper and lower removable clear plastic part-time vs fixed multistrand at 12 months (specifically, "Has your retainer affected your speech?")						
Forde 2018	14	27	4	30	100.0%	3.89 [1.46 , 10.38]
Subtotal (95% CI)		**27**		**30**	**100.0%**	**3.89 [1.46 , 10.38]**
Total events:	14		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.71 (P = 0.007)						



Analysis 1.19. Comparison 1: Removable versus fixed retainers, Outcome 19: Oral health: evidence of caries (dichotomous)

Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Difference M-H, Fixed, 95% CI
1.19.1 Lower removable full-time clear plastic vs fixed multistrand at 12 months						
Millett 2007	0	37	0	47	100.0%	0.00 [-0.05 , 0.05]
Subtotal (95% CI)		**37**		**47**	**100.0%**	**0.00 [-0.05 , 0.05]**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						

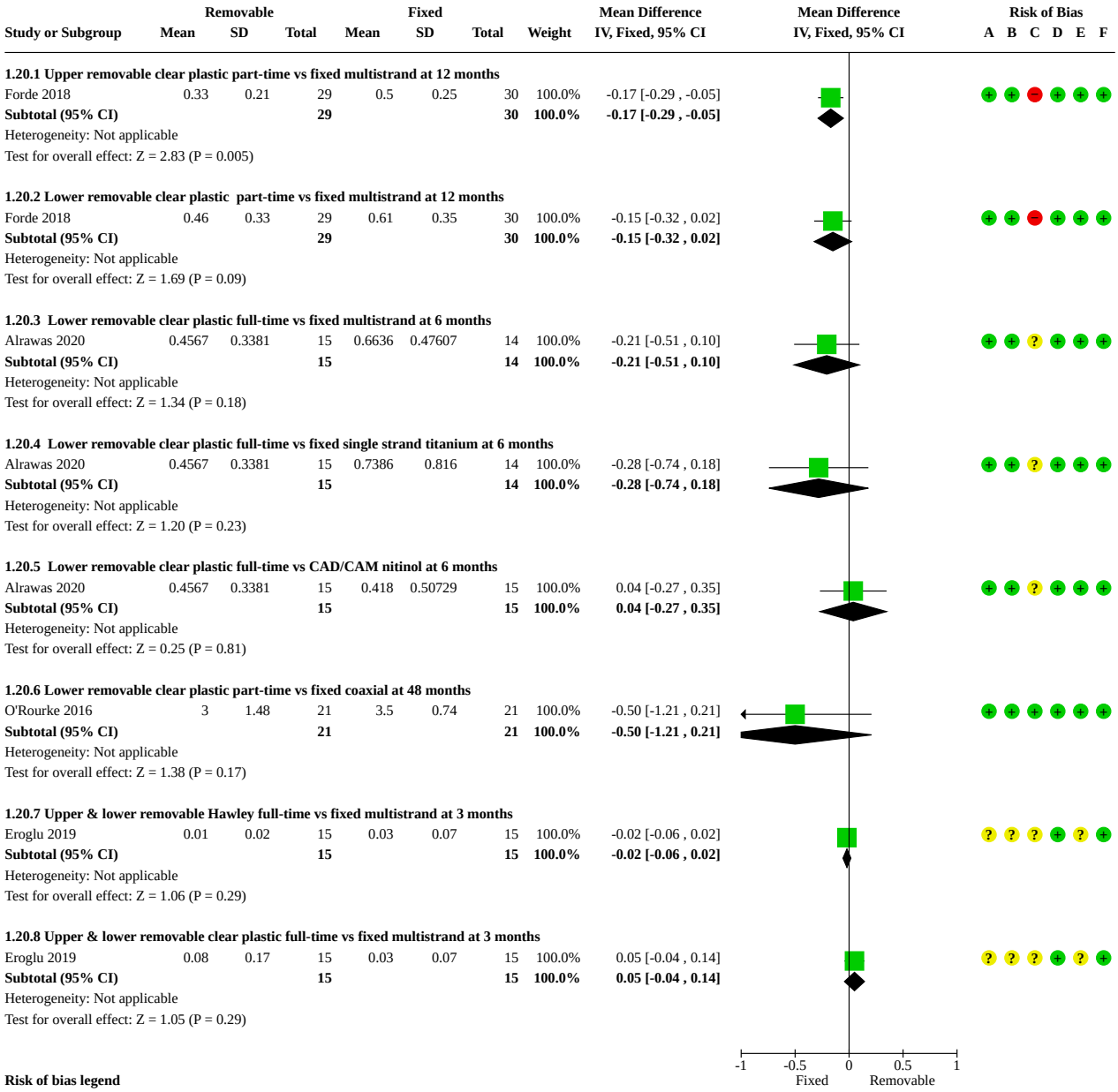




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.20. Comparison 1: Removable versus fixed retainers, Outcome 20: Oral health: Plaque Index

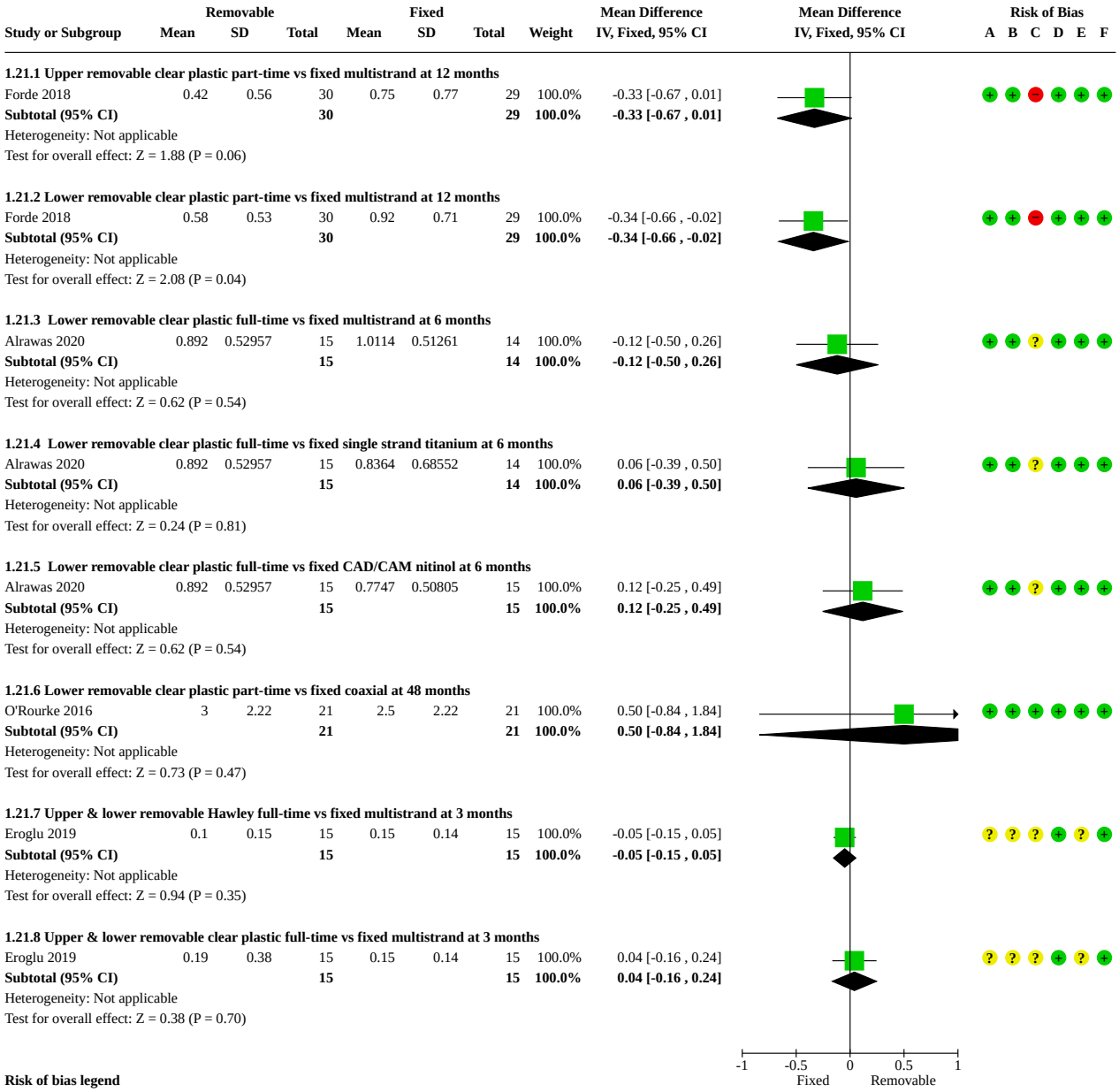


Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.21. Comparison 1: Removable versus fixed retainers, Outcome 21: Oral health: Gingival Index



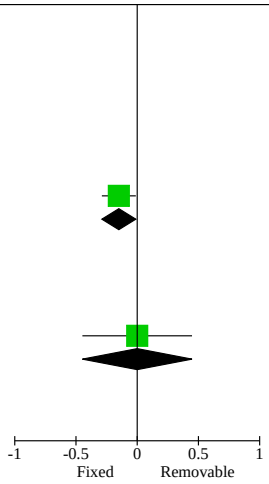
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.22. Comparison 1: Removable versus fixed retainers, Outcome 22: Oral health: Calculus Index

Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F
1.22.1 Upper removable clear plastic part-time vs fixed multistrand at 12 months									
Forde 2018	0	0	30	0	0	29		Not estimable	+ + − + + +
Subtotal (95% CI)			**0**			**0**		**Not estimable**	
Heterogeneity: Not applicable									
Test for overall effect: Not applicable									
1.22.2 Lower removable clear plastic part-time vs fixed multistrand at 12 months									
Forde 2018	0.08	0.21	30	0.23	0.32	29	100.0%	-0.15 [-0.29 , -0.01]	+ + − + + +
Subtotal (95% CI)			**30**			**29**	**100.0%**	**-0.15 [-0.29 , -0.01]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 2.12 (P = 0.03)									
1.22.3 Lower removable clear plastic part-time vs fixed coaxial at 48 months									
O'Rourke 2016	0	0.74	21	0	0.74	21	100.0%	0.00 [-0.45 , 0.45]	+ + + + + +
Subtotal (95% CI)			**21**			**21**	**100.0%**	**0.00 [-0.45 , 0.45]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.00 (P = 1.00)									

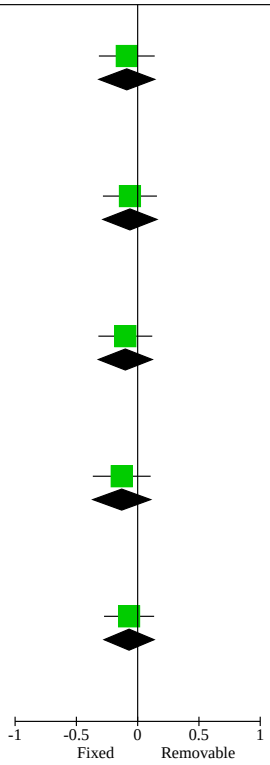


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.23. Comparison 1: Removable versus fixed retainers, Outcome 23: Oral health: probing depth

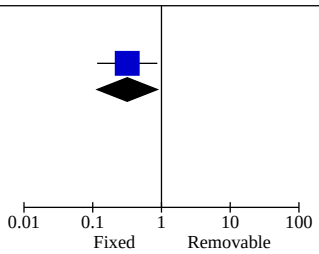
Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F
1.23.1 Lower removable clear plastic full-time vs fixed multistrand at 6 months									
Alrawas 2020	1.6713	0.31713	15	1.76	0.3084	14	100.0%	-0.09 [-0.32 , 0.14]	+ + ? + + +
Subtotal (95% CI)			**15**			**14**	**100.0%**	**-0.09 [-0.32 , 0.14]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.76 (P = 0.45)									
1.23.2 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months									
Alrawas 2020	1.6713	0.31713	15	1.7343	0.28833	14	100.0%	-0.06 [-0.28 , 0.16]	+ + ? + + +
Subtotal (95% CI)			**15**			**14**	**100.0%**	**-0.06 [-0.28 , 0.16]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.56 (P = 0.58)									
1.23.3 Lower removable clear plastic full-time vs CAD/CAM nitinol at 6 months									
Alrawas 2020	1.6713	0.31713	15	1.772	0.2967	15	100.0%	-0.10 [-0.32 , 0.12]	+ + ? + + +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.10 [-0.32 , 0.12]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.90 (P = 0.37)									
1.23.4 Upper & lower removable Hawley full-time vs fixed multistrand at 3 months									
Eroglu 2019	1.46	0.38	15	1.59	0.27	15	100.0%	-0.13 [-0.37 , 0.11]	? ? ? + ? +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.13 [-0.37 , 0.11]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.08 (P = 0.28)									
1.23.5 Upper & lower removable clear plastic full-time vs fixed multistrand at 3 months									
Eroglu 2019	1.52	0.3	15	1.59	0.27	15	100.0%	-0.07 [-0.27 , 0.13]	? ? ? + ? +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.07 [-0.27 , 0.13]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.67 (P = 0.50)									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.24. Comparison 1: Removable versus fixed retainers, Outcome 24: Oral health: periodontal pocketing (dichotomous)

Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F
1.24.1 Lower removable clear full-time plastic vs fixed multistrand at 12 months							
Millett 2007	4	37	16	47	100.0%	0.32 [0.12 , 0.87]	+ + + + − ?
Subtotal (95% CI)		**37**		**47**	**100.0%**	**0.32 [0.12 , 0.87]**	
Total events:	4		16				
Heterogeneity: Not applicable							
Test for overall effect: Z = 2.23 (P = 0.03)							

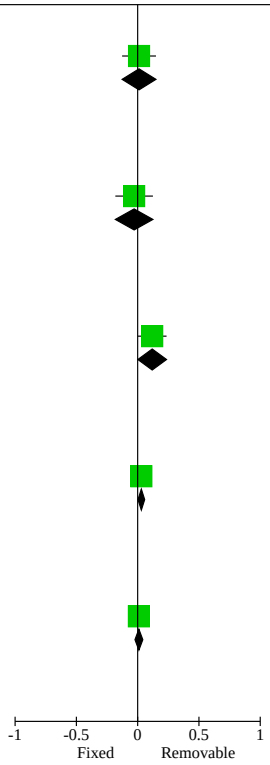


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 1.25. Comparison 1: Removable versus fixed retainers, Outcome 25: Oral health: bleeding on probing

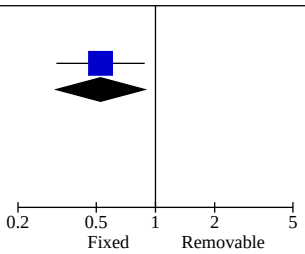
Study or Subgroup	Removable Mean	Removable SD	Removable Total	Fixed Mean	Fixed SD	Fixed Total	Weight	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F
1.25.1 Lower removable clear plastic full-time vs fixed multistrand at 6 months									
Alrawas 2020	0.2433	0.18208	15	0.2321	0.19632	14	100.0%	0.01 [-0.13 , 0.15]	+ + ? + + +
Subtotal (95% CI)			**15**			**14**	**100.0%**	**0.01 [-0.13 , 0.15]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.16 (P = 0.87)									
1.25.2 Lower removable clear plastic full-time vs fixed single strand titanium at 6 months									
Alrawas 2020	0.2433	0.18208	15	0.2721	0.23394	14	100.0%	-0.03 [-0.18 , 0.12]	+ + ? + + +
Subtotal (95% CI)			**15**			**14**	**100.0%**	**-0.03 [-0.18 , 0.12]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.37 (P = 0.71)									
1.25.3 Lower removable clear plastic full-time vs CAD/CAM nitinol at 6 months									
Alrawas 2020	0.2433	0.18208	15	0.1247	0.14242	15	100.0%	0.12 [0.00 , 0.24]	+ + ? + + +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.12 [0.00 , 0.24]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.99 (P = 0.05)									
1.25.4 Upper & lower removable Hawley full-time vs fixed multistrand at 3 months									
Eroglu 2019	0.04	0.05	15	0.01	0.02	15	100.0%	0.03 [0.00 , 0.06]	? ? ? + ? +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.03 [0.00 , 0.06]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 2.16 (P = 0.03)									
1.25.5 Upper & lower removable clear plastic full-time vs fixed multistrand at 3 months									
Eroglu 2019	0.02	0.06	15	0.01	0.02	15	100.0%	0.01 [-0.02 , 0.04]	? ? ? + ? +
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.01 [-0.02 , 0.04]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.61 (P = 0.54)									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 1.26. Comparison 1: Removable versus fixed retainers, Outcome 26: Oral health: gingival bleeding (dichotomous)

Study or Subgroup	Removable Events	Removable Total	Fixed Events	Fixed Total	Weight	Risk Ratio M-H, Fixed, 95% CI	Risk of Bias A B C D E F
1.26.1 Lower removable clear plastic full-time vs fixed multistrand at 12 months							
Millett 2007	12	37	29	47	100.0%	0.53 [0.31 , 0.88]	+ + + + − ?
Subtotal (95% CI)		**37**		**47**	**100.0%**	**0.53 [0.31 , 0.88]**	
Total events:	12		29				
Heterogeneity: Not applicable							
Test for overall effect: Z = 2.44 (P = 0.01)							



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Comparison 2. Fixed versus fixed retainers

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Stability: Irregularity Index	9		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.1.1 Upper CAD/CAM nitinol vs rectangular chain bonded retainers at 6 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.23, 0.43]
2.1.2 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.30 [0.03, 0.57]
2.1.3 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 6 months	3	258	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.12, 0.06]
2.1.4 Lower CAD/CAM nitinol vs rectangular chain bonded retainers at 6 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.12 [-0.30, 0.06]
2.1.5 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.74, 0.66]
2.1.6 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	-0.08 [-0.67, 0.51]
2.1.7 Lower ceramic interlocking bonded vs coaxial wire retainer at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.17 [-0.05, 0.39]
2.1.8 Lower fibre-reinforced composite vs multistrand retainer at 6 months	1	52	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.47, 0.17]
2.1.9 Lower CAD-CAM stainless steel multistrand vs lab stainless steel multistrand at 6 months	1	32	Mean Difference (IV, Fixed, 95% CI)	-0.58 [-1.40, 0.24]
2.1.10 Lower CAD-CAM stainless steel multistrand vs chairside rectangular chain at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.03 [-0.37, 0.43]
2.1.11 Lower lab stainless steel multistrand vs chairside rectangular chain at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.61 [-0.12, 1.34]
2.1.12 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 6 months	1	120	Mean Difference (IV, Fixed, 95% CI)	0.96 [0.57, 1.35]
2.1.13 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 12 months	1	120	Mean Difference (IV, Fixed, 95% CI)	1.15 [0.74, 1.56]
2.1.14 Lower fibre-reinforced composite vs multistrand retainer at 12 months	1	52	Mean Difference (IV, Fixed, 95% CI)	-0.70 [-1.17, -0.23]
2.1.15 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.36 [0.10, 0.63]
2.1.16 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.01 [-0.32, 0.30]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1.17 Lower connected bonding pad retainer vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.38 [0.17, 0.58]
2.1.18 Lower CAD-CAM nitinol vs rectangular braided multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.46 [-0.72, -0.21]
2.1.19 Lower CAD-CAM nitinol vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.21, 0.01]
2.1.20 Lower CAD-CAM nitinol vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.48 [-0.66, -0.29]
2.2 Stability: intercanine width	8		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.2.1 Upper CAD/CAM nitinol vs multistrand retainers at 6 months	1	163	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.60, 0.60]
2.2.2 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.13, 0.73]
2.2.3 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months	1	55	Mean Difference (IV, Fixed, 95% CI)	1.12 [0.80, 1.45]
2.2.4 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 6 months	4	309	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.15, 0.22]
2.2.5 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.62, 0.54]
2.2.6 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.58, 0.52]
2.2.7 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 6 months	1	120	Mean Difference (IV, Fixed, 95% CI)	-0.34 [-1.02, 0.34]
2.2.8 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.47 [0.12, 0.82]
2.2.9 Lower CAD-CAM stainless steel multistrand vs lab stainless steel multistrand at 6 months	1	32	Mean Difference (IV, Fixed, 95% CI)	0.31 [-0.18, 0.80]
2.2.10 Lower CAD-CAM stainless steel multistrand vs chairside rectangular chain at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	1.23 [0.82, 1.64]
2.2.11 Lower lab stainless steel multistrand vs chairside rectangular chain at 6 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.92 [0.49, 1.35]
2.2.12 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 12 months	1	120	Mean Difference (IV, Fixed, 95% CI)	-0.53 [-1.24, 0.18]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.2.13 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.32 [-0.90, 0.26]
2.2.14 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.49, 0.68]
2.2.15 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.41 [-0.19, 1.02]
2.2.16 Lower CAD/CAM nitinol vs rectangular braided multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.29 [-0.23, 0.82]
2.2.17 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.39 [-0.16, 0.94]
2.2.18 Lower CAD/CAM nitinol vs multistrand retainers at 12 months	2	107	Mean Difference (IV, Fixed, 95% CI)	0.18 [-0.28, 0.64]
2.3 Stability: interpremolar width	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.3.1 Upper CAD/CAM nitinol vs multistrand retainers at 6 months	1	163	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.37, 0.97]
2.3.2 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	1	163	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.50, 0.50]
2.4 Stability: intermolar width	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.4.1 Upper CAD/CAM nitinol vs multistrand retainers at 6 months	1	163	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-1.07, 0.67]
2.4.2 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.64, 0.44]
2.4.3 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 6 months	3	243	Mean Difference (IV, Fixed, 95% CI)	0.28 [-0.38, 0.93]
2.4.4 Lower CAD/CAM nitinol vs single strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.28 [-0.96, 1.52]
2.4.5 Lower multistrand stainless steel vs single strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	0.18 [-2.17, 2.53]
2.4.6 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 12 months	1	41	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-1.88, 0.88]
2.5 Stability: anterior arch length	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.5.1 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 6 months	2	80	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.23, 0.24]
2.5.2 Lower CAD/CAM nitinol vs single strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.16 [-0.59, 0.27]
2.5.3 Lower multistrand stainless steel vs single strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.34, 0.34]
2.5.4 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 12 months	1	41	Mean Difference (IV, Fixed, 95% CI)	0.26 [-0.04, 0.56]
2.6 Stability: total arch length	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.6.1 Upper CAD/CAM nitinol vs multistrand retainers at 6 months	1	163	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.74, 0.74]
2.6.2 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-0.47, 0.07]
2.6.3 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	3	280	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.58, 0.53]
2.6.4 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-1.16, 1.12]
2.6.5 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.48 [-1.53, 0.58]
2.6.6 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.46 [-1.71, 0.78]
2.6.7 Lower CAD/CAM nitinol vs rectangular braided multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.06 [-1.18, 1.30]
2.6.8 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.42 [-1.75, 0.92]
2.6.9 Lower CAD/CAM nitinol vs multistrand retainers at 12 months	2	107	Mean Difference (IV, Fixed, 95% CI)	0.26 [-0.61, 1.13]
2.7 Stability: overjet	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.7.1 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.43, 0.43]
2.8 Stability: overbite	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.8.1 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.17, 0.37]
2.9 Stability: maximum rotation	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.9.1 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months	1	60	Mean Difference (IV, Fixed, 95% CI)	2.30 [0.80, 3.80]
2.10 Failure of retainers	20		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.10.1 Upper CAD/CAM nitinol vs multistrand retainers at 6 months	1	163	Risk Ratio (M-H, Fixed, 95% CI)	0.99 [0.57, 1.70]
2.10.2 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months	1	55	Risk Ratio (M-H, Fixed, 95% CI)	2.90 [1.20, 7.03]
2.10.3 Upper fibre-reinforced composite vs multistrand bonded retainer at 18+ months	2	174	Risk Ratio (M-H, Fixed, 95% CI)	1.25 [0.87, 1.78]
2.10.4 Upper conventional vs flowable composite fixed retainer at 24 months	1	600	Risk Ratio (M-H, Fixed, 95% CI)	0.50 [0.28, 0.89]
2.10.5 Upper conventional orthodontic primer vs universal primer fixed retainer at 24 months	1	600	Risk Ratio (M-H, Fixed, 95% CI)	2.23 [1.38, 3.59]
2.10.6 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	2	215	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.67, 1.50]
2.10.7 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months	1	55	Risk Ratio (M-H, Fixed, 95% CI)	1.25 [0.57, 2.77]
2.10.8 Lower ceramic interlocking bonded vs coaxial wire retainer at 6 months	1	40	Risk Ratio (M-H, Fixed, 95% CI)	2.00 [0.41, 9.71]
2.10.9 Lower CAD-CAM stainless steel multistrand vs lab stainless steel multistrand at 6 months	1	32	Risk Ratio (M-H, Fixed, 95% CI)	0.57 [0.21, 1.58]
2.10.10 Lower CAD-CAM stainless steel multistrand vs chairside rectangular chain at 6 months	1	30	Risk Ratio (M-H, Fixed, 95% CI)	1.75 [0.38, 8.15]
2.10.11 Lower lab stainless steel multistrand vs chairside rectangular chain 6 months	1	30	Risk Ratio (M-H, Fixed, 95% CI)	3.06 [0.76, 12.40]
2.10.12 Lower fixed direct coaxial vs indirect multistrand retainer at 12 months	1	106	Risk Ratio (M-H, Fixed, 95% CI)	1.43 [0.59, 3.47]
2.10.13 Lower CAD/CAM nitinol vs multistrand retainers at 12 months	1	41	Risk Ratio (M-H, Fixed, 95% CI)	1.29 [0.67, 2.49]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.10.14 Lower fibre-reinforced composite vs multistrand bonded retainer at 12+ months	7	1337	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.84, 1.21]
2.10.15 Lower fibre-reinforced composite vs twisted ligature wire at 24 months	1	87	Risk Ratio (M-H, Fixed, 95% CI)	2.01 [0.95, 4.24]
2.10.16 Lower spiral wire vs twisted ligature wire at 24 months	1	86	Risk Ratio (M-H, Fixed, 95% CI)	1.51 [0.67, 3.38]
2.10.17 Lower round twisted multistrand versus rectangular braided multistrand at 24 months	1	132	Risk Ratio (M-H, Fixed, 95% CI)	1.16 [0.83, 1.60]
2.10.18 Lower conventional vs flowable composite fixed retainer at 24 months	1	600	Risk Ratio (M-H, Fixed, 95% CI)	0.52 [0.32, 0.87]
2.10.19 Lower conventional orthodontic primer vs universal primer fixed retainer at 24 months	1	600	Risk Ratio (M-H, Fixed, 95% CI)	5.11 [3.17, 8.25]
2.10.20 Lower chemically-cured vs light-cured adhesive at 24 months	1	220	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.64, 1.14]
2.10.21 Lower thick plain wire (canine to canine) vs thin spiral wire (incisors and canines) at 24 months	1	250	Risk Ratio (M-H, Fixed, 95% CI)	0.49 [0.25, 0.98]
2.10.22 Lower thick plain wire (canine to canine) vs thick spiral wire (canine to canine) at 36 months	1	24	Risk Ratio (M-H, Fixed, 95% CI)	0.30 [0.04, 2.27]
2.10.23 Lower thick plain wire (incisors and canines) vs thin spiral wire (incisors and canines) at 36 months	1	24	Risk Ratio (M-H, Fixed, 95% CI)	1.13 [0.32, 3.99]
2.10.24 Lower thick plain wire (canine to canine) vs thin spiral wire (incisors and canines) at 36 months	1	22	Risk Ratio (M-H, Fixed, 95% CI)	0.33 [0.04, 2.73]
2.11 Mean survival time (months)	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.11.1 Lower fixed thick plain titanium wire vs twisted wire retainer at 24 months	1	250	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.50, 0.60]
2.11.2 Lower multistrand spiral wire vs twisted ligature wire at 24 months	1	86	Mean Difference (IV, Fixed, 95% CI)	-0.99 [-1.33, -0.65]
2.11.3 Lower fibre-reinforced composite vs twisted ligature wire at 24 months	1	87	Mean Difference (IV, Fixed, 95% CI)	-2.40 [-2.77, -2.04]
2.11.4 Lower fibre-reinforced composite vs multistrand bonded retainer at 24 months	2	103	Mean Difference (IV, Fixed, 95% CI)	-1.48 [-1.88, -1.08]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.12 Failure risk (Hazard Ratio)	9		Hazard Ratio (IV, Fixed, 95% CI)	Subtotals only
2.12.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months	1	55	Hazard Ratio (IV, Fixed, 95% CI)	2.82 [1.00, 7.97]
2.12.2 Upper multistrand vs CAD/CAM nitinol retainers at 6 months	1	163	Hazard Ratio (IV, Fixed, 95% CI)	0.99 [0.52, 1.89]
2.12.3 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months	1	55	Hazard Ratio (IV, Fixed, 95% CI)	1.51 [0.57, 4.00]
2.12.4 Lower multistrand vs CAD/CAM nitinol retainers at 6 months	1	163	Hazard Ratio (IV, Fixed, 95% CI)	0.93 [0.53, 1.63]
2.12.5 Lower multistrand spiral wire vs fibre-reinforced composite at 12 months	1	87	Hazard Ratio (IV, Fixed, 95% CI)	1.64 [1.04, 2.59]
2.12.6 Lower fibre-reinforced composite vs multistrand bonded retainer at 12 months	1	32	Hazard Ratio (IV, Fixed, 95% CI)	0.77 [0.31, 1.92]
2.12.7 Lower multistrand spiral wire vs twisted ligature wire at 24 months	1	86	Hazard Ratio (IV, Fixed, 95% CI)	1.59 [0.64, 3.96]
2.12.8 Lower fibre-reinforced composite vs twisted ligature wire at 24 months	1	87	Hazard Ratio (IV, Fixed, 95% CI)	2.30 [0.98, 5.44]
2.12.9 Lower round twisted multistrand vs rectangular braided multistrand at 24 months	1	132	Hazard Ratio (IV, Fixed, 95% CI)	0.69 [0.42, 1.13]
2.12.10 Lower light-cured adhesive vs chemically-cured at 24 months	1	220	Hazard Ratio (IV, Fixed, 95% CI)	1.16 [0.78, 1.72]
2.12.11 Upper & lower conventional vs flowable composite fixed retainer at 24 months	1	100	Hazard Ratio (IV, Fixed, 95% CI)	0.50 [0.34, 0.73]
2.12.12 Upper & lower conventional orthodontic primer vs universal primer fixed retainer at 24 months	1	100	Hazard Ratio (IV, Fixed, 95% CI)	3.77 [2.78, 5.11]
2.13 Patient satisfaction: efficiency	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.13.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically,"Did your retainer keep your teeth straight?")	1	62	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.91, 1.10]
2.13.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer keep your teeth straight?")	1	62	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.91, 1.10]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.14 Patient satisfaction: cleaning/hygiene	2		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
2.14.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy to look after?")	1	62	Risk Difference (M-H, Fixed, 95% CI)	0.00 [-0.06, 0.06]
2.14.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy to look after?")	1	62	Risk Difference (M-H, Fixed, 95% CI)	0.03 [-0.05, 0.12]
2.14.3 Lower rigid retainer bonded only 33-43 vs rectangular chain retainer at 6 months (specifically, "Which retainer allowed better cleaning?")	1	30	Risk Difference (M-H, Fixed, 95% CI)	1.00 [0.88, 1.12]
2.15 Patient satisfaction: comfort	2		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
2.15.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy comfortable?")	1	62	Risk Difference (M-H, Fixed, 95% CI)	-0.06 [-0.17, 0.04]
2.15.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy comfortable?")	1	62	Risk Difference (M-H, Fixed, 95% CI)	0.00 [-0.06, 0.06]
2.15.3 Lower rigid retainer bonded only 33-43 vs rectangular chain retainer at 6 months (specifically, "Which retainer was more comfortable in use?")	1	30	Risk Difference (M-H, Fixed, 95% CI)	0.20 [-0.15, 0.55]
2.16 Patient satisfaction: speech	1		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
2.16.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer affect speech?")	1	62	Risk Difference (M-H, Fixed, 95% CI)	-0.03 [-0.12, 0.05]
2.16.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer affect speech?")	1	62	Risk Difference (M-H, Fixed, 95% CI)	0.00 [-0.06, 0.06]
2.17 Patient satisfaction: appointment needed	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.17.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer cause a problem that meant you needed to see your orthodontics?")	1	62	Risk Ratio (M-H, Fixed, 95% CI)	2.67 [0.78, 9.12]
2.17.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer cause a problem that	1	62	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.27, 3.65]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
meant you needed to see your orthodontics?")				
2.18 Patient satisfaction: aesthetics	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.18.1 Lower fibre-reinforced composite vs multistrand bonded retainer at 12 months (specifically, "esthetic result while talking and smiling")	1	32	Mean Difference (IV, Fixed, 95% CI)	1.49 [0.76, 2.22]
2.19 Patient satisfaction: preference for type of retainer	1		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
2.19.1 Lower rigid retainer bonded only 33-43 vs rectangular chain retainer at 6 months (specifically, "preference for type of retainer")	1	30	Risk Difference (M-H, Fixed, 95% CI)	1.00 [0.88, 1.12]
2.20 Oral health: Plaque Index	7		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.20.1 Upper fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	1.00 [0.54, 1.46]
2.20.2 Lower fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	1.09 [0.70, 1.48]
2.20.3 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	4	198	Mean Difference (IV, Fixed, 95% CI)	-0.09 [-0.18, 0.01]
2.20.4 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.32 [-0.82, 0.18]
2.20.5 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	-0.08 [-0.57, 0.42]
2.20.6 Lower ceramic interlocking bonded vs coaxial wire retainer at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	-0.40 [-0.56, -0.24]
2.20.7 Lower CAD/CAM nitinol vs multistrand retainers at 12 months	2	107	Mean Difference (IV, Fixed, 95% CI)	-0.01 [-0.11, 0.10]
2.20.8 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.08, 0.08]
2.20.9 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.09 [-0.20, 0.01]
2.20.10 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.21, 0.02]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.20.11 Lower CAD/CAM nitinol vs rectangular braided multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.12, 0.08]
2.20.12 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.11 [-0.24, 0.02]
2.20.13 Lower round twisted multistrand vs rectangular braided multistrand at 24 months	1	132	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.07, 0.01]
2.21 Oral health: Gingival Index	7		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.21.1 Upper fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.17 [-0.29, 0.63]
2.21.2 Lower fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.59 [0.13, 1.05]
2.21.3 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	4	198	Mean Difference (IV, Fixed, 95% CI)	-0.14 [-0.27, -0.01]
2.21.4 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.50, 0.38]
2.21.5 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	0.18 [-0.27, 0.62]
2.21.6 Lower ceramic interlocking bonded vs coaxial wire retainer at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.10, 0.20]
2.21.7 Lower CAD/CAM nitinol vs multistrand retainers at 12 months	2	107	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.16, 0.16]
2.21.8 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.18, 0.11]
2.21.9 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.20, 0.12]
2.21.10 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.00 [-0.18, 0.18]
2.21.11 Lower CAD-CAM nitinol vs rectangular braided multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.19, 0.11]
2.21.12 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.08 [-0.27, 0.11]
2.21.13 Lower round twisted multistrand vs rectangular braided multistrand at 24 months	1	132	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.14, 0.14]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

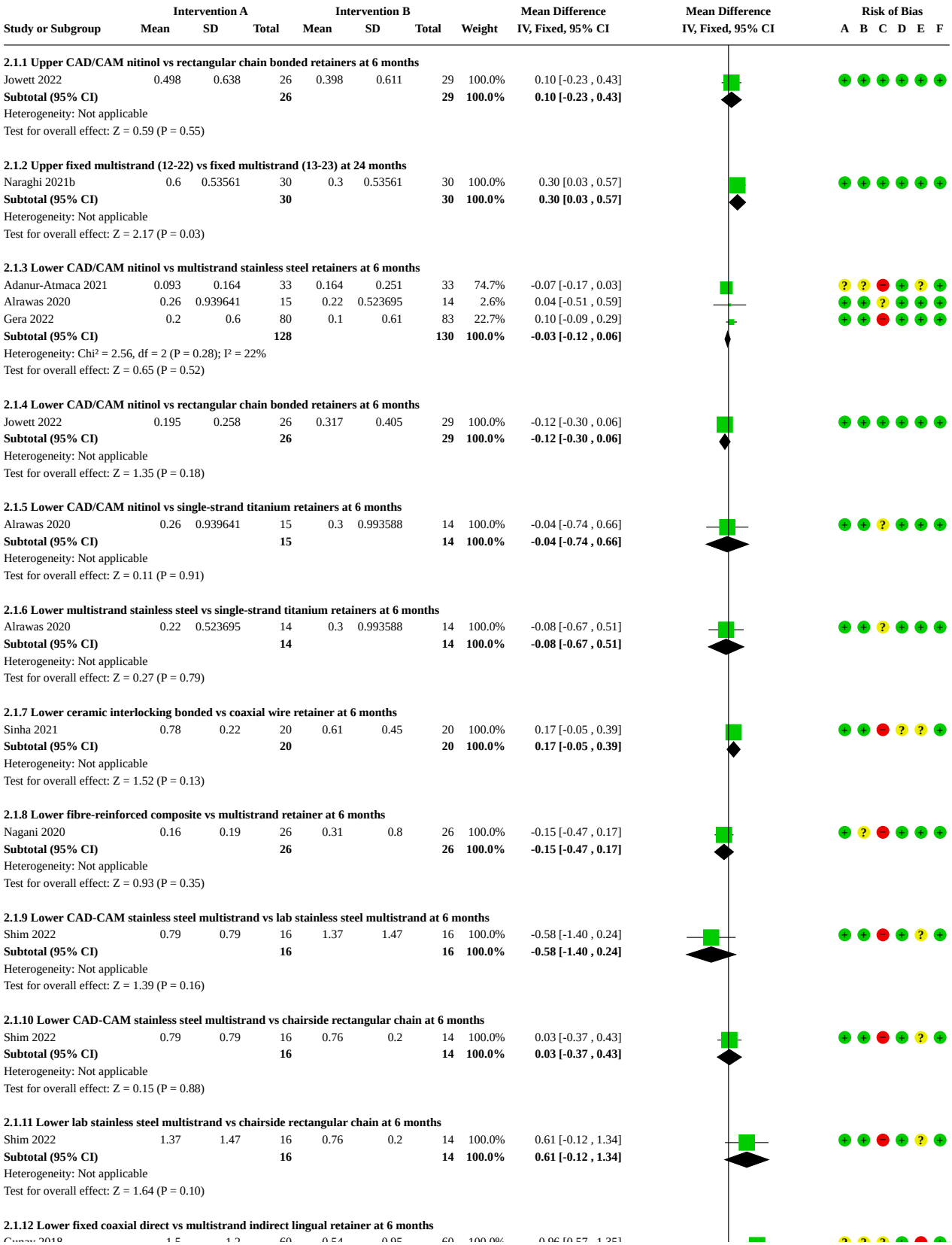
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.22 Oral health: Calculus Index	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.22.1 Upper fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.39, 0.39]
2.22.2 Lower fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.31, 0.31]
2.22.3 Lower CAD/CAM nitinol vs multistrand retainers at 12 months	2	107	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.11, 0.03]
2.22.4 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.09, 0.05]
2.22.5 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.06, 0.07]
2.22.6 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.06, 0.10]
2.22.7 Lower CAD/CAM nitinol vs rectangular braided multistrand at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.10, 0.02]
2.22.8 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months	1	66	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.10, 0.03]
2.23 Oral health: probing depth	3		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.23.1 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	2	81	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.14, 0.02]
2.23.2 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.18, 0.25]
2.23.3 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	0.03 [-0.19, 0.24]
2.23.4 Lower round twisted multistrand versus rectangular braided multistrand at 24 months	1	132	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.09, 0.09]
2.24 Oral health: bleeding on probing	4		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.24.1 Upper fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.46, 0.46]
2.24.2 Lower fibre-reinforced composite vs spiral wire retainers at 6 months	1	40	Mean Difference (IV, Fixed, 95% CI)	0.33 [-0.13, 0.79]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.24.3 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	2	81	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.02, 0.05]
2.24.4 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months	1	29	Mean Difference (IV, Fixed, 95% CI)	-0.15 [-0.29, -0.01]
2.24.5 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months	1	28	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.20, 0.12]
2.24.6 Lower round twisted multistrand vs rectangular braided multistrand at 24 months	1	132	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.06, 0.06]
2.25 Oral health: marginal recession	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.25.1 Lower CAD/CAM nitinol vs multistrand retainers at 6 months	2	103	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-0.11, 0.01]

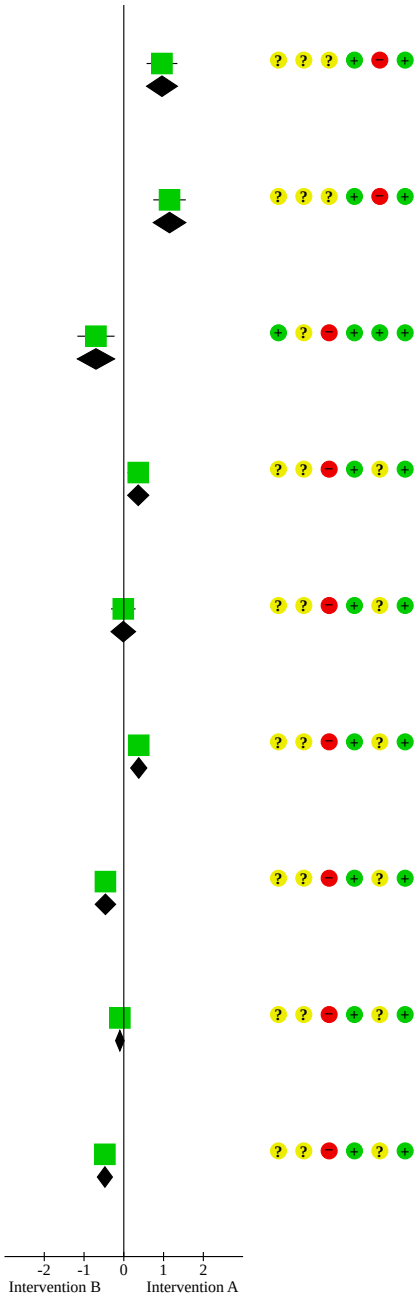
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.1. Comparison 2: Fixed versus fixed retainers, Outcome 1: Stability: Irregularity Index



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.1. (Continued)



2.1.12 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 6 months

Gunay 2018	1.5	1.2	60	0.54	0.95	60	100.0%	0.96 [0.57 , 1.35]
Subtotal (95% CI)			**60**			**60**	**100.0%**	**0.96 [0.57 , 1.35]**

Heterogeneity: Not applicable
Test for overall effect: Z = 4.86 (P < 0.00001)

2.1.13 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 12 months

Gunay 2018	1.97	1.3	60	0.82	0.96	60	100.0%	1.15 [0.74 , 1.56]
Subtotal (95% CI)			**60**			**60**	**100.0%**	**1.15 [0.74 , 1.56]**

Heterogeneity: Not applicable
Test for overall effect: Z = 5.51 (P < 0.00001)

2.1.14 Lower fibre-reinforced composite vs multistrand retainer at 12 months

Nagani 2020	0.27	0.25	26	0.97	1.19	26	100.0%	-0.70 [-1.17 , -0.23]
Subtotal (95% CI)			**26**			**26**	**100.0%**	**-0.70 [-1.17 , -0.23]**

Heterogeneity: Not applicable
Test for overall effect: Z = 2.94 (P = 0.003)

2.1.15 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months

Adanur-Atmaca 2021	0.555	0.732	33	0.192	0.267	33	100.0%	0.36 [0.10 , 0.63]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.36 [0.10 , 0.63]**

Heterogeneity: Not applicable
Test for overall effect: Z = 2.68 (P = 0.007)

2.1.16 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months

Adanur-Atmaca 2021	0.555	0.732	33	0.568	0.53	33	100.0%	-0.01 [-0.32 , 0.30]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.01 [-0.32 , 0.30]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.08 (P = 0.93)

2.1.17 Lower connected bonding pad retainer vs round stainless-steel multistrand at 12 months

Adanur-Atmaca 2021	0.568	0.53	33	0.192	0.267	33	100.0%	0.38 [0.17 , 0.58]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.38 [0.17 , 0.58]**

Heterogeneity: Not applicable
Test for overall effect: Z = 3.64 (P = 0.0003)

2.1.18 Lower CAD-CAM nitinol vs rectangular braided multistrand at 12 months

Adanur-Atmaca 2021	0.093	0.164	33	0.555	0.732	33	100.0%	-0.46 [-0.72 , -0.21]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.46 [-0.72 , -0.21]**

Heterogeneity: Not applicable
Test for overall effect: Z = 3.54 (P = 0.0004)

2.1.19 Lower CAD-CAM nitinol vs round stainless-steel multistrand at 12 months

Adanur-Atmaca 2021	0.093	0.164	33	0.192	0.267	33	100.0%	-0.10 [-0.21 , 0.01]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.10 [-0.21 , 0.01]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.81 (P = 0.07)

2.1.20 Lower CAD-CAM nitinol vs connected bonding pad retainer at 12 months

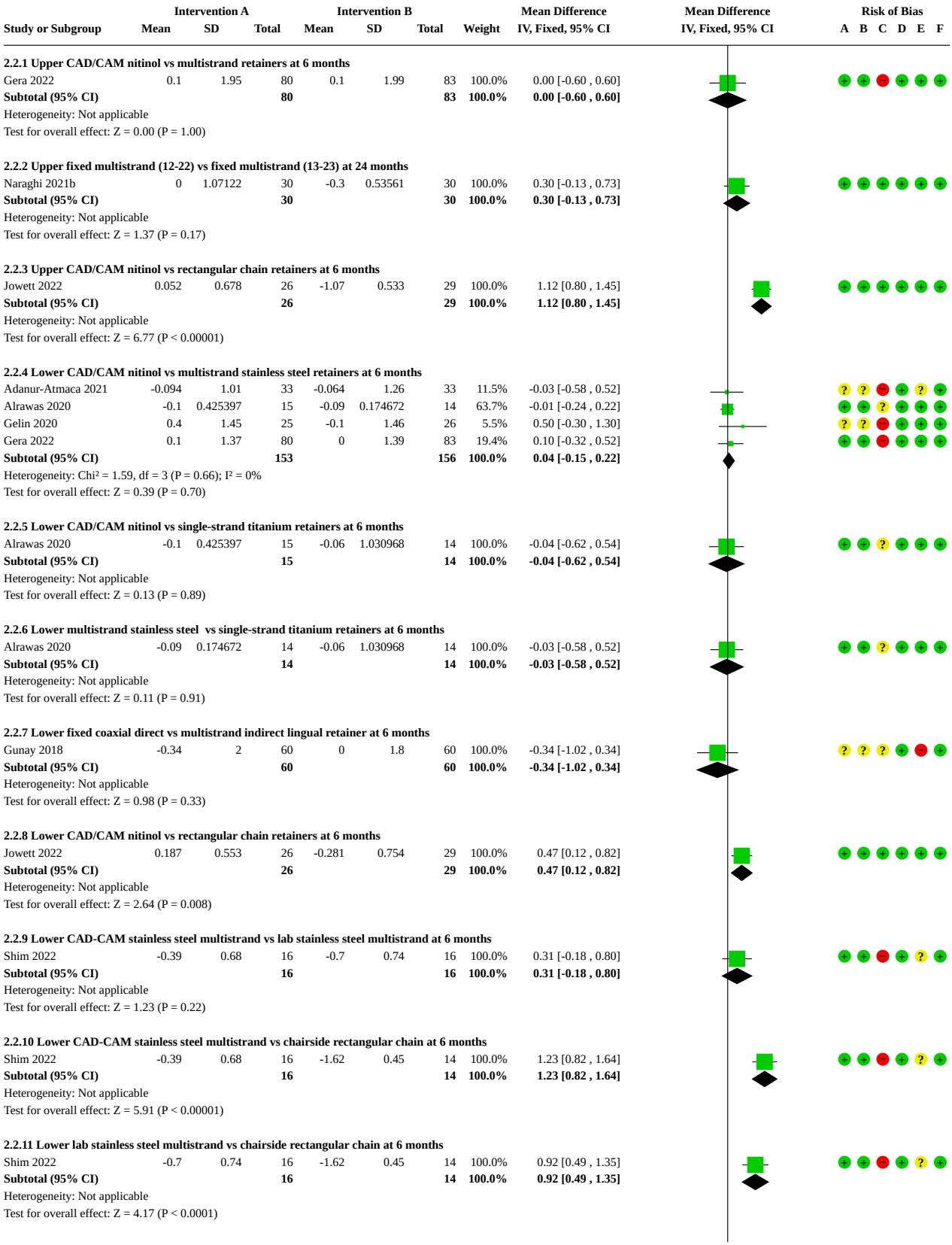
Adanur-Atmaca 2021	0.093	0.164	33	0.568	0.53	33	100.0%	-0.47 [-0.66 , -0.29]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.47 [-0.66 , -0.29]**

Heterogeneity: Not applicable
Test for overall effect: Z = 4.92 (P < 0.00001)

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.2. Comparison 2: Fixed versus fixed retainers, Outcome 2: Stability: intercanine width



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.2. (Continued)

Test for overall effect: Z = 4.17 (P < 0.0001)

2.2.12 Lower fixed coaxial direct vs multistrand indirect lingual retainer at 12 months

Gunay 2018	-0.51	2.15	60	0.02	1.8	60	100.0%	-0.53 [-1.24 , 0.18]
Subtotal (95% CI)			**60**			**60**	**100.0%**	**-0.53 [-1.24 , 0.18]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.46 (P = 0.14)

2.2.13 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months

Adanur-Atmaca 2021	-0.387	1.16	33	-0.07	1.25	33	100.0%	-0.32 [-0.90 , 0.26]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.32 [-0.90 , 0.26]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.07 (P = 0.29)

2.2.14 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months

Adanur-Atmaca 2021	-0.387	1.16	33	-0.485	1.26	33	100.0%	0.10 [-0.49 , 0.68]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.10 [-0.49 , 0.68]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.33 (P = 0.74)

2.2.15 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months

Adanur-Atmaca 2021	-0.07	1.25	33	-0.485	1.26	33	100.0%	0.41 [-0.19 , 1.02]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.41 [-0.19 , 1.02]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.34 (P = 0.18)

2.2.16 Lower CAD/CAM nitinol vs rectangular braided multistrand at 12 months

Adanur-Atmaca 2021	-0.094	1.01	33	-0.387	1.16	33	100.0%	0.29 [-0.23 , 0.82]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.29 [-0.23 , 0.82]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.09 (P = 0.27)

2.2.17 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months

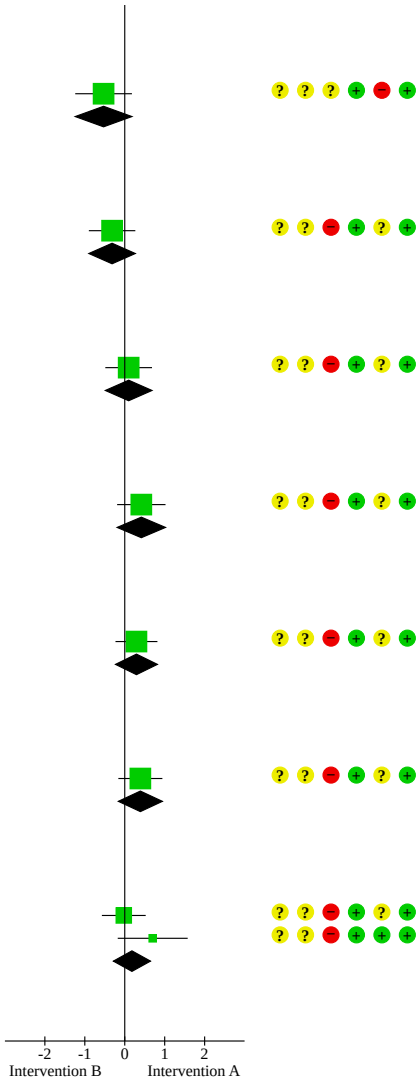
Adanur-Atmaca 2021	-0.094	1.01	33	-0.485	1.26	33	100.0%	0.39 [-0.16 , 0.94]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**0.39 [-0.16 , 0.94]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.39 (P = 0.16)

2.2.18 Lower CAD/CAM nitinol vs multistrand retainers at 12 months

Adanur-Atmaca 2021	-0.094	1.01	33	-0.07	1.25	33	71.8%	-0.02 [-0.57 , 0.52]
Gelin 2020	0.3	1.44	19	-0.4	1.41	22	28.2%	0.70 [-0.18 , 1.58]
Subtotal (95% CI)			**52**			**55**	**100.0%**	**0.18 [-0.28 , 0.64]**

Heterogeneity: Chi² = 1.89, df = 1 (P = 0.17); I² = 47%
Test for overall effect: Z = 0.76 (P = 0.45)

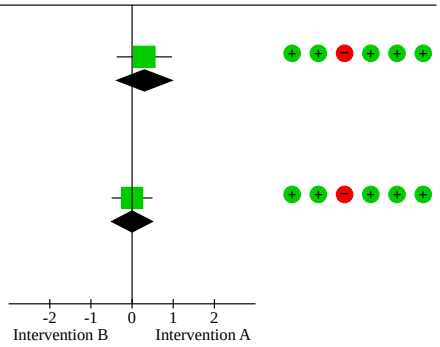


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.3. Comparison 2: Fixed versus fixed retainers, Outcome 3: Stability: interpremolar width

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.3.1 Upper CAD/CAM nitinol vs multistrand retainers at 6 months								
Gera 2022	0.1	2.17	80	-0.2	2.21	83	100.0%	0.30 [-0.37 , 0.97]
Subtotal (95% CI)			**80**			**83**	**100.0%**	**0.30 [-0.37 , 0.97]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.87 (P = 0.38)								
2.3.2 Lower CAD/CAM nitinol vs multistrand retainers at 6 months								
Gera 2022	0.1	1.6	80	0.1	1.63	83	100.0%	0.00 [-0.50 , 0.50]
Subtotal (95% CI)			**80**			**83**	**100.0%**	**0.00 [-0.50 , 0.50]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								

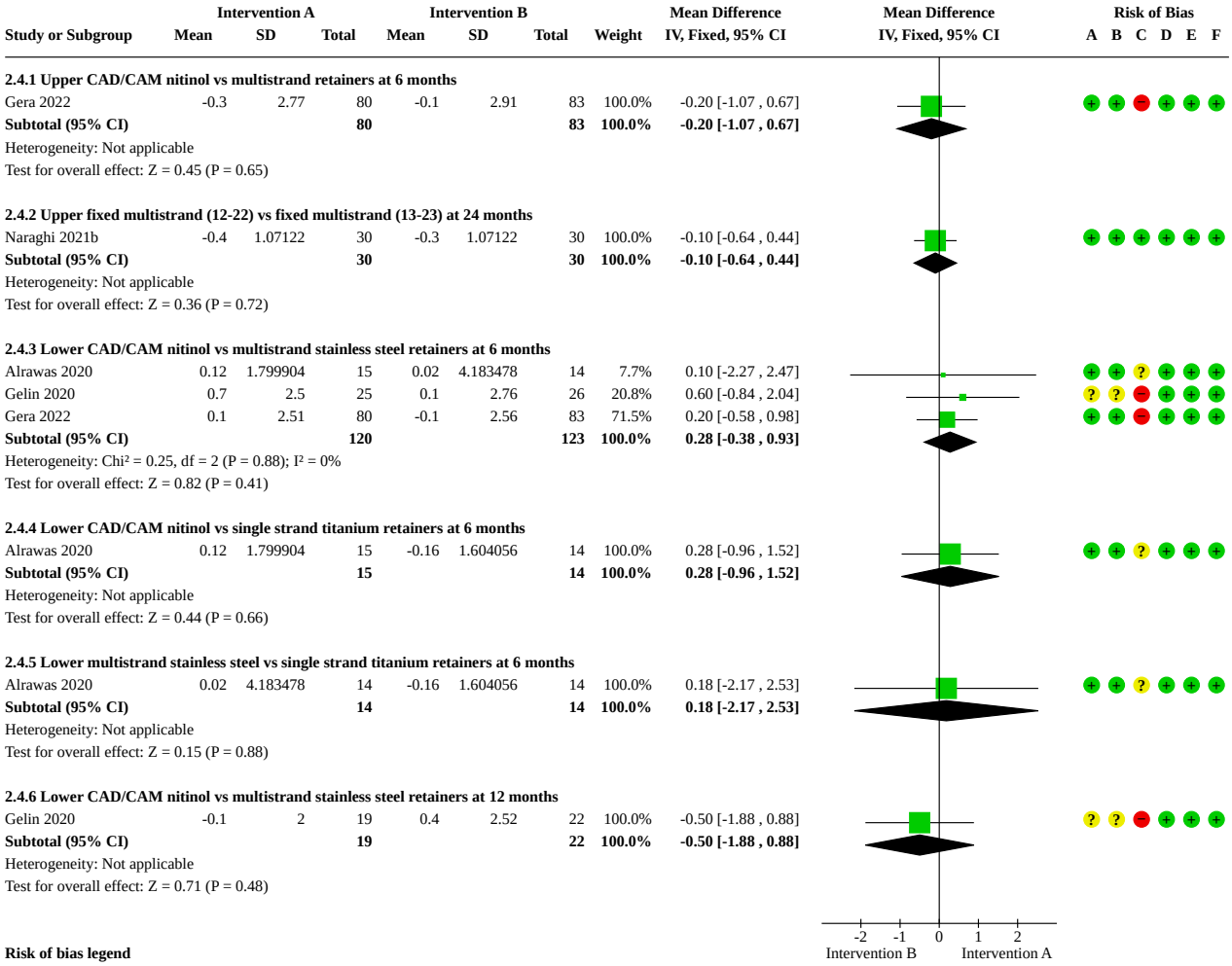


Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.4. Comparison 2: Fixed versus fixed retainers, Outcome 4: Stability: intermolar width



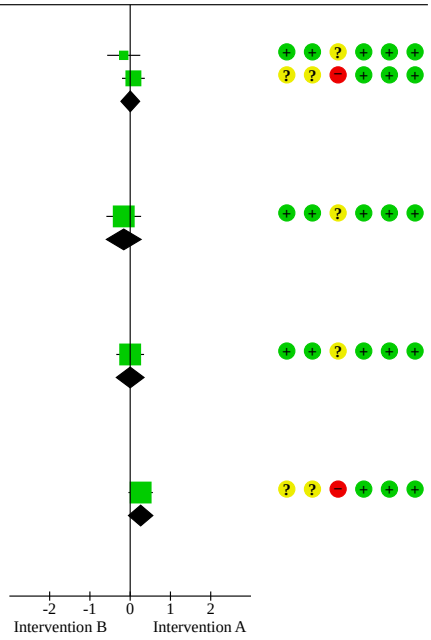
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.5. Comparison 2: Fixed versus fixed retainers, Outcome 5: Stability: anterior arch length

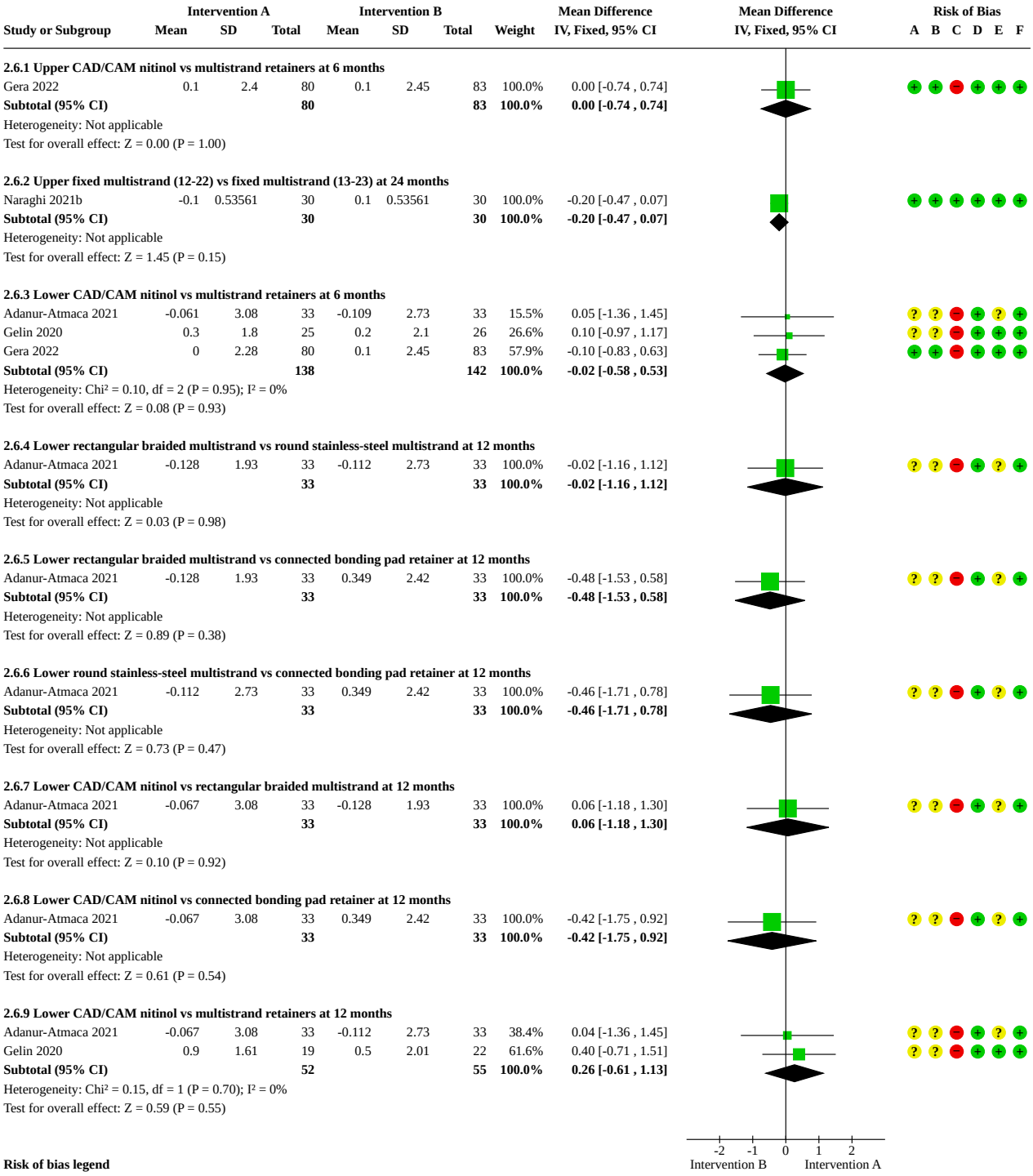
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.5.1 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 6 months								
Alrawas 2020	0.06	0.678311	15	0.22	0.430252	14	32.2%	-0.16 [-0.57 , 0.25]
Gelin 2020	0.02	0.53	25	-0.06	0.5	26	67.8%	0.08 [-0.20 , 0.36]
Subtotal (95% CI)			**40**			**40**	**100.0%**	**0.00 [-0.23 , 0.24]**
Heterogeneity: Chi² = 0.89, df = 1 (P = 0.35); I² = 0%								
Test for overall effect: Z = 0.02 (P = 0.98)								
2.5.2 Lower CAD/CAM nitinol vs single strand titanium retainers at 6 months								
Alrawas 2020	0.06	0.678311	15	0.22	0.497596	14	100.0%	-0.16 [-0.59 , 0.27]
Subtotal (95% CI)			**15**			**14**	**100.0%**	**-0.16 [-0.59 , 0.27]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.73 (P = 0.47)								
2.5.3 Lower multistrand stainless steel vs single strand titanium retainers at 6 months								
Alrawas 2020	0.22	0.430252	14	0.22	0.497596	14	100.0%	0.00 [-0.34 , 0.34]
Subtotal (95% CI)			**14**			**14**	**100.0%**	**0.00 [-0.34 , 0.34]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
2.5.4 Lower CAD/CAM nitinol vs multistrand stainless steel retainers at 12 months								
Gelin 2020	0.08	0.52	19	-0.18	0.46	22	100.0%	0.26 [-0.04 , 0.56]
Subtotal (95% CI)			**19**			**22**	**100.0%**	**0.26 [-0.04 , 0.56]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.68 (P = 0.09)								



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.6. Comparison 2: Fixed versus fixed retainers, Outcome 6: Stability: total arch length

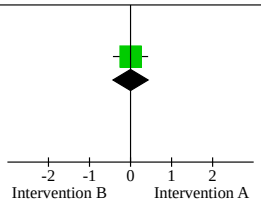


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.7. Comparison 2: Fixed versus fixed retainers, Outcome 7: Stability: overjet

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.7.1 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months								
Naraghi 2021b	0.1	1.07122	30	0.1	0.53561	30	100.0%	0.00 [-0.43 , 0.43]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**0.00 [-0.43 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								

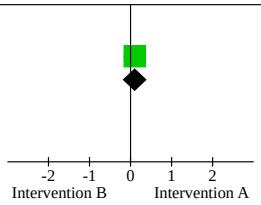




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.8. Comparison 2: Fixed versus fixed retainers, Outcome 8: Stability: overbite

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.8.1 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months								
Naraghi 2021b	0.3	0.53561	30	0.2	0.53561	30	100.0%	0.10 [-0.17 , 0.37]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**0.10 [-0.17 , 0.37]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.72 (P = 0.47)								

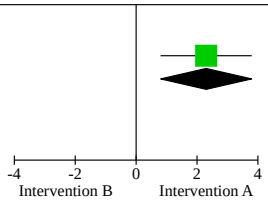




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.9. Comparison 2: Fixed versus fixed retainers, Outcome 9: Stability: maximum rotation

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.9.1 Upper fixed multistrand (12-22) vs fixed multistrand (13-23) at 24 months								
Naraghi 2021b	5.2	2.678049	30	2.9	3.213659	30	100.0%	2.30 [0.80 , 3.80]
Subtotal (95% CI)			**30**			**30**	**100.0%**	**2.30 [0.80 , 3.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.01 (P = 0.003)								

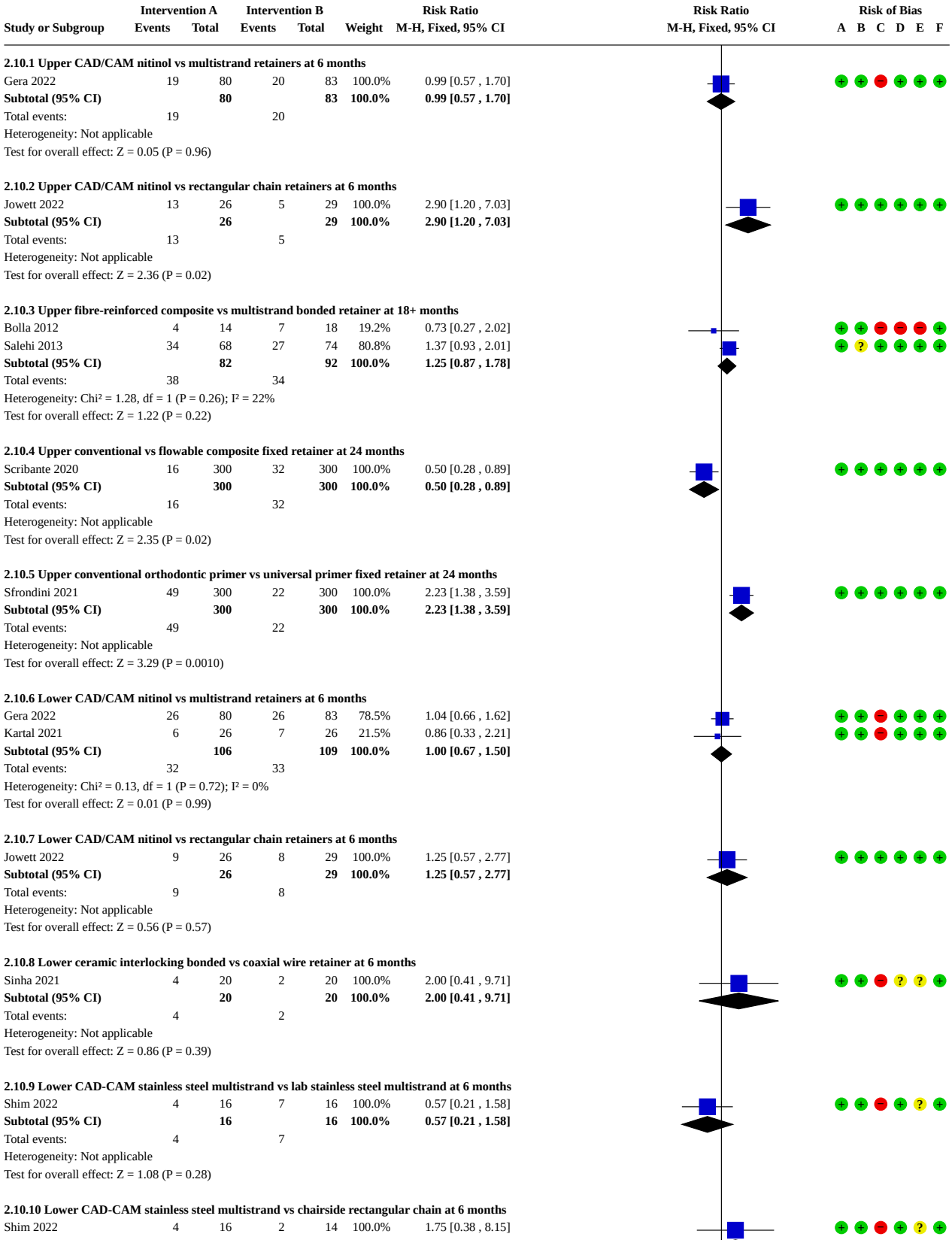




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.10. Comparison 2: Fixed versus fixed retainers, Outcome 10: Failure of retainers



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.10. (Continued)

2.10.10 Lower CAD-CAM stainless steel multistrand vs chairside rectangular chain at 6 months

Shim 2022	4	16	2	14	100.0%	1.75 [0.38 , 8.15]
Subtotal (95% CI)		**16**		**14**	**100.0%**	**1.75 [0.38 , 8.15]**
Total events:	4		2			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.71 (P = 0.48)

2.10.11 Lower lab stainless steel multistrand vs chairside rectangular chain 6 months

Shim 2022	7	16	2	14	100.0%	3.06 [0.76 , 12.40]
Subtotal (95% CI)		**16**		**14**	**100.0%**	**3.06 [0.76 , 12.40]**
Total events:	7		2			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.57 (P = 0.12)

2.10.12 Lower fixed direct coaxial vs indirect multistrand retainer at 12 months

Gunay 2018	10	53	7	53	100.0%	1.43 [0.59 , 3.47]
Subtotal (95% CI)		**53**		**53**	**100.0%**	**1.43 [0.59 , 3.47]**
Total events:	10		7			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.79 (P = 0.43)

2.10.13 Lower CAD/CAM nitinol vs multistrand retainers at 12 months

Gelin 2020	10	19	9	22	100.0%	1.29 [0.67 , 2.49]
Subtotal (95% CI)		**19**		**22**	**100.0%**	**1.29 [0.67 , 2.49]**
Total events:	10		9			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.75 (P = 0.45)

2.10.14 Lower fibre-reinforced composite vs multistrand bonded retainer at 12+ months

Bolla 2012	7	34	10	32	6.2%	0.66 [0.29 , 1.52]
Nagani 2020	67	156	49	156	29.6%	1.37 [1.02 , 1.83]
Rose 2002	5	10	1	10	0.6%	5.00 [0.70 , 35.50]
Salehi 2013	29	68	28	74	16.2%	1.13 [0.75 , 1.68]
Scribante 2011	13	90	23	102	13.0%	0.64 [0.35 , 1.19]
Sfrondini 2014	27	240	50	282	27.7%	0.63 [0.41 , 0.98]
Soboutia 2016	15	42	11	41	6.7%	1.33 [0.70 , 2.55]
Subtotal (95% CI)		**640**		**697**	**100.0%**	**1.01 [0.84 , 1.21]**
Total events:	163		172			

Heterogeneity: Chi² = 15.12, df = 6 (P = 0.02); I² = 60%
Test for overall effect: Z = 0.07 (P = 0.95)

2.10.15 Lower fibre-reinforced composite vs twisted ligature wire at 24 months

Soboutia 2016	15	42	8	45	100.0%	2.01 [0.95 , 4.24]
Subtotal (95% CI)		**42**		**45**	**100.0%**	**2.01 [0.95 , 4.24]**
Total events:	15		8			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.83 (P = 0.07)

2.10.16 Lower spiral wire vs twisted ligature wire at 24 months

Soboutia 2016	11	41	8	45	100.0%	1.51 [0.67 , 3.38]
Subtotal (95% CI)		**41**		**45**	**100.0%**	**1.51 [0.67 , 3.38]**
Total events:	11		8			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.00 (P = 0.32)

2.10.17 Lower round twisted multistrand versus rectangular braided multistrand at 24 months

Węgrodzka 2021	37	66	32	66	100.0%	1.16 [0.83 , 1.60]
Subtotal (95% CI)		**66**		**66**	**100.0%**	**1.16 [0.83 , 1.60]**
Total events:	37		32			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.87 (P = 0.39)

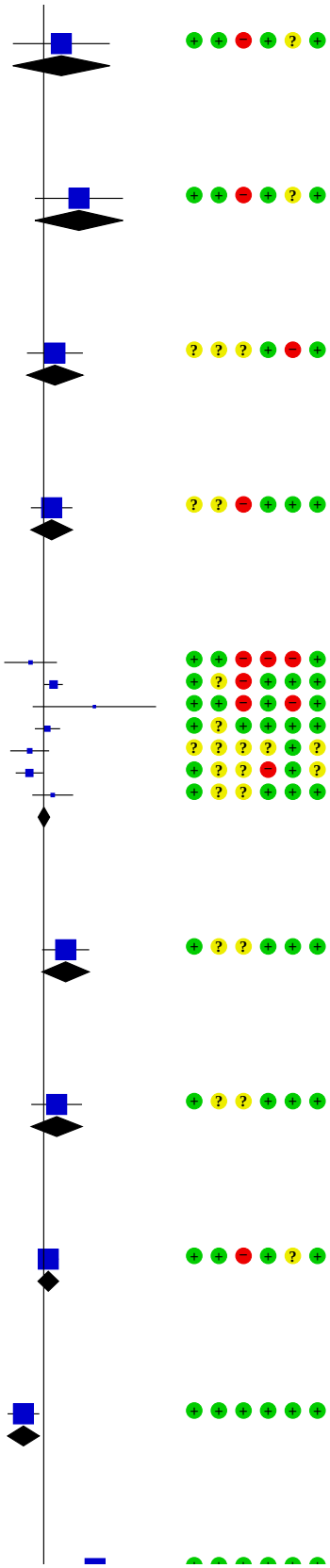
2.10.18 Lower conventional vs flowable composite fixed retainer at 24 months

Scribante 2020	21	300	40	300	100.0%	0.53 [0.32 , 0.87]
Subtotal (95% CI)		**300**		**300**	**100.0%**	**0.53 [0.32 , 0.87]**
Total events:	21		40			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.51 (P = 0.01)

2.10.19 Lower conventional orthodontic primer vs universal primer fixed retainer at 24 months

Sfrondini 2021 92 300 18 300 100.0% 5.11 [3.17 , 8.25]



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.10. (Continued)

2.10.19 Lower conventional orthodontic primer vs universal primer fixed retainer at 24 months

Sfrondini 2021	92	300	18	300	100.0%	5.11 [3.17 , 8.25]
Subtotal (95% CI)		**300**		**300**	**100.0%**	**5.11 [3.17 , 8.25]**
Total events:	92		18			

Heterogeneity: Not applicable
Test for overall effect: Z = 6.67 (P < 0.00001)

2.10.20 Lower chemically-cured vs light-cured adhesive at 24 months

Pandis 2013	47	110	55	110	100.0%	0.85 [0.64 , 1.14]
Subtotal (95% CI)		**110**		**110**	**100.0%**	**0.85 [0.64 , 1.14]**
Total events:	47		55			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.08 (P = 0.28)

2.10.21 Lower thick plain wire (canine to canine) vs thin spiral wire (incisors and canines) at 24 months

Arash 2020	10	112	25	138	100.0%	0.49 [0.25 , 0.98]
Subtotal (95% CI)		**112**		**138**	**100.0%**	**0.49 [0.25 , 0.98]**
Total events:	10		25			

Heterogeneity: Not applicable
Test for overall effect: Z = 2.01 (P = 0.04)

2.10.22 Lower thick plain wire (canine to canine) vs thick spiral wire (canine to canine) at 36 months

Årtun 1987	1	11	4	13	100.0%	0.30 [0.04 , 2.27]
Subtotal (95% CI)		**11**		**13**	**100.0%**	**0.30 [0.04 , 2.27]**
Total events:	1		4			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.17 (P = 0.24)

2.10.23 Lower thick plain wire (incisors and canines) vs thin spiral wire (incisors and canines) at 36 months

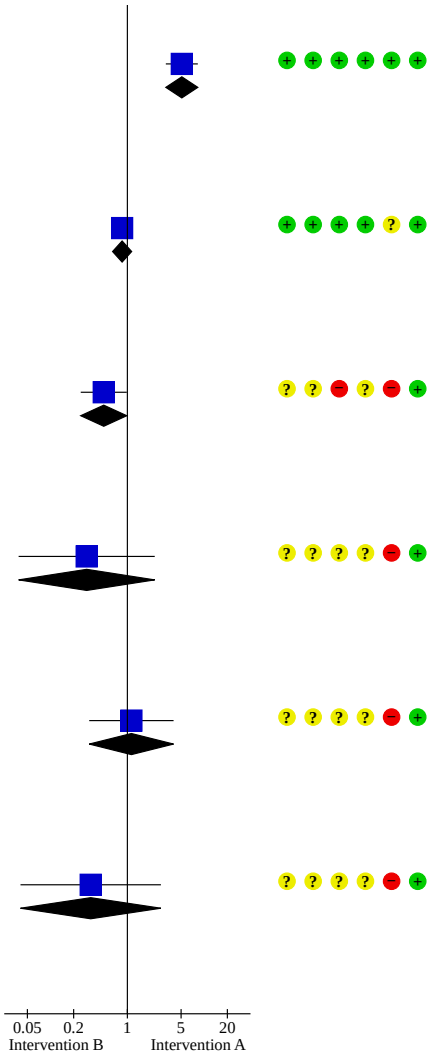
Årtun 1987	4	13	3	11	100.0%	1.13 [0.32 , 3.99]
Subtotal (95% CI)		**13**		**11**	**100.0%**	**1.13 [0.32 , 3.99]**
Total events:	4		3			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.19 (P = 0.85)

2.10.24 Lower thick plain wire (canine to canine) vs thin spiral wire (incisors and canines) at 36 months

Årtun 1987	1	11	3	11	100.0%	0.33 [0.04 , 2.73]
Subtotal (95% CI)		**11**		**11**	**100.0%**	**0.33 [0.04 , 2.73]**
Total events:	1		3			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.02 (P = 0.31)

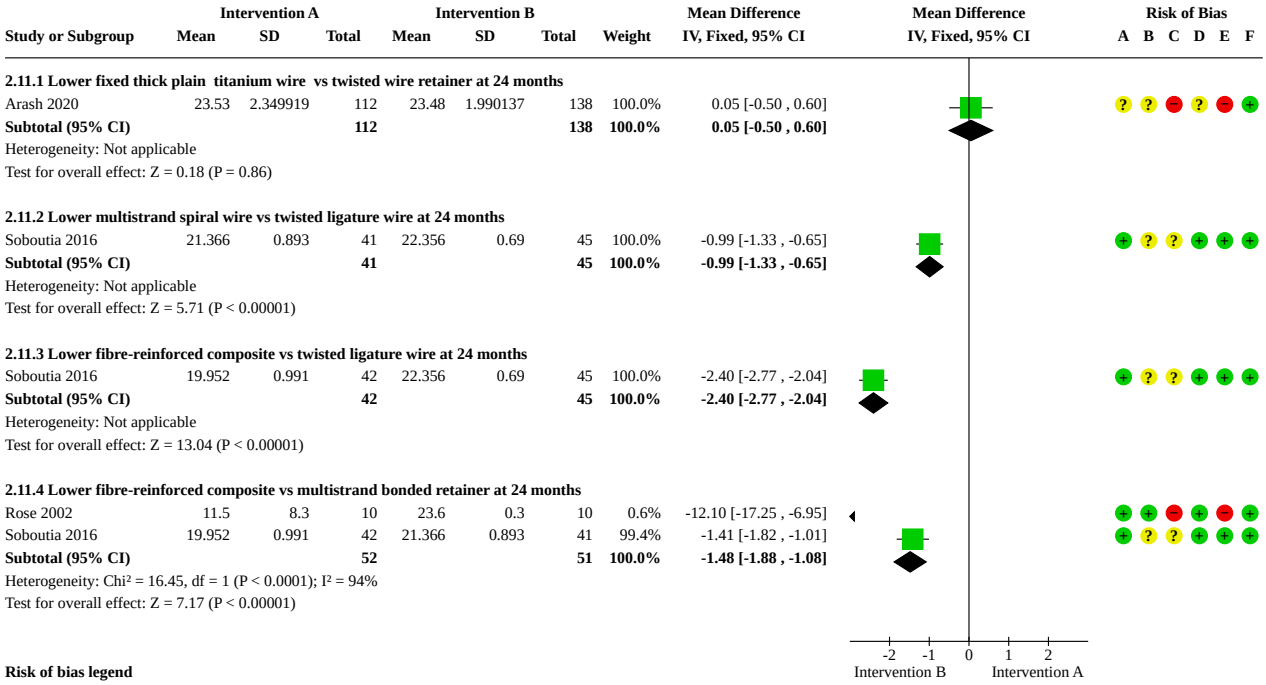


Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.11. Comparison 2: Fixed versus fixed retainers, Outcome 11: Mean survival time (months)



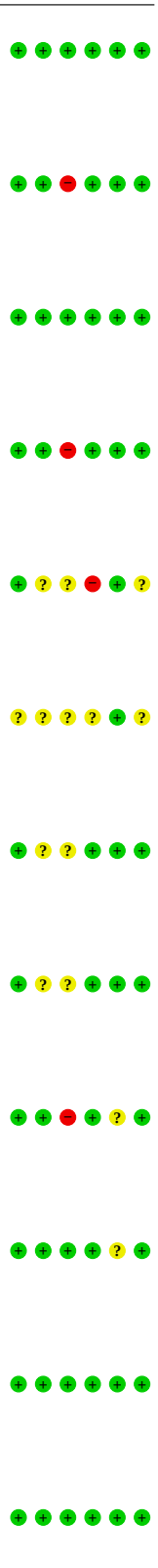
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.12. Comparison 2: Fixed versus fixed retainers, Outcome 12: Failure risk (Hazard Ratio)

Study or Subgroup	log[Hazard Ratio]	SE	Intervention A Total	Intervention B Total	Weight	Hazard Ratio IV, Fixed, 95% CI
2.12.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months						
Jowett 2022	1.036737	0.53016	26	29	100.0%	2.82 [1.00 , 7.97]
Subtotal (95% CI)			**26**	**29**	**100.0%**	**2.82 [1.00 , 7.97]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.96 (P = 0.05)						
2.12.2 Upper multistrand vs CAD/CAM nitinol retainers at 6 months						
Gera 2022	-0.01005	0.330562	83	80	100.0%	0.99 [0.52 , 1.89]
Subtotal (95% CI)			**83**	**80**	**100.0%**	**0.99 [0.52 , 1.89]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.03 (P = 0.98)						
2.12.3 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months						
Jowett 2022	0.41211	0.496415	26	29	100.0%	1.51 [0.57 , 4.00]
Subtotal (95% CI)			**26**	**29**	**100.0%**	**1.51 [0.57 , 4.00]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.83 (P = 0.41)						
2.12.4 Lower multistrand vs CAD/CAM nitinol retainers at 6 months						
Gera 2022	-0.072571	0.286602	83	80	100.0%	0.93 [0.53 , 1.63]
Subtotal (95% CI)			**83**	**80**	**100.0%**	**0.93 [0.53 , 1.63]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.25 (P = 0.80)						
2.12.5 Lower multistrand spiral wire vs fibre-reinforced composite at 12 months						
Sfrondini 2014	0.494696	0.23275	40	47	100.0%	1.64 [1.04 , 2.59]
Subtotal (95% CI)			**40**	**47**	**100.0%**	**1.64 [1.04 , 2.59]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.13 (P = 0.03)						
2.12.6 Lower fibre-reinforced composite vs multistrand bonded retainer at 12 months						
Scribante 2011	-0.261365	0.466514	17	15	100.0%	0.77 [0.31 , 1.92]
Subtotal (95% CI)			**17**	**15**	**100.0%**	**0.77 [0.31 , 1.92]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.56 (P = 0.58)						
2.12.7 Lower multistrand spiral wire vs twisted ligature wire at 24 months						
Soboutia 2016	0.464363	0.464747	41	45	100.0%	1.59 [0.64 , 3.96]
Subtotal (95% CI)			**41**	**45**	**100.0%**	**1.59 [0.64 , 3.96]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.00 (P = 0.32)						
2.12.8 Lower fibre-reinforced composite vs twisted ligature wire at 24 months						
Soboutia 2016	0.834647	0.43789	42	45	100.0%	2.30 [0.98 , 5.44]
Subtotal (95% CI)			**42**	**45**	**100.0%**	**2.30 [0.98 , 5.44]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.91 (P = 0.06)						
2.12.9 Lower round twisted multistrand vs rectangular braided multistrand at 24 months						
Węgrodzka 2021	-0.371064	0.250216	66	66	100.0%	0.69 [0.42 , 1.13]
Subtotal (95% CI)			**66**	**66**	**100.0%**	**0.69 [0.42 , 1.13]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.48 (P = 0.14)						
2.12.10 Lower light-cured adhesive vs chemically-cured at 24 months						
Pandis 2013	0.14842	0.201735	110	110	100.0%	1.16 [0.78 , 1.72]
Subtotal (95% CI)			**110**	**110**	**100.0%**	**1.16 [0.78 , 1.72]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.74 (P = 0.46)						
2.12.11 Upper & lower conventional vs flowable composite fixed retainer at 24 months						
Scribante 2020	-0.693147	0.191408	50	50	100.0%	0.50 [0.34 , 0.73]
Subtotal (95% CI)			**50**	**50**	**100.0%**	**0.50 [0.34 , 0.73]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.62 (P = 0.0003)						
2.12.12 Upper & lower conventional orthodontic primer vs universal primer fixed retainer at 24 months						
Sfrondini 2021	1.327075	0.155194	50	50	100.0%	3.77 [2.78 , 5.11]
Subtotal (95% CI)			**50**	**50**	**100.0%**	**3.77 [2.78 , 5.11]**
Heterogeneity: Not applicable						



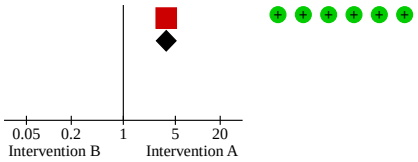


Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.12. (Continued)

Sfrondini 2021	1.327075	0.155194	50	50	100.0%	3.77 [2.78 , 5.11]	
Subtotal (95% CI)			**50**	**50**	**100.0%**	**3.77 [2.78 , 5.11]**	

Heterogeneity: Not applicable
Test for overall effect: Z = 8.55 (P < 0.00001)

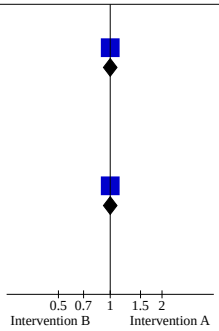


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.13. Comparison 2: Fixed versus fixed retainers, Outcome 13: Patient satisfaction: efficiency

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.13.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically,"Did your retainer keep your teeth straight?")						
Jowett 2022	30	31	30	31	100.0%	1.00 [0.91 , 1.10]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**1.00 [0.91 , 1.10]**
Total events:	30		30			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						
2.13.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specificially, "Did your retainer keep your teeth straight?")						
Jowett 2022	30	31	30	31	100.0%	1.00 [0.91 , 1.10]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**1.00 [0.91 , 1.10]**
Total events:	30		30			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						

Risk Ratio M-H, Fixed, 95% CI



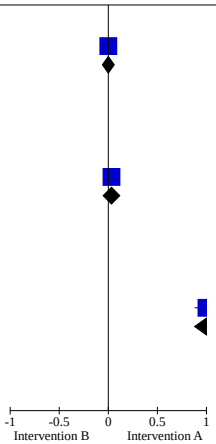
Risk of Bias A B C D E F

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.14. Comparison 2: Fixed versus fixed retainers, Outcome 14: Patient satisfaction: cleaning/hygiene

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.14.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy to look after?")						
Jowett 2022	31	31	31	31	100.0%	0.00 [-0.06 , 0.06]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**0.00 [-0.06 , 0.06]**
Total events:	31		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						
2.14.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy to look after?")						
Jowett 2022	31	31	30	31	100.0%	0.03 [-0.05 , 0.12]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**0.03 [-0.05 , 0.12]**
Total events:	31		30			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.74 (P = 0.46)						
2.14.3 Lower rigid retainer bonded only 33-43 vs rectangular chain retainer at 6 months (specifically, "Which retainer allowed better cleaning?")						
Ferreira 2019	15	15	0	15	100.0%	1.00 [0.88 , 1.12]
Subtotal (95% CI)		**15**		**15**	**100.0%**	**1.00 [0.88 , 1.12]**
Total events:	15		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 16.26 (P < 0.00001)						

Risk Difference M-H, Fixed, 95% CI



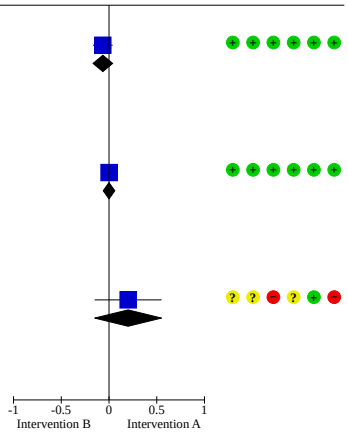
Risk of Bias A B C D E F

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.15. Comparison 2: Fixed versus fixed retainers, Outcome 15: Patient satisfaction: comfort

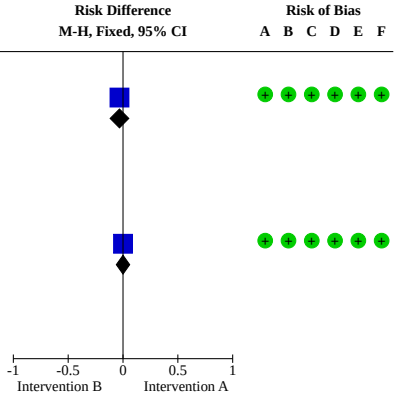
Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.15.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy comfortable?")						
Jowett 2022	29	31	31	31	100.0%	-0.06 [-0.17 , 0.04]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**-0.06 [-0.17 , 0.04]**
Total events:	29		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.23 (P = 0.22)						
2.15.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Was your retainer easy comfortable?")						
Jowett 2022	31	31	31	31	100.0%	0.00 [-0.06 , 0.06]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**0.00 [-0.06 , 0.06]**
Total events:	31		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						
2.15.3 Lower rigid retainer bonded only 33-43 vs rectangular chain retainer at 6 months (specifically, "Which retainer was more comfortable in use?")						
Ferreira 2019	9	15	6	15	100.0%	0.20 [-0.15 , 0.55]
Subtotal (95% CI)		**15**		**15**	**100.0%**	**0.20 [-0.15 , 0.55]**
Total events:	9		6			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.12 (P = 0.26)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.16. Comparison 2: Fixed versus fixed retainers, Outcome 16: Patient satisfaction: speech

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.16.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer affect speech?")						
Jowett 2022	30	31	31	31	100.0%	-0.03 [-0.12 , 0.05]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**-0.03 [-0.12 , 0.05]**
Total events:	30		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.74 (P = 0.46)						
2.16.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer affect speech?")						
Jowett 2022	31	31	31	31	100.0%	0.00 [-0.06 , 0.06]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**0.00 [-0.06 , 0.06]**
Total events:	31		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						

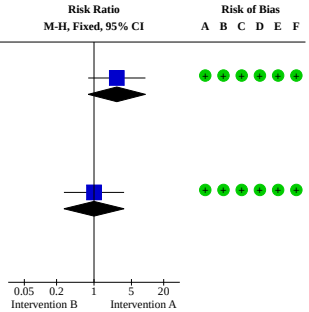


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.17. Comparison 2: Fixed versus fixed retainers, Outcome 17: Patient satisfaction: appointment needed

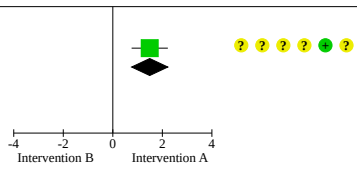
Study or Subgroup	Intervention A Events	Total	Intervention B Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
2.17.1 Upper CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer cause a problem that meant you needed to see your orthodontics?")						
Jowett 2022	8	31	3	31	100.0%	2.67 [0.78 , 9.12]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**2.67 [0.78 , 9.12]**
Total events:	8		3			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.56 (P = 0.12)						
2.17.2 Lower CAD/CAM nitinol vs rectangular chain retainers at 6 months (specifically, "Did your retainer cause a problem that meant you needed to see your orthodontics?")						
Jowett 2022	4	31	4	31	100.0%	1.00 [0.27 , 3.65]
Subtotal (95% CI)		**31**		**31**	**100.0%**	**1.00 [0.27 , 3.65]**
Total events:	4		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.18. Comparison 2: Fixed versus fixed retainers, Outcome 18: Patient satisfaction: aesthetics

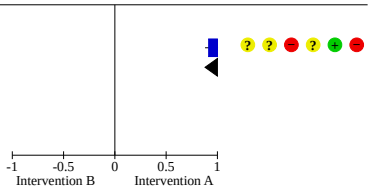
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.18.1 Lower fibre-reinforced composite vs multistrand bonded retainer at 12 months (specifically, "esthetic result while talking and smiling")								
Scribante 2011	9.73	0.42	17	8.24	1.39	15	100.0%	1.49 [0.76 , 2.22]
Subtotal (95% CI)			**17**			**15**	**100.0%**	**1.49 [0.76 , 2.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.99 (P < 0.0001)								



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.19. Comparison 2: Fixed versus fixed retainers, Outcome 19: Patient satisfaction: preference for type of retainer

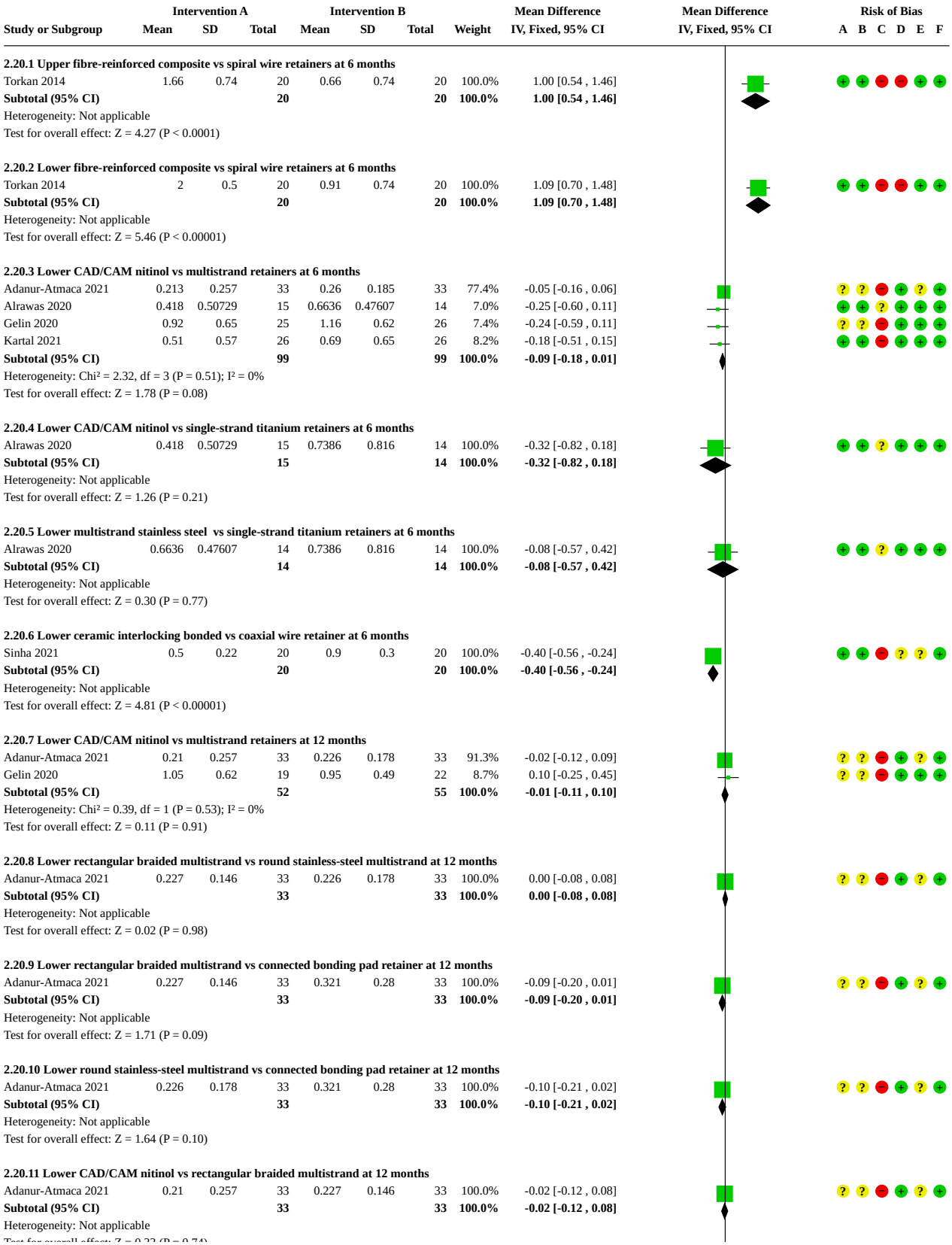
Study or Subgroup	Intervention A Events	Total	Intervention B Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
2.19.1 Lower rigid retainer bonded only 33-43 vs rectangular chain retainer at 6 months (specifically, "preference for type of retainer")						
Ferreira 2019	15	15	0	15	100.0%	1.00 [0.88 , 1.12]
Subtotal (95% CI)		**15**		**15**	**100.0%**	**1.00 [0.88 , 1.12]**
Total events:	15		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 16.26 (P < 0.00001)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

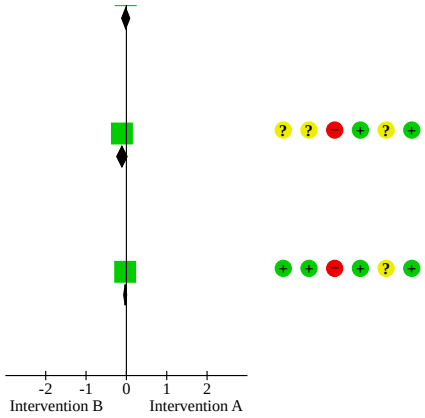
Analysis 2.20. Comparison 2: Fixed versus fixed retainers, Outcome 20: Oral health: Plaque Index



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.20. (Continued)

Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.02 [-0.12 , 0.08]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.33 (P = 0.74)								
2.20.12 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months								
Adanur-Atmaca 2021	0.21	0.257	33	0.321	0.28	33	100.0%	-0.11 [-0.24 , 0.02]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.11 [-0.24 , 0.02]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.68 (P = 0.09)								
2.20.13 Lower round twisted multistrand vs rectangular braided multistrand at 24 months								
Węgrodzka 2021	0.08	0.09	66	0.11	0.12	66	100.0%	-0.03 [-0.07 , 0.01]
Subtotal (95% CI)			**66**			**66**	**100.0%**	**-0.03 [-0.07 , 0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.62 (P = 0.10)								



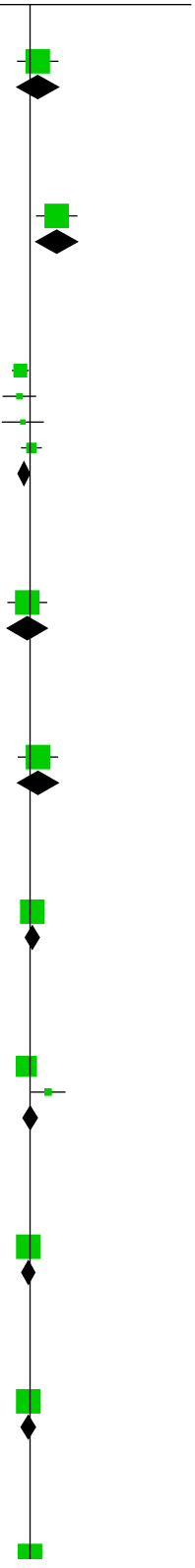
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.21. Comparison 2: Fixed versus fixed retainers, Outcome 21: Oral health: Gingival Index

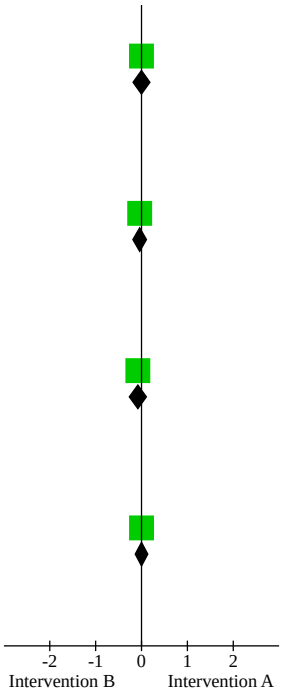
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.21.1 Upper fibre-reinforced composite vs spiral wire retainers at 6 months								
Torkan 2014	1	0.74	20	0.83	0.74	20	100.0%	0.17 [-0.29 , 0.63]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**0.17 [-0.29 , 0.63]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.73 (P = 0.47)								
2.21.2 Lower fibre-reinforced composite vs spiral wire retainers at 6 months								
Torkan 2014	1	0.74	20	0.41	0.74	20	100.0%	0.59 [0.13 , 1.05]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**0.59 [0.13 , 1.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.52 (P = 0.01)								
2.21.3 Lower CAD/CAM nitinol vs multistrand retainers at 6 months								
Adanur-Atmaca 2021	0.225	0.361	33	0.44	0.412	33	48.4%	-0.21 [-0.40 , -0.03]
Alrawas 2020	0.7747	0.50805	15	1.0114	0.51261	14	12.2%	-0.24 [-0.61 , 0.14]
Gelin 2020	0.88	0.9	25	1.04	0.79	26	7.8%	-0.16 [-0.63 , 0.31]
Kartal 2021	0.44	0.45	26	0.41	0.4	26	31.6%	0.03 [-0.20 , 0.26]
Subtotal (95% CI)			**99**			**99**	**100.0%**	**-0.14 [-0.27 , -0.01]**
Heterogeneity: Chi² = 2.95, df = 3 (P = 0.40); I² = 0%								
Test for overall effect: Z = 2.05 (P = 0.04)								
2.21.4 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months								
Alrawas 2020	0.7747	0.50805	15	0.8364	0.68552	14	100.0%	-0.06 [-0.50 , 0.38]
Subtotal (95% CI)			**15**			**14**	**100.0%**	**-0.06 [-0.50 , 0.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.27 (P = 0.78)								
2.21.5 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months								
Alrawas 2020	1.0114	0.51261	14	0.8364	0.68552	14	100.0%	0.18 [-0.27 , 0.62]
Subtotal (95% CI)			**14**			**14**	**100.0%**	**0.18 [-0.27 , 0.62]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.76 (P = 0.44)								
2.21.6 Lower ceramic interlocking bonded vs coaxial wire retainer at 6 months								
Sinha 2021	0.75	0.25	20	0.7	0.25	20	100.0%	0.05 [-0.10 , 0.20]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**0.05 [-0.10 , 0.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.63 (P = 0.53)								
2.21.7 Lower CAD/CAM nitinol vs multistrand retainers at 12 months								
Adanur-Atmaca 2021	0.248	0.368	33	0.327	0.357	33	83.0%	-0.08 [-0.25 , 0.10]
Gelin 2020	0.63	0.76	19	0.23	0.43	22	17.0%	0.40 [0.01 , 0.79]
Subtotal (95% CI)			**52**			**55**	**100.0%**	**0.00 [-0.16 , 0.16]**
Heterogeneity: Chi² = 4.91, df = 1 (P = 0.03); I² = 80%								
Test for overall effect: Z = 0.03 (P = 0.97)								
2.21.8 Lower rectangular braided multistrand vs round stainless-steel multistrand at 12 months								
Adanur-Atmaca 2021	0.288	0.232	33	0.327	0.357	33	100.0%	-0.04 [-0.18 , 0.11]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.04 [-0.18 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.60)								
2.21.9 Lower rectangular braided multistrand vs connected bonding pad retainer at 12 months								
Adanur-Atmaca 2021	0.288	0.232	33	0.328	0.404	33	100.0%	-0.04 [-0.20 , 0.12]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.04 [-0.20 , 0.12]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.49 (P = 0.62)								
2.21.10 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months								
Adanur-Atmaca 2021	0.327	0.357	33	0.328	0.404	33	100.0%	-0.00 [-0.18 , 0.18]



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

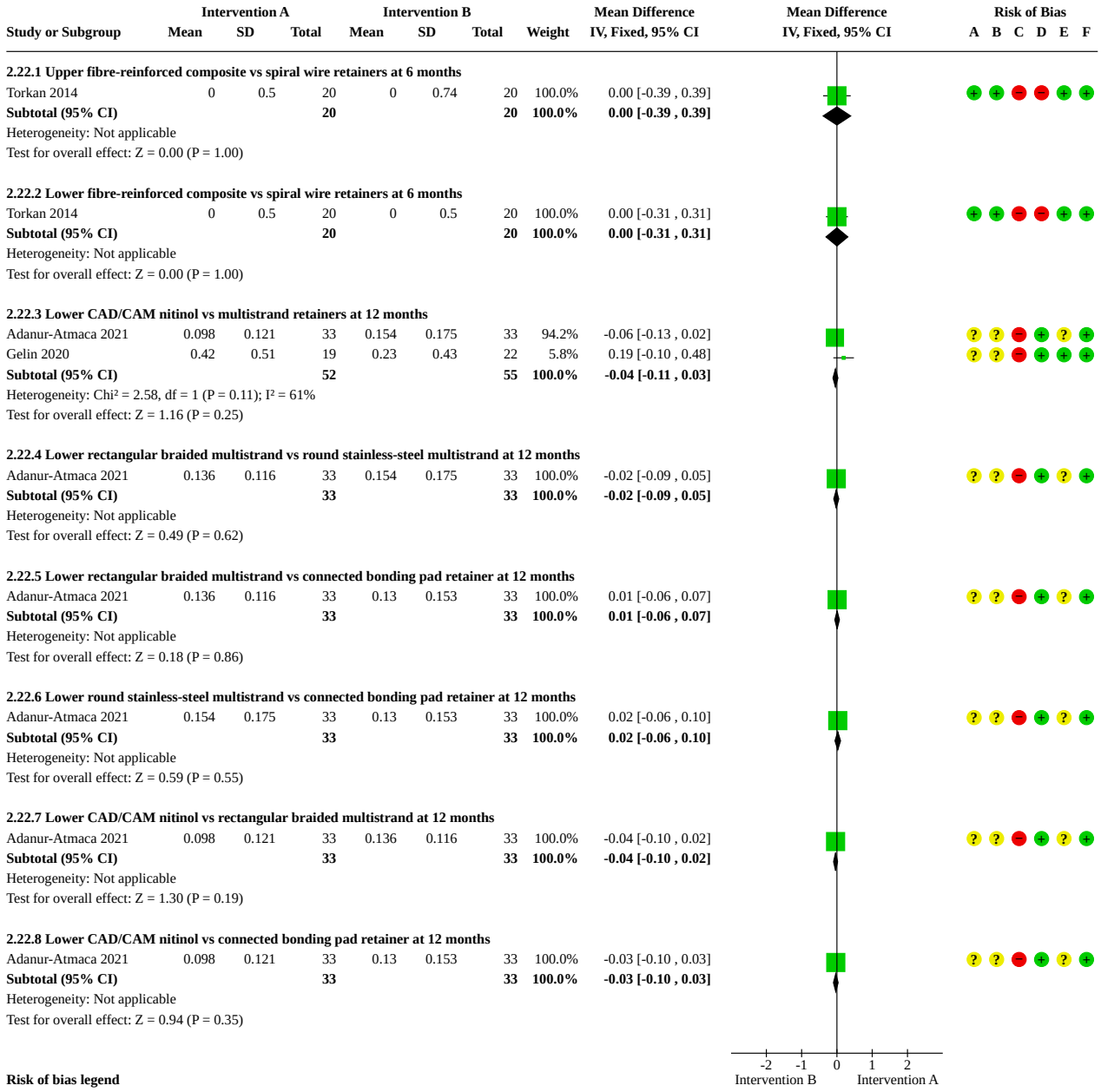
Analysis 2.21. (Continued)

2.21.10 Lower round stainless-steel multistrand vs connected bonding pad retainer at 12 months								
Adanur-Atmaca 2021	0.327	0.357	33	0.328	0.404	33	100.0%	-0.00 [-0.18 , 0.18]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.00 [-0.18 , 0.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.01 (P = 0.99)								
2.21.11 Lower CAD-CAM nitinol vs rectangular braided multistrand at 12 months								
Adanur-Atmaca 2021	0.248	0.368	33	0.288	0.232	33	100.0%	-0.04 [-0.19 , 0.11]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.04 [-0.19 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.60)								
2.21.12 Lower CAD/CAM nitinol vs connected bonding pad retainer at 12 months								
Adanur-Atmaca 2021	0.248	0.368	33	0.328	0.404	33	100.0%	-0.08 [-0.27 , 0.11]
Subtotal (95% CI)			**33**			**33**	**100.0%**	**-0.08 [-0.27 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.84 (P = 0.40)								
2.21.13 Lower round twisted multistrand vs rectangular braided multistrand at 24 months								
Węgrodzka 2021	0.58	0.37	66	0.58	0.43	66	100.0%	0.00 [-0.14 , 0.14]
Subtotal (95% CI)			**66**			**66**	**100.0%**	**0.00 [-0.14 , 0.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.22. Comparison 2: Fixed versus fixed retainers, Outcome 22: Oral health: Calculus Index



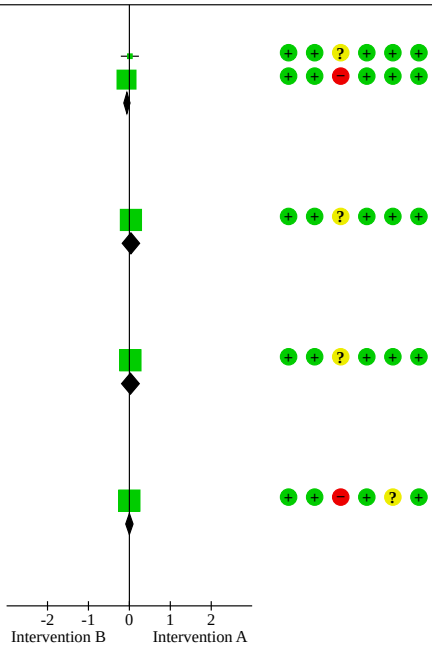
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.23. Comparison 2: Fixed versus fixed retainers, Outcome 23: Oral health: probing depth

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.23.1 Lower CAD/CAM nitinol vs multistrand retainers at 6 months								
Alrawas 2020	1.772	0.2967	15	1.76	0.3084	14	11.9%	0.01 [-0.21 , 0.23]
Kartal 2021	1.1	0.11	26	1.17	0.18	26	88.1%	-0.07 [-0.15 , 0.01]
Subtotal (95% CI)			**41**			**40**	**100.0%**	**-0.06 [-0.14 , 0.02]**
Heterogeneity: Chi² = 0.47, df = 1 (P = 0.49); I² = 0%								
Test for overall effect: Z = 1.55 (P = 0.12)								
2.23.2 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months								
Alrawas 2020	1.772	0.2967	15	1.7343	0.28833	14	100.0%	0.04 [-0.18 , 0.25]
Subtotal (95% CI)			**15**			**14**	**100.0%**	**0.04 [-0.18 , 0.25]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.73)								
2.23.3 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months								
Alrawas 2020	1.76	0.3084	15	1.7343	0.28833	14	100.0%	0.03 [-0.19 , 0.24]
Subtotal (95% CI)			**15**			**14**	**100.0%**	**0.03 [-0.19 , 0.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.23 (P = 0.82)								
2.23.4 Lower round twisted multistrand versus rectangular braided multistrand at 24 months								
Węgrodzka 2021	1	0.01	66	1	0.37	66	100.0%	0.00 [-0.09 , 0.09]
Subtotal (95% CI)			**66**			**66**	**100.0%**	**0.00 [-0.09 , 0.09]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								



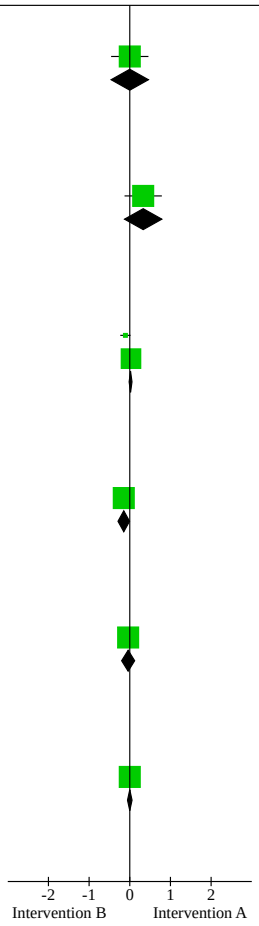
Risk of bias legend

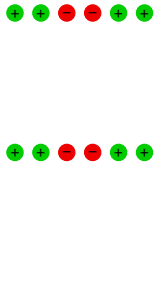
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 2.24. Comparison 2: Fixed versus fixed retainers, Outcome 24: Oral health: bleeding on probing

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.24.1 Upper fibre-reinforced composite vs spiral wire retainers at 6 months								
Torkan 2014	0.5	0.74	20	0.5	0.74	20	100.0%	0.00 [-0.46 , 0.46]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**0.00 [-0.46 , 0.46]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
2.24.2 Lower fibre-reinforced composite vs spiral wire retainers at 6 months								
Torkan 2014	0.66	0.74	20	0.33	0.74	20	100.0%	0.33 [-0.13 , 0.79]
Subtotal (95% CI)			**20**			**20**	**100.0%**	**0.33 [-0.13 , 0.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.41 (P = 0.16)								
2.24.3 Lower CAD/CAM nitinol vs multistrand retainers at 6 months								
Alrawas 2020	0.1247	0.14242	15	0.2321	0.19632	14	7.8%	-0.11 [-0.23 , 0.02]
Kartal 2021	0.03	0.09	26	0	0.03	26	92.2%	0.03 [-0.01 , 0.07]
Subtotal (95% CI)			**41**			**40**	**100.0%**	**0.02 [-0.02 , 0.05]**
Heterogeneity: Chi² = 4.24, df = 1 (P = 0.04); I² = 76%								
Test for overall effect: Z = 1.08 (P = 0.28)								
2.24.4 Lower CAD/CAM nitinol vs single-strand titanium retainers at 6 months								
Alrawas 2020	0.1247	0.14242	15	0.2721	0.23394	14	100.0%	-0.15 [-0.29 , -0.01]
Subtotal (95% CI)			**15**			**14**	**100.0%**	**-0.15 [-0.29 , -0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.03 (P = 0.04)								
2.24.5 Lower multistrand stainless steel vs single-strand titanium retainers at 6 months								
Alrawas 2020	0.2321	0.19632	14	0.2721	0.23394	14	100.0%	-0.04 [-0.20 , 0.12]
Subtotal (95% CI)			**14**			**14**	**100.0%**	**-0.04 [-0.20 , 0.12]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.49 (P = 0.62)								
2.24.6 Lower round twisted multistrand vs rectangular braided multistrand at 24 months								
Węgrodzka 2021	0	0.19	66	0	0.13	66	100.0%	0.00 [-0.06 , 0.06]
Subtotal (95% CI)			**66**			**66**	**100.0%**	**0.00 [-0.06 , 0.06]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								

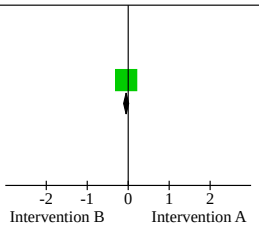




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 2.25. Comparison 2: Fixed versus fixed retainers, Outcome 25: Oral health: marginal recession

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
2.25.1 Lower CAD/CAM nitinol vs multistrand retainers at 6 months								
Gelin 2020	0	0	25	0	0	26		Not estimable
Kartal 2021	0.03	0.05	26	0.08	0.15	26	100.0%	-0.05 [-0.11 , 0.01]
Subtotal (95% CI)			**51**			**52**	**100.0%**	**-0.05 [-0.11 , 0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.61 (P = 0.11)								





Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Comparison 3. Removable versus removable retainers

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Stability: Irregularity Index	7		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.1.1 Upper Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.07 [-0.58, 0.72]
3.1.2 Lower Hawley full-time vs clear plastic full-time 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [0.02, 0.18]
3.1.3 Upper Hawley full-time vs clear plastic part-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.60, 0.70]
3.1.4 Lower Hawley full-time vs clear plastic part-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.05 [-0.03, 0.13]
3.1.5 Upper clear plastic full-time vs clear plastic part-time at 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.10, 0.06]
3.1.6 Lower clear plastic full-time vs clear plastic part-time at 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-0.14, 0.04]
3.1.7 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.38 [0.22, 0.54]
3.1.8 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.17 [0.03, 0.31]
3.1.9 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.45 [0.29, 0.61]
3.1.10 Lower Hawley part-time (after 4 monthd full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.21 [0.07, 0.35]
3.1.11 Upper clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.07 [-0.06, 0.20]
3.1.12 Lower clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.10, 0.18]
3.1.13 Upper Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months	1	310	Mean Difference (IV, Fixed, 95% CI)	0.25 [0.08, 0.42]
3.1.14 Lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months	1	310	Mean Difference (IV, Fixed, 95% CI)	0.42 [0.23, 0.61]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1.15 Upper clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.45, 0.39]
3.1.16 Lower clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.38, 0.34]
3.1.17 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.47 [-1.06, 0.12]
3.1.18 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.65, 0.67]
3.1.19 Upper clear plastic with mobile app vs clear plastic without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.03 [-0.14, 0.20]
3.1.20 Lower clear plastic with mobile app vs clear plastic without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.09 [-0.01, 0.19]
3.1.21 Upper Hawley part-time 12 months vs Hawley full-time 6 months + part-time 6 months at 12 monthsm	1	52	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.41, 0.41]
3.1.22 Lower Hawley part-time 12 months vs Hawley full-time 6 months + part-time 6 months at 12 months	1	52	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.28, 0.68]
3.1.23 Upper & lower Begg full-time vs clear plastic full-time at 6 months	1	224	Mean Difference (IV, Fixed, 95% CI)	0.25 [0.19, 0.31]
3.2 Stability: intercanine width	5		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.2.1 Upper Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-1.35, 0.75]
3.2.2 Lower Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.88, 1.08]
3.2.3 Upper Hawley full-time vs clear plastic part-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-1.41, 0.81]
3.2.4 Lower Hawley full-time vs clear plastic part-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.88, 0.88]
3.2.5 Upper clear plastic full-time vs clear plastic part-time at 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.00 [-1.12, 1.12]
3.2.6 Lower clear plastic full-time vs clear plastic part-time 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.99, 0.79]
3.2.7 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.70 [-1.76, 0.36]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.2.8 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.90, 1.10]
3.2.9 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.40 [-1.51, 0.71]
3.2.10 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.90, 0.90]
3.2.11 Upper clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.83, 1.43]
3.2.12 Lower clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-1.01, 0.81]
3.2.13 Upper Hawley full-time vs modified clear plastic full-time at 6 months	1	31	Mean Difference (IV, Fixed, 95% CI)	0.46 [-1.13, 2.05]
3.2.14 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months	1	24	Mean Difference (IV, Fixed, 95% CI)	0.31 [-1.56, 2.18]
3.2.15 Upper Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months	1	310	Mean Difference (IV, Fixed, 95% CI)	0.06 [-0.05, 0.17]
3.2.16 Lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time 6 months	1	310	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.03, 0.11]
3.2.17 Upper clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.30, 0.31]
3.2.18 Lower clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-0.32, 0.22]
3.2.19 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.22 [-1.69, 1.25]
3.2.20 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-1.36, 1.24]
3.3 Stability: interpremolar width	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.3.1 Upper Hawley full-time vs modified clear plastic full-time at 6 months	1	31	Mean Difference (IV, Fixed, 95% CI)	0.20 [-1.17, 1.57]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.3.2 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months	1	24	Mean Difference (IV, Fixed, 95% CI)	0.11 [-1.46, 1.68]
3.4 Stability: intermolar width	5		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.4.1 Upper Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.95, 0.75]
3.4.2 Lower Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.72, 1.32]
3.4.3 Upper Hawley full-time vs clear plastic part-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-1.06, 0.66]
3.4.4 Lower Hawley full-time vs clear plastic part-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.98, 1.18]
3.4.5 Upper clear plastic full-time vs clear plastic part-time at 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.94, 0.74]
3.4.6 Lower clear plastic full-time vs clear plastic part-time at 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-1.19, 0.79]
3.4.7 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.67, 1.07]
3.4.8 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.85, 1.25]
3.4.9 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.77, 0.97]
3.4.10 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-1.16, 0.96]
3.4.11 Upper clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.95, 0.75]
3.4.12 Lower clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-1.32, 0.72]
3.4.13 Upper Hawley full-time vs modified clear plastic full-time 6 months (mesiobuccal cusps)	1	31	Mean Difference (IV, Fixed, 95% CI)	0.34 [-1.12, 1.80]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.4.14 Upper Hawley full-time vs modified clear plastic full-time at 6 months (distobuccal cusps)	1	31	Mean Difference (IV, Fixed, 95% CI)	0.08 [-1.61, 1.77]
3.4.15 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months (mesiobuccal cusps)	1	24	Mean Difference (IV, Fixed, 95% CI)	0.14 [-1.53, 1.81]
3.4.16 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months (distobuccal cusps)	1	24	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-2.00, 1.94]
3.4.17 Upper Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months	1	310	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.14, 0.06]
3.4.18 Lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months	1	310	Mean Difference (IV, Fixed, 95% CI)	0.06 [-0.03, 0.15]
3.4.19 Upper clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	-0.12 [-0.31, 0.06]
3.4.20 Lower clear plastic full-time vs clear plastic part-time at 6 months	1	56	Mean Difference (IV, Fixed, 95% CI)	0.11 [-0.38, 0.59]
3.4.21 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.74 [-3.06, 1.58]
3.4.22 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-2.27, 2.17]
3.5 Stability: arch length	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.5.1 Upper Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.40 [-1.73, 0.93]
3.5.2 Lower Hawley full-time vs clear plastic full-time at 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-1.28, 1.08]
3.5.3 Upper Hawley full-time vs clear plastic part-time 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-1.58, 0.98]
3.5.4 Lower Hawley full-time vs clear plastic part-time 4 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.00 [-1.24, 1.24]
3.5.5 Upper clear plastic full-time vs clear plastic part-time at 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.10 [-1.20, 1.40]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.5.6 Lower clear plastic full-time vs clear plastic part-time 4 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.10 [-1.07, 1.27]
3.5.7 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.70 [-2.01, 0.61]
3.5.8 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months full-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-1.64, 0.64]
3.5.9 Upper Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.60 [-1.89, 0.69]
3.5.10 Lower Hawley part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-1.69, 0.69]
3.5.11 Upper clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	55	Mean Difference (IV, Fixed, 95% CI)	0.10 [-1.16, 1.36]
3.5.12 Lower clear plastic part-time (after 4 months full-time) vs clear plastic part-time (after 4 months part-time) at 8 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.00 [-1.17, 1.17]
3.5.13 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.14 [-4.77, 4.49]
3.5.14 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.18 [-3.96, 3.60]
3.6 Stability: overjet	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.6.1 Upper & lower clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.48, 0.35]
3.6.2 Upper & lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.19 [-0.32, 0.70]
3.7 Stability: overbite	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.7.1 Upper & lower clear plastic full-time vs clear plastic part-time at 6 months	1	57	Mean Difference (IV, Fixed, 95% CI)	-0.34 [-0.69, 0.01]
3.7.2 Upper & lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.24 [-0.96, 0.48]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.8 Stability: maintaining corrected rotations in the upper	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.8.1 Upper Hawley full-time for 3 months then part-time vs clear plastic part-time at 9 months	1	75	Risk Ratio (M-H, Fixed, 95% CI)	4.88 [1.13, 21.07]
3.9 Stability: maintaining quality of finish (PAR)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.9.1 Upper & lower Begg full-time vs clear plastic full-time at 6 months	1	224	Mean Difference (IV, Fixed, 95% CI)	1.71 [1.44, 1.98]
3.10 Stability: maintaining quality of finish (ABO Grading)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.10.1 Upper Hawley vs clear plastic 1.5 mm at 6 months	1	44	Mean Difference (IV, Fixed, 95% CI)	-2.13 [-5.70, 1.44]
3.10.2 Upper Hawley vs clear plastic 1 mm at 6 months	1	44	Mean Difference (IV, Fixed, 95% CI)	0.51 [-3.40, 4.42]
3.10.3 Upper clear plastic 1.5 mm vs clear plastic 1 mm at 6 months	1	44	Mean Difference (IV, Fixed, 95% CI)	2.64 [-1.39, 6.67]
3.11 Failures of retainers: total number of failures	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.11.1 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	0.87 [0.63, 1.19]
3.11.2 Lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	0.60 [0.43, 0.83]
3.11.3 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	222	Risk Ratio (M-H, Fixed, 95% CI)	0.72 [0.57, 0.90]
3.11.4 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time 12 months	1	72	Risk Ratio (M-H, Fixed, 95% CI)	0.59 [0.30, 1.17]
3.12 Failure of retainers: broken retainers	3		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.12.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs 1.5 mm clear plastic part-time at 6 months	1	346	Risk Ratio (M-H, Fixed, 95% CI)	3.21 [1.55, 6.64]
3.12.2 Upper & lower Hawley full-time vs 0.75mm clear plastic full-time at 12 months	1	222	Risk Ratio (M-H, Fixed, 95% CI)	0.76 [0.53, 1.09]
3.12.3 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	1.16 [0.68, 1.96]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.12.4 Lower Hawley full-time vs 0.75mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	0.51 [0.30, 0.86]
3.12.5 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months	1	72	Risk Ratio (M-H, Fixed, 95% CI)	0.23 [0.05, 1.01]
3.13 Failure of retainers: lost retainers	3		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.13.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs 1.5 mm clear plastic part-time at 6 months	1	346	Risk Ratio (M-H, Fixed, 95% CI)	1.66 [0.70, 3.97]
3.13.2 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time 12 months	1	222	Risk Ratio (M-H, Fixed, 95% CI)	0.55 [0.31, 0.97]
3.13.3 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	0.52 [0.24, 1.13]
3.13.4 Lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	0.57 [0.24, 1.35]
3.13.5 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months	1	72	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.25, 2.90]
3.14 Failure of retainers: nonfitting retainers	3		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.14.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs 1.5 mm clear plastic part-time at 6 months	1	346	Risk Ratio (M-H, Fixed, 95% CI)	0.74 [0.13, 4.37]
3.14.2 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	222	Risk Ratio (M-H, Fixed, 95% CI)	2.95 [0.61, 14.28]
3.14.3 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	1.96 [0.18, 21.04]
3.14.4 Lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months	1	111	Risk Ratio (M-H, Fixed, 95% CI)	3.93 [0.45, 34.05]
3.14.5 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months	1	72	Risk Ratio (M-H, Fixed, 95% CI)	1.59 [0.28, 8.93]
3.15 Failure of retainers: serious abrasion	2		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
3.15.1 Upper & lower Hawley full-time vs 0.75mm clear plastic full-time 12 months	1	222	Risk Difference (M-H, Fixed, 95% CI)	-0.02 [-0.05, 0.01]
3.15.2 Upper Hawley full-time vs 0.75 mm clear plastic full-time 12 months	1	111	Risk Difference (M-H, Fixed, 95% CI)	-0.02 [-0.07, 0.03]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.15.3 Lower Hawley full-time vs 0.75mm clear plastic full-time 12 months	1	111	Risk Difference (M-H, Fixed, 95% CI)	-0.02 [-0.07, 0.03]
3.15.4 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months	1	72	Risk Difference (M-H, Fixed, 95% CI)	0.00 [-0.05, 0.05]
3.16 Mean survival time (days)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.16.1 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months	1	72	Mean Difference (IV, Fixed, 95% CI)	46.50 [-9.00, 102.00]
3.17 Failure risk (Hazard Ratio)	2		Hazard Ratio (IV, Fixed, 95% CI)	Subtotals only
3.17.1 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time 12 months	1	222	Hazard Ratio (IV, Fixed, 95% CI)	0.69 [0.45, 1.05]
3.17.2 Upper (Hawley + clear plastic) full-time vs lower (Hawley + clear plastic) full-time at 12 months	1	222	Hazard Ratio (IV, Fixed, 95% CI)	0.71 [0.47, 1.08]
3.17.3 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months	1	72	Hazard Ratio (IV, Fixed, 95% CI)	0.77 [0.48, 1.24]
3.18 Patient satisfaction: acceptance	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.18.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "appearance")	1	86	Mean Difference (IV, Fixed, 95% CI)	-2.77 [-3.24, -2.30]
3.18.2 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "self-confidence")	1	86	Mean Difference (IV, Fixed, 95% CI)	-2.18 [-2.56, -1.80]
3.19 Patient satisfaction: acceptance (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.19.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically, "embarrassed to wear retainer?")	1	348	Risk Ratio (M-H, Fixed, 95% CI)	2.42 [1.30, 4.49]
3.19.2 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically, "worse or much worse than wearing fixed appliances?")	1	349	Risk Ratio (M-H, Fixed, 95% CI)	9.37 [3.80, 23.10]
3.20 Patient satisfaction: compliance (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.20.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically: "able to wear retainer as instructed?")	1	347	Risk Ratio (M-H, Fixed, 95% CI)	3.03 [1.46, 6.30]
3.20.2 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically, "able to wear retainers away from home?")	1	344	Risk Ratio (M-H, Fixed, 95% CI)	1.22 [0.87, 1.70]
3.21 Patient satisfaction: cleaning/hygiene	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.21.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "hygiene")	1	86	Mean Difference (IV, Fixed, 95% CI)	-0.42 [-0.89, 0.05]
3.21.2 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "gingival irritation")	1	86	Mean Difference (IV, Fixed, 95% CI)	-1.28 [-1.73, -0.83]
3.22 Patient satisfaction: comfort	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.22.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "comfort")	1	86	Mean Difference (IV, Fixed, 95% CI)	-1.86 [-2.19, -1.53]
3.22.2 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "teeth biting: closing teeth with retainers, not chewing food")	1	86	Mean Difference (IV, Fixed, 95% CI)	0.91 [0.44, 1.38]
3.22.3 Upper & lower Hawley full-time vs clear plastic full-time at 6 months(specifically, "fitting")	1	86	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.44, 0.48]
3.22.4 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "swallowing")	1	86	Mean Difference (IV, Fixed, 95% CI)	-1.28 [-1.68, -0.88]
3.23 Patient satisfaction: discomfort/pain (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.23.1 Upper & lower Hawley full-time for 3 months + part-time for 3 months vs clear plastic part-time at 6 months (specifically, "never or only on occasion")	1	349	Risk Ratio (M-H, Fixed, 95% CI)	1.43 [0.77, 2.65]
3.24 Patient satisfaction: speech	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.24.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months	1	86	Mean Difference (IV, Fixed, 95% CI)	-1.25 [-1.69, -0.81]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

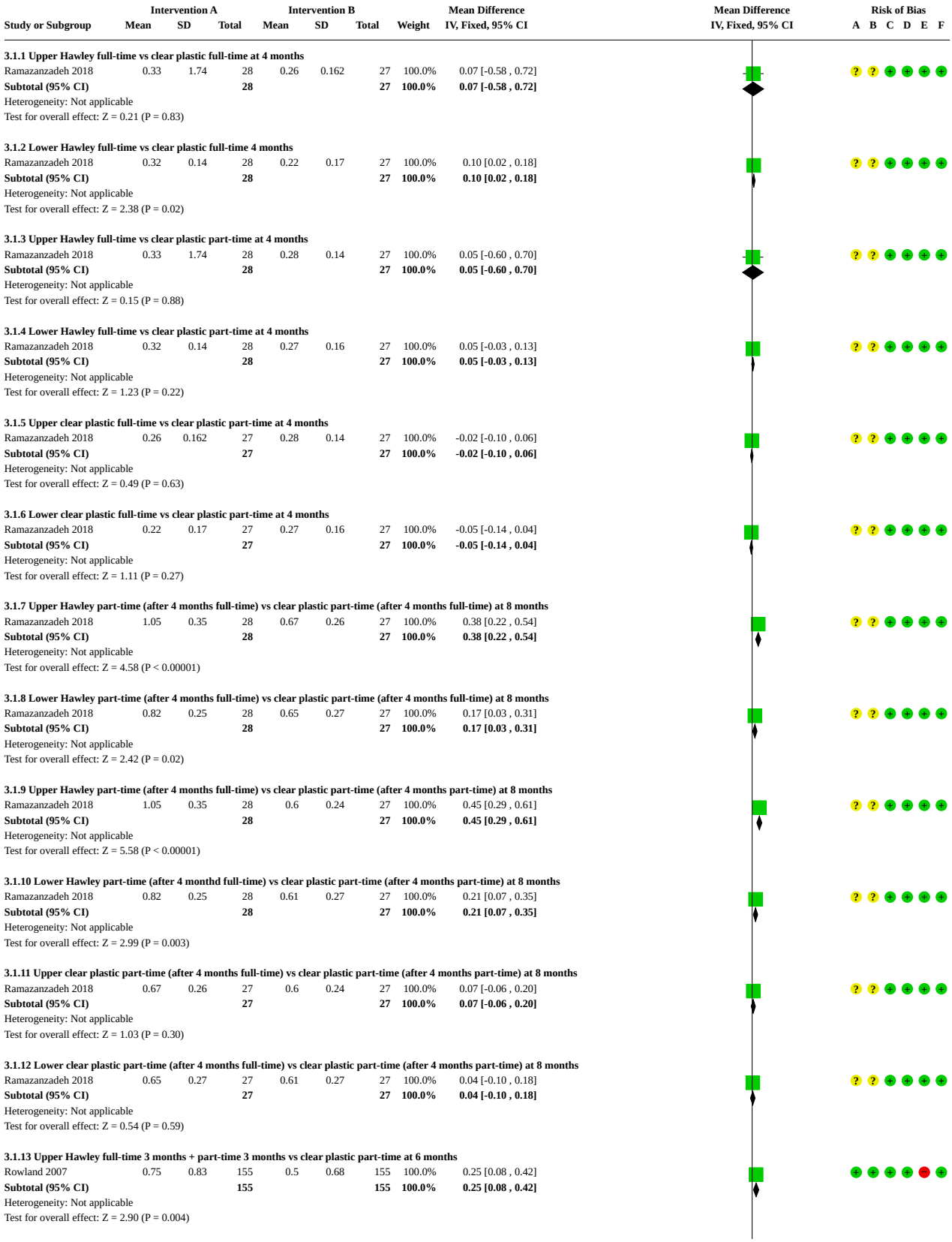
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.25 Patient satisfaction: speech (dichotomous)	1		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
3.25.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months(specifically, "number of sound distortions")	1	180	Risk Ratio (M-H, Fixed, 95% CI)	3.00 [0.12, 72.68]
3.26 Patient satisfaction: durability	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.26.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months	1	86	Mean Difference (IV, Fixed, 95% CI)	1.83 [1.07, 2.59]
3.27 Oral health: Plaque Index	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.27.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.07 [-0.16, 0.02]
3.27.2 Upper clear plastic with mobile app vs clear plastic without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.03 [-0.11, 0.05]
3.27.3 Lower clear plastic with mobile app vs clear plastic without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.08, 0.10]
3.28 Oral health: Gingival Index	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.28.1 Upper Hawley full-time vs clear plastic full-time 1.5 mm at 6 months	1	44	Mean Difference (IV, Fixed, 95% CI)	-0.05 [-0.28, 0.18]
3.28.2 Upper Hawley full-time vs clear plastic full-time 1 mm at 6 months	1	44	Mean Difference (IV, Fixed, 95% CI)	-0.21 [-0.49, 0.07]
3.28.3 Upper clear plastic full-time 1.5 mm vs clear plastic full-time 1 mm at 6 months	1	44	Mean Difference (IV, Fixed, 95% CI)	-0.16 [-0.43, 0.11]
3.28.4 Upper & lower Hawley full-time vs clear plastic full-time at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.09 [-0.30, 0.12]
3.29 Oral health: probing depth	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.29.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	-0.06 [-0.31, 0.19]
3.29.2 Upper clear plastic with mobile app vs without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.04 [-0.13, 0.05]
3.29.3 Lower clear plastic with mobile app vs without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.03 [-0.04, 0.10]
3.30 Oral health: bleeding on probing	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.30.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months	1	30	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.02, 0.06]
3.30.2 Upper clear plastic with mobile app vs without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	-0.02 [-0.06, 0.02]
3.30.3 Lower clear plastic with mobile app vs without mobile app at 12 months	1	54	Mean Difference (IV, Fixed, 95% CI)	0.01 [-0.03, 0.05]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.1. Comparison 3: Removable versus removable retainers, Outcome 1: Stability: Irregularity Index



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.1. (Continued)

Heterogeneity: Not applicable
Test for overall effect: Z = 2.90 (P = 0.004)

3.1.14 Lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months

Rowland 2007	1.2	0.98	155	0.78	0.72	155	100.0%	0.42 [0.23 , 0.61]
Subtotal (95% CI)			**155**			**155**	**100.0%**	**0.42 [0.23 , 0.61]**

Heterogeneity: Not applicable
Test for overall effect: Z = 4.30 (P < 0.0001)

3.1.15 Upper clear plastic full-time vs clear plastic part-time at 6 months

Gill 2007	0.138	0.76	28	0.168	0.84	29	100.0%	-0.03 [-0.45 , 0.39]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**-0.03 [-0.45 , 0.39]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.14 (P = 0.89)

3.1.16 Lower clear plastic full-time vs clear plastic part-time at 6 months

Gill 2007	0.29	0.57	28	0.31	0.79	29	100.0%	-0.02 [-0.38 , 0.34]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**-0.02 [-0.38 , 0.34]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.11 (P = 0.91)

3.1.17 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

Kaya 2019	0	0.99	15	0.47	0.62	15	100.0%	-0.47 [-1.06 , 0.12]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.47 [-1.06 , 0.12]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.56 (P = 0.12)

3.1.18 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

Kaya 2019	0.73	1	15	0.72	0.84	15	100.0%	0.01 [-0.65 , 0.67]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.01 [-0.65 , 0.67]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.03 (P = 0.98)

3.1.19 Upper clear plastic with mobile app vs clear plastic without mobile app at 12 months

Al-Moghrabi 2020	0.37	0.31	29	0.34	0.33	25	100.0%	0.03 [-0.14 , 0.20]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**0.03 [-0.14 , 0.20]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.34 (P = 0.73)

3.1.20 Lower clear plastic with mobile app vs clear plastic without mobile app at 12 months

Al-Moghrabi 2020	0.38	0.2	29	0.29	0.19	25	100.0%	0.09 [-0.01 , 0.19]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**0.09 [-0.01 , 0.19]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.69 (P = 0.09)

3.1.21 Upper Hawley part-time 12 months vs Hawley full-time 6 months + part-time 6 months at 12 monthsm

Shawesh 2010	2	0.7	24	2	0.8	28	100.0%	0.00 [-0.41 , 0.41]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**0.00 [-0.41 , 0.41]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.00 (P = 1.00)

3.1.22 Lower Hawley part-time 12 months vs Hawley full-time 6 months + part-time 6 months at 12 months

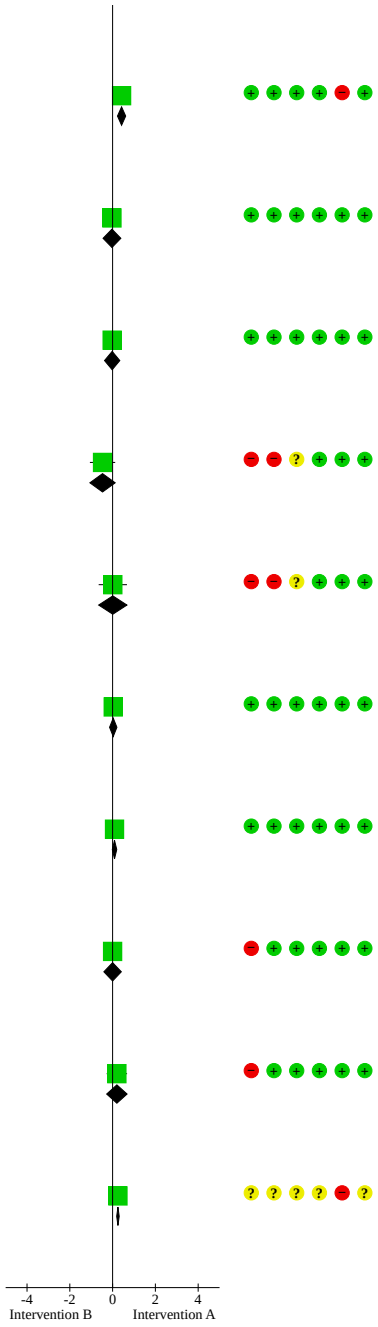
Shawesh 2010	2	1	24	1.8	0.7	28	100.0%	0.20 [-0.28 , 0.68]
Subtotal (95% CI)			**24**			**28**	**100.0%**	**0.20 [-0.28 , 0.68]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.82 (P = 0.41)

3.1.23 Upper & lower Begg full-time vs clear plastic full-time at 6 months

Kumar 2011	0.37	0.29	112	0.12	0.12	112	100.0%	0.25 [0.19 , 0.31]
Subtotal (95% CI)			**112**			**112**	**100.0%**	**0.25 [0.19 , 0.31]**

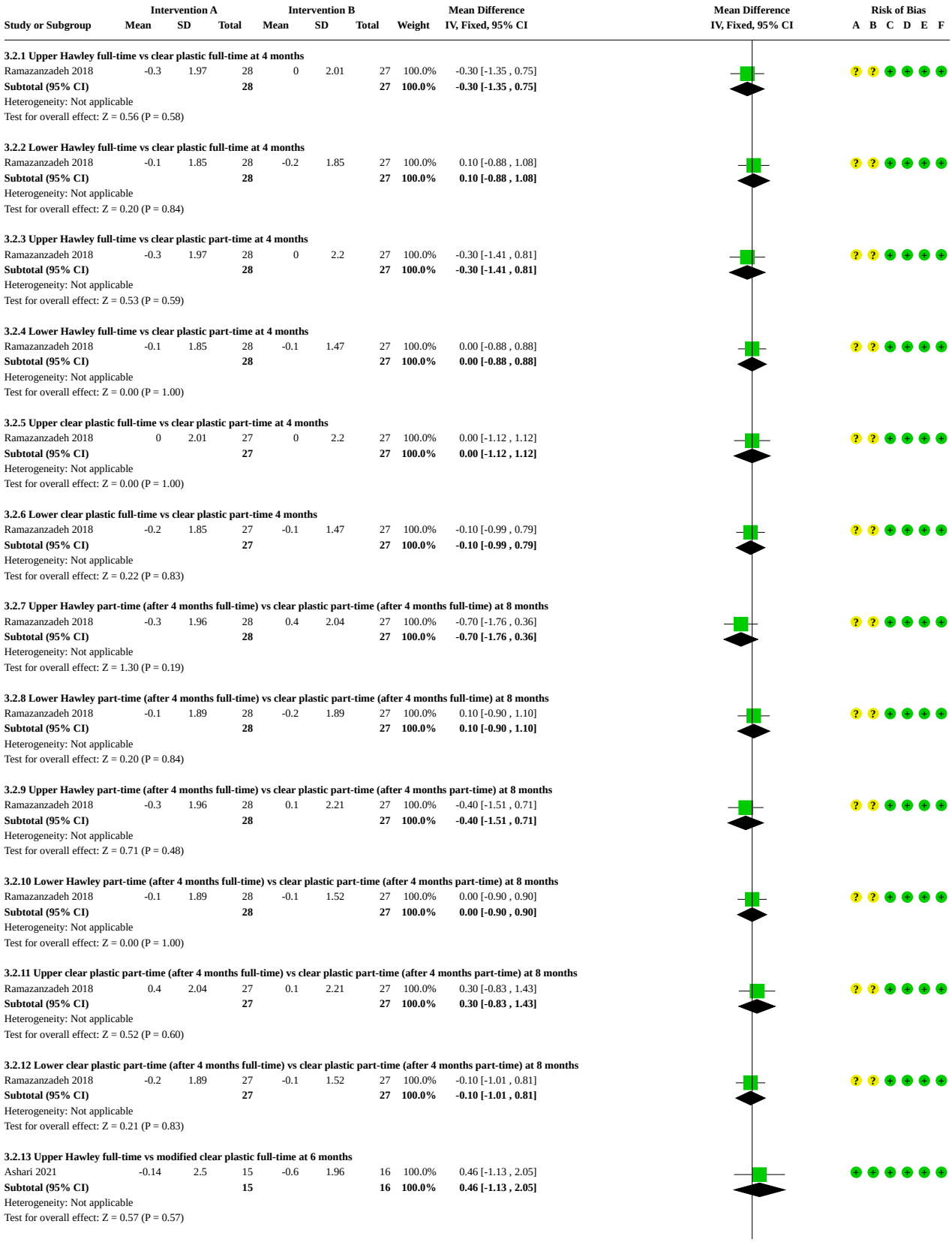
Heterogeneity: Not applicable
Test for overall effect: Z = 8.43 (P < 0.00001)



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.2. Comparison 3: Removable versus removable retainers, Outcome 2: Stability: intercanine width



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.2. (Continued)

Heterogeneity: Not applicable
Test for overall effect: Z = 0.57 (P = 0.57)

3.2.14 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months

Ashari 2021	-0.4	2.63	12	-0.71	2.01	12	100.0%	0.31 [-1.56 , 2.18]
Subtotal (95% CI)			**12**			**12**	**100.0%**	**0.31 [-1.56 , 2.18]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.32 (P = 0.75)

3.2.15 Upper Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months

Rowland 2007	0.5	0.53	155	0.44	0.41	155	100.0%	0.06 [-0.05 , 0.17]
Subtotal (95% CI)			**155**			**155**	**100.0%**	**0.06 [-0.05 , 0.17]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.11 (P = 0.26)

3.2.16 Lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time 6 months

Rowland 2007	0.32	0.35	155	0.28	0.29	155	100.0%	0.04 [-0.03 , 0.11]
Subtotal (95% CI)			**155**			**155**	**100.0%**	**0.04 [-0.03 , 0.11]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.10 (P = 0.27)

3.2.17 Upper clear plastic full-time vs clear plastic part-time at 6 months

Gill 2007	-0.17	0.56	28	-0.172	0.61	29	100.0%	0.00 [-0.30 , 0.31]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**0.00 [-0.30 , 0.31]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.01 (P = 0.99)

3.2.18 Lower clear plastic full-time vs clear plastic part-time at 6 months

Gill 2007	0.107	0.56	28	0.158	0.467	29	100.0%	-0.05 [-0.32 , 0.22]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**-0.05 [-0.32 , 0.22]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.37 (P = 0.71)

3.2.19 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

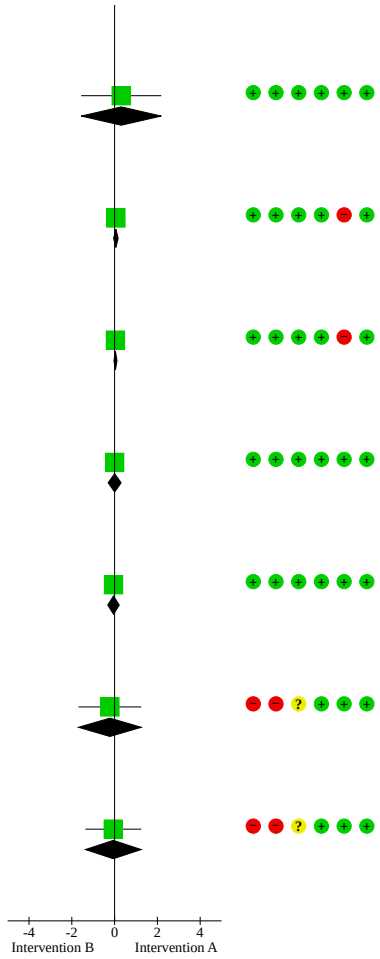
Kaya 2019	0.17	1.78	15	0.39	2.29	15	100.0%	-0.22 [-1.69 , 1.25]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.22 [-1.69 , 1.25]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.29 (P = 0.77)

3.2.20 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

Kaya 2019	-0.4	1.51	15	-0.34	2.09	15	100.0%	-0.06 [-1.36 , 1.24]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.06 [-1.36 , 1.24]**

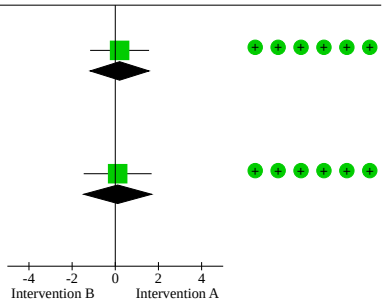
Heterogeneity: Not applicable
Test for overall effect: Z = 0.09 (P = 0.93)



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.3. Comparison 3: Removable versus removable retainers, Outcome 3: Stability: interpremolar width

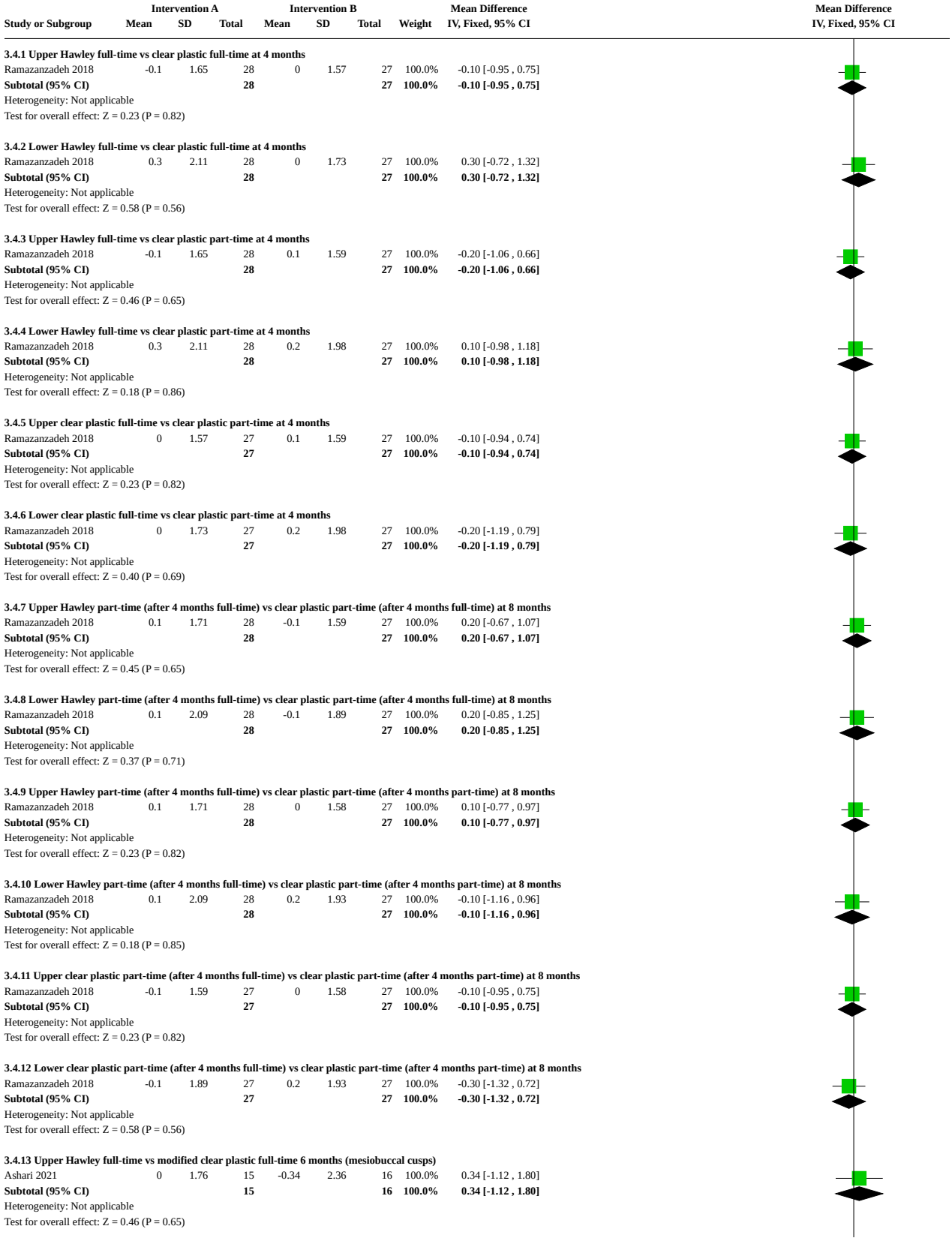
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI	Risk of Bias A B C D E F
3.3.1 Upper Hawley full-time vs modified clear plastic full-time at 6 months									
Ashari 2021	-0.13	1.91	15	-0.33	1.97	16	100.0%	0.20 [-1.17 , 1.57]	
Subtotal (95% CI)			**15**			**16**	**100.0%**	**0.20 [-1.17 , 1.57]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.29 (P = 0.77)									
3.3.2 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months									
Ashari 2021	-0.59	1.9	12	-0.7	2.03	12	100.0%	0.11 [-1.46 , 1.68]	
Subtotal (95% CI)			**12**			**12**	**100.0%**	**0.11 [-1.46 , 1.68]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.14 (P = 0.89)									



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.4. Comparison 3: Removable versus removable retainers, Outcome 4: Stability: intermolar width



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.4. (Continued)

Heterogeneity: Not applicable
Test for overall effect: Z = 0.46 (P = 0.65)

3.4.14 Upper Hawley full-time vs modified clear plastic full-time at 6 months (distobuccal cusps)

Ashari 2021	-0.07	2.12	15	-0.15	2.66	16	100.0%	0.08 [-1.61 , 1.77]
Subtotal (95% CI)			**15**			**16**	**100.0%**	**0.08 [-1.61 , 1.77]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.09 (P = 0.93)

3.4.15 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months (mesiobuccal cusps)

Ashari 2021	-0.28	1.78	12	-0.42	2.36	12	100.0%	0.14 [-1.53 , 1.81]
Subtotal (95% CI)			**12**			**12**	**100.0%**	**0.14 [-1.53 , 1.81]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.16 (P = 0.87)

3.4.16 Upper Hawley part-time (after 6 months full-time) vs modified clear plastic part-time (after 6 months full-time) at 12 months (distobuccal cusps)

Ashari 2021	-0.25	2.2	12	-0.22	2.69	12	100.0%	-0.03 [-2.00 , 1.94]
Subtotal (95% CI)			**12**			**12**	**100.0%**	**-0.03 [-2.00 , 1.94]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.03 (P = 0.98)

3.4.17 Upper Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months

Rowland 2007	0.34	0.46	155	0.38	0.45	155	100.0%	-0.04 [-0.14 , 0.06]
Subtotal (95% CI)			**155**			**155**	**100.0%**	**-0.04 [-0.14 , 0.06]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.77 (P = 0.44)

3.4.18 Lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months

Rowland 2007	0.43	0.42	155	0.37	0.35	155	100.0%	0.06 [-0.03 , 0.15]
Subtotal (95% CI)			**155**			**155**	**100.0%**	**0.06 [-0.03 , 0.15]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.37 (P = 0.17)

3.4.19 Upper clear plastic full-time vs clear plastic part-time at 6 months

Gill 2007	-0.21	0.44	28	-0.085	0.22	29	100.0%	-0.12 [-0.31 , 0.06]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**-0.12 [-0.31 , 0.06]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.35 (P = 0.18)

3.4.20 Lower clear plastic full-time vs clear plastic part-time at 6 months

Gill 2007	0.23	0.38	28	0.124	1.25	28	100.0%	0.11 [-0.38 , 0.59]
Subtotal (95% CI)			**28**			**28**	**100.0%**	**0.11 [-0.38 , 0.59]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.43 (P = 0.67)

3.4.21 Upper Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

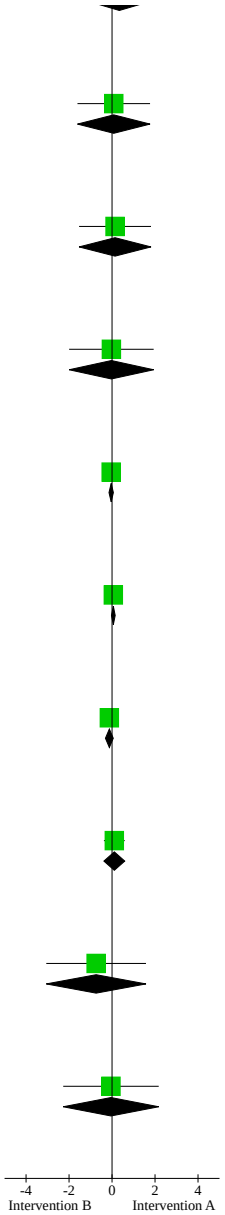
Kaya 2019	0.28	3.68	15	1.02	2.73	15	100.0%	-0.74 [-3.06 , 1.58]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.74 [-3.06 , 1.58]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.63 (P = 0.53)

3.4.22 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

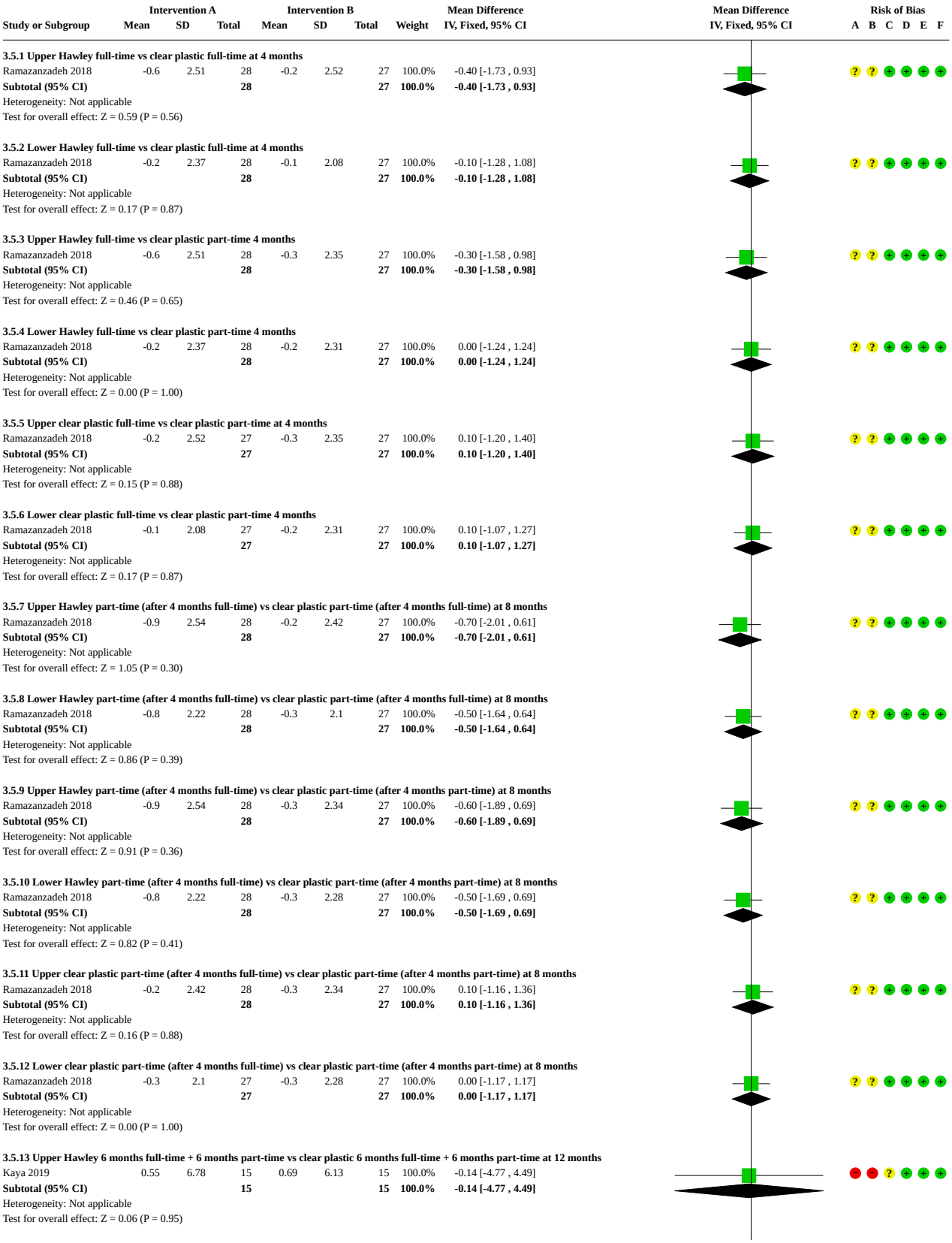
Kaya 2019	0.92	4.05	15	0.97	1.69	15	100.0%	-0.05 [-2.27 , 2.17]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.05 [-2.27 , 2.17]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.04 (P = 0.96)



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.5. Comparison 3: Removable versus removable retainers, Outcome 5: Stability: arch length



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

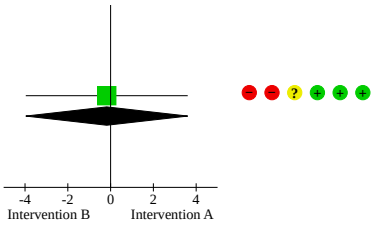
Analysis 3.5. (Continued)

Heterogeneity: Not applicable
Test for overall effect: Z = 0.06 (P = 0.95)

3.5.14 Lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Kaya 2019	-0.2	5.58	15	-0.02	4.98	15	100.0%	-0.18 [-3.96 , 3.60]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.18 [-3.96 , 3.60]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.09 (P = 0.93)



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.6. Comparison 3: Removable versus removable retainers, Outcome 6: Stability: overjet

3.6.1 Upper & lower clear plastic full-time vs clear plastic part-time at 6 months

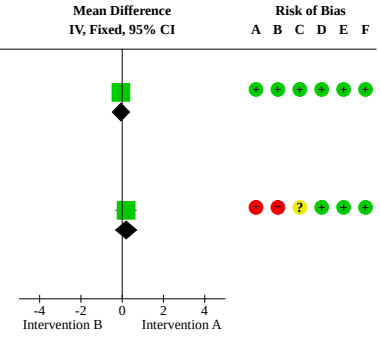
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Gill 2007	0.1	0.66	28	0.162	0.92	29	100.0%	-0.06 [-0.48 , 0.35]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**-0.06 [-0.48 , 0.35]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.29 (P = 0.77)

3.6.2 Upper & lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Kaya 2019	0.3	0.71	15	0.11	0.71	15	100.0%	0.19 [-0.32 , 0.70]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.19 [-0.32 , 0.70]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.73 (P = 0.46)



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.7. Comparison 3: Removable versus removable retainers, Outcome 7: Stability: overbite

3.7.1 Upper & lower clear plastic full-time vs clear plastic part-time at 6 months

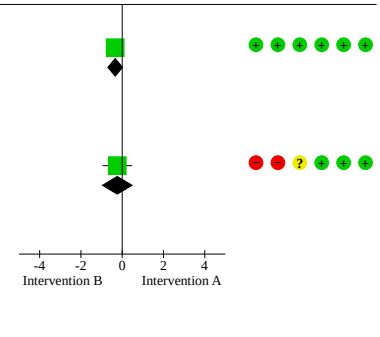
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Gill 2007	0.02	0.66	28	0.36	0.704	29	100.0%	-0.34 [-0.69 , 0.01]
Subtotal (95% CI)			**28**			**29**	**100.0%**	**-0.34 [-0.69 , 0.01]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.88 (P = 0.06)

3.7.2 Upper & lower Hawley 6 months full-time + 6 months part-time vs clear plastic 6 months full-time + 6 months part-time at 12 months

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Kaya 2019	0.19	1.11	15	0.43	0.88	15	100.0%	-0.24 [-0.96 , 0.48]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.24 [-0.96 , 0.48]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.66 (P = 0.51)

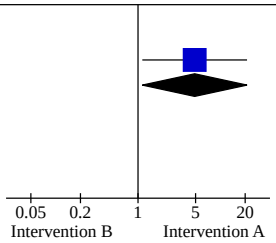


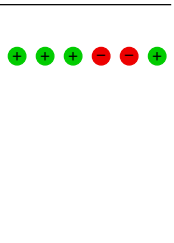
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.8. Comparison 3: Removable versus removable retainers, Outcome 8: Stability: maintaining corrected rotations in the upper

Study or Subgroup	Intervention A Events	Total	Intervention B Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.8.1 Upper Hawley full-time for 3 months then part-time vs clear plastic part-time at 9 months						
Rohaya 2006	9	36	2	39	100.0%	4.88 [1.13 , 21.07]
Subtotal (95% CI)		**36**		**39**	**100.0%**	**4.88 [1.13 , 21.07]**
Total events:	9		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.12 (P = 0.03)						

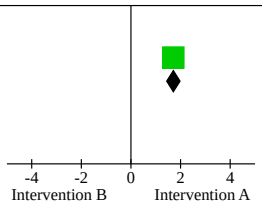


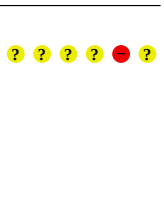


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.9. Comparison 3: Removable versus removable retainers, Outcome 9: Stability: maintaining quality of finish (PAR)

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.9.1 Upper & lower Begg full-time vs clear plastic full-time at 6 months								
Kumar 2011	2.5	1.35	112	0.79	0.54	112	100.0%	1.71 [1.44 , 1.98]
Subtotal (95% CI)			**112**			**112**	**100.0%**	**1.71 [1.44 , 1.98]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 12.45 (P < 0.00001)								



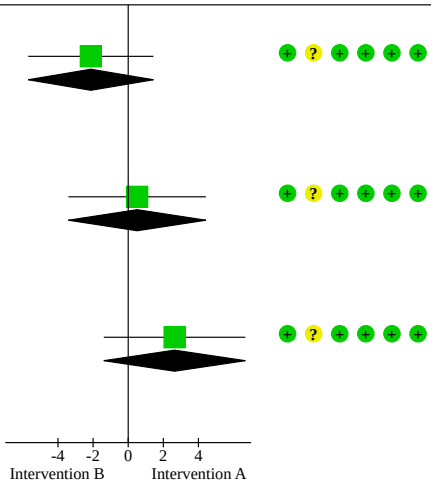


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.10. Comparison 3: Removable versus removable retainers, Outcome 10: Stability: maintaining quality of finish (ABO Grading)

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.10.1 Upper Hawley vs clear plastic 1.5 mm at 6 months								
Moslemzadeh 2018	-0.61	5.81	22	1.52	6.25	22	100.0%	-2.13 [-5.70 , 1.44]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**-2.13 [-5.70 , 1.44]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.17 (P = 0.24)								
3.10.2 Upper Hawley vs clear plastic 1 mm at 6 months								
Moslemzadeh 2018	-0.61	5.81	22	-1.12	7.35	22	100.0%	0.51 [-3.40 , 4.42]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**0.51 [-3.40 , 4.42]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.26 (P = 0.80)								
3.10.3 Upper clear plastic 1.5 mm vs clear plastic 1 mm at 6 months								
Moslemzadeh 2018	1.52	6.25	22	-1.12	7.35	22	100.0%	2.64 [-1.39 , 6.67]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**2.64 [-1.39 , 6.67]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.28 (P = 0.20)								

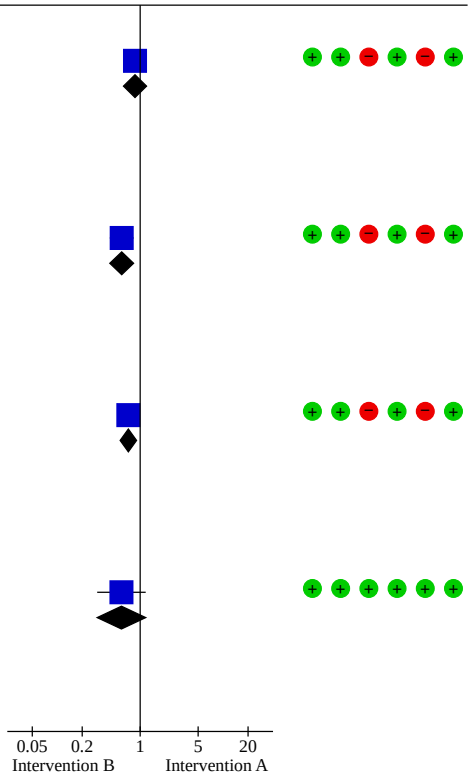


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.11. Comparison 3: Removable versus removable retainers, Outcome 11: Failures of retainers: total number of failures

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.11.1 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	30	56	34	55	100.0%	0.87 [0.63 , 1.19]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**0.87 [0.63 , 1.19]**
Total events:	30		34			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.88 (P = 0.38)						
3.11.2 Lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	25	56	41	55	100.0%	0.60 [0.43 , 0.83]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**0.60 [0.43 , 0.83]**
Total events:	25		41			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.05 (P = 0.002)						
3.11.3 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	55	112	75	110	100.0%	0.72 [0.57 , 0.90]
Subtotal (95% CI)		**112**		**110**	**100.0%**	**0.72 [0.57 , 0.90]**
Total events:	55		75			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.82 (P = 0.005)						
3.11.4 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time 12 months						
Zhu 2017	9	35	16	37	100.0%	0.59 [0.30 , 1.17]
Subtotal (95% CI)		**35**		**37**	**100.0%**	**0.59 [0.30 , 1.17]**
Total events:	9		16			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.51 (P = 0.13)						

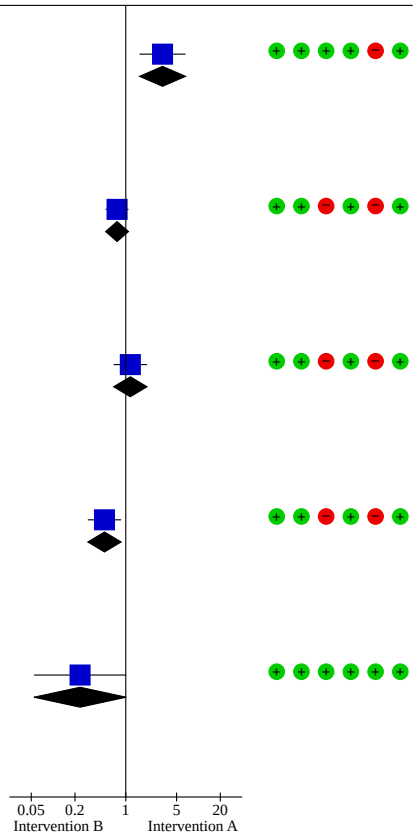


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.12. Comparison 3: Removable versus removable retainers, Outcome 12: Failure of retainers: broken retainers

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.12.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs 1.5 mm clear plastic part-time at 6 months						
Rowland 2007	26	164	9	182	100.0%	3.21 [1.55 , 6.64]
Subtotal (95% CI)		**164**		**182**	**100.0%**	**3.21 [1.55 , 6.64]**
Total events:	26		9			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.14 (P = 0.002)						
3.12.2 Upper & lower Hawley full-time vs 0.75mm clear plastic full-time at 12 months						
Sun 2011	34	112	44	110	100.0%	0.76 [0.53 , 1.09]
Subtotal (95% CI)		**112**		**110**	**100.0%**	**0.76 [0.53 , 1.09]**
Total events:	34		44			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.49 (P = 0.14)						
3.12.3 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	20	56	17	55	100.0%	1.16 [0.68 , 1.96]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**1.16 [0.68 , 1.96]**
Total events:	20		17			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.54 (P = 0.59)						
3.12.4 Lower Hawley full-time vs 0.75mm clear plastic full-time at 12 months						
Sun 2011	14	56	27	55	100.0%	0.51 [0.30 , 0.86]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**0.51 [0.30 , 0.86]**
Total events:	14		27			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.51 (P = 0.01)						
3.12.5 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months						
Zhu 2017	2	35	9	37	100.0%	0.23 [0.05 , 1.01]
Subtotal (95% CI)		**35**		**37**	**100.0%**	**0.23 [0.05 , 1.01]**
Total events:	2		9			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.94 (P = 0.05)						



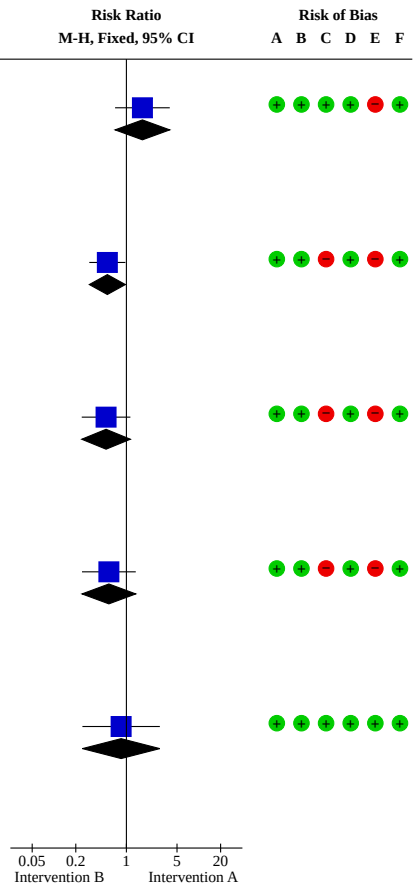
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.13. Comparison 3: Removable versus removable retainers, Outcome 13: Failure of retainers: lost retainers

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.13.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs 1.5 mm clear plastic part-time at 6 months						
Rowland 2007	12	164	8	182	100.0%	1.66 [0.70 , 3.97]
Subtotal (95% CI)		**164**		**182**	**100.0%**	**1.66 [0.70 , 3.97]**
Total events:	12		8			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.15 (P = 0.25)						
3.13.2 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time 12 months						
Sun 2011	15	112	27	110	100.0%	0.55 [0.31 , 0.97]
Subtotal (95% CI)		**112**		**110**	**100.0%**	**0.55 [0.31 , 0.97]**
Total events:	15		27			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.07 (P = 0.04)						
3.13.3 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	8	56	15	55	100.0%	0.52 [0.24 , 1.13]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**0.52 [0.24 , 1.13]**
Total events:	8		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.64 (P = 0.10)						
3.13.4 Lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	7	56	12	55	100.0%	0.57 [0.24 , 1.35]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**0.57 [0.24 , 1.35]**
Total events:	7		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.28 (P = 0.20)						
3.13.5 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months						
Zhu 2017	4	35	5	37	100.0%	0.85 [0.25 , 2.90]
Subtotal (95% CI)		**35**		**37**	**100.0%**	**0.85 [0.25 , 2.90]**
Total events:	4		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.27 (P = 0.79)						

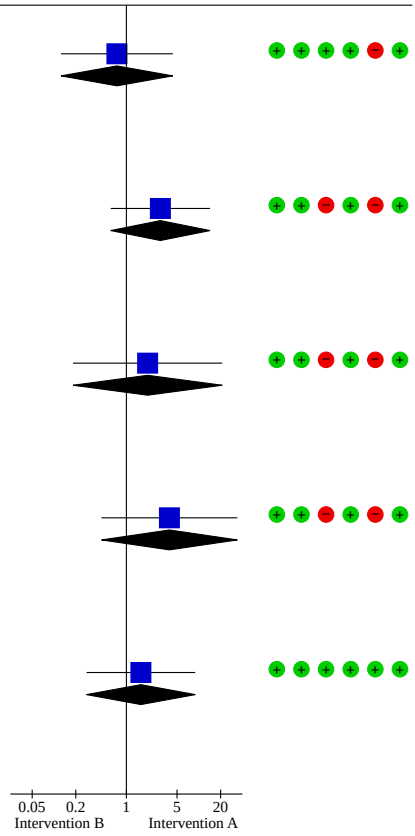


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.14. Comparison 3: Removable versus removable retainers, Outcome 14: Failure of retainers: nonfitting retainers

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.14.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs 1.5 mm clear plastic part-time at 6 months						
Rowland 2007	2	164	3	182	100.0%	0.74 [0.13 , 4.37]
Subtotal (95% CI)		**164**		**182**	**100.0%**	**0.74 [0.13 , 4.37]**
Total events:	2		3			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.33 (P = 0.74)						
3.14.2 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	6	112	2	110	100.0%	2.95 [0.61 , 14.28]
Subtotal (95% CI)		**112**		**110**	**100.0%**	**2.95 [0.61 , 14.28]**
Total events:	6		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.34 (P = 0.18)						
3.14.3 Upper Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	2	56	1	55	100.0%	1.96 [0.18 , 21.04]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**1.96 [0.18 , 21.04]**
Total events:	2		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.56 (P = 0.58)						
3.14.4 Lower Hawley full-time vs 0.75 mm clear plastic full-time at 12 months						
Sun 2011	4	56	1	55	100.0%	3.93 [0.45 , 34.05]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**3.93 [0.45 , 34.05]**
Total events:	4		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.24 (P = 0.21)						
3.14.5 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months						
Zhu 2017	3	35	2	37	100.0%	1.59 [0.28 , 8.93]
Subtotal (95% CI)		**35**		**37**	**100.0%**	**1.59 [0.28 , 8.93]**
Total events:	3		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.52 (P = 0.60)						

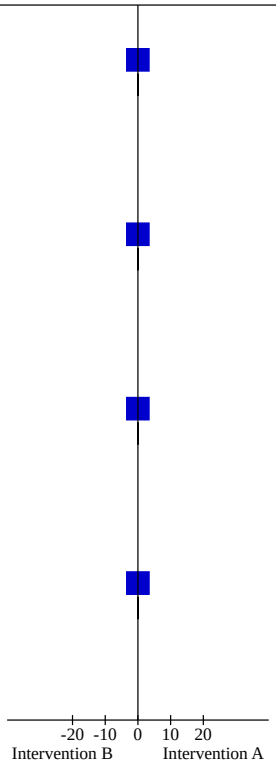


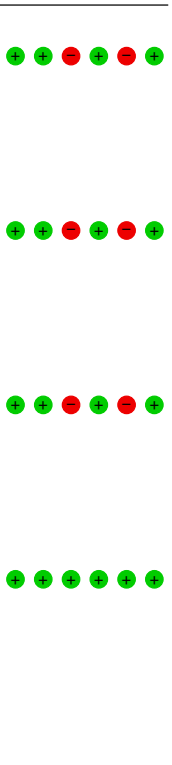
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.15. Comparison 3: Removable versus removable retainers, Outcome 15: Failure of retainers: serious abrasion

Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Difference M-H, Fixed, 95% CI
3.15.1 Upper & lower Hawley full-time vs 0.75mm clear plastic full-time 12 months						
Sun 2011	0	112	2	110	100.0%	-0.02 [-0.05 , 0.01]
Subtotal (95% CI)		**112**		**110**	**100.0%**	**-0.02 [-0.05 , 0.01]**
Total events:	0		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.18 (P = 0.24)						
3.15.2 Upper Hawley full-time vs 0.75 mm clear plastic full-time 12 months						
Sun 2011	0	56	1	55	100.0%	-0.02 [-0.07 , 0.03]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**-0.02 [-0.07 , 0.03]**
Total events:	0		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.73 (P = 0.46)						
3.15.3 Lower Hawley full-time vs 0.75mm clear plastic full-time 12 months						
Sun 2011	0	56	1	55	100.0%	-0.02 [-0.07 , 0.03]
Subtotal (95% CI)		**56**		**55**	**100.0%**	**-0.02 [-0.07 , 0.03]**
Total events:	0		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.73 (P = 0.46)						
3.15.4 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months						
Zhu 2017	0	35	0	37	100.0%	0.00 [-0.05 , 0.05]
Subtotal (95% CI)		**35**		**37**	**100.0%**	**0.00 [-0.05 , 0.05]**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						

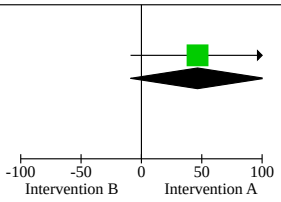




Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.16. Comparison 3: Removable versus removable retainers, Outcome 16: Mean survival time (days)

Study or Subgroup	Intervention A Mean	Intervention A SD	Intervention A Total	Intervention B Mean	Intervention B SD	Intervention B Total	Weight	Mean Difference IV, Fixed, 95% CI
3.16.1 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months								
Zhu 2017	307.5	109.9	35	261	130	37	100.0%	46.50 [-9.00 , 102.00]
Subtotal (95% CI)			**35**			**37**	**100.0%**	**46.50 [-9.00 , 102.00]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.64 (P = 0.10)								



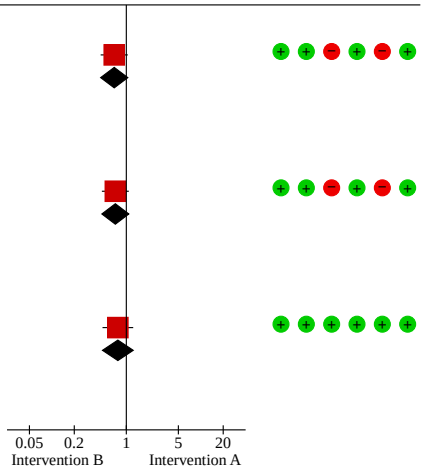


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.17. Comparison 3: Removable versus removable retainers, Outcome 17: Failure risk (Hazard Ratio)

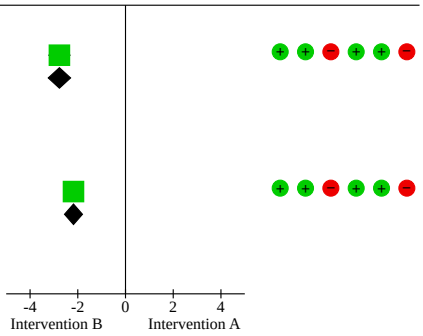
Study or Subgroup	log[Hazard Ratio]	SE	Intervention A Total	Intervention B Total	Weight	Hazard Ratio IV, Fixed, 95% CI
3.17.1 Upper & lower Hawley full-time vs 0.75 mm clear plastic full-time 12 months						
Sun 2011	-0.373966	0.214046	112	110	100.0%	0.69 [0.45 , 1.05]
Subtotal (95% CI)			**112**	**110**	**100.0%**	**0.69 [0.45 , 1.05]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.75 (P = 0.08)						
3.17.2 Upper (Hawley + clear plastic) full-time vs lower (Hawley + clear plastic) full-time at 12 months						
Sun 2011	-0.339677	0.212302	112	110	100.0%	0.71 [0.47 , 1.08]
Subtotal (95% CI)			**112**	**110**	**100.0%**	**0.71 [0.47 , 1.08]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.60 (P = 0.11)						
3.17.3 Upper & lower 1 mm clear plastic full-time vs 0.75 mm clear plastic full-time at 12 months						
Zhu 2017	-0.261365	0.242117	35	37	100.0%	0.77 [0.48 , 1.24]
Subtotal (95% CI)			**35**	**37**	**100.0%**	**0.77 [0.48 , 1.24]**
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.08 (P = 0.28)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.18. Comparison 3: Removable versus removable retainers, Outcome 18: Patient satisfaction: acceptance

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.18.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "appearance")								
Saleh 2017	5.84	1.19	41	8.61	1	45	100.0%	-2.77 [-3.24 , -2.30]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-2.77 [-3.24 , -2.30]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 11.63 (P < 0.00001)								
3.18.2 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "self-confidence")								
Saleh 2017	6.93	1.01	41	9.11	0.73	45	100.0%	-2.18 [-2.56 , -1.80]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-2.18 [-2.56 , -1.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 11.38 (P < 0.00001)								

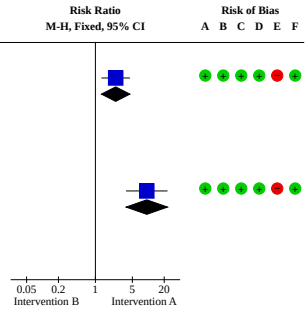


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.19. Comparison 3: Removable versus removable retainers, Outcome 19: Patient satisfaction: acceptance (dichotomous)

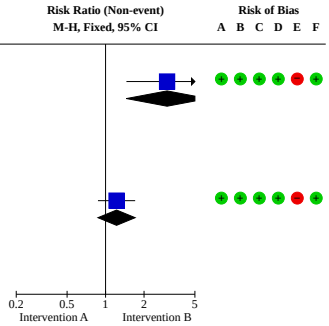
Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.19.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically, "embarrassed to wear retainer?")						
Rowland 2007	29	167	13	181	100.0%	2.42 [1.30 , 4.49]
Subtotal (95% CI)		**167**		**181**	**100.0%**	**2.42 [1.30 , 4.49]**
Total events:	29		13			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.79 (P = 0.005)						
3.19.2 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically, "worse or much worse than wearing fixed appliances?")						
Rowland 2007	43	167	5	182	100.0%	9.37 [3.80 , 23.10]
Subtotal (95% CI)		**167**		**182**	**100.0%**	**9.37 [3.80 , 23.10]**
Total events:	43		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 4.86 (P < 0.00001)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.20. Comparison 3: Removable versus removable retainers, Outcome 20: Patient satisfaction: compliance (dichotomous)

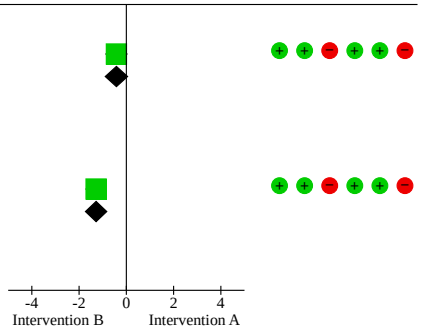
Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio (Non-event) M-H, Fixed, 95% CI
3.20.1 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically: "able to wear retainer as instructed?")						
Rowland 2007	141	166	172	181	100.0%	3.03 [1.46 , 6.30]
Subtotal (95% CI)		**166**		**181**	**100.0%**	**3.03 [1.46 , 6.30]**
Total events:	141		172			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.97 (P = 0.003)						
3.20.2 Upper & lower Hawley full-time 3 months + part-time 3 months vs clear plastic part-time at 6 months (specifically, "able to wear retainers away from home?")						
Rowland 2007	111	162	135	182	100.0%	1.22 [0.87 , 1.70]
Subtotal (95% CI)		**162**		**182**	**100.0%**	**1.22 [0.87 , 1.70]**
Total events:	111		135			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.16 (P = 0.25)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.21. Comparison 3: Removable versus removable retainers, Outcome 21: Patient satisfaction: cleaning/hygiene

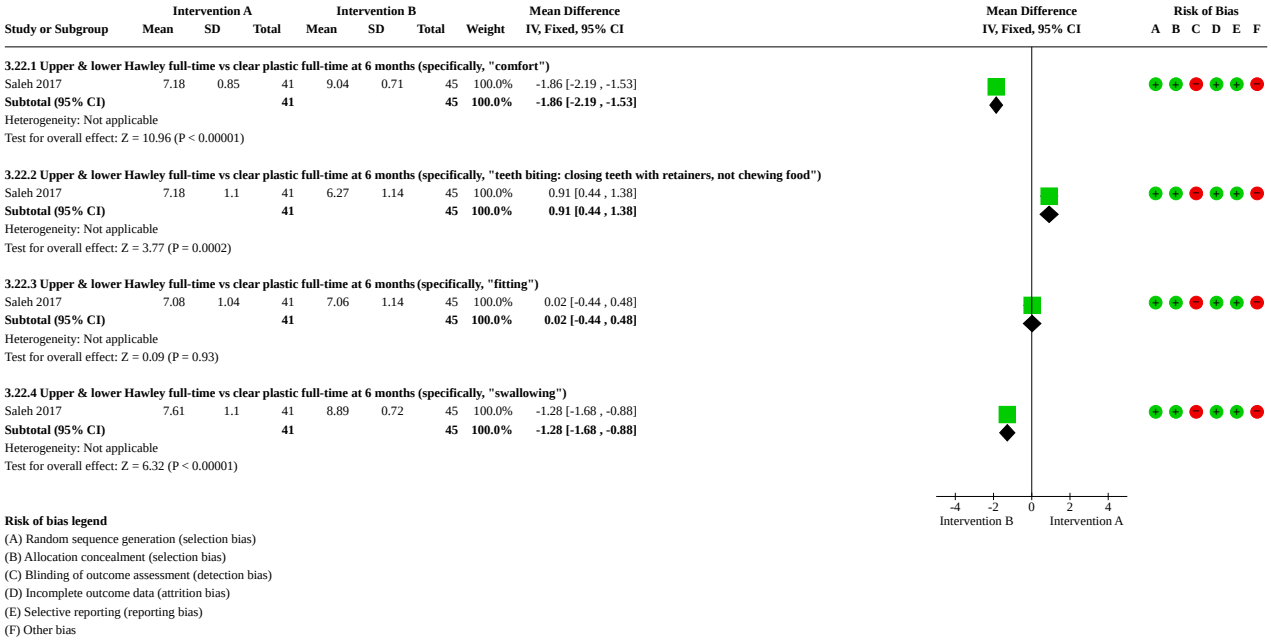
Study or Subgroup	Intervention A Mean	Intervention A SD	Intervention A Total	Intervention B Mean	Intervention B SD	Intervention B Total	Weight	Mean Difference IV, Fixed, 95% CI
3.21.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "hygiene")								
Saleh 2017	7.56	0.99	41	7.98	1.24	45	100.0%	-0.42 [-0.89 , 0.05]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-0.42 [-0.89 , 0.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.74 (P = 0.08)								
3.21.2 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "gingival irritation")								
Saleh 2017	7.71	1.25	41	8.99	0.82	45	100.0%	-1.28 [-1.73 , -0.83]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-1.28 [-1.73 , -0.83]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.56 (P < 0.00001)								



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

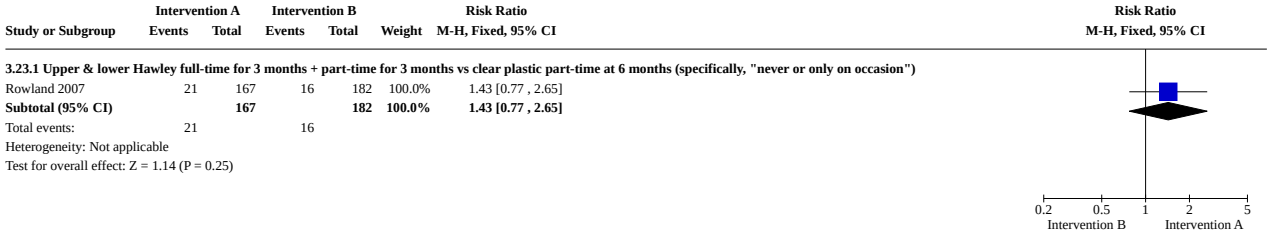
Analysis 3.22. Comparison 3: Removable versus removable retainers, Outcome 22: Patient satisfaction: comfort



Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.22.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "comfort")								
Saleh 2017	7.18	0.85	41	9.04	0.71	45	100.0%	-1.86 [-2.19 , -1.53]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-1.86 [-2.19 , -1.53]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 10.96 (P < 0.00001)								
3.22.2 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "teeth biting: closing teeth with retainers, not chewing food")								
Saleh 2017	7.18	1.1	41	6.27	1.14	45	100.0%	0.91 [0.44 , 1.38]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**0.91 [0.44 , 1.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.77 (P = 0.0002)								
3.22.3 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "fitting")								
Saleh 2017	7.08	1.04	41	7.06	1.14	45	100.0%	0.02 [-0.44 , 0.48]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**0.02 [-0.44 , 0.48]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.09 (P = 0.93)								
3.22.4 Upper & lower Hawley full-time vs clear plastic full-time at 6 months (specifically, "swallowing")								
Saleh 2017	7.61	1.1	41	8.89	0.72	45	100.0%	-1.28 [-1.68 , -0.88]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-1.28 [-1.68 , -0.88]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.32 (P < 0.00001)								

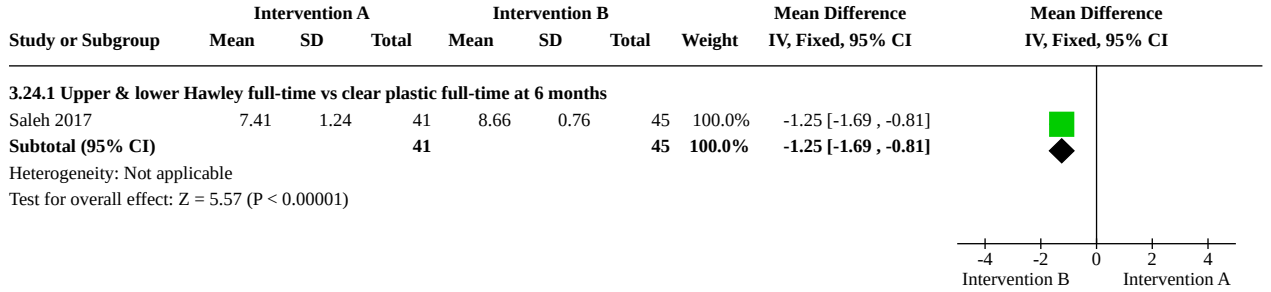
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.23. Comparison 3: Removable versus removable retainers, Outcome 23: Patient satisfaction: discomfort/pain (dichotomous)



Study or Subgroup	Intervention A Events	Total	Intervention B Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.23.1 Upper & lower Hawley full-time for 3 months + part-time for 3 months vs clear plastic part-time at 6 months (specifically, "never or only on occasion")						
Rowland 2007	21	167	16	182	100.0%	1.43 [0.77 , 2.65]
Subtotal (95% CI)		**167**		**182**	**100.0%**	**1.43 [0.77 , 2.65]**
Total events:	21		16			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.14 (P = 0.25)						

Analysis 3.24. Comparison 3: Removable versus removable retainers, Outcome 24: Patient satisfaction: speech

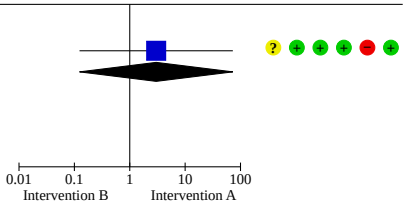


Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.24.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months								
Saleh 2017	7.41	1.24	41	8.66	0.76	45	100.0%	-1.25 [-1.69 , -0.81]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**-1.25 [-1.69 , -0.81]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.57 (P < 0.00001)								

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.25. Comparison 3: Removable versus removable retainers, Outcome 25: Patient satisfaction: speech (dichotomous)

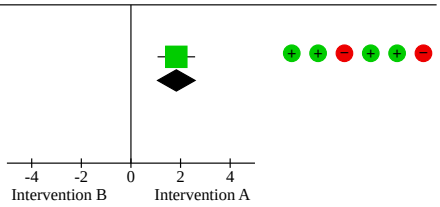
Study or Subgroup	Intervention A Events	Intervention A Total	Intervention B Events	Intervention B Total	Weight	Risk Ratio M-H, Fixed, 95% CI
3.25.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months (specifically, "number of sound distortions")						
Wan 2017	1	90	0	90	100.0%	3.00 [0.12 , 72.68]
Subtotal (95% CI)		**90**		**90**	**100.0%**	**3.00 [0.12 , 72.68]**
Total events:	1		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.68 (P = 0.50)						



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.26. Comparison 3: Removable versus removable retainers, Outcome 26: Patient satisfaction: durability

Study or Subgroup	Intervention A Mean	Intervention A SD	Intervention A Total	Intervention B Mean	Intervention B SD	Intervention B Total	Weight	Mean Difference IV, Fixed, 95% CI
3.26.1 Upper & lower Hawley full-time vs clear plastic full-time at 6 months								
Saleh 2017	8.12	1.04	41	6.29	2.35	45	100.0%	1.83 [1.07 , 2.59]
Subtotal (95% CI)			**41**			**45**	**100.0%**	**1.83 [1.07 , 2.59]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.74 (P < 0.00001)								

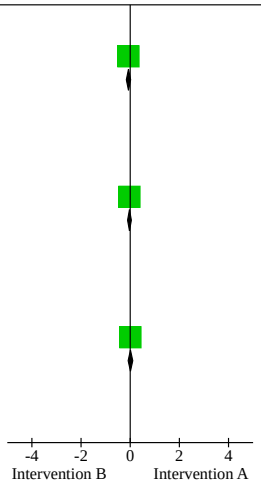


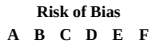
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 3.27. Comparison 3: Removable versus removable retainers, Outcome 27: Oral health: Plaque Index

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.27.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months								
Eroglu 2019	0.01	0.02	15	0.08	0.17	15	100.0%	-0.07 [-0.16 , 0.02]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.07 [-0.16 , 0.02]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.58 (P = 0.11)								
3.27.2 Upper clear plastic with mobile app vs clear plastic without mobile app at 12 months								
Al-Moghrabi 2020	0.67	0.18	29	0.7	0.1	25	100.0%	-0.03 [-0.11 , 0.05]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**-0.03 [-0.11 , 0.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.77 (P = 0.44)								
3.27.3 Lower clear plastic with mobile app vs clear plastic without mobile app at 12 months								
Al-Moghrabi 2020	0.7	0.19	29	0.69	0.16	25	100.0%	0.01 [-0.08 , 0.10]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**0.01 [-0.08 , 0.10]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.21 (P = 0.83)								

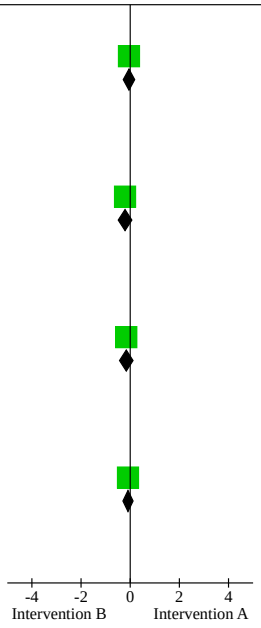


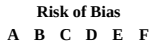


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 3.28. Comparison 3: Removable versus removable retainers, Outcome 28: Oral health: Gingival Index

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.28.1 Upper Hawley full-time vs clear plastic full-time 1.5 mm at 6 months								
Moslemzadeh 2018	-0.29	0.4	22	-0.24	0.39	22	100.0%	-0.05 [-0.28 , 0.18]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**-0.05 [-0.28 , 0.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.42 (P = 0.67)								
3.28.2 Upper Hawley full-time vs clear plastic full-time 1 mm at 6 months								
Moslemzadeh 2018	-0.29	0.4	22	-0.08	0.53	22	100.0%	-0.21 [-0.49 , 0.07]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**-0.21 [-0.49 , 0.07]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.48 (P = 0.14)								
3.28.3 Upper clear plastic full-time 1.5 mm vs clear plastic full-time 1 mm at 6 months								
Moslemzadeh 2018	-0.24	0.39	22	-0.08	0.53	22	100.0%	-0.16 [-0.43 , 0.11]
Subtotal (95% CI)			**22**			**22**	**100.0%**	**-0.16 [-0.43 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.14 (P = 0.25)								
3.28.4 Upper & lower Hawley full-time vs clear plastic full-time at 3 months								
Eroglu 2019	0.1	0.15	15	0.19	0.38	15	100.0%	-0.09 [-0.30 , 0.12]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.09 [-0.30 , 0.12]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.85 (P = 0.39)								



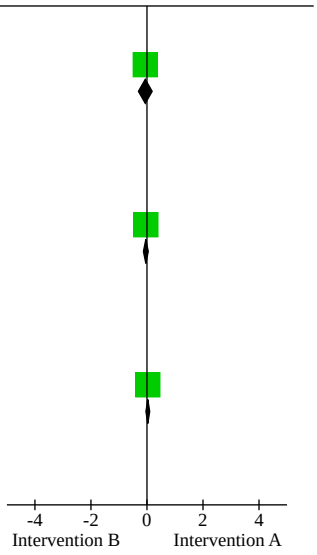


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

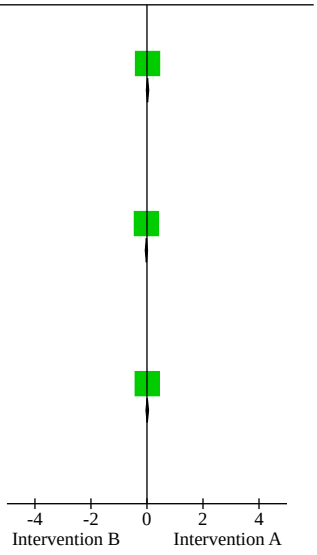
Analysis 3.29. Comparison 3: Removable versus removable retainers, Outcome 29: Oral health: probing depth

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.29.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months								
Eroglu 2019	1.46	0.38	15	1.52	0.3	15	100.0%	-0.06 [-0.31 , 0.19]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**-0.06 [-0.31 , 0.19]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.48 (P = 0.63)								
3.29.2 Upper clear plastic with mobile app vs without mobile app at 12 months								
Al-Moghrabi 2020	2.11	0.15	29	2.15	0.17	25	100.0%	-0.04 [-0.13 , 0.05]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**-0.04 [-0.13 , 0.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.91 (P = 0.36)								
3.29.3 Lower clear plastic with mobile app vs without mobile app at 12 months								
Al-Moghrabi 2020	2.06	0.12	29	2.03	0.13	25	100.0%	0.03 [-0.04 , 0.10]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**0.03 [-0.04 , 0.10]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.88 (P = 0.38)								



Analysis 3.30. Comparison 3: Removable versus removable retainers, Outcome 30: Oral health: bleeding on probing

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
3.30.1 Upper & lower Hawley full-time vs clear plastic full-time at 3 months								
Eroglu 2019	0.04	0.05	15	0.02	0.06	15	100.0%	0.02 [-0.02 , 0.06]
Subtotal (95% CI)			**15**			**15**	**100.0%**	**0.02 [-0.02 , 0.06]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.99 (P = 0.32)								
3.30.2 Upper clear plastic with mobile app vs without mobile app at 12 months								
Al-Moghrabi 2020	0.08	0.07	29	0.1	0.07	25	100.0%	-0.02 [-0.06 , 0.02]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**-0.02 [-0.06 , 0.02]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.05 (P = 0.30)								
3.30.3 Lower clear plastic with mobile app vs without mobile app at 12 months								
Al-Moghrabi 2020	0.13	0.09	29	0.12	0.07	25	100.0%	0.01 [-0.03 , 0.05]
Subtotal (95% CI)			**29**			**25**	**100.0%**	**0.01 [-0.03 , 0.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.46 (P = 0.65)								



Comparison 4. Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1 Stability: upper irregularity index	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.1.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.45, 0.65]
4.1.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	-0.40 [-0.95, 0.15]
4.1.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-0.83, 0.43]
4.1.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	-0.60 [-1.27, 0.07]
4.1.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.93, 0.33]
4.1.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-0.94, 0.54]
4.2 Stability: upper intercanine width	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.2.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.46, 0.26]
4.2.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.59, 0.39]
4.2.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.90 [-0.12, 1.92]
4.2.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	0.80 [0.10, 1.50]
4.2.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	1.00 [-0.01, 2.01]
4.2.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	0.90 [0.17, 1.63]
4.3 Stability: upper intermolar width	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.3.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.77, 0.77]
4.3.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.70, 0.90]
4.3.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.60, 1.00]
4.3.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.48, 1.08]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.3.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.40, 0.80]
4.3.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.53, 0.93]
4.4 Stability: upper arch length	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.4.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.26, 0.66]
4.4.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.29, 0.69]
4.4.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.00 [-1.11, 1.11]
4.4.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-1.30, 1.10]
4.4.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-1.23, 0.83]
4.4.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-1.46, 0.86]
4.5 Stability: lower irregularity index	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.5.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-0.62, 0.22]
4.5.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.73, 0.13]
4.5.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-1.01, 0.01]
4.5.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	-1.00 [-1.65, -0.35]
4.5.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.84, 0.24]
4.5.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	-0.70 [-1.37, -0.03]
4.6 Stability: lower intercanine width	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.6.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	1.00 [0.78, 1.22]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.6.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	1.20 [0.74, 1.66]
4.6.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.70 [0.26, 1.14]
4.6.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	1.30 [0.76, 1.84]
4.6.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.70 [0.26, 1.14]
4.6.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.55, 0.75]
4.7 Stability: lower intermolar width	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.7.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	0.60 [-0.23, 1.43]
4.7.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	0.70 [-0.28, 1.68]
4.7.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.80 [0.06, 1.54]
4.7.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	1.00 [0.07, 1.93]
4.7.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.82, 1.22]
4.7.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.20, 0.80]
4.8 Stability: lower arch length	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.8.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	0.70 [0.20, 1.20]
4.8.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	0.50 [-0.07, 1.07]
4.8.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.31, 0.71]
4.8.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	0.40 [-0.15, 0.95]
4.8.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-0.95, -0.05]

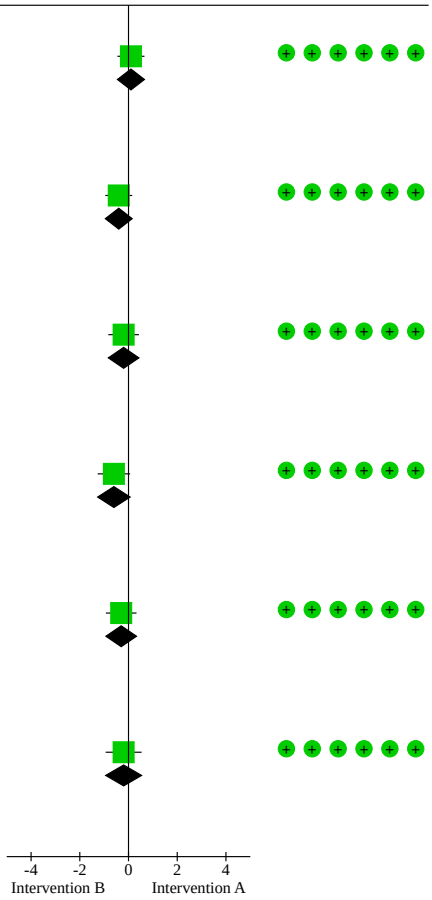
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
4.8.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.60, 0.40]
4.9 Stability: overjet	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.9.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	50	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.83, 0.23]
4.9.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	-0.80 [-1.43, -0.17]
4.9.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.81, 0.21]
4.9.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	-0.70 [-1.31, -0.09]
4.9.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.00 [-0.54, 0.54]
4.9.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.51, 0.71]
4.10 Stability: overbite	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.10.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year	1	47	Mean Difference (IV, Fixed, 95% CI)	-0.60 [-1.17, -0.03]
4.10.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years	1	47	Mean Difference (IV, Fixed, 95% CI)	-0.80 [-1.37, -0.23]
4.10.3 Upper CP and lower CTC vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	-0.20 [-0.72, 0.32]
4.10.4 Upper CP and lower CTC vs positioner at 2 years	1	46	Mean Difference (IV, Fixed, 95% CI)	-0.60 [-1.30, 0.10]
4.10.5 Upper CP and lower IPR vs positioner at 1 year	1	48	Mean Difference (IV, Fixed, 95% CI)	0.40 [-0.09, 0.89]
4.10.6 Upper CP and lower IPR vs positioner at 2 years	1	45	Mean Difference (IV, Fixed, 95% CI)	0.20 [-0.46, 0.86]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.1. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 1: Stability: upper irregularity index

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.1.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0.6	1	25	0.5	1	25	100.0%	0.10 [-0.45 , 0.65]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**0.10 [-0.45 , 0.65]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.72)								
4.1.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	0.5	0.8	24	0.9	1.1	23	100.0%	-0.40 [-0.95 , 0.15]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**-0.40 [-0.95 , 0.15]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.42 (P = 0.16)								
4.1.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	0.6	1	25	0.8	1.2	23	100.0%	-0.20 [-0.83 , 0.43]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.20 [-0.83 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.62 (P = 0.53)								
4.1.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	0.5	0.8	24	1.1	1.4	22	100.0%	-0.60 [-1.27 , 0.07]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**-0.60 [-1.27 , 0.07]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.76 (P = 0.08)								
4.1.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	0.5	1	25	0.8	1.2	23	100.0%	-0.30 [-0.93 , 0.33]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.30 [-0.93 , 0.33]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.94 (P = 0.35)								
4.1.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	0.9	1.1	23	1.1	1.4	22	100.0%	-0.20 [-0.94 , 0.54]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**-0.20 [-0.94 , 0.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.60)								

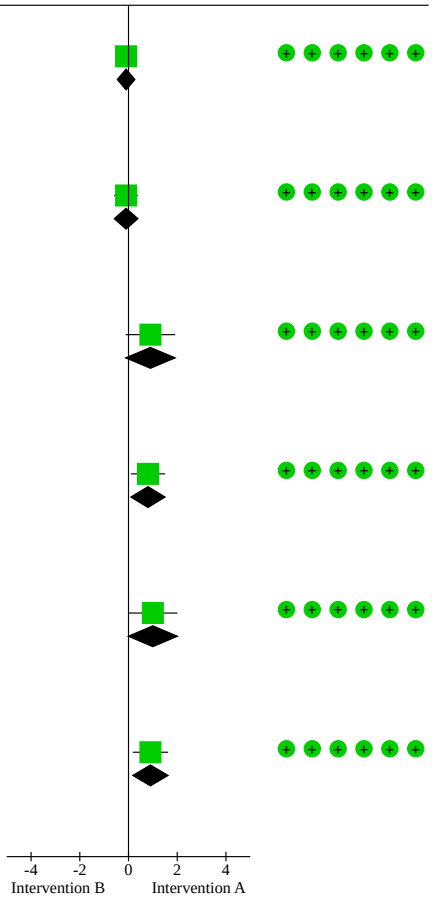


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.2. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 2: Stability: upper intercanine width

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.2.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	-0.7	0.7	25	-0.6	0.6	25	100.0%	-0.10 [-0.46 , 0.26]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**-0.10 [-0.46 , 0.26]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.54 (P = 0.59)								
4.2.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	-1	0.8	24	-0.9	0.9	23	100.0%	-0.10 [-0.59 , 0.39]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**-0.10 [-0.59 , 0.39]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.40 (P = 0.69)								
4.2.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	-0.7	0.7	25	-1.6	2.4	23	100.0%	0.90 [-0.12 , 1.92]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.90 [-0.12 , 1.92]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.73 (P = 0.08)								
4.2.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	-1	0.8	24	-1.8	1.5	22	100.0%	0.80 [0.10 , 1.50]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**0.80 [0.10 , 1.50]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.23 (P = 0.03)								
4.2.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	-0.6	0.6	25	-1.6	2.4	23	100.0%	1.00 [-0.01 , 2.01]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**1.00 [-0.01 , 2.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.94 (P = 0.05)								
4.2.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	-0.9	0.9	23	-1.8	1.5	22	100.0%	0.90 [0.17 , 1.63]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**0.90 [0.17 , 1.63]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.43 (P = 0.02)								

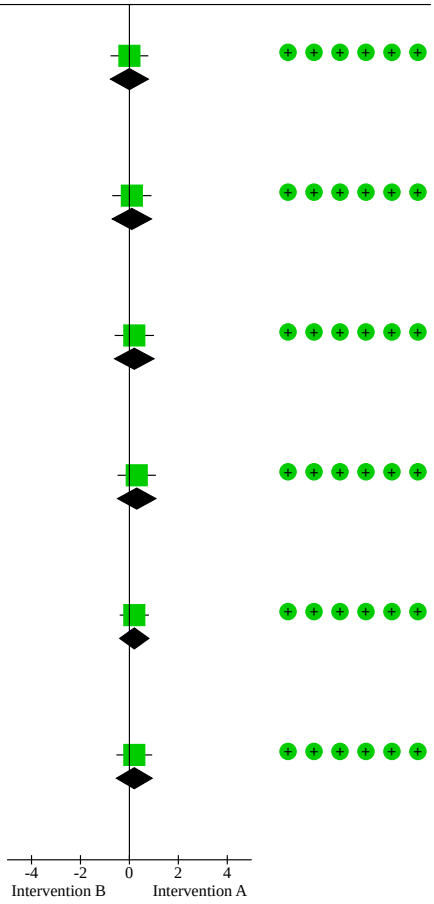


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.3. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 3: Stability: upper intermolar width

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.3.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	-0.8	1.7	25	-0.8	1	25	100.0%	0.00 [-0.77 , 0.77]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**0.00 [-0.77 , 0.77]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
4.3.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	-1.1	1.5	24	-1.2	1.3	23	100.0%	0.10 [-0.70 , 0.90]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**0.10 [-0.70 , 0.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
4.3.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	-0.8	1.7	25	-1	1.1	23	100.0%	0.20 [-0.60 , 1.00]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.20 [-0.60 , 1.00]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.49 (P = 0.63)								
4.3.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	-1.1	1.5	24	-1.4	1.2	22	100.0%	0.30 [-0.48 , 1.08]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**0.30 [-0.48 , 1.08]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								
4.3.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	-0.8	1	25	-1	1.1	23	100.0%	0.20 [-0.40 , 0.80]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.20 [-0.40 , 0.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.66 (P = 0.51)								
4.3.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	-1.2	1.3	23	-1.4	1.2	22	100.0%	0.20 [-0.53 , 0.93]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**0.20 [-0.53 , 0.93]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.54 (P = 0.59)								

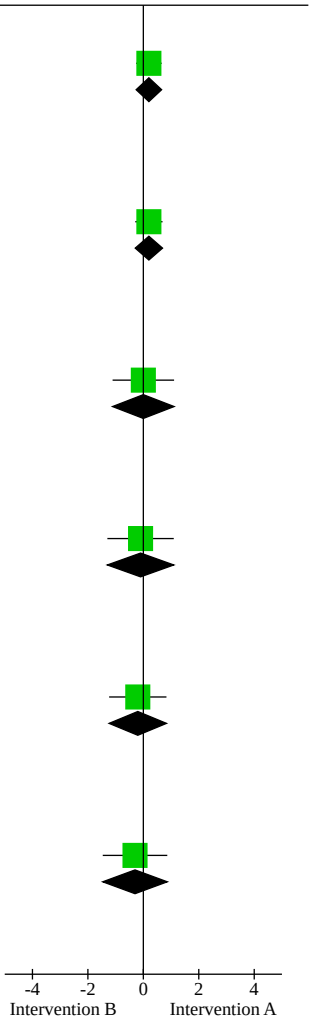


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.4. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 4: Stability: upper arch length

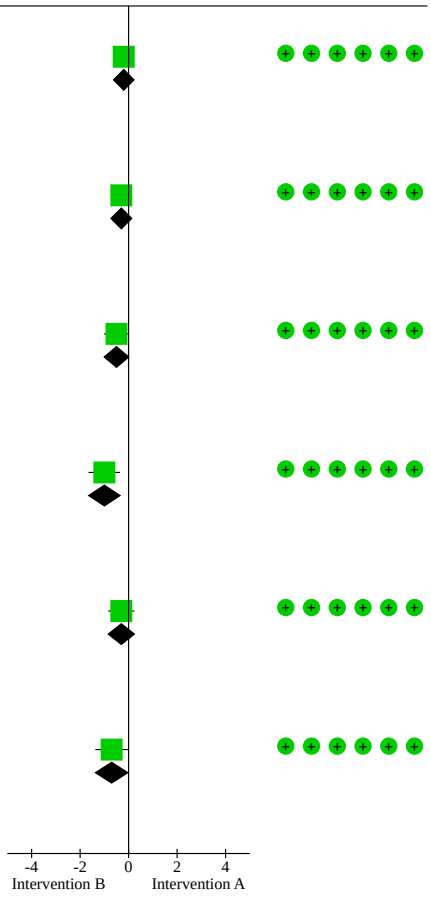
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.4.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0.3	1.1	25	0.1	0.4	25	100.0%	0.20 [-0.26 , 0.66]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**0.20 [-0.26 , 0.66]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.85 (P = 0.39)								
4.4.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	0	1	24	-0.2	0.7	23	100.0%	0.20 [-0.29 , 0.69]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**0.20 [-0.29 , 0.69]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.80 (P = 0.43)								
4.4.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	0.3	1.1	25	0.3	2.5	23	100.0%	0.00 [-1.11 , 1.11]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.00 [-1.11 , 1.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
4.4.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	0	1	24	0.1	2.7	22	100.0%	-0.10 [-1.30 , 1.10]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**-0.10 [-1.30 , 1.10]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.16 (P = 0.87)								
4.4.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	0.1	0.4	25	0.3	2.5	23	100.0%	-0.20 [-1.23 , 0.83]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.20 [-1.23 , 0.83]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.38 (P = 0.70)								
4.4.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	-0.2	0.7	23	0.1	2.7	22	100.0%	-0.30 [-1.46 , 0.86]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**-0.30 [-1.46 , 0.86]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.51 (P = 0.61)								



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.5. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 5: Stability: lower irregularity index

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.5.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0.6	0.6	25	0.8	0.9	25	100.0%	-0.20 [-0.62 , 0.22]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**-0.20 [-0.62 , 0.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.92 (P = 0.36)								
4.5.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	0.6	0.7	24	0.9	0.8	23	100.0%	-0.30 [-0.73 , 0.13]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**-0.30 [-0.73 , 0.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.37 (P = 0.17)								
4.5.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	1	0.6	25	1.5	1.1	23	100.0%	-0.50 [-1.01 , 0.01]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.50 [-1.01 , 0.01]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.93 (P = 0.05)								
4.5.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	0.6	0.7	24	1.6	1.4	22	100.0%	-1.00 [-1.65 , -0.35]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**-1.00 [-1.65 , -0.35]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.02 (P = 0.003)								
4.5.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	0.8	0.9	25	1.1	1	23	100.0%	-0.30 [-0.84 , 0.24]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.30 [-0.84 , 0.24]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.09 (P = 0.28)								
4.5.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	0.9	0.8	23	1.6	1.4	22	100.0%	-0.70 [-1.37 , -0.03]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**-0.70 [-1.37 , -0.03]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.05 (P = 0.04)								

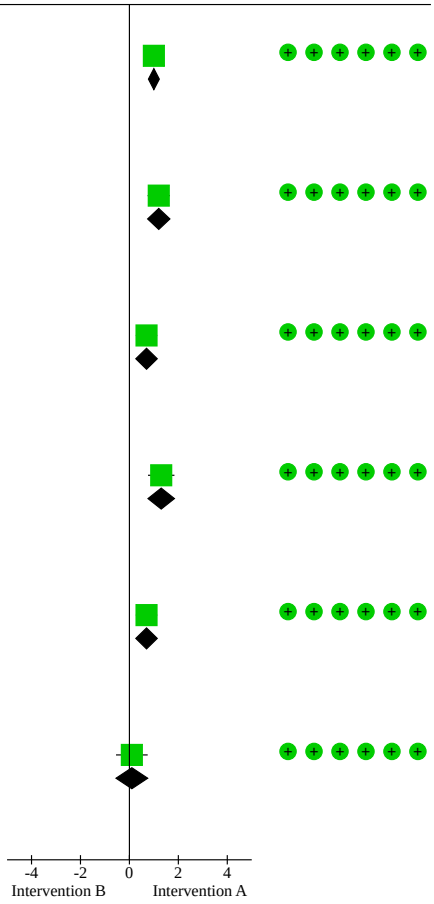


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.6. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 6: Stability: lower intercanine width

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.6.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0	0.4	25	-1	0.4	25	100.0%	1.00 [0.78 , 1.22]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**1.00 [0.78 , 1.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 8.84 (P < 0.00001)								
4.6.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	0.2	0.5	24	-1	1	23	100.0%	1.20 [0.74 , 1.66]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**1.20 [0.74 , 1.66]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.17 (P < 0.00001)								
4.6.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	0	0.4	25	-0.7	1	23	100.0%	0.70 [0.26 , 1.14]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.70 [0.26 , 1.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.13 (P = 0.002)								
4.6.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	0.2	0.5	24	-1.1	1.2	22	100.0%	1.30 [0.76 , 1.84]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**1.30 [0.76 , 1.84]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.72 (P < 0.00001)								
4.6.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	0	0.4	25	-0.7	1	23	100.0%	0.70 [0.26 , 1.14]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.70 [0.26 , 1.14]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.13 (P = 0.002)								
4.6.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	-1	1	23	-1.1	1.2	22	100.0%	0.10 [-0.55 , 0.75]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**0.10 [-0.55 , 0.75]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.30 (P = 0.76)								

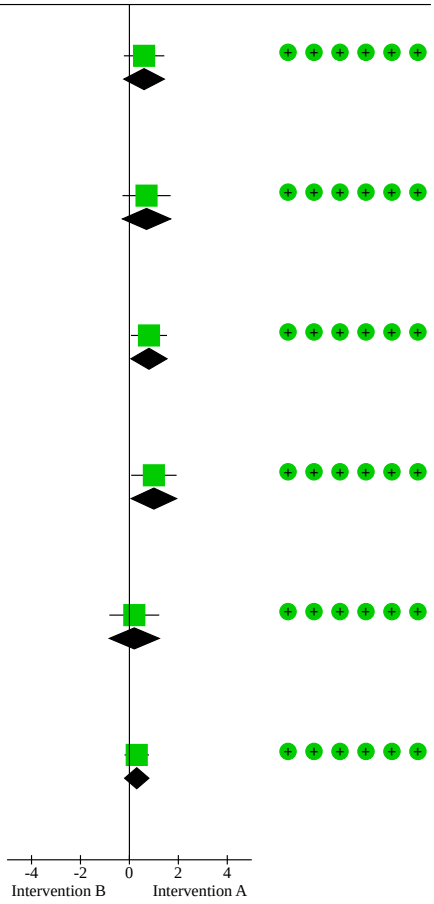


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.7. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 7: Stability: lower intermolar width

Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.7.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0	1.1	25	-0.6	1.8	25	100.0%	0.60 [-0.23 , 1.43]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**0.60 [-0.23 , 1.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.42 (P = 0.15)								
4.7.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	-0.3	1.2	24	-1	2.1	23	100.0%	0.70 [-0.28 , 1.68]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**0.70 [-0.28 , 1.68]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.40 (P = 0.16)								
4.7.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	0	0.1	25	-0.8	1.8	23	100.0%	0.80 [0.06 , 1.54]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.80 [0.06 , 1.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.13 (P = 0.03)								
4.7.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	-0.3	1.2	24	-1.3	1.9	22	100.0%	1.00 [0.07 , 1.93]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**1.00 [0.07 , 1.93]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.11 (P = 0.03)								
4.7.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	-0.6	1.8	25	-0.8	1.8	23	100.0%	0.20 [-0.82 , 1.22]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.20 [-0.82 , 1.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.38 (P = 0.70)								
4.7.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	-1	0.9	23	-1.3	0.8	22	100.0%	0.30 [-0.20 , 0.80]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**0.30 [-0.20 , 0.80]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.18 (P = 0.24)								

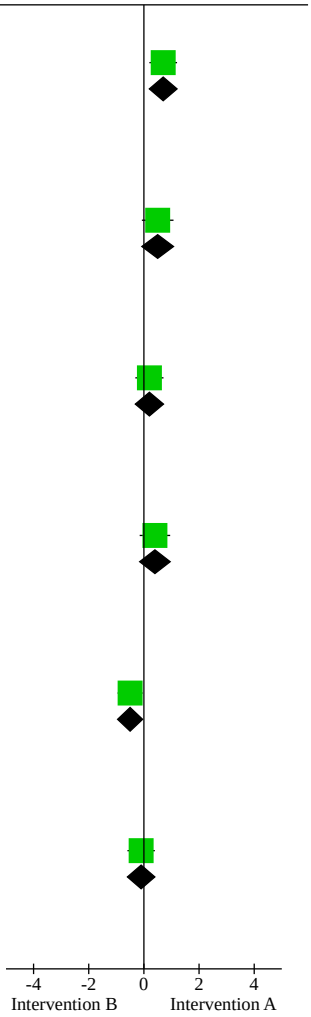


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.8. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 8: Stability: lower arch length

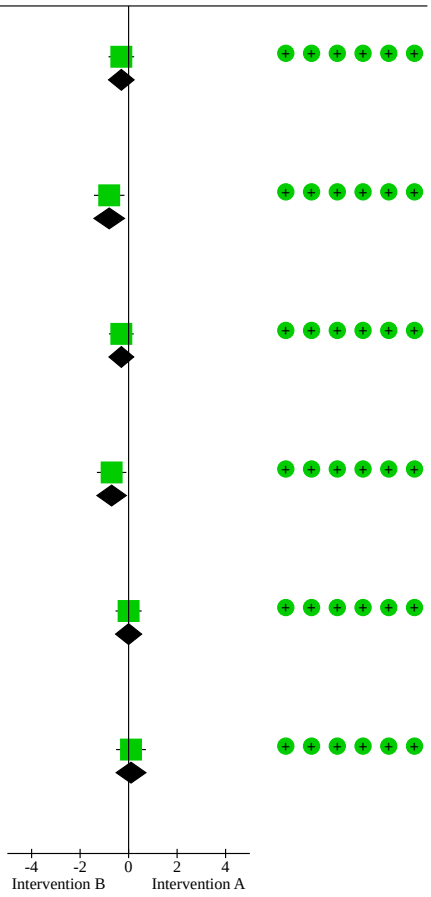
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.8.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0.7	1	25	0	0.8	25	100.0%	0.70 [0.20 , 1.20]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**0.70 [0.20 , 1.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.73 (P = 0.006)								
4.8.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	0.4	1.1	24	-0.1	0.9	23	100.0%	0.50 [-0.07 , 1.07]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**0.50 [-0.07 , 1.07]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.71 (P = 0.09)								
4.8.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	0.7	1	25	0.5	0.8	23	100.0%	0.20 [-0.31 , 0.71]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.20 [-0.31 , 0.71]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.77 (P = 0.44)								
4.8.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	0.4	1.1	24	0	0.8	22	100.0%	0.40 [-0.15 , 0.95]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**0.40 [-0.15 , 0.95]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.42 (P = 0.16)								
4.8.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	0	0.8	25	0.5	0.8	23	100.0%	-0.50 [-0.95 , -0.05]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.50 [-0.95 , -0.05]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.16 (P = 0.03)								
4.8.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	-0.1	0.9	23	0	0.8	22	100.0%	-0.10 [-0.60 , 0.40]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**-0.10 [-0.60 , 0.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.39 (P = 0.69)								



Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.9. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 9: Stability: overjet

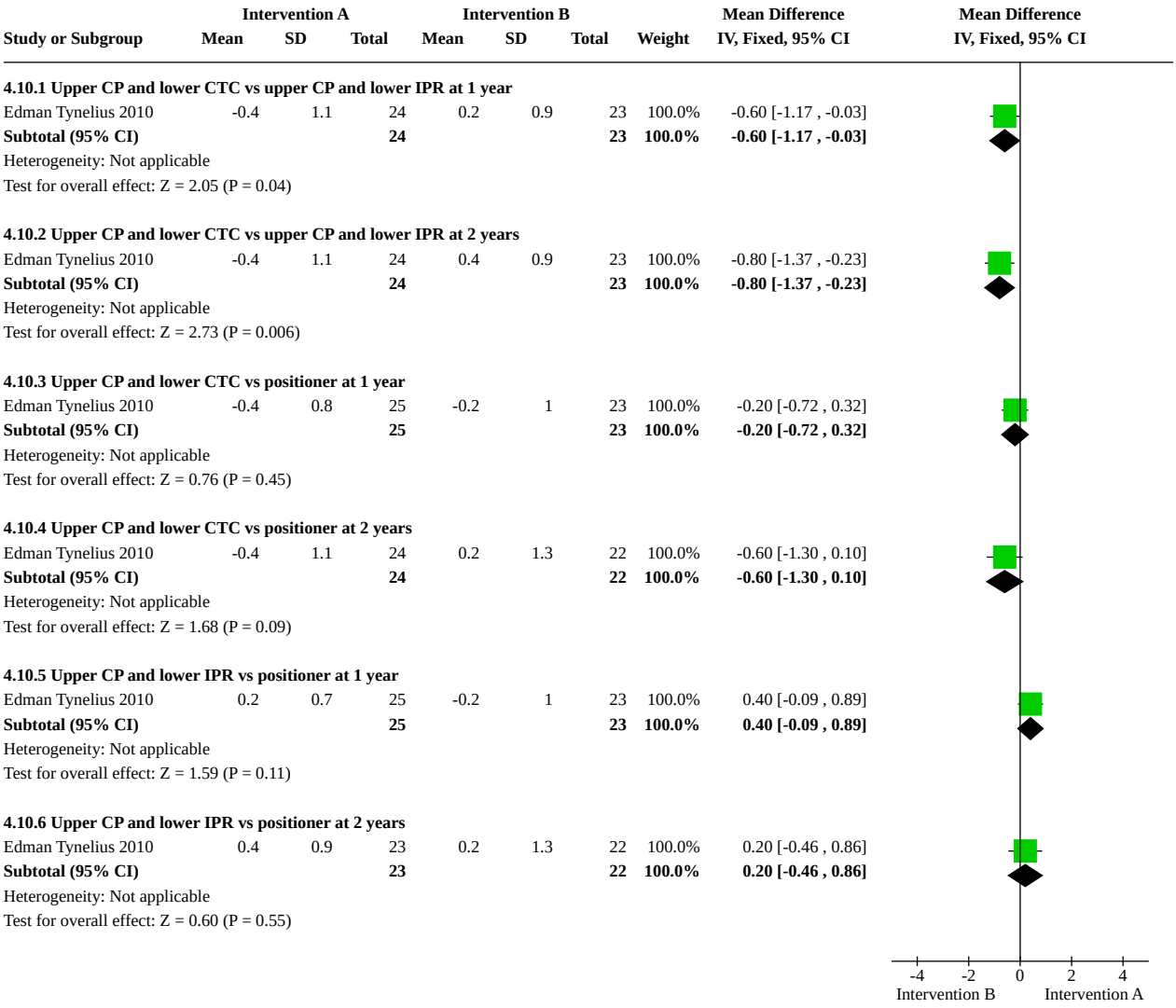
Study or Subgroup	Intervention A Mean	SD	Total	Intervention B Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
4.9.1 Upper CP and lower CTC vs upper CP and lower IPR at 1 year								
Edman Tynelius 2010	0	0.9	25	0.3	1	25	100.0%	-0.30 [-0.83 , 0.23]
Subtotal (95% CI)			**25**			**25**	**100.0%**	**-0.30 [-0.83 , 0.23]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.11 (P = 0.26)								
4.9.2 Upper CP and lower CTC vs upper CP and lower IPR at 2 years								
Edman Tynelius 2010	-0.3	1.1	24	0.5	1.1	23	100.0%	-0.80 [-1.43 , -0.17]
Subtotal (95% CI)			**24**			**23**	**100.0%**	**-0.80 [-1.43 , -0.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.49 (P = 0.01)								
4.9.3 Upper CP and lower CTC vs positioner at 1 year								
Edman Tynelius 2010	0	0.9	25	0.3	0.9	23	100.0%	-0.30 [-0.81 , 0.21]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**-0.30 [-0.81 , 0.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.15 (P = 0.25)								
4.9.4 Upper CP and lower CTC vs positioner at 2 years								
Edman Tynelius 2010	-0.3	1.1	24	0.4	1	22	100.0%	-0.70 [-1.31 , -0.09]
Subtotal (95% CI)			**24**			**22**	**100.0%**	**-0.70 [-1.31 , -0.09]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.26 (P = 0.02)								
4.9.5 Upper CP and lower IPR vs positioner at 1 year								
Edman Tynelius 2010	0.3	1	25	0.3	0.9	23	100.0%	0.00 [-0.54 , 0.54]
Subtotal (95% CI)			**25**			**23**	**100.0%**	**0.00 [-0.54 , 0.54]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
4.9.6 Upper CP and lower IPR vs positioner at 2 years								
Edman Tynelius 2010	0.5	1.1	23	0.4	1	22	100.0%	0.10 [-0.51 , 0.71]
Subtotal (95% CI)			**23**			**22**	**100.0%**	**0.10 [-0.51 , 0.71]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.32 (P = 0.75)								



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 4.10. Comparison 4: Upper clear plastic (CP) + lower fixed canine-canine (CTC) versus upper CP + lower interproximal reduction (IPR) (stripping) vs positioner, Outcome 10: Stability: overbite



Comparison 5. Upper retainer versus no retention after treatment of maxillary impacted canines

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Stability: upper irregularity index	1	63	Mean Difference (IV, Fixed, 95% CI)	-0.90 [-1.33, -0.47]
5.2 Stability: upper intercanine width	1	63	Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.73, 0.13]
5.3 Stability: upper intermolar width	1	63	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-0.53, 0.33]
5.4 Stability: upper arch length	1	63	Mean Difference (IV, Fixed, 95% CI)	0.30 [0.03, 0.57]

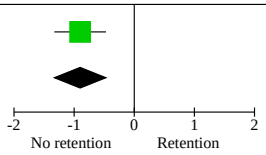
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

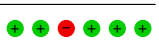
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.5 Stability: upper maximum rotation	1	63	Mean Difference (IV, Fixed, 95% CI)	-1.70 [-3.20, -0.20]

Analysis 5.1. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 1: Stability: upper irregularity index

Study or Subgroup	Retention Mean	SD	Total	No retention Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Naraghi 2021a	0.4	0.554726	32	1.3	1.090504	31	100.0%	-0.90 [-1.33 , -0.47]
Total (95% CI)			**32**			**31**	**100.0%**	**-0.90 [-1.33 , -0.47]**

Heterogeneity: Not applicable
Test for overall effect: Z = 4.11 (P < 0.0001)
Test for subgroup differences: Not applicable



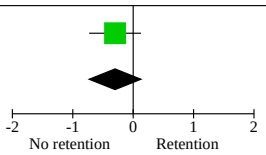


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 5.2. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 2: Stability: upper intercanine width

Study or Subgroup	Retention Mean	SD	Total	No retention Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Naraghi 2021a	0.1	0.554726	32	0.4	1.090504	31	100.0%	-0.30 [-0.73 , 0.13]
Total (95% CI)			**32**			**31**	**100.0%**	**-0.30 [-0.73 , 0.13]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.37 (P = 0.17)
Test for subgroup differences: Not applicable



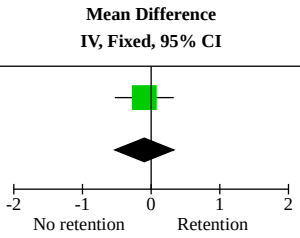


Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 5.3. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 3: Stability: upper intermolar width

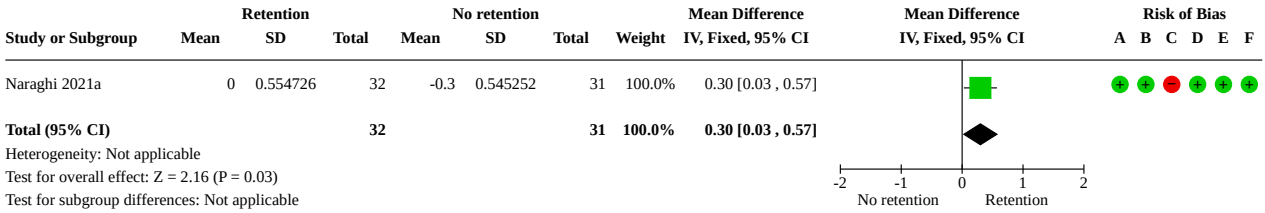
Study or Subgroup	Retention Mean	SD	Total	No retention Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Naraghi 2021a	-0.1	0.554726	32	0	1.090504	31	100.0%	-0.10 [-0.53 , 0.33]
Total (95% CI)			**32**			**31**	**100.0%**	**-0.10 [-0.53 , 0.33]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.46 (P = 0.65)
Test for subgroup differences: Not applicable



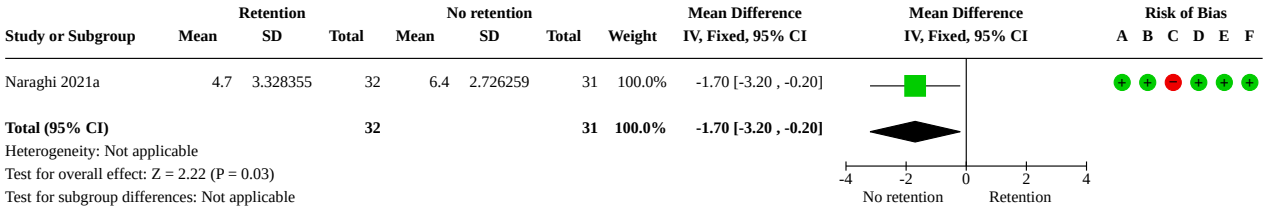
Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Analysis 5.4. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 4: Stability: upper arch length



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

Analysis 5.5. Comparison 5: Upper retainer versus no retention after treatment of maxillary impacted canines , Outcome 5: Stability: upper maximum rotation



Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of outcome assessment (detection bias)
(D) Incomplete outcome data (attrition bias)
(E) Selective reporting (reporting bias)
(F) Other bias

ADDITIONAL TABLES

Table 1. Forde 2018: patient satisfaction by retainer group

Question		Bonded retainer (n = 30)	Clear plastic retainer (n = 30)	P-value (Pearson Chi²)
Do you wear your retainers as instructed?	All the time	30 (100%)	10 (33.3%)	P < 0.001[a]
	Most of the time	0 (0%)	11 (36.7%)	
	Some of the time	0 (0%)	5 (16.7%)	
	Never	0 (0%)	1 (3.3%)	
	Missing data	0 (0%)	3 (10%)	
Has your retainer broken or come off since it was fitted?	Yes more than once	5 (16.7%)	1 (3.3%)	P = 0.001[a]

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 1. Forde 2018: patient satisfaction by retainer group *(Continued)*

	Yes once	12 (40%)	2 (6.7%)	
	No	13 (43.3%)	24 (80%)	
	Missing data	0 (0%)	3 (10%)	
Have you lost your retainer since it was fitted?	Yes more than once	0 (0%)	2 (6.7%)	P = 0.73
	Yes once	1 (3.3%)	3 (10%)	
	No	29 (96.7%)	22 (73.3%)	
	Missing data	0 (0%)	3 (10%)	
Are you embarrassed to wear your retainer?	Yes most of the time	0 (0%)	0 (0%)	P = 0.065
	Yes on occasion	0 (0%)	2 (6.7%)	
	No	30 (100%)	25 (83.3%)	
	Missing data	0 (0%)	3 (10%)	
Have you experienced any discomfort with your retainers?	Yes most of the time	0 (0%)	1 (3.3%)	P = 0.002[a]
	Yes on occasion	1 (3.3%)	10 (33.3%)	
	No	29 (96.7%)	16 (53.3%)	
	Missing data	0 (0%)	3 (10%)	
Has your retainer affected your speech?	Yes most of the time	0 (0%)	2 (6.7%)	P = 0.004[a]
	Yes on occasion	4 (13.3%)	12 (40%)	
	No	26 (86.7%)	13 (43.3%)	
	Missing data	0 (0%)	3 (10%)	
How have you found wearing your retainer in comparison to your fixed appliances?	Better	17 (56.7%)	16 (53.3%)	P = 0.096
	The same	12 (40%)	7 (23.3%)	
	Worse	0 (0%)	4 (13.3%)	
	Missing data	1 (3.3%)	3 (10%)	
How have you found cleaning your retainers?	Very easy	2 (6.7%)	12 (40%)	P = 0.001[a]
	Easy	18 (60%)	13 (43.3%)	
	Difficult	10 (33.3%)	2 (6.7%)	
	Missing data	0 (0%)	3 (10%)	

[a]Responses to patient satisfaction questionnaires following 12 months of retention.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Table 2. Thickett 2010: stability results (medians only) for full-time clear plastic vs part-time at 1 year

Outcome	Full-time at debond	Part-time at debond	P value at debond	Full-time after 1 year	Part-time after 1 year	P value after 1 year
Lower LII	0.04	0.14	0.46	0.71	0.89	0.5
Upper LII	0.21	0.14	0.14	1.08	1.09	0.8
Lower intercanine width	27.24	27.07	0.22	27.07	26.47	0.65
Lower intermolar width	34.09	33.32	0.52	34.35	33.6	0.61
Lower arch length	21.1	21.02	0.44	21.28	20.14	0.06
Upper intercanine width	35.49	34.79	0.08	34.93	34.56	0.52
Upper intermolar width	40.23	39.57	0.35	40.34	39.39	0.68
Upper arch length	24.98	24.19	0.22	25.07	25.15	0.97
Overjet	2.54	2.36	0.6	2.76	2.39	0.37
Overbite	2.86	3.31	0.14	3.14	3.74	0.05 (P < 0.05)

LII: Little's Irregularity Index in mm; **P:** probability; **vs:** versus

APPENDICES

Appendix 1. Glossary of terms relating to retention

Removable retainers in this review

Hawley retainer

This is the oldest reported removable retainer, first described in 1919. It consists of an acrylic baseplate that covers the palate or lies behind the lower teeth, and is usually held in place by attachments known as Adam's clasps that clip onto the first molar tooth that is present in each quadrant. There is typically a labial wire that crosses the labial (external) surfaces of the anterior teeth, which holds the retainer in place, and maintains the stability of the front teeth.

Clear plastic retainer

This is a clear plastic removable retainer that typically covers all the teeth. It is formed by heating up a sheet of plastic and moulding this around a model of the patient's teeth after their orthodontic treatment is complete. This moulding can be achieved by 'vacuum-forming', or by pressure. It can be modified to change the amount of occlusal coverage.

Begg retainer

This removable retainer is similar to a Hawley retainer in that it has an acrylic baseplate that covers the palate, or lies behind the lower teeth. However, it does not have clasps covering the occlusal (biting) surfaces. It is retained principally by the labial bow wire, which runs on the outside of all the teeth. The aim is to allow more occlusal settling, but retention of the appliance may not be as good as Hawley retainers due to the lack of specific clasps attached to teeth.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Positioner

A positioner is an appliance that is used to improve the interdigitation of the teeth at the end of treatment. It is typically a single piece of plastic that has been created in the laboratory to guide a person's upper and lower teeth into the ideal position. Although it is really used to 'position' the teeth, it has been used to hold the teeth in position after it has completed its positioning role. The person usually wears it only at night for a defined period.

Fixed retainers in this review

Multistrand wire fixed retainer

Multistrand wire is a flexible wire made up of stainless steel wires of smaller diameter wound together like a piece of rope. It is typically bonded to every tooth in their lingual (internal) aspect, It is usually bonded with composite resin adhesive.

Thick canine to canine retainer

This fixed retainer is made of stainless steel, but is of a thicker diameter than the multistrand and, therefore, more rigid. It is typically bonded only to the canines (hence the name). It allows easier cleaning of the incisors. The other potential advantage that is claimed is that if either attachment fails, the wearer will know (which may not be the case if one of the attachments on the multistrand fails). However, it does not bond to the incisors, so some operators feel there may be the potential for relapse of these teeth.

Fibre-reinforced composite retainer

This is a plasma-treated polyethylene woven ribbon that can be adapted to the lingual (internal) surfaces of the teeth, and then bonded to teeth using composite resin. Once composite resin adhesive is added, the retainer becomes more rigid and helps to splint the teeth together. It has the potential advantages of being easily adapted to the lingual surfaces of the teeth, and aesthetically pleasing. Some operators worry that the increased rigidity of this type of retainer may make it more prone to failure as teeth move differentially in normal function.

CAD/CAM nitinol retainer

This CAD/CAM (computer-aided designed/computer-aided manufacturing) fabricated lingual retainer wire is made from a custom-cut nickel-titanium sheet or block. This method of manufacture is used to minimise the number of bends in the material to make it less susceptible to fracture of the wire. The wire is brightened, cleaned, smoothed, polished, and made corrosion resistant. Manufacturers claim that the wires become less susceptible to microbial colonisation after polishing and they can be digitally positioned to reduce the risk of undesirable occlusal contacts.

Adjunctive procedures to reduce relapse

Interproximal reduction

Also known as interproximal stripping, this is a process of reshaping the teeth where they contact adjacent teeth to make them broader and flatter. It is hoped that this will make the teeth more stable and less likely to relapse. The reshaping can be undertaken with small abrasive strips that are placed between the teeth, either by hand or on a dental handpiece. It is usually not painful for the patient and no anaesthetic is required before, during or after. A minimal amount of enamel is removed.

Pericision

Also known as circumferential supracrestal fibreotomy, this is a surgical procedure where the fibres around the neck of the teeth are released to reduce the chances of relapse. The fibres are cut with a sharp blade under local anaesthetic.

Measures of treatment success

Little's Irregularity Index (LII)

This is an index described by a UK orthodontist called Bob Little (Little 1981). It is used to describe the irregularity of the anterior six teeth. The index measures the distance, in millimetres, between the contact points of crooked teeth, and then adds them together. Therefore, the Irregularity Index is the sum of all the displaced contacts between the anterior teeth (canine to canine).

Peer Assessment Rating (PAR)

This is an index used to provide an objective measurement of orthodontic treatment outcome. Described by Richmond 1992, it provides a score measuring any abnormalities in the occlusion - the smaller the score, the better the quality of the treatment result.

Appendix 2. Cochrane Oral Health Trials Register search strategy

Cochrane Oral Health's Trials Register is available via the Cochrane Register of Studies. For information on how the register is compiled, see https://oralhealth.cochrane.org/trials.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

From January 2013, searches of the Cochrane Oral Health Trials Register were conducted via the Cochrane Register of Studies using the search strategy below:

#1 (orthodontic*) AND (INREGISTER)
#2 ((retention or retain* or relaps* or fraenectom* or frenectom* or fiberotom* or fibreotom* or "interproximal stripping" or pericision or reproximation)) AND (INREGISTER)
#3 (#1 AND #2) AND (INREGISTER)

We undertook a previous search of the Oral Health Group Trials Register in February 2012 using the Procite software and the search strategy below:

(orthodontic* and (retention or retainer* or relapse* or "duration of retention" OR (retention and ((gingiv* or periodont*) and surg*)) or fraenectom* or frenectom* or fiberotom* or fibreotom* or "interproximal stripping" or pericision or reproximation)

Appendix 3. The Cochrane Central Register of Controlled Clinical Trials (CENTRAL) search strategy

#1 ORTHODONTICS explode all trees
#2 orthodontic*
#3 (#1 or #2)
#4 (retention or retain*)
#5 (stabilis* or stabiliz*)
#6 (fraenectom* or frenectom*)
#7 (fiberotom* or fibreotom*)
#8 (interproximal next strip*)
#9 pericision
#10 reproximat*
#11((gingiv* near surg*) or (periodont* near surg*))
#12 (retain* or retention)
#13 (#11 and #12)
#14 (#4 or #5 or #6 or #7 or #8 or #9 or #10)
#15 (#13 or #14)
#16 (#3 and #15)

Appendix 4. MEDLINE (Ovid) search strategy

1. exp ORTHODONTICS/
2. orthodontic$.mp.
3. or/1-2
4. (retention or retain$).mp.
5. (stabilise$ or stabilize$).mp.
6. (fraenectom$ or frenectom$).mp.
7. (fiberotom$ or fibreotom$).mp.
8. "interproximal stripping".mp.
9. pericision.mp.
10. reproximat$.mp.
11. ((gingiv$ or periodont$) adj4 surg$).mp.
12. (retain or retention).mp.
13. 11 and 12
14. or/4-10
15. 13 or 14
16. 3 and 15

The above subject search was linked with the highly sensitive search strategy designed by Cochrane for identifying randomised controlled trials and controlled clinical trials in MEDLINE (as described in Lefebvre 2022, box 3c).

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. groups.ab.
9. or/1-8

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

10. exp animals/ not humans.sh.
11. 9 not 10

Appendix 5. Embase (Ovid) search strategy

1. exp ORTHODONTICS/
2. orthodontic$.mp.
3. or/1-2
4. (retention or retain$).mp.
5. (stabilise$ or stabilize$).mp.
6. (fraenectom$ or frenectom$).mp.
7. (fiberotom$ or fibreotom$).mp.
8. "interproximal stripping".mp.
9. pericision.mp.
10. reproximat$.mp.
11. ((gingiv$ or periodont$) adj4 surg$).mp.
12. (retain or retention).mp.
13. 11 and 12
14. or/4-10
15. 13 or 14
16. 3 and 15

The above subject search was linked with the highly sensitive search strategy designed by Cochrane for identifying randomised controlled trials and controlled clinical trials in Embase (as described in Lefebvre 2022, box 3e).

1. Randomized controlled trial/
2. Controlled clinical study/
3. random$.ti,ab.
4. randomization/
5. intermethod comparison/
6. placebo.ti,ab.
7. (compare or compared or comparison).ti.
8. ((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab.
9. (open adj label).ti,ab.
10. ((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab.
11. double blind procedure/
12. parallel group$1.ti,ab.
13. (crossover or cross over).ti,ab.
14. ((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant $1)).ti,ab.
15. (assigned or allocated).ti,ab.
16. (controlled adj7 (study or design or trial)).ti,ab.
17. (volunteer or volunteers).ti,ab.
18. human experiment/
19. trial.ti.
20. or/1-19
21. random$ adj sampl$ adj7 ("cross section$" or questionnaire$1 or survey$ or database$1)).ti,ab. not (comparative study/ or controlled study/ or randomi?ed controlled.ti,ab. or randomly assigned.ti,ab.)
22. Cross-sectional study/ not (randomized controlled trial/ or controlled clinical study/ or controlled study/ or randomi?ed controlled.ti,ab. or control group$1.ti,ab.)
23. (((case adj control$) and random$) not randomi?ed controlled).ti,ab.
24. (Systematic review not (trial or study)).ti.
25. (nonrandom$ not random$).ti,ab.
26. "Random field$".ti,ab.
27. (random cluster adj3 sampl$).ti,ab.
28. (review.ab. and review.pt.) not trial.ti.
29. "we searched".ab. and (review.ti. or review.pt.)

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

30."update review".ab.
31.(databases adj4 searched).ab.
32.(rat or rats or mouse or mice or swine or porcine or murine or sheep or lambs or pigs or piglets or rabbit or rabbits or cat or cats or dog or dogs or cattle or bovine or monkey or monkeys or trout or marmoset$1).ti. and animal experiment/
33.Animal experiment/ not (human experiment/ or human/)
34.or/21-33
35.20 not 34

Appendix 6. Grey Literature Database: OpenGrey (www.opengrey.eu)

orthodontic and retain
orthodontic and retention

Appendix 7. The US National Institutes of Health Trials Register (ClinicalTrials.gov) search strategy

orthodontic and retain
orthodontic and retention

Appendix 8. World Health Organization International Clinical Trials Registry Platform search strategy

orthodontic and retain
orthodontic and retention

WHAT'S NEW

Date	Event	Description
22 May 2023	New citation required and conclusions have changed	30 new studies added
22 May 2023	New search has been performed	New search. Author order changed and two new authors added.

HISTORY

Protocol first published: Issue 3, 2000
Review first published: Issue 1, 2004

Date	Event	Description
26 January 2016	New citation required but conclusions have not changed	Search re-run. Review now includes 13 new studies, making a total of 15 included studies. There are 4 ongoing studies and 4 studies awaiting classification.
26 January 2016	New search has been performed	Search updated. It was decided to only include randomised controlled trials. Quasi-randomised controlled trials that had been included in previous reviews were no longer included, to reduce risk of bias and to increase confidence in the results. As a result, three studies that were previously included have now been excluded.
19 August 2008	Amended	Converted to new review format.
11 November 2005	New search has been performed	This update of the review includes one more study bringing the total number up to five. The search (May 2005) was conducted 28 months after the initial search (January 2003). The results of the review remain unchanged.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

CONTRIBUTIONS OF AUTHORS

2023 update: two new authors joined the team: Conchita Martin (CM) and Alvaro Limones (AL). CM co-ordinated the update. CM and AL independently screened new titles and abstracts and assessed eligibility, with SL resolving disagreements. CM, AL, SL, BD, DB and DM extracted the data. AL wrote the letters to the study authors. CM and AL wrote the first draft of this update and uploaded the new data into Review Manager. All review authors supervised the update of the review.

Previous versions: Simon Littlewood (SL), Declan Millett (DM), Bridget Doubleday (BD) and David Bearn (DB) wrote the protocol. All review authors contributed to writing the review. SL co-ordinated the review and wrote the letters to the authors and wrote the first draft of the review. SL worked with Cochrane Oral Health to devise the search strategy. SL, DM, BD, DB and Helen V Worthington (HW) independently assessed the eligibility of the studies, extracted data and assessed the risk of bias in the studies. HW assessed the suitability of the data from the study reports for inclusion, and advised on and helped undertake statistical analysis.

DECLARATIONS OF INTEREST

Conchita Martin: none known
Alvaro Limones: none known
Simon J Littlewood: I have previously given educational orthodontic lectures for 3M Unitek. This in no way compromises the integrity of my involvement in this review, but is disclosed for complete transparency. I was a co-author on the study Jowett 2022, which received funding for retainer materials from Forestadent and CA Digital.
Declan T Millett: none known
Bridget Doubleday: none known
David R Bearn: I received fees from Ormco Europe for lectures not directly related to this review topic.
Helen V Worthington: none known. I am an Editor with Cochrane Oral Health and was previously one of the Co-ordinating Editors. I was not involved in the editorial processing of this review.
Four of the review authors (SL, DM, BD and DB) have an interest or have been involved in studies investigating retention procedures, but this has in no way influenced the quality or objectivity of this review.
Three of the review authors (SL, DM and BD) are co-authors of included studies (Jowett 2022; Millett 2007; Rohaya 2006), so these review authors did not undertake the analysis of these studies during the review process.

SOURCES OF SUPPORT

Internal sources

- St Luke's Hospital, Bradford, UK

  Support to author Simon Littlewood
- Glasgow Dental Hospital, Glasgow, UK

  Support to authors of this review
- University of Cork, Cork, Ireland

  Support to author Declan Millett
- School of Dentistry, The University of Manchester, UK

  Support to Cochrane Oral Health and author Helen Worthington

External sources

- National Institute for Health Research (NIHR), UK

  This project was supported by the NIHR, via Cochrane Infrastructure funding to Cochrane Oral Health. The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the Evidence Synthesis Programme, the NIHR, the NHS or the Department of Health and Social Care.
- Cochrane Oral Health Global Alliance, Other

  The production of Cochrane Oral Health reviews has been supported financially by our Global Alliance since 2011 (ohg.cochrane.org/partnerships-alliances). Contributors in recent years have been: British Association for the Study of Community Dentistry, UK; British Society of Paediatric Dentistry, UK; the Canadian Dental Hygienists Association, Canada; Centre for Dental Education and Research at All India Institute of Medical Sciences, India; National Center for Dental Hygiene Research & Practice, USA; New York University College of Dentistry, USA; NHS Education for Scotland, UK; Swiss Society for Endodontology, Switzerland.

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

2023 update

- Bonding systems for retainers are considered part of the retention, so this intervention, which was not considered before, has been added in this update. Adjunctive use of a mobile application to increase the use of removable retainers was also an intervention considered in this update.
- Settling of the occlusion was excluded as an outcome in this update since it is not directly related to stability. Therefore, one study included in the previous version of this review has now been excluded (Aslan 2013).
- We changed the term 'polyethylene ribbon retainer' used in the previous version to 'fibre-reinforced composite retainer'.
- We changed the term 'thermoplastic removable retainer' used in the previous version to 'clear plastic removable retainer'.
- In the next update, we plan to expand the review to include aligners.

2016 update

- After careful consideration, we decided to exclude quasi-randomised studies to reduce the risk of bias. This resulted in the exclusion of three previously included studies (Edwards 1988; Sauget 1997; Taner 2000).
- We changed the term 'quality of life' from the original protocol to 'patient satisfaction'. This is because it better reflects the outcome being described.
- In addition, we decided to exclude studies included in the previous review that essentially compared different methods of placing fixed retainers (Alaa ElDeen 2021; Bovali 2014).

INDEX TERMS

Medical Subject Headings (MeSH)

Dental Care; *Drug-Related Side Effects and Adverse Reactions; *Gingivitis; *Orthodontic Brackets; *Periodontal Diseases

MeSH check words

Adult; Child; Humans

Copyright © 2023 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ANEXO 4.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II).

CUARTO ARTÍCULO: (Martin et al., 2023)			
Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]	**Respuesta**
Pregunta 1	-	Pregunta de investigación y criterios de inclusión: ¿Se establecieron preguntas de investigación y criterios de inclusión que incluyeron los componentes de PICO (Población, Intervención, Grupo comparador, Resultado)?	Si
Pregunta 2	**CRÍTICA**	**Protocolo registrado antes del inicio de la revisión:** ¿Se registró o publicó el protocolo de la revisión antes de empezar su realización y se han justificado las desviaciones significativas del protocolo?	**Si**
Pregunta 3	-	Diseño de estudios seleccionados: ¿Se explica por los autores cual es el diseño de estudio que se ha seleccionado para confeccionar la revisión?	Si
Pregunta 4	**CRÍTICA**	**Búsqueda exhaustiva de la literatura:** ¿Los autores de la revisión utilizaron una estrategia de búsqueda exhaustiva?	**Si**
Pregunta 5	-	Duplicación en la selección de estudios: ¿Los autores de la revisión realizaron la selección de estudios por duplicado?	Si
Pregunta 6		Duplicación en la extracción de datos: ¿Los autores de la revisión realizaron la extracción de datos por duplicado?	Si
Pregunta 7	**CRÍTICA**	**Justificación para excluir estudios individuales:** ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?	**Si**
Pregunta 8	-	Descripción de estudios incluidos: ¿Los autores de la revisión describieron los estudios incluidos con suficiente detalle?	Si
Pregunta 9	**CRÍTICA**	**Evaluación del riesgo de sesgo de los estudios primarios incluidos:** ¿Utilizaron los autores de la revisión una técnica satisfactoria para evaluar el riesgo de sesgo (RoB) en los estudios individuales que se incluyeron en la revisión?	**Si**

© Álvaro Limones Burgos, 2024

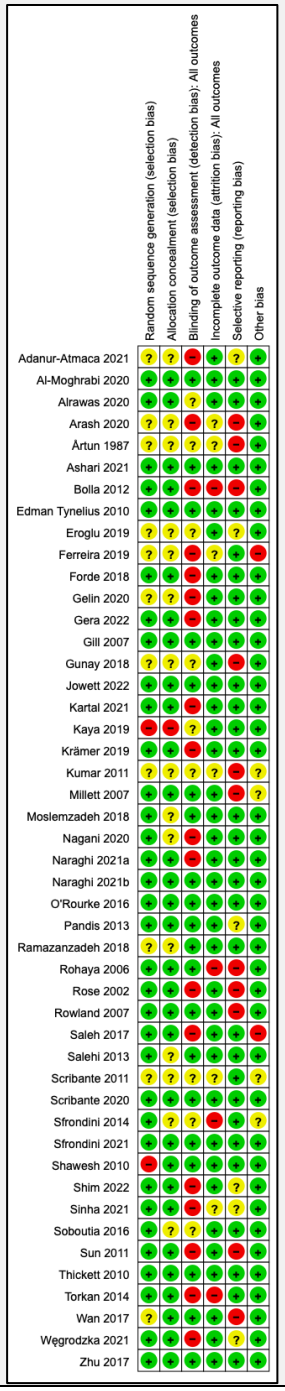
Pregunta 10	-	Fuentes de financiación: ¿Los autores de la revisión informaron sobre las fuentes de financiación de los estudios incluidos en la revisión?	Si
Pregunta 11	**CRÍTICA**	**Adecuados métodos analíticos:** Si se realizó un metaanálisis, ¿los autores de la revisión utilizaron métodos apropiados para la combinación estadística de resultados?	**Si**
Pregunta 12	-	Impacto del RoB en la metaanálisis: Si se realizó un metaanálisis, ¿los autores de la revisión evaluaron el impacto potencial de RoB de los estudios individuales sobre los resultados del metaanálisis u otra síntesis de evidencia?	Si
Pregunta 13	**CRÍTICA**	**Impacto del RoB en los resultados y discusión:** ¿Los revisores tuvieron en cuenta el RoB de los estudios primarios al interpretar/discutir los resultados de la revisión?	**Si**
Pregunta 14	-	Evaluación de la Heterogeneidad: ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?	SI
Pregunta 15	**CRÍTICA**	**Evaluación de la presencia e impacto del sesgo de publicación:** Si se realizó una síntesis cuantitativa, ¿se llevó a cabo una investigación adecuada del sesgo de publicación (sesgo de estudios pequeños) y se discutió su probable impacto en los resultados de la revisión?	**No aplica** (no fue posible estudiarlo por la escasa inclusión de estudios homogéneos)
Pregunta 16	-	Conflictos de interés: ¿Se informaron las posibles fuentes de conflicto de interés, incluyendo cualquier financiamiento recibido para llevar a cabo la revisión?	Si
VALORACIÓN FINAL DEL RIESGO DE SESGO: **ALTA CONFIANZA**			

ANEXO 4.2: Resultados de la evaluación del riesgo de sesgo (ROBIS)

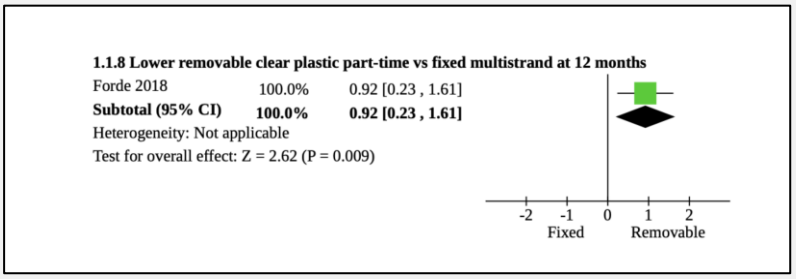
1ª FASE	2ª FASE				3ª FASE
Identificación de la pregunta de investigación y sus componentes	**Criterios de elegibilidad (Dominio 1)**	**Identificación y selección de estudios (Dominio 2)**	**Recopilación de datos y evaluación de estudios (Dominio 3)**	**Síntesis y conclusiones (Dominio 4)**	**Riesgo de sesgo global**
¿La pregunta abordada en la revisión está correctamente definida de acuerdo a sus objetivos? S/P S/PN/N/NI* Revisiones de intervención: *P: Población* *I: Intervención* *C: Comparador* *O: Variables* Revisiones de pruebas de precisión diagnóstica: *P: Población* *I: Test diagnóstico* *C: Estándar de referencia* *O: Enfermedad objetivo* Revisiones de asociaciones etiológicas o factores de riesgo: *P: Población* *I: Exposición* *C: Comparador* *O: Variables* Revisiones de estudios pronóstico: *P: Población* *I: Factor pronóstico o modelo predictivo* *O: Desenlace a predecir* Revisiones de variables e instrumentos de medición: *P: Población/Campo de conocimiento* *O: Variables e instrumentos de medición*	1.1 ¿Se ajustó la revisión a los objetivos y criterios de elegibilidad predefinidos? S/I S/PN/N/NI* 1.2 ¿Los criterios de elegibilidad eran apropiados para la pregunta de la revisión? S/I S/PN/N/NI* 1.3 ¿Eran inequívocos los criterios de elegibilidad? S/ S/PN/N/NI* 1.4 ¿Eran apropiadas las restricciones de los criterios de elegibilidad basadas en las características del estudio (por ejemplo, fecha, tamaño de la muestra, calidad del estudio, resultados medidos, etc.)? S/ S/PN/N/NI* 1.5 ¿Era adecuada alguna restricción en los criterios de elegibilidad basada en las fuentes de fuentes de información (por ejemplo, estado o formato de publicación, idioma, disponibilidad de datos)? S/I S/PN/N/NI*	2.1 ¿Incluyó la búsqueda una gama apropiada de bases de datos/fuentes electrónicas para informes publicados y no publicados? S/ S/PN/N/NI* 2.2 ¿Se utilizaron métodos adicionales a la búsqueda en bases de datos para identificar informes relevantes? S/ S/PN/N/NI* 2.3 ¿Los términos y la estructura de la estrategia de búsqueda permitieron recuperar el mayor número posible de estudios elegibles? S/I S/PN/N/NI* 2.4 ¿Fueron apropiadas las restricciones basadas en la fecha, el formato de publicación o el idioma? S/ S/PN/N/NI* 2.5 ¿Se realizaron esfuerzos para minimizar los errores en la selección de los estudios? S/ S/PN/N/NI*	3.1 ¿Se hicieron esfuerzos para minimizar los errores en la recogida de datos? S/ S/PN/N/NI* 3.2 ¿Se dispuso de suficientes características de los estudios para que tanto los autores de la revisión como los lectores pudieran interpretar los resultados? S/ S/PN/N/NI* 3.3 ¿Se recogieron todos los resultados relevantes de los estudios para su uso en la síntesis? S/ S/PN/N/NI* 3.4 ¿Se evaluó formalmente el riesgo de sesgo (o la calidad metodológica) utilizando criterios apropiados? S/P S/PN/N/NI* 3.5 ¿Se hicieron esfuerzos para minimizar los errores en la evaluación del riesgo de sesgo? S/ S/PN/N/NI*	4.1 ¿Incluyó la síntesis todos los estudios que debía? S/P S/PN/N/NI* 4.2 ¿Se comunicaron todos los análisis predefinidos o se explicaron las desviaciones? S/ S/PN/N/NI* 4.3 ¿Fue adecuada la síntesis dada la naturaleza y la similitud de las preguntas de investigación, los diseños de los estudios y los resultados de los estudios incluidos? S/ S/PN/N/NI* 4.4 ¿La variación entre estudios (heterogeneidad) fue mínima o se abordó en la síntesis? S/I S/PN/N/NI* 4.5 ¿Fueron sólidos los hallazgos, por ejemplo, como se demostró mediante el gráfico en embudo o los análisis de sensibilidad? S/P S/PN/N/NI* 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? S/ S/PN/N/NI*	A. ¿Se han abordado en la interpretación de los resultados todos los problemas señalados en los dominios del uno al cuatro? S/ S/PN/N/NI* B. ¿Se consideró adecuadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? S/ S/PN/N/NI* C. ¿Evitaron los revisores destacar los resultados en función de su significación estadística? S/ S/PN/N/NI*
¿Coincide la pregunta abordada por la revisión con la pregunta objetivo? SI/ NCIERTO/NO	**Preocupación por la especificación de los criterios de elegibilidad del estudio** ALTA/INCIERTA BAJA	**Preocupación por los métodos utilizados para identificar y/o seleccionar los estudios** ALTA/INCIERTA BAJA	**Preocupación por los métodos de recogida de datos y evaluación de los estudios** ALTA/INCIERT /BAJA	**Preocupaciones sobre la síntesis y los resultados** ALTA/INCIERT /BAJA	**Riesgo de sesgo global considerado** ALTO/INCIER BAJ /INEXISTENTE

© Álvaro Limones Burgos, 2024

ANEXO 4.3: Resultados de la evaluación de la certeza de la evidencia científica (GRADE).

Dominios	**Consideraciones** [Para los ejemplos de esta evaluación fue utilizada la variable principal – Índice de irregularidad, evaluada a 6 meses de seguimiento]	**Variación del nivel de certeza de la evidencia**
Riesgo de sesgo [Dominio 1]	 La herramienta original para evaluar el riesgo de sesgo (Higgins et al., 2011) no reveló riesgos significativos de sesgo en las dimensiones evaluadas. Los resultados de esta evaluación se presentan en la siguiente figura: Los altos riesgos de sesgo más freuentemente identificados fueron por la realización de asignaciones sin el adecuado cegamiento, la ausencia de equivalencia entre los grupos de intervención antes del tratamiento, evaluadores de resultados que no estaban cegados y que podrían haber influenciado la interpretación de los datos. Además, se detectaron problemas en el reporte de resultados, tales como la omisión de resultados preespecificados, la adición de resultados no especificados inicialmente y el reporte parcial de datos. Otros factores que contribuyeron al sesgo incluyeron el no informe de seguimiento de las directrices CONSORT. Por todo lo cual, se decide degradar la evidencia un nivel por riesgo de sesgo.	↓ 1 nivel

© Álvaro Limones Burgos, 2024

Inconsistencia [Dominio 2]	No fue posible estudiar la heterogeneidad cuando solo había un único estudio incluido en cualquiera de las variables reportadas y consideradas críticas, como se puede ver por ejemplo en la siguiente figura:  	No aplicable
Carácter indirecto de la evidencia [Dominio 3]	No se detectaron caracteres indirectos de la evidencia en ninguno de los aspectos del formato PICO: 1. La población de estudio reflejo exactamente la población de interés sobre la que se realizaron inferencias. 2. Las intervenciones fueron reproducibles en los entornos clínicos sobre los que se realizaron las inferencias. 3. Los resultados fueron evaluados a través de variables que tenían un significado directo y práctico para los pacientes (estabilidad, supervivencia, satisfacción del paciente y efectos adversos). 4. Los estudios incluidos en el metaanálisis eran controlados, ofreciendo comparaciones directas entre las intervenciones.	No aplica
Imprecisión [Dominio 4]	El estudio no alcanzó el tamaño óptimo de información fijado en 33 participantes por grupo. Esto se calculó basándose en los resultados del estudio de Forde et al., 2018, que encontró una diferencia de medias de 0.92 ± 0.69, con una relación de muestreo de 1:1, un nivel de significancia (alfa) del 5%, y un poder estadístico del 80%. Debido a esto, se redujo un nivel por imprecisión. Solo fue incluido un único estudio.	↓ 2 nivel en todas las variables críticas
Sesgo de Publicación [Dominio 5]	No es posible estudiar el sesgo de publicación cuando solo hay un único estudio incluido.	No aplica
Gradiente dosis-respuesta [Dominio 6]	No aplicable.	No aplicable
Gran magnitud del efecto	La magnitud combinada del efecto no presento una gran magnitud del efecto.	No aplicable

© Álvaro Limones Burgos, 2024

[Dominio 7]		
Factores de confusión plausibles [Dominio 8]	No se detectaron factores de confusión plausible que aumentase nuestra confianza en las estimaciones.	No aplicable
Computo total de la evaluación	Disminución de la certeza en tres niveles debido al riesgo de sesgo e imprecisión.	↓ 3 nivel en todas las variables críticas ⊕◯◯◯ MUY BAJA

© Álvaro Limones Burgos, 2024

ANEXO 4.4: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA).

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
TÍTULO			
Título	1	Identificar en el título del manuscrito que se trata de una revisión sistemática.	p. 1
RESUMEN			
Objetivos	2.1	Proporcionar una declaración explícita de los principales objetivos de la revisión.	p.1
Criterios de Inclusión y exclusion	2.2	Especificar los criterios de inclusión y exclusión de la revisión.	p.1
Fuentes de información	2.3	Especificar las fuentes de información (por ejemplo, las bases de datos o registros) utilizados para identificar los estudios y la fecha de la última búsqueda.	p.1
Riesgo de sesgo	2.4	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos.	/
Métodos de la síntesis de los resultados	2.5	Especifique los métodos utilizados para presentar y sintetizar los resultados.	p. 1-2
Estudios incluidos	2.6	Dar el número total de estudios y participantes incluidos y resumir las características pertinentes de los estudios.	p. 2
Síntesis de los resultados	2.7	Indicar las variables de estudio. Si se ha realizado un metaanálisis, notifique la estimación del efecto o exposición y su IC. Si se comparan dos o más grupos, indíquen la dirección del efecto (es decir, que grupo es favorecido).	p. 2
Degradación de la evidencia	2.8	Proporcionar un breve resumen de la degradación de la evidencia (por ejemplo, por riesgo de sesgo, incoherencia o impresión.	p. 2
Interpretación	2.9	Proporcionar una interpretación general de los resultados e implicaciones importantes.	p. 2-3
Financiación	2.10	Especificar la principal fuente de financiación de la revisión.	/
Registro	2.11	Proporcionar el nombre del registro y el número de registro.	p. 1
INTRODUCIÓN			
Fundamento	3	Describir la justificación de la revisión en el contexto de los conocimientos existentes.	p. 15
Objetivos	4	Proporcionar una declaración explícita de los objetivos de la revisión.	p. 15
MÉTODOS			
Criterios de elegibilidad	5	Especificar los criterios de inclusión y exclusión de la revisión y cómo se agruparon los estudios para la síntesis.	p. 15-16
Fuentes de información	6	Especificar todas las bases de datos, registros, sitios web, organizaciones, listas de referencia y otras fuentes consultadas para identificar estudios. Especificar la fecha de la última revisión.	p. 16
Estrategia de búsqueda	7	Presentar las estrategias de búsqueda completa para al menos la base de datos principal, indicar filtros u otras estrategias de búsqueda utilizadas.	p.16
Proceso de selección	8	Especificar los métodos utilizados para la selección de estudios.	p. 16-17

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Recogida de datos	9	Especifique los métodos utilizados para recopilar datos de los estudios, incluyendo cuántos revisores recogieron datos de cada informe.	p. 17
Datos extraidos	10a	Listar y definir todas las variables críticas.	p. 17-18
	10b	Listar y definir todas las demás variables secundarias. Describa cualquier suposición sobre cualquier información que falta o no esté clara.	p. 17-18
Métodos de evaluación del riesgo de sesgo	11	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos, incluyendo detalles de la herramienta o herramientas utilizadas, cuántos revisores evaluaron cada estudio y si trabajaban independientemente, y si se utilizaron procesos de automatización.	p. 18
Medidas del efecto	12	Especifícar para cada variable la medida(s) del efecto (por ejemplo. relación de riesgo, diferencia media) utilizada en la síntesis o presentación de resultados.	p p. 17-18
Métodos de síntesis	13a	Describir los procesos utilizados para decidir qué estudios eran elegibles para cada síntesis (por ejemplo: tabular las características de intervención del estudio y comparar con los grupos previstos para cada síntesis).	p. 57-110
	13b	Describir cualquier método necesario para preparar los datos de presentación o síntesis, como el tratamiento de las estadísticas de datos faltantes o las conversiones de datos.	p. 17-18
	13c	Describir cualquier método utilizado para tabular o visualizar los resultados de los estudios y la síntesis individuales.	p. 17-18
	13d	Describir cualquier método utilizado para sintetizar los resultados y proporcionar una justificación. Si se realizó un metaanálisis, describa el(s) modelo(s), método(s) para identificar la presencia y extensión de la heterogeneidad estadística, y paquetes de software utilizado.	p. 17-18
	13e	Describa cualquier método utilizado para explorar posibles causas de heterogeneidad entre los resultados del estudio (por ejemplo, análisis de subgrupos, metaregresión).	p. 18
	13f	Describir cualquier análisis de sensibilidad realizado para evaluar la solidez de los resultados sintetizados.	p. 18
Implicaciones del riesgo de sesgo por datos faltantes	14	Describir cualquier método han utilizado para evaluar el riesgo de sesgo debido a la falta de datos (a partir de sesgos informativos).	p. 18
Evaluación de la certeza de la evidencia	15	Describir cualquier método utilizado para evaluar la certeza (o la confianza) de la evidencia.	p. 18
RESULTADOS			
Selección de estudios	16a	Describir los resultados del proceso de búsqueda y selección, idealmente utilizando un diagrama de flujo.	p. 19
	16b	Presentar los estudios excluidos y sus razones.	p. 23, 110-113
Características de los estudios incluidos	17	Presentar cada estudio incluido y sus características.	p. 57-110
Riesgo de sesgo en los estudios incluidos	18	Presentar el riesgo de sesgo para cada estudio incluido.	p. 23-27

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Resultados de los estudios individuales	19	Para todas las variables presentar la estimación del efecto y su IC.	p. 22-23
Resultados de la síntesis	20a	Resumir brevemente las características y el riesgo de sesgo de los estudios incluidos.	p. 23-27
	20b	Si se ha realizado un metaanálisis, presente para cada una de las estimaciones medidas de heterogeneidad estadística. Si se comparan los grupos, describa la dirección del efecto.	p. 27-40
	20c	Informar de posibles causas de heterogeneidad.	p. 32, 42-44
	20d	Informar de los análisis de sensibilidad realizados para evaluar la solidez de los resultados sintetizados.	p. 44
Evaluación del riesgo de sesgo por datos faltantes	21	Informar de la evaluación del riesgo de sesgo debido a datos faltantes.	p. 23-27
Certeza de la evidencia	22	Evaluaciones de la certeza (o confianza) de la evidencia para cada variable evaluada.	p. 5-14
DISCUSIÓN			
Discusión	23a	Proporcionar una interpretación general de los resultados en el contexto científico actual.	p. 40-45
	23b	Examinar cualquier limitación de la evidencia incluida en la revisión.	p. 44
	23c	Debatir las limitaciones de los procesos de revisión utilizados.	p. 44
	23d	Debatir las implicaciones de los resultados de la práctica, la política y la futura investigación.	p. 45
OTRA INFORMACIÓN			
Registro y protocolo	24a	Proporcionar información del registro de la revisión, incluyendo el nombre del registro y el número de registro, o afirmar que la revisión no está registrada.	p. 1
	24b	Indicar dónde puede accederse al protocolo de revisión o afirmar que no se ha elaborado un protocolo.	p. 1
	24c	Describir y explicar las modificaciones de la información facilitada en el registro o en el protocolo.	p. 254
Apoyo	25	Describa las fuentes de ayuda financiera o no financiera para la revisión, y el papel de los patrocinadores en la revisión.	p. 253
Intereses de competencia	26	Declare cualquier interés competitivo de los autores de revisión.	p. 253
Disponibilidad de datos, código y otros materiales	27	Informe sobre cuál de los siguientes está disponible públicamente y dónde se pueden encontrar: plantillas de recogida de datos; datos extraídos de estudios incluidos; datos utilizados para todos los análisis; o cualquier otro material utilizado en la revisión.	p. 1 y 15

© Álvaro Limones Burgos, 2024

ANEXO V: Manuscrito de la quinta revisión sistemática

(Limones et al., 2022)

© Álvaro Limones Burgos, 2024

THE JOURNAL OF PROSTHETIC DENTISTRY

SYSTEMATIC REVIEW

Outcome measurements and quality of randomized controlled clinical trials of tooth-supported fixed dental prostheses: A systematic review and qualitative analysis

Alvaro Limones, DDS, MS, PhD,[a] Alicia Celemín-Viñuela, DDS, PhD,[b] Marta Romeo-Rubio, DDS, PhD,[c] Raquel Castillo-Oyagüe, DDS, PhD,[d] Miguel Gómez-Polo, DDS, PhD,[e] and Juan Antonio Martínez Vázquez de Parga, DDS, PhD[f]

ABSTRACT

Statement of problem. The lack of consensus regarding a standardized set of outcome measurements and noncompliance with current reporting guidelines in clinical trials of tooth-supported fixed dental prostheses (FDPs) hamper interstudy comparability, compromise scientific evidence, and waste research effort and resources in prosthetic dentistry.

Purpose. The primary objective of this systematic review was to identify all primary and secondary outcome measurements assessed in randomized controlled trials (RCTs) of tooth-supported FDPs. Secondary objectives were to assess their methodological quality by using the Cochrane Collaboration's risk of bias tool (RoB, v2.0) and their reporting quality by means of a standardized 16-item CONSORT assessment tool through published reports.

Material and methods. An electronic search was conducted in MEDLINE, EMBASE, and Cochrane library to identify all RCT-related articles published in the past 10 years. Differences in RoB were tested with the Pearson chi-square test, and those in CONSORT score with the Student *t* test.

Results. A total of 64 RCTs from 79 publications were deemed eligible. The diversity of outcome measures used in the field is apparent. Twenty percent of the included studies had a low RoB, 79% showed some concerns, and 1% had a high RoB. The mean ±standard deviation CONSORT compliance score was 22.56 ±3.17. Trials adhered to the CONSORT statement reported lower RoB than those that did not adhere ($P<.001$). RCTs with a low RoB reported more comprehensive adherence to CONSORT guidelines than those with some concerns (MD 4 [95% CI 1.52-6.48]; $P=.004$).

Conclusions. A standardized core outcome reporting set in clinical research on tooth-supported FDPs remains evident. Adherence to the CONSORT statement continues to be low, with some RoB concerns that can be improved. (J Prosthet Dent 2022;■:■-■)

Evidence-based dentistry (EBD) has been adopted worldwide to inform daily clinical practice, education, and policy for the past 30 years driving new research and delivering relevant, trustworthy, and impactful clinical decision-making.[1] With the advent of the digital era, new materials, dental techniques, and workflows for tooth-supported fixed dental prostheses (FDPs) are expected to be assessed by high-quality randomized controlled trials (RCTs), the standard in clinical research because of their ability to directly assess cause-effect relationships with minimal bias or confounding factors and to provide evidence for safe and efficient in-

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
[a]Student, Assistant Professor, Department of Conservative & Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain.
[b]Professor, Department of Conservative & Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain.
[c]Professor, Department of Conservative & Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain.
[d]Cathedratic Professor, Department of Conservative & Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain.
[e]Professor, Department of Conservative & Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain.
[f]Professor, Department of Conservative & Prosthetic Dentistry, Faculty of Dentistry, Complutense University of Madrid (UCM), Madrid, Spain.

Clinical Implications

Adherence to the CONSORT statement when reporting a clinical trial of tooth-supported FDPs has been proven to reduce the risk of bias.

terventions.[2] Nevertheless, sometimes the lack of homogeneity, clarity, transparency, and/or scientific rigor according to existing standards can lead to confusing, inaccurate, or misleading conclusions in the field.[3]

Over the past 2 decades, the development of international registers for clinical trial protocols and reporting guidance, such as the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2013 Declaration[2] and the 2010 Consolidated Standards of Reporting Trials (CONSORT) Statement,[4,5] has reduced waste in biomedical research and created an open research environment of transparency and accountability.[3] A prospective registration of the protocol on an international public platform recognized by the World Health Organization (WHO) and the International Committee of Medical Journal Editors (ICMJE) becomes mandatory to avoid the risks of selectively reporting positive results (scientific reasons) and to promote patient safety (ethical reasons).[6] For example, registration when submitting a manuscript for publication does not serve either of these purposes.

Particularly in the field of tooth-supported FDPs, a standardized core outcome reporting set (SCORES) in clinical research is still lacking, causing inconsistencies in reporting and difficulties in comparing and combining the findings in meta-analyses.[7,8] So far, the most comprehensive clinical criteria used to evaluate direct or indirect restorations have been the United States Public Health Service (USPHS)/Ryge criteria,[9] the California Dental Association (CDA) criteria,[10] and the World Dental Federation (FDI)/Hickel criteria.[11,12] Modifications of the USPHS[13] and CDA criteria[14] have been used to include more categories in the assessment. Cagidiaco et al[15] recently proposed a novel prosthodontic functional index for teeth (FIT) for overall assessment. However, many researchers have selected specialized outcome measurements in their particular area.[16-21]

This systematic review aimed to identify issues related to outcome measurements, study design, and reporting in the field of clinical research on tooth-supported FDPs to ensure the highest methodological quality, homogeneity, and representativeness of the daily clinical practice. Multiple hypotheses were generated to explore associations. Null hypotheses were that no difference in RoB would be identified between RCTs following CONSORT statement and RCTs not following CONSORT statement and that no difference in compliance with CONSORT statement items would be identified between RCTs with low RoB and those with some RoB concerns.

MATERIAL AND METHODS

The full trial protocol was registered at the Open Science Framework database (DOI: 10.17605/OSF.IO/CW5F7). This systematic review was reported by the 2020 PRISMA statement for reporting systematic reviews (Supplemental Appendix 1, available online).[22-25]

The primary objective of this systematic review was to identify all primary and secondary outcome measurements assessed in the RCTs of tooth-supported FDPs. Secondary objectives were to assess their methodological quality by using the Cochrane Collaboration RoB tool (v2.0)[26] and their reporting quality through a standardized 16-item CONSORT assessment tool designed by Loguercio et al[27] (Supplemental Appendix 1, available online) in those studies that reported on its use.

A researcher (A.L.) experienced in database searching, supported by a medical librarian, designed a comprehensive search strategy for published studies and following reporting guidelines for systematic reviews[28] by using a validated research methodology filter for RCTs (with 97% specificity and 93% sensitivity) (Supplemental Appendix 1, available online).[29] The search engines used were MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials (CCRCT) from the inception of each database to March 15, 2022. A search of the non–peer-reviewed literature was performed in the OpenGrey database. An email alert service was used for MEDLINE to identify newly published studies that matched the search strategy. Systematic reviews of RCTs were scrutinized to identify previously unidentified articles.

Inclusion criteria were RCT-related articles published in high-impact dental journals (quartiles Q1 to Q2) indexed in the 2019 Journal Citation Reports (JCR) category "Dentistry, Oral Surgery & Medicine" (Supplemental Appendix 1, available online) in the past 10 years, whose participants had undergone any tooth-supported FDP, corresponding to crowns, fixed partial dentures, veneers, or indirect inlays or onlays that compared different materials, prosthetic designs, operative dental techniques, or workflows.

The eligibility criteria, data extraction, RoB assessment, and CONSORT compliance were independently peer-reviewed (A.L., M.R.-R.) by using a systematic review software program (DistillerSR; Evidence Partners Inc). Any ambiguities were resolved by the coauthors through discussion. Agreement was measured and Kappa scores calculated. Data were introduced into a pspreadsheet (Excel, v15.17; Microsoft Corp). Missing data relevant to this review were requested from the corresponding authors via email. When studies had

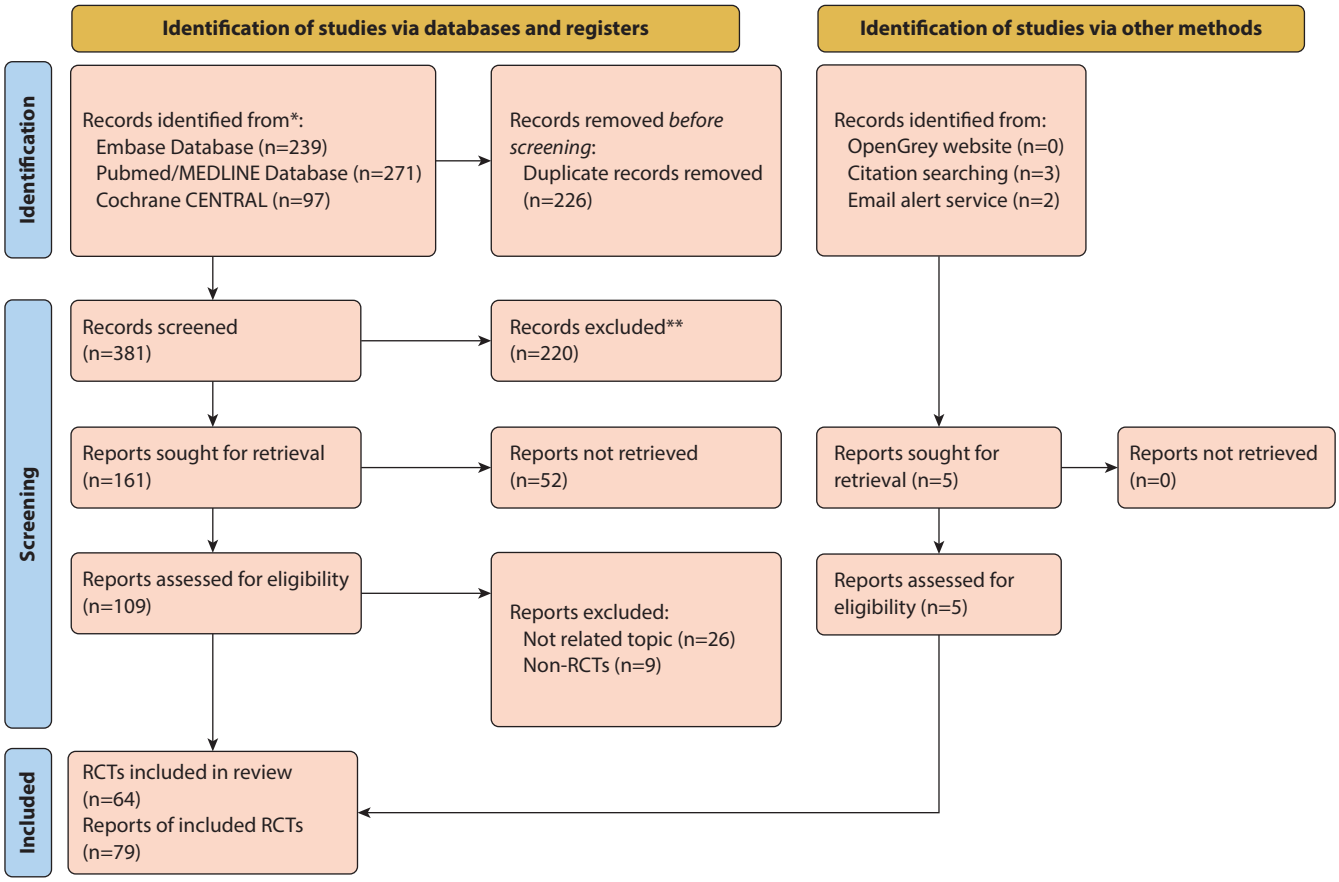


Figure 1. Search and article-selection process.

multiple publications, a unique identifier based on the first publication was considered. It was determined that none of the studies included in this review were retracted because of error or fraud. The following elements were recorded: first author name, year of publication, scientific journal, adherence to CONSORT statement, ethics committee approval, registration status, trial design, setting, sample size calculations, primary outcomes, random sequence generation, allocation concealment, blinding, intention-to-treat or per-protocol analysis, dropouts, funding disclosure, and conflict of interest.

Statistical analysis was performed with a statistical software program (IBM SPSS Statistics, v27; IBM Corp). Categorical variables were presented as frequencies by percentages, while quantitative variables were presented as mean ±standard deviation (SD). Differences were assessed with the Pearson chi-square test and the Student *t* test for independent samples (α=.05). A Shapiro-Wilks test was applied to evaluate the normality of the sample.

RESULTS

The overall numbers of records screened, identified, selected, excluded, and included are presented in a PRISMA flow diagram (Fig. 1). A total of 64 potentially relevant trials from 79 publications (agreement=94.50%; kappa=.867; 95% CI=0.764 to 0.970) were included.[15-21,30-101] The outcome measurements identified in RCTs to assess tooth-supported FDPs are shown in Table 1. The frequency of current clinical criteria being used to assess the quality of FDPs is shown in Table 2. The characteristics of the included studies, the RoB assessment, and the adherence to the CONSORT statement are shown in the Supplemental Appendix 1, available online. A total of 20% of the studies had a low RoB, 79% had some concerns, and 1% had high RoB (Supplemental Appendix 1, available online). Characteristics of the included trials in terms of their methodological design are shown in Table 3. A total of 72% of the studies did not mention using the CONSORT reporting statement, 69% did not describe a sample size calculation, and 50% did not describe the primary outcome of the research. The Shapiro-Wilks test revealed a normal distribution of the data for the low-RoB group (P=.075) and in the some-RoB-concerns group (P=.068). The Pearson chi-square test indicated significant differences in RoB among trials adhering to the CONSORT statement, reporting lower RoB, and trials not adhering to the CONSORT statement (P<.001; Table 4). The Student *t* test revealed that

Table 1. Outcome measurements reported in randomized controlled trials of tooth-supported fixed dental prostheses

Outcome Domains	Outcome Subdomains	Outcome Measurements	Criteria	Outcome Classification	Number of Times Reported in RCTs (*n*=64)
Success rate	/	Permanence on site without complications	Success/Not success	Primary outcome	29
Survival rate	/	Permanence on site without or with clinically acceptable complications	Survival/Failure	Primary outcome	44
Technical complications	A. Restoration integrity	Framework fracture and gross fracture	Modified USPHS/Ryge criteria Modified CDA criteria	Secondary outcome	18
		Chipping	Modified USPHS/Ryge criteria Modified CDA criteria FDI/Hickel criteria	Secondary outcome	33
		Proximal contact point and proximal and food impaction	USPHS/Ryge criteria CDA criteria FDI/Hickel criteria Copenhagen/Hosseini Index Score	Secondary outcome	25
		Occlusal contour Occlusal wear	Modified USPHS/Ryge criteria Modified CDA criteria FDI/Hickel criteria	Primary outcome	21
		Marginal adaptation and marginal integrity	USPHS/Ryge criteria CDA criteria FIT/Ferrari-Cagidiaco criteria Copenhagen/Hosseini Index Score Scanning Electron Microscopy (SEM)	Secondary outcome	40
		Occlusion and static and dynamic occlusion	USPHS/Ryge criteria CDA criteria FIT/Ferrari-Cagidiaco criteria Copenhagen/Hosseini Index Score Scanning electron microscopy (SEM)	Secondary outcome	4
		Marginal Fit	Replica technique (Boening KW 2000)	Primary outcome	4
	B. Decementation	Loss of retention and debonding	Modified USPHS/Ryge criteria Modified CDA criteria	Secondary outcome	19
Biologic complications	A. Dental abutment	Loss of tooth vitality and endodontic pathology	None	Primary outcome	15
		Abutment tooth fracture and tooth fracture	Modified USPHS/Ryge criteria	Primary outcome	16
		Tooth integrity (enamel cracks)	Modified USPHS/Ryge criteria FDI/Hickel criteria	Secondary outcome	2
		Postoperative hypersensitivity	Modified USPHS/Ryge criteria Modified CDA criteria FDI/Hickel criteria Visual Analog Scale (VAS)	Primary outcome	27
		Secondary caries and recurrence of initial pathology (caries, erosion, abfraction)	USPHS/Ryge criteria CDA criteria FDI/Hickel criteria	Secondary outcome	32
	B. Periodontium	Bleeding on probing (BoP)	Sulcus Bleeding Index (Mühlemann and Son 1971) Gingival Bleeding Index (Ainamo and Bay 1975) FIT/Ferrari-Cagidiaco criteria	Primary outcome	11
		Clinical attachment loss (CAL) (mm)	None	Secondary outcome	1
		Probing attachment level (PAL) (mm)	None	Secondary outcome	3
		Probing pocket depth (PPD) (mm)	None	Secondary outcome	19
		Plaque control record (PCR) (%)	ÓLeary, 1972	Secondary outcome	3
		Plaque Index (PI)	Silness and Löe, 1964 FIT/Ferrari-Cagidiaco criteria	Secondary outcome	14
		Gingival Index (GI)	Löe and Silness, 1963	Secondary outcome	21
		Periodontal response (always compared with reference tooth) and tissue health	FDI/Hickel criteria Modified CDA criteria	Secondary outcome	4
		Quality and quantity of mucosa and gingival marginal stability	FIT/Ferrari-Cagidiaco criteria Agustín-Panadero, 2021	Secondary outcome	1
		Adjacent mucosa	FDI/Hickel criteria	Secondary outcome	3
		Radiographic examination and bone level in radiograph	FDI/Hickel criteria FIT/Ferrari-Cagidiaco criteria	Secondary outcome	21

(*continued on next page*)

Table 1. (*Continued*) Outcome measurements reported in randomized controlled trials of tooth-supported fixed dental prostheses

Outcome Domains	Outcome Subdomains	Outcome Measurements	Criteria	Outcome Classification	Number of Times Reported in RCTs (*n*=64)
Esthetic complications	/	Surface texture and gloss and luster and roughness	Modified USPHS/Ryge criteria Modified CDA criteria FDI/Hickel criteria	Secondary outcome	18
		Cavosurface marginal discoloration and surface or marginal staining	USPHS/Ryge criteria CDA criteria FDI/Hickel criteria Copenhagen/Hosseini Index Score	Secondary outcome	25
		Color match, color stability, translucency	USPHS/Ryge criteria CDA criteria FDI/Hickel criteria FIT/Ferrari-Cagidiaco criteria Copenhagen/Hosseini Index Score CIELab/CIEch system	Primary outcome	28
		Anatomic form and interproximal contact and papillae and papilla volumen score	FDI/Hickel criteria PES/WES Score FIT/Ferrari-Cagidiaco criteria Copenhagen/Hosseini Index Score	Secondary outcome	10
Cost-effectiveness	A. Total costs	Clinical cost ($/€ per h)	Cost of professional time per h	Primary outcome	1
		Laboratory cost	None	Secondary outcome	1
	B. Appointments	Number of appointments	None	Secondary outcome	1
Time efficiency	A. Clinical time	Clinical impressions	Work time needed to achieve a clinically acceptable implant impression	Secondary outcome	6
		Occlusal adjustment	Zhang et al, 2018	Primary outcome	2
	B. Laboratory time	Laboratory manufacture	Work time needed to design and fabricate the restoration	Secondary outcome	3
Patient satisfaction	General assessment	Patient satisfaction with restoration and patient's opinion	FDI/Hickel criteria Visual Analog Scale (VAS)	Primary outcome	20
		Operator's satisfaction	Visual Analog Scale (VAS)	Secondary outcome	5
		QoLFAST-10 questionnaire OHIP-14 questionnaire OHIP-20 questionnaire OHIP-49 questionnaire	Castillo-Oyague et al, 2016 Slade et al, 1997 Allen and Locker, 2002 Slade and Spencer, 1994	Primary outcome	6
		Other criteria	Researcher's own criteria	Secondary outcome	5
Oral health	General assessment	Oral and general health	FDI/Hickel criteria	Secondary outcome	0

trials with low RoB reported more comprehensive adherence to the CONSORT statement than those with some concerns (MD=4; 95% CI=1.52-6.48; *P*=.004; Table 5).

DISCUSSION

The present systematic review found that adherence to the CONSORT statement is associated with lower RoB and better reporting quality in RCTs on tooth-supported FDPs; therefore, the null hypotheses were rejected. The diverse range of outcome measurements used has been documented, with no improvement since the publication of the most recent review in 2013.[102] The need of a SCORES in clinical research remains evident. Most authors use success and survival rate as primary outcomes in clinical trials.[42,47,49,60] Success has been commonly defined as the permanence of the prosthesis in situ and intact, without complications or alterations in patient satisfaction. Survival has been routinely defined as the permanence of the prosthesis in situ even with complications that can be considered acceptable, without the need for replacement.[103] It is essential to record the details of these criteria, otherwise comparability and transparency may be compromised. The analysis showed that FDP assessment involved 8 domains of outcome measurement: success rate, survival rate, technical/mechanical properties, biological properties, esthetic/cosmetic properties, time-effectiveness, cost-effectiveness, and patient-reported outcome measures. The modified USPHS/Ryge criteria,[9] the modified CDA criteria,[10] and the FDI/Hickel criteria to assess technical properties and complications operate in a similar manner. These criteria yield comparable results for the chipping, marginal adaptation, occlusal wear, proximal anatomic form, occlusion, and debonding.

Table 2. Frequency of current clinical assessment criteria used to assess the quality of restorations

Standardized Criteria	Frequency, N (%)
Modified USPHS/Ryge criteria	23 (35.9)
Modified CDA criteria	9 (14.1)
FDI/Hickel criteria	4 (6.2)
FIT/Ferrari-Cagidiaco criteria	1 (1.6)
Copenhagen/Hosseini Index Score	2 (1.6)
None	26 (40.6)
Total	64 (100)

The main biological properties and complications analyzed have been tooth vitality, secondary caries, postoperative hypersensitivity, and the clinical reaction of periodontal tissues to the prosthetic material. However, numerous different outcome measurements have been reported to clinically assess the relationship between tooth-supported FDPs and the periodontium.[104] This heterogeneity poses difficulties in synthesizing data and comparing studies.[7] Prosthodontic studies tend to be more focused on testing the materials' resistance and overall mechanical performance rather than on evaluating their interactions with soft tissues. The most frequently selected outcome measurements, consistent with those used in periodontal research, include bleeding on probing (BoP), clinical attachment loss (CAL), probing pocket depth (PPD), and plaque control record (PCR).[105-109]

In research into dental esthetics, the objective index that has demonstrated the highest repeatability is the pink esthetic score and white esthetic score (PES/WES) described by Chen et al[110] for single-unit implant-supported anterior FDPs. This index has also been adapted for single-unit tooth-supported anterior FDPs.[111] The modified USPHS/Ryge criteria,[9] the modified CDA criteria,[10] and the FDI/Hickel criteria also assess esthetic properties and complications in a similar manner through surface staining, color, and esthetic anatomic form.[11,12] Hosseini and Gotfredsen[111] validated the Copenhagen Index score (CIS) to be applied for implant-supported FDPs in the premolar region. This index has also been adapted for tooth-supported FDPs.[49] Currently, the most objective and rapid instrument for assessing tooth color are spectrophotometers that use the Commission Internationale de l'Eclairage (CIE) Lab or CIECh color scale, which identifies color according to black/white (L* value), green/red (a* value), and yellow/blue dimensions (b* value) or lightness (L* value), chroma (C* value), and hue (h* value). This technique has been extensively used in quantifying tooth color and tooth color differences from the equation $\Delta Eab=([\Delta L]^2+[\Delta a]^2+[\Delta b]^2)^{1/2}$.[112]

Major benefits of digitization are standardized protocols, improved predictability, and reduced clinical and laboratory time.[83,84] Therefore, evaluating the time-efficiency of digital and conventional technologies to develop FDPs has become increasingly necessary.

The distinguishing feature of patient satisfaction outcome measurements is that it attempts to evaluate the subjectivity of each patient at specific time points (pretreatment or posttreatment at 1, 3, 5, 10, or more years). The use of patient-reported outcome measures (PROMs) has emerged in dentistry, leading to a shift to patient-centered care.[113] Both the visual analog scale (VAS) and the oral satisfaction assessment (OSS) have been generalized for a rapid and overall evaluation of patient satisfaction,[114] while the 14- and 20-item oral health impact profiles (OHIP-14[115] and OHIP-20[116]), both being short versions of the original 49-item OHIP-49,[117] have been widely used instruments in the last decades for a more comprehensive assessment after therapeutic dental interventions. The oral impacts on daily performances (OIDP) is another frequently used valid instrument to describe the impact of oral conditions on quality of life.[118] Recently, among other advances in the field, a specific, complete, and short (10-item) questionnaire, Quality of Life related to Function, Aesthetics, Socialization, and Thoughts about health-behavioral habits (QoLFAST-10), that includes all the possible areas through which the quality-of-life (QoL) construct may be evaluated has been validated. The QoLFAST-10 index, which was initially designed for implant-supported prostheses, may be useful to improve the efficiency of clinical trials of tooth-supported FDPs that consider patient preferences and sociodemographic profiles to help predict treatment success.[119,120] According to a recent meta-analysis by Wittneben et al,[113] evidence on tooth-supported FDPs is still scarce, justifying additional clinical trials.

The findings of this systematic review also showed that the use of the CONSORT reporting statement, as well as better compliance with its items, decreases the RoB, consistent with similar reviews on other types of oral health interventions.[27,121,122] The CONSORT statement has been reported to be effective in assisting researchers in the reporting process.[123] However, considering the findings of this research, it is not yet widely used by investigators in the field, and consequently misinformation about certain critical aspects related to methodology has appeared in several trials in the field.[43,55,56,61,65,94]

The sample size calculation is a fundamental requirement that should be included in the report of a trial. Except for new analyses based on artificial intelligence (AI) that can work with large sample sizes,[124-126] it is impractical to study the entire population in any investigation. Therefore, a subset of participants is selected, which adequately represents the population from which conclusions may be drawn. Thereby, true

Table 3. Design characteristics of trials included in review

Methodological issues	*N*	(%)
Number of RCTs	64	
Number of published reports	79	
Design		
Split-mouth	12	18.7
Cross-over	5	7.8
Parallel	45	70.3
Centers		
Single-center	56	87.5
Multicenter	8	12.5
Setting		
Private practice	11	17.2
Public health service	6	9.4
University setting	47	73.4
Ethics committee approval	59	92.2
Registered on accessible international platform	22	34.4
Adherence to the CONSORT 2010 Statement	18	28.1
Explicit mention of primary outcome	32	50.0
Sample size calculation	20	31.2
Randomization sequence generation		
Not mentioned	24	37.5
Nonrandom sequence generation	0	0.0
Coin-tossing, throwing dice, or dealing shuffled cards	10	15.6
Computer-generated randomization list	30	46.9
Randomization method		
Not mentioned	52	81.2
Simple randomization	6	9.4
Block randomization	5	7.8
Stratified randomization	1	1.6
Minimization	0	0.0
Allocation concealment method		
Not mentioned	40	62.5
Central randomization	5	7.8
Sequentially numbered, opaque and sealed envelopes	19	29.7
Patient blinding	17	26.6
Caregiver blinding	0	0.0
Examiner blinding	43	67.2
Direction of the trial		
Not mentioned	62	96.9
Noninferiority	0	0.0
Superiority	2	3.1
Equivalence	0	0.0
Appropriate analysis		
Intention-to-treat analysis (ITT)	35	54.7
Modified intention-to-treat analysis (mITT)	29	45.3
Per-protocol analysis	0	0.0
Lost data as missing at random (MAR) or completely at random (MCAR)	28	43.7
Relevant lost data as missing not at random (MNAR)	1	1.6
Sensitivity analysis	0	0.0
Funding statement reported	39	60.9
Internally	13	20.3
Externally	26	40.6
Governmental organization	6	9.4
For-profit private organization	20	31.2
Nonprofit private organization	0	0.0
Both	6	9.4
Description of funder's involvement and conflict of interest	26	40.6

inferences can be made by using minimum resources and maximum efficiency.

Randomization is mandatory to generate balanced groups and to avoid confounding factors that may overestimate or underestimate values in the conclusions. In prosthetic clinical trials, the involvement of at least 1 prosthodontist and 1 dental laboratory technician makes the interventions complex and large sample sizes difficult to obtain. In prosthetic clinical trials, stratified randomization or minimization better ensures equitable groups than simple or block randomization by balancing potential confounding factors.[4] Concealment of the assignment, often confused with masking,[4,5] safeguards the possible manipulation of participants before they are assigned to intervention groups. Methods used to conceal the allocation include central allocation by a responsible third party or tamper-proof methods like sealed, opaque, sequentially numbered envelopes opened consecutively after the participant has been irreversibly allocated.

Deviations from the protocol may occur in clinical trials, where participants do not always follow instructions, drop out of studies, or are lost to follow-up, which may lead to selection bias or dropout bias in the results. Therefore, to fully preserve the benefit of randomization, all randomized participants should be included in the statistical process, known as intention-to-treat (ITT) analysis, a method that has been widely recommended as the preferred analysis strategy.[4,5] ITT analysis requires identifying and fully reporting protocol deviations and reasons for losses to classify the lost data as missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) to establish an appropriate method for handling lost data (such as multiple imputations, inverse probability weighting, or maximum likelihood estimation) and conducting sensitivity analyses.[127] Conversely, if missing data are omitted from the statistical assessment, a per protocol (PP) analysis will be performed, thus compromising randomization, reducing the sample size and power, introducing bias, and providing an altered estimation of the results. This occurs when the frequency or causes of the losses differ among the intervention groups.[4,5]

To cover the economic costs involved in studies, funding from governmental or private organizations (for-profit or nonprofit),[127] support in exchange for giving visibility to brands in publications (sponsorship), obtaining support from research groups, or collaboration with corporate social responsibility (CSR) campaigns in favor of trading houses, among others, have been used. However, financial interests must always be disclosed in the paper. It is critical that the protocol explicitly describes the participation and responsibilities of those funders in designing, conducting, analyzing, and

Table 4. Statistical results of first null hypothesis testing

CONSORT Statement	Total Number of Studies	Low ROB N (%)	Some Concerns N (%)	High ROB N (%)	*P*
Studies that follow 2010 CONSORT Statement	18	9 (50.0)	9 (50.0)	0 (0.00)	<.001
Studies that do not follow 2010 CONSORT Statement	46	4 (8.4)	41 (89.1)	1 (2.2)	

Table 5. Statistical results of second null hypothesis testing

CONSORT Statement	Low ROB		Some Concerns		Mean Difference *P*
	Number of Studies	Mean (±SD)	Number of Studies	Mean (±SD)	
Compliance with 2010 CONSORT Statement	9	24.56 (2.13)	9	20.56 (2.79)	4.00 (1.52 - 6.48) *P*=.004

interpreting the data or in disseminating the study results. Authors must control the final decision regarding these issues. Research ethics, objectivity, and common sense must prevail in the contractual agreements signed between the funders and the researchers.

Limitations of this systematic review included that the search, although systematic and reproducible, was not comprehensive as it only considered RCTs published in the highest-ranking dental journals (quartiles: Q1 and Q2), a limitation that could compromise the external validity of the results. Another limitation was that, although the studies were careful to ensure that the judgments were reliable and reproducible, another group of researchers may arrive at different judgments for some of the trials included in the sample.

Future studies should explore the development of a SCORES in clinical research on tooth-supported FDPs through a Delphi survey and a consensus meeting according to the COMET initiative.[128] Future RCTs should endeavor to register their studies prospectively in international registries; adhere to the CONSORT statement; make sample size calculations; report their primary results; describe randomization, allocation concealment, and blinding in detail; and explain reasons for dropouts. Even though most journals have adopted the CONSORT statement in their instructions to authors, it has not yet been widely used in published reports.

CONCLUSIONS

Based on the findings of this systematic review, the following conclusions were drawn:

1. The evaluation of clinical trials on tooth-supported FDPs led to a variety of outcome measures.
2. Standardized criteria should be developed covering the following areas: (i) success rate, (ii) survival rate, (iii) technical complications, (iv) biological complications, (v) cosmetic complications, (vi) cost-effectiveness, (vii) time efficiency, and (viii) patient-reported outcome measures.
3. Most RCTs on tooth-supported FDPs did not follow the CONSORT statement, and some concerns were identified in the RoB analysis.
4. Those studies with higher compliance with the CONSORT statement items had a lower RoB.
5. Future studies should improve their methodologies and reporting quality.

REFERENCES

1. Clarkson J, Worthington H. Leadership in evidence-based dentistry. J Dent 2019;87:16-9.
2. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: Defining standard protocol items for clinical trials. Ann Intern Med 2013;158:200-7.
3. Pandis N, Fleming PS, Katsaros C, Ioannidis JPA. Dental research waste in design, analysis, and reporting: A systematic review. J Dent Res 2021;100:245-52.
4. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.
5. Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010;152:726-32.
6. Smaïl-Faugeron V, Fron-Chabouis H, Durieux P. Clinical trial registration in oral health journals. J Dent Res 2015;94:8S-13S.
7. Limones A, Molinero-Mourelle P, Azevedo L, Romeo-Rubio M, Correia A, Gómez-Polo M. Zirconia-ceramic versus metal-ceramic posterior multiunit tooth-supported fixed dental prostheses: A systematic review and meta-analysis of randomized controlled trials. J Am Dent Assoc 2020;151:230-8.e7.
8. Poggio CE, Ercoli C, Rispoli L, Maiorana C, Esposito M. Metal-free materials for fixed prosthodontic restorations. Cochrane Database Syst Rev 2017;12:CD009606.
9. Cvar JF, Ryge G. Reprint of criteria for the clinical evaluation of dental restorative materials. 1971. Clin Oral Investig 2005;9:215-32.
10. California Dental Association. Guidelines for the assessment of clinical quality and professional performance. Los Angeles: CDA; 1977. p. 53-5.
11. Hickel R, Peschke A, Tyas M, Mjör I, Bayne S, Peters M, et al. FDI World Dental Federation: Clinical criteria for the evaluation of direct and indirect restorations-update and clinical examples. Clin Oral Investig 2010;14:349-66.
12. Hickel R, Peschke A, Tyas M, Mjör I, Bayne S, Peters M, et al. FDI World Dental Federation - clinical criteria for the evaluation of direct and indirect restorations. Update and clinical examples. J Adhes Dent 2010;12:259-72.
13. Ryge G. Clinical criteria. Int Dent J 1980;30:347-58.
14. Allander L, Birkhed D, Bratthall D. Quality evaluation of anterior restorations in private practice. Swed Dent J 1989;13:141-50.

15. Cagidiaco EF, Grandini S, Goracci C, Joda T. A pilot trial on lithium disilicate partial crowns using a novel prosthodontic functional index for teeth (FIT). BMC Oral Health 2019;19:276.
16. Agustín-Panadero R, Serra-Pastor B, Loi I, Suárez MJ, Pelaez J, Solá-Ruíz F. Clinical behavior of posterior fixed partial dentures with a biologically oriented preparation technique: A 5-year randomized controlled clinical trial. J Prosthet Dent 2021;125:870-6.
17. Ahrberg D, Lauer HC, Ahrberg M, Weigl P. Evaluation of fit and efficiency of CAD/CAM fabricated all-ceramic restorations based on direct and indirect digitalization: A double-blinded, randomized clinical trial. Clin Oral Investig 2016;20:291-300.
18. Akın A, Toksavul S, Toman M. Clinical marginal and internal adaptation of maxillary anterior single all-ceramic crowns and 2-year randomized controlled clinical trial. J Prosthodont 2015;24:345-50.
19. Makarouna M, Ullmann K, Lazarek K, Boening KW. Six-year clinical performance of lithium disilicate fixed partial dentures. Int J Prosthodont 2011;24:204-6.
20. Metiner C, Türker ŞB. Evaluation of clinical success and marginal adaptations of inlay and onlay restorations manufactured after conventional and digital impressions: A prospective randomized controlled clinical study. Int J Comput Dent 2021;24:207-23.
21. Vanoorbeek S, Vandamme K, Lijnen I, Naert I. Computer-aided designed/computer-assisted manufactured composite resin versus ceramic single-tooth restorations: A 3-year clinical study. Int J Prosthodont 2010;23:223-30.
22. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
23. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ 2009;339:b2535.
24. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. BMJ 2009;339:b2700.
25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Med 2009;6:e1000100.
26. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: A revised tool for assessing RoB in randomised trials. BMJ 2019;366:l4898.
27. Loguercio AD, Maran BM, Hanzen TA, Paula AM, Perdigão J, Reis A. Randomized clinical trials of dental bleaching - Compliance with the CONSORT Statement: A systematic review. Braz Oral Res 2017;31:e60.
28. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: An extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev 2021;10:39.
29. Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: Analytical survey. BMJ 2004;328:1040.
30. Benic GI, Mühlemann S, Fehmer V, Hämmerle CH, Sailer I. Randomized controlled within-subject evaluation of digital and conventional workflows for the fabrication of lithium disilicate single crowns. Part I: Digital versus conventional unilateral impressions. J Prosthet Dent 2016;116:777-82.
31. Sailer I, Benic GI, Fehmer V, Hämmerle CHF, Mühlemann S. Randomized controlled within-subject evaluation of digital and conventional workflows for the fabrication of lithium disilicate single crowns. Part II: CAD-CAM versus conventional laboratory procedures. J Prosthet Dent 2017:11843-8.
32. Zeltner M, Sailer I, Mühlemann S, Özcan M, Hämmerle CH, Benic GI. Randomized controlled within-subject evaluation of digital and conventional workflows for the fabrication of lithium disilicate single crowns. Part III: Marginal and internal fit. J Prosthet Dent 2017;117:354-62.
33. Bömicke W, Rammelsberg P, Zenthöfer A, Ohlmann B. Clinical performance of zirconia-ceramic cantilever fixed partial dentures-Longitudinal nine-year results from a prospective, randomized, controlled pilot study. J Prosthodont Res 2019;63:334-9.
34. Bömicke W, Rathmann F, Pilz M, Bermejo JL, Waldecker M, Ohlmann B, et al. Clinical performance of posterior inlay-retained and wing-retained monolithic zirconia resin-bonded fixed partial dentures: Stage one results of a randomized controlled trial. J Prosthodont 2021;30:384-93.
35. Botelho MG, Chan AWK, Leung NCH, Lam WYH. Long-term evaluation of cantilevered versus fixed-fixed resin-bonded fixed partial dentures for missing maxillary incisors. J Dent 2016;45:59-66.
36. Botelho MG, Yon MJY, Mak KCK, Lam WYH. A randomised controlled trial of two-unit cantilevered or three-unit fixed-movable resin-bonded fixed partial dentures replacing missing molars. J Dent 2020;103:103519.
37. van den Breemer C, Gresnigt M, Özcan M, Kerdijk W, Cune MS. Prospective randomized clinical trial on the survival of lithium disilicate posterior partial crowns bonded using immediate or delayed dentin sealing: Short-term results on tooth sensitivity and patient satisfaction. Oper Dent 2019;44:E212-22.
38. van den Breemer CRG, Cune MS, Özcan M, Naves LZ, Kerdijk W, Gresnigt MMM. Randomized clinical trial on the survival of lithium disilicate posterior partial restorations bonded using immediate or delayed dentin sealing after 3 years of function. J Dent 2019;85:1-10.
39. Cagidiaco EF, Discepoli N, Goracci C, Carboncini F, Vigolo P, Ferrari M. Randomized clinical trial on single zirconia crowns with feather-edge vs chamfer finish lines: Four-year results. Int J Periodontics Restorative Dent 2019;39:817-26.
40. Cheng C-W, Ye S-Y, Chien C-H, Chen C-J, Papaspyridakos P, Ko C-C. Randomized clinical trial of a conventional and a digital workflow for the fabrication of interim crowns: An evaluation of treatment efficiency, fit, and the effect of clinician experience. J Prosthet Dent 2021;125:73-81.
41. Christensen RP, Ploeger BJ. A clinical comparison of zirconia, metal and alumina fixed-prosthesis frameworks veneered with layered or pressed ceramic: A three-year report. J Am Dent Assoc 2010;141:1317-29.
42. Crins LAMJ, Opdam NJM, Kreulen CM, Bronkhorst EM, Sterenborg BAMM, Huysmans MCDNJM, et al. Randomized controlled trial on the performance of direct and indirect composite restorations in patients with severe tooth wear. Dent Mater 2021;37:1645-54.
43. El-Ma'aita AA, Al-Rabab'ah M, Abu-Awwad M, Hattar S, Devlin H. Endocrowns Clinical performance and patient satisfaction: A randomized clinical trial of three monolithic ceramic restorations. J Prosthodont 2022;31:30-7.
44. Eltoukhy RI, Elkaffas AA, Ali AI, Mahmoud SH. Indirect Resin Composite Inlays Cemented with a self-adhesive, self-etch or a conventional resin cement luting agent: A 5 years prospective clinical evaluation. J Dent 2021;112:103740.
45. Esquivel-Upshaw J, Rose W, Oliveira E, Yang M, Clark AE, Anusavice K. Randomized, controlled clinical trial of bilayer ceramic and metal-ceramic crown performance. J Prosthodont 2013;22:166-73.
46. Etman MK, Woolford MJ. Three-year clinical evaluation of two ceramic crown systems: A preliminary study. J Prosthet Dent 2010;103:80-90.
47. Ferrari M, Ferrari Cagidiaco E, Goracci C, Sorrentino R, Zarone F, Grandini S, et al. Posterior partial crowns out of lithium disilicate (LS2) with or without posts: A randomized controlled prospective clinical trial with a 3-year follow up. J Dent 2019;83:12-7.
48. Gjelvold B, Chrcanovic BR, Korduner E-K, Collin-Bagewitz I, Kisch J. Intraoral digital impression technique compared to conventional impression technique. a randomized clinical trial. J Prosthodont 2016;25:282-7.
49. Gotfredsen K, Alyass N, Hagen M. A 5-Year randomized clinical trial on 3-unit fiber-reinforced vs 3- or 2-unit metal-ceramic resin-bonded fixed dental prostheses. Int J Prosthodont 2021;34:703-11.
50. Gresnigt MMM, Kalk W, Özcan M. Randomized controlled split-mouth clinical trial of direct laminate veneers with two micro-hybrid resin composites. J Dent 2012;40:766-75.
51. Gresnigt MMM, Kalk W, Özcan M, Gresnigt M. Randomized clinical trial of indirect resin composite and ceramic veneers: Up to 3-year follow-up. J Adhes Dent 2013;15:181-90.
52. Gresnigt MMM, Cune MS, Jansen K, van der Made SAM, Özcan M. Randomized clinical trial on indirect resin composite and ceramic laminate veneers: Up to 10-year findings. J Dent 2019;86:102-9.
53. Grohmann P, Bindl A, Hämmerle C, Mehl A, Sailer I. Three-unit posterior zirconia-ceramic fixed dental prostheses (FDPs) veneered with layered and milled (CAD-on) veneering ceramics: 1-year follow-up of a randomized controlled clinical trial. Quintessence Int 2015;46:871-80.
54. Haddadi Y, Bahrami G, Isidor F. Evaluation of operating time and patient perception using conventional impression taking and intraoral scanning for crown manufacture: A split-mouth, randomized clinical study. Int J Prosthodont 2018;31:55-9.
55. Halawani M, Aboushady Y, Kader S, Solimei L, Benedicenti S. Evaluating the marginal integrity and clinical outcome of posterior zirconia inlay-retained fixed dental prostheses: A randomized clinical trial. Int J Prosthodont 2021;34:324-33.
56. Hammoudi W, Trulsson M, Svensson P, Smedberg JI. Long-term results of a randomized clinical trial of 2 types of ceramic crowns in participants with extensive tooth wear. J Prosthet Dent 2022;127:248-57.
57. Hu J, Zhu Q. Effect of immediate dentin sealing on preventive treatment for postcementation hypersensitivity. Int J Prosthodont 2010;23:49-52.
58. Huettig F, Prutscher A, Goldammer C, Kreutzer CA, Weber H. First clinical experiences with CAD/CAM-fabricated PMMA-based fixed dental prostheses as long-term temporaries. Clin Oral Investig 2016;20:161-8.
59. Izgi AD, Eskimez S, Kale E, Değer Y. Directly fabricated inlay-retained glass- and polyethylene fiber-reinforced composite fixed dental prostheses in posterior single missing teeth: A short-term clinical observation. J Adhes Dent 2011;13:383-91.
60. Juloski J, Fadda GM, Monticelli F, Fajó-Pascual M, Goracci C, Ferrari M. Four-year survival of endodontically treated premolars restored with fiber posts. J Dent Res 2014;93:52S-8S.
61. Jurema AB, Bresciani E, Caneppele TMF. Influence of glass fiber posts on the fracture susceptibility of endodontically treated maxillary anterior teeth

with direct veneers: Preliminary results of a randomized clinical trial. J Esthet Restor Dent 2021;33:613-20.
62. Klur T, Hasan I, Ottersbach K, Stark H, Fichte M, Dirk C, et al. PEKK-made indirect temporary crowns and bridges: A clinical pilot study. Clin Oral Investig 2019;23:771-7.
63. Kozmacs C, Schaper K, Lauer HC, Piwowarczyk A. Evaluation of hypersensitivity after the placement of metal-ceramic crowns cemented with two luting agents: Long-term results of a prospective clinical study. J Prosthet Dent 2017;118:347-52.
64. Malmstrom H, Dellanzo-Savu A, Xiao J, Feng C, Jabeen A, Romero M, et al. Success, clinical performance and patient satisfaction of direct fibre-reinforced composite fixed partial dentures-a two-year clinical study. J Oral Rehabil 2015;42:906-13.
65. Marchionatti AME, Wandscher VF, May MM, Bottino MA, May LG. Color stability of ceramic laminate veneers cemented with light-polymerizing and dual-polymerizing luting agent: A split-mouth randomized clinical trial. J Prosthet Dent 2017;118:604-10.
66. McKenna G, Allen F, Woods N, O'Mahony D, Cronin M, DaMata C, et al. Cost-effectiveness of tooth replacement strategies for partially dentate elderly: A randomized controlled clinical trial. Community Dent Oral Epidemiol 2014;42:366-74.
67. Monaco C, Llukacej A, Baldissara P, Arena A, Scotti R. Zirconia-based versus metal-based single crowns veneered with overpressing ceramic for restoration of posterior endodontically treated teeth: 5-year results of a randomized controlled clinical study. J Dent 2017;65:56-63.
68. Naenni N, Bindl A, Sax C, Hämmerle C, Sailer I. A randomized controlled clinical trial of 3-unit posterior zirconia-ceramic fixed dental prostheses (FDP) with layered or pressed veneering ceramics: 3-year results. J Dent 2015;43:1365-70.
69. Nicolaisen M, Bahrami G, Schropp L, Isidor F. Comparison of metal-ceramic and all-ceramic three-unit posterior fixed dental prostheses: A 3-year randomized clinical trial. Int J Prosthodont 2016;29:259-64.
70. Nicolaisen M, Bahrami G, Schropp L, Isidor F. Functional and esthetic comparison of metal-ceramic and all-ceramic posterior three-unit fixed dental prostheses. Int J Prosthodont 2016;29:473-81.
71. Paniz G, Nart J, Gobbato L, Mazzocco F, Stellini E, De Simone G, et al. Clinical periodontal response to anterior all-ceramic crowns with either chamfer or feather-edge subgingival tooth preparations: Six-month results and patient perception. Int J Periodontics Restorative Dent 2017;37:61-8.
72. Peláez J, Cogolludo PG, Serrano B, Lozano JFL, Suárez MJ. A prospective evaluation of zirconia posterior fixed dental prostheses: Three-year clinical results. J Prosthet Dent 2012;107:373-9.
73. Pelaez J, Cogolludo PG, Serrano B, Serrano JFL, Suarez MJ. A four-year prospective clinical evaluation of zirconia and metal-ceramic posterior fixed dental prostheses. Int J Prosthodont 2012;25:451-8.
74. Peumans M, Voet M, de Munck J, van Landuyt K, van Ende A, van Meerbeek B. Four-year clinical evaluation of a self-adhesive luting agent for ceramic inlays. Clin Oral Investig 2013;17:739-50.
75. Pousette Lundgren G, Morling Vestlund GI, Trulsson M, Dahllöf G. A randomized controlled trial of crown therapy in young individuals with amelogenesis imperfecta. J Dent Res 2015;94:1041-7.
76. Lundgren GP, Vestlund GIM, Dahllöf G. Crown therapy in young individuals with amelogenesis imperfecta: Long term follow-up of a randomized controlled trial. J Dent 2018;76:102-8.
77. Reitemeier B, Hänsel K, Kastner C, Weber A, Walter MH. A prospective 10-year study of metal ceramic single crowns and fixed dental prosthesis retainers in private practice set tings. J Prosthet Dent 2013;109:149-55.
78. Reitemeier B, Hänsel K, Range U, Walter MH. Prospective study on metal ceramic crowns in private practice settings: 20-year results. Clin Oral Investig 2019;23:1823-8.
79. Rinke S, Zuck T, Hausdörfer T, Leha A, Wassmann T, Ziebolz D. Prospective clinical evaluation of chairside-fabricated zirconia-reinforced lithium silicate ceramic partial crowns 5-year results. Clin Oral Investig 2022;26:1593-603.
80. Sailer I, Gottnerb J, Kanelb S, Hammerle CHF. Randomized controlled clinical trial of zirconia-ceramic and metal-ceramic posterior fixed dental prostheses: A 3-year follow-up. Int J Prosthodont 2009;22:553-60.
81. Sailer I, Balmer M, Jürg H, Hämmerle CHF, Känel S, Thoma DS. Comparison of fixed dental prostheses with zirconia and metal frameworks: Five-year results of a randomized controlled clinical trial. Int J Prosthodont 2017;30:426-8.
82. Sailer I, Balmer M, Hüsler J, Hämmerle CHF, Känel S, Thoma DS. 10-year randomized trial (RCT) of zirconia-ceramic and metal-ceramic fixed dental prostheses. J Dent 2018;76:32-9.
83. Sailer I, Mühlemann S, Fehmer V, Hämmerle CHF, Benic GI. Randomized controlled clinical trial of digital and conventional workflows for the fabrication of zirconia-ceramic fixed partial dentures. Part I: Time efficiency of complete-arch digital scans versus conventional impressions. J Prosthet Dent 2019;121:69-75.
84. Mühlemann S, Benic GI, Fehmer V, Hämmerle CHF, Sailer I. Randomized controlled clinical trial of digital and conventional workflows for the fabrication of zirconia-ceramic posterior fixed partial dentures. Part II: Time efficiency of CAD-CAM versus conventional laboratory procedures. J Prosthet Dent 2019;121:252-7.
85. Benic GI, Sailer I, Zeltner M, Gütermann JN, Özcan M, Mühlemann S. Randomized controlled clinical trial of digital and conventional workflows for the fabrication of zirconia-ceramic fixed partial dentures. Part III: Marginal and internal fit. J Prosthet Dent 2019;121:426-31.
86. Seidel A, Belli R, Breidebach N, Wichmann M, Matta RE. The occlusal wear of ceramic fixed dental prostheses: 3-Year results in a randomized controlled clinical trial with split-mouth design. J Dent 2020;103:103500.
87. Selz CF, Strub JR, Vach K, Guess PC. Long-term performance of posterior InCeram Alumina crowns cemented with different luting agents: A prospective, randomized clinical split-mouth study over 5 years. Clin Oral Investig 2014;18:1695-703.
88. Seydler B, Schmitter M. Clinical performance of two different CAD/CAM-fabricated ceramic crowns: 2-Year results. J Prosthet Dent 2015;114:212-6.
89. Shetty R, Bhat A, Mehta D, Finger W. Effect of a calcium phosphate desensitizer on pre- and postcementation sensitivity of teeth prepared for full-coverage restorations: A randomized, placebo-controlled clinical study. Int J Prosthodont 2017:38-42.
90. Schmitter M, Hamadi K, Rammelsberg P. Survival of two post systems-five-year results of a randomized clinical trial. Quintessence Int 2011;42:843-50.
91. Soares-Rusu I, Villavicencio-Espinoza CA, de Oliveira NA, Wang L, Honório HM, Rubo JH, et al. clinical evaluation of lithium disilicate veneers manufactured by CAD/CAM Compared with heat-pressed methods: Randomized controlled clinical trial. Oper Dent 2021;46:4-14.
92. Sterzenbach G, Franke A, Naumann M. Rigid versus flexible dentine-like endodontic posts-clinical testing of a biomechanical concept: Seven-year results of a randomized controlled clinical pilot trial on endodontically treated abutment teeth with severe hard tissue loss. J Endod 2012;38:1557-63.
93. Suarez MJ, Perez C, Pelaez J, Lopez-Suarez C, Gonzalo E. A randomized clinical trial comparing zirconia and metal-ceramic three-unit posterior fixed partial dentures: A 5-year follow-up. J Prosthodont 2019;28:750-6.
94. Taschner M, Krämer N, Lohbauer U, Pelka M, Breschi L, Petschelt A, et al. Leucite-reinforced glass ceramic inlays luted with self-adhesive resin cement: A 2-year in vivo study. Dent Mater 2012;28:535-40.
95. Vigolo P, Mutinelli S. Evaluation of zirconium-oxide-based ceramic single-unit posterior fixed dental prostheses (FDPs) generated with two CAD/CAM systems compared to porcelain-fused-to-metal single-unit posterior FDPs: A 5-year clinical prospective study. J Prosthodont 2012;21:265-9.
96. Wolfart S, Müller F, Gerß J, Heyedcke G, Marré B, Böning K, et al. The randomized shortened dental arch study: Oral health-related quality of life. Clin Oral Investig 2014;18:525-33.
97. Sasse M, Kern M, Marré B, Walter MH. Clinical performance of cantilevered fixed dental prostheses abutments in the shortened dental arch. J Dent 2014;42:373-6.
98. Walter MH, Dreyhaupt J, Mundt T, Kohal R, Kern M, Rauch A, et al. Periodontal health in shortened dental arches: A 10-year RCT. J Prosthodont Res 2020;64:498-505.
99. Walter MH, Marré B, Dreyhaupt J, Heydecke G, Rauch A, Mundt T, et al. Rehabilitation of shortened dental arches: A fifteen-year randomised trial. J Oral Rehabil 2021;48:738-44.
100. Zenthöfer A, Ohlmann B, Rammelsberg P, Bömicke W. Performance of zirconia ceramic cantilever fixed dental prostheses: 3-year results from a prospective, randomized, controlled pilot study. J Prosthet Dent 2015;114:34-9.
101. Zhang R, Ding Q, Sun Y, Zhang L, Xie Q. Assessment of CAD-CAM zirconia crowns designed with 2 different methods: A self-controlled clinical trial. J Prosthet Dent 2018;120:686-92.
102. Patel DR, O'Brien T, Petrie A, Petridis H. A systematic review of outcome measurements and quality of studies evaluating fixed tooth-supported restorations. J Prosthodont 2014;23:421-33.
103. Anusavice KJ. Standardizing failure, success, and survival decisions in clinical studies of ceramic and metal-ceramic fixed dental prostheses. Dent Mater 2012;28:102-11.
104. Ercoli C, Tarnow D, Poggio CE, Tsigarida A, Ferrari M, Caton JG, et al. The relationships between tooth-supported fixed dental prostheses and restorations and the periodontium. J Prosthodont 2021;30:305-17.
105. Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, et al. Periodontitis: Consensus report of workgroup 2 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. J Periodontol 2018;89:S173-82.
106. Chapple ILC, Mealey BL, van Dyke TE, Bartold PM, Dommisch H, Eickholz P, et al. Periodontal health and gingival diseases and conditions on an intact and a reduced periodontium: Consensus report of workgroup 1 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. J Periodontol 2018;89:S74-84.
107. Lang NP, Bartold PM. Periodontal health. J Periodontol 2018;89:S9-16.
108. Trombelli L, Farina R, Silva CO, Tatakis DN. Plaque-induced gingivitis: Case definition and diagnostic considerations. J Periodontol 2018;89:S46-73.
109. Belser UC, Grütter L, Vailati F, Bornstein MM, Weber H-P, Buser D. Outcome evaluation of early placed maxillary anterior single-tooth implants

ARTICLE IN PRESS

using objective esthetic criteria: A cross-sectional, retrospective study in 45 patients with a 2- to 4-year follow-up using pink and white esthetic scores. J Periodontol 2009;80:140-51.
110. Chen J, Chiang C, Zhang Y. Esthetic evaluation of natural teeth in anterior maxilla using the pink and white esthetic scores. Clin Implant Dent Relat Res 2018;20:770-7.
111. Hosseini M, Gotfredsen K. A feasible, aesthetic quality evaluation of implant-supported single crowns: An analysis of validity and reliability. Clin Oral Implants Res 2012;23:453-8.
112. Gómez-Polo C, Montero J, Gómez-Polo M, de Parga JA, Celemin-Viñuela A. Natural tooth color estimation based on age and gender. J Prosthodont 2017;26:107-14.
113. Wittneben J-G, Wismeijer D, Brägger U, Joda T, Abou-Ayash S. Patient-reported outcome measures focusing on aesthetics of implant- and tooth-supported fixed dental prostheses: A systematic review and meta-analysis. Clin Oral Implants Res 2018;29:224-40.
114. Montero J, Bravo M, Albaladejo A. Validation of two complementary oral-health related quality of life indicators (OIDP and OSS 0-10) in two qualitatively distinct samples of the Spanish population. Health Qual Life Outcomes 2008;6:101.
115. Slade GD. Derivation and validation of a short-form oral health impact profile. Community Dent Oral Epidemiol 1997;25:284-90.
116. Allen F, Locker D. A modified short version of the oral health impact profile for assessing health-related quality of life in edentulous adults. Int J Prosthodont 2002;15:446-50.
117. Montero J, López JF, Vicente MP, Galindo MP, Albaladejo A, Bravo M. Comparative validity of the OIDP and OHIP-14 in describing the impact of oral health on quality of life in a cross-sectional study performed in Spanish adults. Med Oral Patol Oral Cir Bucal 2011;16:e816-21.
118. Castillo-Oyagüe R, Suárez-García M-J, Perea C, Río J del, Lynch CD, Gonzalo E, et al. Validation of a new, specific, complete, and short OHR-QoL scale (QoLFAST-10) for wearers of implant overdentures and fixed-detachable hybrid prostheses. J Dent 2016;49:22-32.
119. Castillo-Oyagüe R, Perea C, Suárez-García M-J, Río J del, Lynch CD, Preciado A. Validation of the "Quality of Life related to function, aesthetics, socialization, and thoughts about health-behavioural habits (QoLFAST-10)" scale for wearers of implant-supported fixed partial dentures. J Dent 2016;55:82-91.
120. Ortiz MIG, Ribeiro MES, Lima DANL, Silva CM, Loretto SC, da Silva E Souza Júnior MH. Compliance of randomized clinical trials on dental caries prevention methods with the consort statement: A systematic review. J Evid Based Dent Pract 2021;21:101542.
121. Sponchiado-Júnior EC, Vieira W de A, Frozoni M, Herkrath FJ, de-Jesus-Soares A. CONSORT compliance in randomized clinical trials of regenerative endodontic treatments of necrotic immature teeth: A systematic review. J Endod 2021;47:1751-66.
122. Sarkis-Onofre R, Poletto-Neto V, Cenci MS, Pereira-Cenci T, Moher D. Impact of the CONSORT Statement endorsement in the completeness of reporting of randomized clinical trials in restorative dentistry. J Dent 2017;58:54-9.
123. Joda T, Yeung AWK, Hung K, Zitzmann NU, Bornstein MM. Disruptive innovation in dentistry: What it is and what could be next. J Dent Res 2021;100:448-53.
124. Shan T, Tay FR, Gu L. Application of artificial intelligence in dentistry. J Dent Res 2021;100:232-44.
125. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. Nat Med 2020;26:1364-74.
126. Bhaskaran K, Smeeth L. What is the difference between missing completely at random and missing at random? Int J Epidemiol 2014;43:1336-9.
127. Hakoum MB, Jouni N, Abou-Jaoude EA, Hasbani DJ, Abou-Jaoude EA, Lopes LC, et al. Characteristics of funding of clinical trials: Cross-sectional survey and proposed guidance. BMJ Open 2017;7:e015997.
128. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: Version 1.0. Trials 2017;18:280.

Corresponding author:
Dr Alvaro Limones
Department of Conservative & Prosthetic Dentistry
Faculty of Dentistry
Complutense University of Madrid (U.C.M.)
Pza. Ramón y Cajal, s/n, 28040, Madrid
SPAIN
Email: alimones@ucm.es

Copyright © 2022 by the Editorial Council for *The Journal of Prosthetic Dentistry*.
https://doi.org/10.1016/j.prosdent.2022.04.022

ANEXO 5.1: Resultados de la evaluación de la calidad metodológica (AMSTAR-II).

QUINTA REVISIÓN SISTEMÁTICA: (Limones et al., 2022)			
Pregunta	**Relevancia**	**Dominio de evaluación** [Pregunta explícita (Shea et al., 2017)]	**Respuesta**
Pregunta 1	-	Pregunta de investigación y criterios de inclusión: ¿Se establecieron preguntas de investigación y criterios de inclusión que incluyeron los componentes de PICO (Población, Intervención, Grupo comparador, Resultado)?	Si (Los criterios específicos de una RS de instrumentos y variables de medición)
Pregunta 2	**CRÍTICA**	**Protocolo registrado antes del inicio de la revisión:** ¿Se registró o publicó el protocolo de la revisión antes de empezar su realización y se han justificado las desviaciones significativas del protocolo?	**Si**
Pregunta 3	-	Diseño de estudios seleccionados: ¿Se explica por los autores cual es el diseño de estudio que se ha seleccionado para confeccionar la revisión?	Si
Pregunta 4	**CRÍTICA**	**Búsqueda exhaustiva de la literatura:** ¿Los autores de la revisión utilizaron una estrategia de búsqueda exhaustiva?	**Si**
Pregunta 5	-	Duplicación en la selección de estudios: ¿Los autores de la revisión realizaron la selección de estudios por duplicado?	Si
Pregunta 6		Duplicación en la extracción de datos: ¿Los autores de la revisión realizaron la extracción de datos por duplicado?	Si
Pregunta 7	**CRÍTICA**	**Justificación para excluir estudios individuales:** ¿Los autores de la revisión proporcionaron una lista de estudios excluidos y justificaron las exclusiones?	**Parcialmente** (Incluida en el diagrama de flujo y descriptivamente en la sección de resultados)
Pregunta 8	-	Descripción de estudios incluidos: ¿Los autores de la revisión describieron los estudios incluidos con suficiente detalle?	Si (En material suplementario)
Pregunta 9	**CRÍTICA**	**Evaluación del riesgo de sesgo de los estudios primarios incluidos:** ¿Utilizaron los autores de la revisión una técnica satisfactoria	**Si**

© Álvaro Limones Burgos, 2024

		para evaluar el riesgo de sesgo (RoB) en los estudios individuales que se incluyeron en la revisión?	
Pregunta 10	-	Fuentes de financiación: ¿Los autores de la revisión informaron sobre las fuentes de financiación de los estudios incluidos en la revisión?	Si
Pregunta 11	**CRÍTICA**	**Adecuados métodos análiticos:** Si se realizó un metaanálisis, ¿los autores de la revisión utilizaron métodos apropiados para la combinación estadística de resultados?	**No aplica**
Pregunta 12	-	Impacto del RoB en la metaanálisis: Si se realizó un metaanálisis, ¿los autores de la revisión evaluaron el impacto potencial de RoB de los estudios individuales sobre los resultados del metaanálisis u otra síntesis de evidencia?	No aplica
Pregunta 13	**CRÍTICA**	**Impacto del RoB en los resultados y discusión:** ¿Los revisores tuvieron en cuenta el RoB de los estudios primarios al interpretar/discutir los resultados de la revisión?	**Si**
Pregunta 14	-	Evaluación de la Heterogeneidad: ¿Los autores de la revisión proporcionaron una explicación satisfactoria y una discusión sobre cualquier heterogeneidad observada en los resultados de la revisión?	SI
Pregunta 15	**CRÍTICA**	**Evaluación de la presencia e impacto del sesgo de publicación:** Si se realizó una síntesis cuantitativa, ¿se llevó a cabo una investigación adecuada del sesgo de publicación (sesgo de estudios pequeños) y se discutió su probable impacto en los resultados de la revisión?	**Si**
Pregunta 16	-	Conflictos de interés: ¿Se informaron las posibles fuentes de conflicto de interés, incluyendo cualquier financiamiento recibido para llevar a cabo la revisión?	Si
VALORACIÓN FINAL DEL RIESGO DE SESGO: **ALTA CONFIANZA**			

© Álvaro Limones Burgos, 2024

ANEXO 5.2: Resultados de la evaluación del riesgo de sesgo (ROBIS).

1ª FASE	2ª FASE				3ª FASE
Identificación de la pregunta de investigación y sus componentes	**Criterios de elegibilidad (Dominio 1)**	**Identificación y selección de estudios (Dominio 2)**	**Recopilación de datos y evaluación de estudios (Dominio 3)**	**Síntesis y conclusiones (Dominio 4)**	**Riesgo de sesgo global**
¿La pregunta abordada en la revisión está correctamente definida de acuerdo a sus objetivos? S/PS/PN/N/NI* Revisiones de intervención: *P: Población* *I: Intervención* *C: Comparador* *O: Variables* Revisiones de pruebas de precisión diagnóstica: *P: Población* *I: Test diagnóstico* *C: Estándar de referencia* *O: Enfermedad objetivo* Revisiones de asociaciones etiológicas o factores de riesgo: *P: Población* *I: Exposición* *C: Comparador* *O: Variables* Revisiones de estudios pronóstico: *P: Población* *I: Factor pronóstico o modelo predictivo* *O: Desenlace a predecir* Revisiones de variables e instrumentos de medición: *P: Población/Campo de conocimiento* *O: Variables e instrumentos de medición*	1.1 ¿Se ajustó la revisión a los objetivos y criterios de elegibilidad predefinidos? S/PS/PN/N/NI* 1.2 ¿Los criterios de elegibilidad eran apropiados para la pregunta de la revisión? S/PS/PN/N/NI* 1.3 ¿Eran inequívocos los criterios de elegibilidad? S/PS/PN/N/NI* 1.4 ¿Eran apropiadas las restricciones de los criterios de elegibilidad basadas en las características del estudio (por ejemplo, fecha, tamaño de la muestra, calidad del estudio, resultados medidos, etc.)? S/PS/PN/N/NI* 1.5 ¿Era adecuada alguna restricción en los criterios de elegibilidad basada en las fuentes de fuentes de información (por ejemplo, estado o formato de publicación, idioma, disponibilidad de datos)? S/PS/PN/N/NI*	2.1 ¿Incluyó la búsqueda una gama apropiada de bases de datos/fuentes electrónicas para informes publicados y no publicados? S/PS/PN/N/NI* 2.2 ¿Se utilizaron métodos adicionales a la búsqueda en bases de datos para identificar informes relevantes? S/PS/PN/N/NI* 2.3 ¿Los términos y la estructura de la estrategia de búsqueda permitieron recuperar el mayor número posible de estudios elegibles? S/PS/PN/N/NI* 2.4 ¿Fueron apropiadas las restricciones basadas en la fecha, el formato de publicación o el idioma? S/PS/PN/N/NI* 2.5 ¿Se realizaron esfuerzos para minimizar los errores en la selección de los estudios? S/PS/PN/N/NI*	3.1 ¿Se hicieron esfuerzos para minimizar los errores en la recogida de datos? S/PS/PN/N/NI* 3.2 ¿Se dispuso de suficientes características de los estudios para que tanto los autores de la revisión como los lectores pudieran interpretar los resultados? S/PS/PN/N/NI* 3.3 ¿Se recogieron todos los resultados relevantes de los estudios para su uso en la síntesis? S/PS/PN/N/NI* 3.4 ¿Se evaluó formalmente el riesgo de sesgo (o la calidad metodológica) utilizando criterios apropiados? S/PS/PN/N/NI* 3.5 ¿Se hicieron esfuerzos para minimizar los errores en la evaluación del riesgo de sesgo? S/PS/PN/N/NI*	4.1 ¿Incluyó la síntesis todos los estudios que debía? S/PS/PN/N/NI* 4.2 ¿Se comunicaron todos los análisis predefinidos o se explicaron las desviaciones? S/PS/PN/N/NI* 4.3 ¿Fue adecuada la síntesis dada la naturaleza y la similitud de las preguntas de investigación, los diseños de los estudios y los resultados de los estudios incluidos? S/PS/PN/N/NI* 4.4 ¿La variación entre estudios (heterogeneidad) fue mínima o se abordó en la síntesis? S/PS/PN/N/NI* No aplica 4.5 ¿Fueron sólidos los hallazgos, por ejemplo, como se demostró mediante el gráfico en embudo o los análisis de sensibilidad? S/PS/PN/N/NI* No aplica 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? S/PS/PN/N/NI*	A. ¿Se han abordado en la interpretación de los resultados todos los problemas señalados en los dominios del uno al cuatro? S/PS/PN/N/NI* B. ¿Se consideró adecuadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? S/PS/PN/N/NI* C. ¿Evitaron los revisores destacar los resultados en función de su significación estadística? S/PS/PN/N/NI*
¿Coincide la pregunta abordada por la revisión con la pregunta objetivo? SI/INCIERTO/NO	**Preocupación por la especificación de los criterios de elegibilidad del estudio** ALTA/INCIERTA/BAJA	**Preocupación por los métodos utilizados para identificar y/o seleccionar los estudios** ALTA/INCIERTA/BAJA	**Preocupación por los métodos de recogida de datos y evaluación de los estudios** ALTA/INCIERTA/BAJA	**Preocupaciones sobre la síntesis y los resultados** ALTA/INCIERTA/BAJA	**Riesgo de sesgo global considerado** ALTA/BAJA/INEXISTENTE

© Álvaro Limones Burgos, 2024

ANEXO 5.3: Resultados de la evaluación del cumplimiento de la guía de reporte (PRISMA).

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
TÍTULO			
Título	1	Identificar en el título del manuscrito que se trata de una revisión sistemática.	p. 1
RESUMEN			
Objetivos	2.1	Proporcionar una declaración explícita de los principales objetivos de la revisión.	p.1
Criterios de Inclusión y exclusion	2.2	Especificar los criterios de inclusión y exclusión de la revisión.	p.1
Fuentes de información	2.3	Especificar las fuentes de información (por ejemplo, las bases de datos o registros) utilizados para identificar los estudios y la fecha de la última búsqueda.	p.1
Riesgo de sesgo	2.4	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos.	p. 1
Métodos de la síntesis de los resultados	2.5	Especifique los métodos utilizados para presentar y sintetizar los resultados.	p. 1
Estudios incluidos	2.6	Dar el número total de estudios y participantes incluidos y resumir las características pertinentes de los estudios.	p. 1
Síntesis de los resultados	2.7	Indicar las variables de estudio. Si se ha realizado un metaanálisis, notifique la estimación del efecto o exposición y su IC. Si se comparan dos o más grupos, indíquen la dirección del efecto (es decir, que grupo es favorecido).	p. 1
Degradación de la evidencia	2.8	Proporcionar un breve resumen de la degradación de la evidencia (por ejemplo, por riesgo de sesgo, incoherencia o impresión.	p. 1
Interpretación	2.9	Proporcionar una interpretación general de los resultados e implicaciones importantes.	p. 1
Financiación	2.10	Especificar la principal fuente de financiación de la revisión.	/
Registro	2.11	Proporcionar el nombre del registro y el número de registro.	/
INTRODUCIÓN			
Fundamento	3	Describir la justificación de la revisión en el contexto de los conocimientos existentes.	p. 1-2
Objetivos	4	Proporcionar una declaración explícita de los objetivos de la revisión.	p. 1
MÉTODOS			
Criterios de elegibilidad	5	Especificar los criterios de inclusión y exclusión de la revisión y cómo se agruparon los estudios para la síntesis.	p. 2
Fuentes de información	6	Especificar todas las bases de datos, registros, sitios web, organizaciones, listas de referencia y otras fuentes consultadas para identificar estudios. Especificar la fecha de la última revisión.	p. 2
Estrategia de búsqueda	7	Presentar las estrategias de búsqueda completa para al menos la base de datos principal, indicar filtros u otras estrategias de búsqueda utilizadas.	p. 2
Proceso de selección	8	Especificar los métodos utilizados para la selección de estudios.	p. 2

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
Recogida de datos	9	Especifique los métodos utilizados para recopilar datos de los estudios, incluyendo cuántos revisores recogieron datos de cada informe.	p. 2-3
Datos extraidos	10a	Listar y definir todas las variables críticas.	p. 2-3
	10b	Listar y definir todas las demás variables secundarias. Describa cualquier suposición sobre cualquier información que falta o no esté clara.	p. 2-3
Métodos de evaluación del riesgo de sesgo	11	Especifique los métodos utilizados para evaluar el riesgo de sesgos en los estudios incluidos, incluyendo detalles de la herramienta o herramientas utilizadas, cuántos revisores evaluaron cada estudio y si trabajaban independientemente, y si se utilizaron procesos de automatización.	p. 2
Medidas del efecto	12	Especifícar para cada variable la medida(s) del efecto (por ejemplo. relación de riesgo, diferencia media) utilizada en la síntesis o presentación de resultados.	p. 3
Métodos de síntesis	13a	Describir los procesos utilizados para decidir qué estudios eran elegibles para cada síntesis (por ejemplo: tabular las características de intervención del estudio y comparar con los grupos previstos para cada síntesis).	p. 2 y Material Suplementario
	13b	Describir cualquier método necesario para preparar los datos de presentación o síntesis, como el tratamiento de las estadísticas de datos faltantes o las conversiones de datos.	p. 2-3
	13c	Describir cualquier método utilizado para tabular o visualizar los resultados de los estudios y la síntesis individuales.	p. 2-3
	13d	Describir cualquier método utilizado para sintetizar los resultados y proporcionar una justificación. Si se realizó un metaanálisis, describa el(s) modelo(s), método(s) para identificar la presencia y extensión de la heterogeneidad estadística, y paquetes de software utilizado.	p. 2-3
	13e	Describa cualquier método utilizado para explorar posibles causas de heterogeneidad entre los resultados del estudio (por ejemplo, análisis de subgrupos, metaregresión).	No aplica
	13f	Describir cualquier análisis de sensibilidad realizado para evaluar la solidez de los resultados sintetizados.	No aplica
Implicaciones del riesgo de sesgo por datos faltantes	14	Describir cualquier método han utilizado para evaluar el riesgo de sesgo debido a la falta de datos (a partir de sesgos informativos).	p. 2-3
Evaluación de la certeza de la evidencia	15	Describir cualquier método utilizado para evaluar la certeza (o la confianza) de la evidencia.	p. 2-3
RESULTADOS			
Selección de estudios	16a	Describir los resultados del proceso de búsqueda y selección, idealmente utilizando un diagrama de flujo.	p. 3
	16b	Presentar los estudios excluidos y sus razones.	p. 3
Características de los estudios incluidos	17	Presentar cada estudio incluido y sus características.	Material Suplementario
Riesgo de sesgo en los estudios incluidos	18	Presentar el riesgo de sesgo para cada estudio incluido.	p. 3 y 5
Resultados de los estudios	19	Para todas las variables presentar la estimación del efecto y su IC.	p. 4 y 5

© Álvaro Limones Burgos, 2024

Sección y Tema	Número	Aspectos de comprobación	Ubicación donde se informa en el artículo
individuales			
Resultados de la síntesis	20a	Resumir brevemente las características y el riesgo de sesgo de los estudios incluidos.	p. 6-8
	20b	Si se ha realizado un metaanálisis, presente para cada una de las estimaciones medidas de heterogeneidad estadística. Si se comparan los grupos, describa la dirección del efecto.	p. 8
	20c	Informar de posibles causas de heterogeneidad.	No aplica
	20d	Informar de los análisis de sensibilidad realizados para evaluar la solidez de los resultados sintetizados.	No aplica
Evaluación del riesgo de sesgo por datos faltantes	21	Informar de la evaluación del riesgo de sesgo debido a datos faltantes.	p. 3 y 5
Certeza de la evidencia	22	Evaluaciones de la certeza (o confianza) de la evidencia para cada variable evaluada.	No aplica
DISCUSIÓN			
Discusión	23a	Proporcionar una interpretación general de los resultados en el contexto científico actual.	p. 5-8
	23b	Examinar cualquier limitación de la evidencia incluida en la revisión.	p. 8
	23c	Debatir las limitaciones de los procesos de revisión utilizados.	p. 8
	23d	Debatir las implicaciones de los resultados de la práctica, la política y la futura investigación.	p. 8
OTRA INFORMACIÓN			
Registro y protocolo	24a	Proporcionar información del registro de la revisión, incluyendo el nombre del registro y el número de registro, o afirmar que la revisión no está registrada.	p. 2
	24b	Indicar dónde puede accederse al protocolo de revisión o afirmar que no se ha elaborado un protocolo.	p. 2
	24c	Describir y explicar las modificaciones de la información facilitada en el registro o en el protocolo.	p. 2
Apoyo	25	Describa las fuentes de ayuda financiera o no financiera para la revisión, y el papel de los patrocinadores en la revisión.	p. 1
Intereses de competencia	26	Declare cualquier interés competitivo de los autores de revisión.	p. 1
Disponibilidad de datos, código y otros materiales	27	Informe sobre cuál de los siguientes está disponible públicamente y dónde se pueden encontrar: plantillas de recogida de datos; datos extraídos de estudios incluidos; datos utilizados para todos los análisis; o cualquier otro material utilizado en la revisión.	OSF: 10.17605/OSF.IO/CW5F7

12. REFERENCIAS

Agha, R. A., Sohrabi, C., Mathew, G., Franchi, T., Kerwan, A., & O'Neill, N. (2020). The PROCESS 2020 Guideline: Updating Consensus Preferred Reporting Of CasESeries in Surgery (PROCESS) Guidelines. *Int J Surg*, *84*, 231-235. https://doi.org/10.1016/j.ijsu.2020.11.005

Allen, F., & Locker, D. (2002). A modified short version of the oral health impact profile for assessing health-related quality of life in edentulous adults. *Int J Prosthodont*, *15*(5), 446-450.

Alonso-Coello, P., Schünemann, H. J., Moberg, J., Brignardello-Petersen, R., Akl, E. A., Davoli, M., Treweek, S., Mustafa, R. A., Rada, G., Rosenbaum, S., Morelli, A., Guyatt, G. H., & Oxman, A. D. (2018). [GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction]. *Gac Sanit*, *32*(2), 166.e161-166.e110. https://doi.org/10.1016/j.gaceta.2017.02.010 (Marcos GRADE de la evidencia a la decisión (EtD): un enfoque sistemático y transparente para tomar decisiones sanitarias bien informadas. 1: Introducción.)

Altman, D. G., & Bland, J. M. (1996). Detecting skewness from summary information. *Bmj*, *313*(7066), 1200. https://doi.org/10.1136/bmj.313.7066.1200

Andrade, C. (2021). HARKing, Cherry-Picking, P-Hacking, Fishing Expeditions, and Data Dredging and Mining as Questionable Research Practices. *J Clin Psychiatry*, *82*(1). https://doi.org/10.4088/JCP.20f13804

Andrews, J., Guyatt, G., Oxman, A. D., Alderson, P., Dahm, P., Falck-Ytter, Y., Nasser, M., Meerpohl, J., Post, P. N., Kunz, R., Brozek, J., Vist, G., Rind, D., Akl, E. A., & Schünemann, H. J. (2013). GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol*, *66*(7), 719-725. https://doi.org/10.1016/j.jclinepi.2012.03.013

Andrews, J. C., Schünemann, H. J., Oxman, A. D., Pottie, K., Meerpohl, J. J., Coello, P. A., Rind, D., Montori, V. M., Brito, J. P., Norris, S., Elbarbary, M., Post, P., Nasser, M., Shukla, V., Jaeschke, R., Brozek, J., Djulbegovic, B., & Guyatt, G. (2013). GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *J Clin Epidemiol*, *66*(7), 726-735. https://doi.org/10.1016/j.jclinepi.2013.02.003

Anzures-Cabrera, J., Sarpatwari, A., & Higgins, J. P. (2011). Expressing findings from meta-analyses of continuous outcomes in terms of risks. *Stat Med*, *30*(25), 2967-2985. https://doi.org/10.1002/sim.4298

© Álvaro Limones Burgos, 2024

Arbuthnot, J. (1710). II. An argument for divine providence, taken from the constant regularity observ'd in the births of both sexes. By Dr. John Arbuthnott, Physitian in Ordinary to Her Majesty, and Fellow of the College of Physitians and the Royal Society. *Philosophical Transactions of the Royal Society of London*, *27*(328), 186-190.

Archibald L. Cochrane. (1973). *Effectiveness and efficiency: Random Reflections on health services*. Nuffield Provincial Hospitals Trust, London. ISBN: 0900574178.

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19-32. https://doi.org/10.1080/1364557032000119616

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid Based Healthc*, *13*(3), 132-140. https://doi.org/10.1097/xeb.0000000000000055

Asimov, I. (2014). *Historia y cronología de la ciencia y los descubrimientos: Cómo la ciencia ha dado forma a nuestro mundo* (Ariel, Ed.).

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., Vist, G. E., Falck-Ytter, Y., Meerpohl, J., Norris, S., & Guyatt, G. H. (2011). GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*, *64*(4), 401-406. https://doi.org/10.1016/j.jclinepi.2010.07.015

Barker, T. H., Stone, J. C., Sears, K., Klugar, M., Tufanaru, C., Leonardi-Bee, J., Aromataris, E., & Munn, Z. (2023). The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. *JBI Evid Synth*, *21*(3), 494-506. https://doi.org/10.11124/jbies-22-00430

Barry, E. S., Merkebu, J., & Varpio, L. (2022). State-of-the-art literature review methodology: A six-step approach for knowledge synthesis. *Perspect Med Educ*, *11*(5), 281-288. https://doi.org/10.1007/s40037-022-00725-9

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088-1101.

Beller, E. M., Glasziou, P. P., Altman, D. G., Hopewell, S., Bastian, H., Chalmers, I., Gøtzsche, P. C., Lasserson, T., & Tovey, D. (2013). PRISMA for Abstracts: reporting systematic reviews in journal and conference abstracts. *PLoS Med*, *10*(4), e1001419. https://doi.org/10.1371/journal.pmed.1001419

Borges Migliavaca, C., Stein, C., Colpani, V., Barker, T. H., Munn, Z., & Falavigna, M. (2020). How are systematic reviews of prevalence conducted? A methodological study. *BMC Med Res Methodol*, *20*(1), 96. https://doi.org/10.1186/s12874-020-00975-3

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C., Kressel, H. Y., Rifai, N., Golub, R. M., Altman,

© Álvaro Limones Burgos, 2024

D. G., Hooft, L., Korevaar, D. A., & Cohen, J. F. (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Bmj*, *351*, h5527. https://doi.org/10.1136/bmj.h5527

Bradburn, M. J., Deeks, J. J., Berlin, J. A., & Russell Localio, A. (2007). Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med*, *26*(1), 53-77. https://doi.org/10.1002/sim.2528

Brignardello-Petersen, R., Guyatt, G. H., Mustafa, R. A., Chu, D. K., Hultcrantz, M., Schünemann, H. J., & Tomlinson, G. (2021). GRADE guidelines 33: Addressing imprecision in a network meta-analysis. *J Clin Epidemiol*, *139*, 49-56. https://doi.org/10.1016/j.jclinepi.2021.07.011

Brouwers, M. C., Kerkvliet, K., & Spithoff, K. (2016). The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *Bmj*, *352*, i1152. https://doi.org/10.1136/bmj.i1152

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., Fervers, B., Graham, I. D., Grimshaw, J., Hanna, S. E., Littlejohns, P., Makarski, J., & Zitzelsberger, L. (2012). The Global Rating Scale complements the AGREE II in advancing the quality of practice guidelines. *Journal of Clinical Epidemiology*, *65*(5), 526-534. https://doi.org/https://doi.org/10.1016/j.jclinepi.2011.10.008

Brozek, J. L., Canelo-Aybar, C., Akl, E. A., Bowen, J. M., Bucher, J., Chiu, W. A., Cronin, M., Djulbegovic, B., Falavigna, M., Guyatt, G. H., Gordon, A. A., Hilton Boon, M., Hutubessy, R. C. W., Joore, M. A., Katikireddi, V., LaKind, J., Langendam, M., Manja, V., Magnuson, K., . . . Schünemann, H. J. (2021). GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence-An overview in the context of health decision-making. *J Clin Epidemiol*, *129*, 138-150. https://doi.org/10.1016/j.jclinepi.2020.09.018

Brunetti, M., Shemilt, I., Pregno, S., Vale, L., Oxman, A. D., Lord, J., Sisk, J., Ruiz, F., Hill, S., Guyatt, G. H., Jaeschke, R., Helfand, M., Harbour, R., Davoli, M., Amato, L., Liberati, A., & Schünemann, H. J. (2013). GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol*, *66*(2), 140-150. https://doi.org/10.1016/j.jclinepi.2012.04.012

California-Dental-Association. (1977). Guidelines for the assessment of clinical quality and professional performance. *Los Angeles: CDA*, 53-55.

Calvert, M., Kyte, D., Mercieca-Bebber, R., Slade, A., Chan, A. W., King, M. T., Hunn, A., Bottomley, A., Regnault, A., Chan, A. W., Ells, C., O'Connor, D., Revicki, D., Patrick, D., Altman, D., Basch, E., Velikova, G., Price, G., Draper, H., . . . Groves, T. (2018). Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *Jama*, *319*(5), 483-494. https://doi.org/10.1001/jama.2017.21903

Campbell, F., Tricco, A. C., Munn, Z., Pollock, D., Saran, A., Sutton, A., White, H., & Khalil, H. (2023). Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): the same

© Álvaro Limones Burgos, 2024

but different— the "Big Picture" review family. *Systematic Reviews*, *12*(1). https://doi.org/10.1186/s13643-023-02178-5

Campbell, J. M., Kulgar, M., Ding, S., Carmody, D. P., Hakonsen, S. J., Jadotte, Y. T., & Ws, C. (2020). Diagnostic test accuracy systematic reviews. *JBI manual for evidence synthesis*, *1*.

Carrasco-Labra, A., Brignardello-Petersen, R., Santesso, N., Neumann, I., Mustafa, R. A., Mbuagbaw, L., Etxeandia Ikobaltzeta, I., De Stio, C., McCullagh, L. J., Alonso-Coello, P., Meerpohl, J. J., Vandvik, P. O., Brozek, J. L., Akl, E. A., Bossuyt, P., Churchill, R., Glenton, C., Rosenbaum, S., Tugwell, P., . . . Schünemann, H. J. (2016). Improving GRADE evidence tables part 1: a randomized trial shows improved understanding of content in summary of findings tables with a new format. *J Clin Epidemiol*, *74*, 7-18. https://doi.org/10.1016/j.jclinepi.2015.12.007

Castillo-Oyagüe, R., Perea, C., Suárez-García, M. J., Río, J. D., Lynch, C. D., & Preciado, A. (2016). Validation of the "Quality of Life related to function, aesthetics, socialization, and thoughts about health-behavioural habits (QoLFAST-10)" scale for wearers of implant-supported fixed partial dentures. *J Dent*, *55*, 82-91. https://doi.org/10.1016/j.jdent.2016.10.004

Castillo-Oyagüe, R., Suárez-García, M. J., Perea, C., Río, J. D., Lynch, C. D., Gonzalo, E., Torres-Lagares, D., & Preciado, A. (2016). Validation of a new, specific, complete, and short OHRQoL scale (QoLFAST-10) for wearers of implant overdentures and fixed-detachable hybrid prostheses. *J Dent*, *49*, 22-32. https://doi.org/10.1016/j.jdent.2016.04.011

Catalá-López, F., & Tobías, A. (2013). Clinical evidence synthesis and network meta-analysis with indirect-treatment comparisons. *Med Clin (Barc)*, *140*(4), 182-187. https://doi.org/10.1016/j.medcli.2012.09.013 (Síntesis de la evidencia clínica y metaanálisis en red con comparaciones indirectas.)

Chan, A. W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P. C., Krleža-Jerić, K., Hróbjartsson, A., Mann, H., Dickersin, K., Berlin, J. A., Doré, C. J., Parulekar, W. R., Summerskill, W. S., Groves, T., Schulz, K. F., Sox, H. C., Rockhold, F. W., Rennie, D., & Moher, D. (2013). SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*, *158*(3), 200-207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Chen, J., Chiang, C., & Zhang, Y. (2018). Esthetic evaluation of natural teeth in anterior maxilla using the pink and white esthetic scores. *Clin Implant Dent Relat Res*, *20*(5), 770-777. https://doi.org/10.1111/cid.12631

Chen, Y., Yang, K., Marušić, A., Qaseem, A., Meerpohl, J. J., Flottorp, S., Akl, E. A., Schünemann, H. J., Chan, E. S., & Falck-Ytter, Y. (2017). A reporting tool for practice guidelines in health care: the RIGHT statement. *Annals of internal medicine*, *166*(2), 128-132.

Clarkson, J., & Worthington, H. (2019). Leadership in evidence based dentistry. *J Dent*, *87*, 16-19. https://doi.org/10.1016/j.jdent.2019.05.012

© Álvaro Limones Burgos, 2024

Cohen, J. F., Deeks, J. J., Hooft, L., Salameh, J. P., Korevaar, D. A., Gatsonis, C., Hopewell, S., Hunt, H. A., Hyde, C. J., Leeflang, M. M., Macaskill, P., McGrath, T. A., Moher, D., Reitsma, J. B., Rutjes, A. W. S., Takwoingi, Y., Tonelli, M., Whiting, P., Willis, B. H., . . . McInnes, M. D. F. (2021). Preferred reporting items for journal and conference abstracts of systematic reviews and meta-analyses of diagnostic test accuracy studies (PRISMA-DTA for Abstracts): checklist, explanation, and elaboration. *Bmj*, *372*, n265. https://doi.org/10.1136/bmj.n265

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *J Clin Epidemiol*, *68*(2), 134-143. https://doi.org/10.1016/j.jclinepi.2014.11.010

Cruz Rivera, S., Liu, X., Chan, A. W., Denniston, A. K., & Calvert, M. J. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit Health*, *2*(10), e549-e560. https://doi.org/10.1016/s2589-7500(20)30219-3

Cuello-Garcia, C. A., Santesso, N., Morgan, R. L., Verbeek, J., Thayer, K., Ansari, M. T., Meerpohl, J., Schwingshackl, L., Katikireddi, S. V., Brozek, J. L., Reeves, B., Murad, M. H., Falavigna, M., Mustafa, R., Regidor, D. L., Alexander, P. E., Garner, P., Akl, E. A., Guyatt, G., & Schünemann, H. J. (2022). GRADE guidance 24 optimizing the integration of randomized and non-randomized studies of interventions in evidence syntheses and health guidelines. *J Clin Epidemiol*, *142*, 200-208. https://doi.org/10.1016/j.jclinepi.2021.11.026

Cvar, J. F., & Ryge, G. (2005). Reprint of criteria for the clinical evaluation of dental restorative materials. 1971. *Clin Oral Investig*, *9*(4), 215-232. https://doi.org/10.1007/s00784-005-0018-z

Damen, J. A. A., Moons, K. G. M., van Smeden, M., & Hooft, L. (2023). How to conduct a systematic review and meta-analysis of prognostic model studies. *Clin Microbiol Infect*, *29*(4), 434-440. https://doi.org/10.1016/j.cmi.2022.07.019

De Vries, R. B., Hooijmans, C. R., Langendam, M. W., van Luijk, J., Leenaars, M., Ritskes-Hoitinga, M., & Wever, K. E. (2015). A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. *Evidence-based Preclinical Medicine*, *2*(1), 1-9.

Deeks, J. J., Bossuyt, P. M., Leeflang, M. M., & Takwoingi, Y. (2023). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. *Cochrane, 2023*, *Version 2.0 (updated July 2023)*, Available from https://training.cochrane.org/handbook-diagnostic-test-accuracy/current.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials*, *7*(3), 177-188. https://doi.org/10.1016/0197-2456(86)90046-2

© Álvaro Limones Burgos, 2024

Dobbins, M. (2017). Rapid Review Guidebook: Steps for conducting a rapid review. National Collaborating Centre for Methods and Tools (NCCMT). *McMaster University, Canada*, Available from https://www.nccmt.ca/capacity-development/rapid-review-guidebook.

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*, *6*(12), e011458. https://doi.org/10.1136/bmjopen-2016-011458

Dreyer, N. A., Bryant, A., & Velentgas, P. (2016). The GRACE Checklist: A Validated Assessment Tool for High Quality Observational Studies of Comparative Effectiveness. *J Manag Care Spec Pharm*, *22*(10), 1107-1113. https://doi.org/10.18553/jmcp.2016.22.10.1107

Dwan, K., Li, T., Altman, D. G., & Elbourne, D. (2019). CONSORT 2010 statement: extension to randomised crossover trials. *Bmj*, *366*, l4378. https://doi.org/10.1136/bmj.l4378

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

Eliades, T., Behrents, R. G., Lindauer, S. J., & Rice, D. P. (2024). Reducing the quality of our evidence base by publishing at any cost. *Eur J Orthod*, *46*(3). https://doi.org/10.1093/ejo/cjae015

Elsman, E. B. M., Butcher, N. J., Mokkink, L. B., Terwee, C. B., Tricco, A., Gagnier, J. J., Aiyegbusi, O. L., Barnett, C., Smith, M., Moher, D., & Offringa, M. (2022). Study protocol for developing, piloting and disseminating the PRISMA-COSMIN guideline: a new reporting guideline for systematic reviews of outcome measurement instruments. *Syst Rev*, *11*(1), 121. https://doi.org/10.1186/s13643-022-01994-5

Evidence-Based Medicine Working Group. (1992). Evidence-based medicine. A new approach to teaching the practice of medicine. *Jama*, *268*(17), 2420-2425. https://doi.org/10.1001/jama.1992.03490170092032

Feng, Y., Liang, S., Zhang, Y., Chen, S., Wang, Q., Huang, T., Sun, F., Liu, X., Zhu, H., & Pan, H. (2022). Automated medical literature screening using artificial intelligence: a systematic review and meta-analysis. *J Am Med Inform Assoc*, *29*(8), 1425-1432. https://doi.org/10.1093/jamia/ocac066

Fire, M., & Guestrin, C. (2019). Over-optimization of academic publishing metrics: observing Goodhart's Law in action. *Gigascience*, *8*(6). https://doi.org/10.1093/gigascience/giz053

Fisher, R. A. (1925). Statistical methods for research workers. *Oliver and Boyd: Edinburgh*

, *11th ed.*

Foroutan, F., Guyatt, G., Zuk, V., Vandvik, P. O., Alba, A. C., Mustafa, R., Vernooij, R., Arevalo-Rodriguez, I., Munn, Z., Roshanov, P., Riley, R., Schandelmaier, S., Kuijpers, T.,

© Álvaro Limones Burgos, 2024

Siemieniuk, R., Canelo-Aybar, C., Schunemann, H., & Iorio, A. (2020). GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. *J Clin Epidemiol*, *121*, 62-70. https://doi.org/10.1016/j.jclinepi.2019.12.023

Gagnier, J. J., Kienle, G., Altman, D. G., Moher, D., Sox, H., & Riley, D. (2014). The CARE guidelines: consensus-based clinical case report guideline development. *J Clin Epidemiol*, *67*(1), 46-51. https://doi.org/10.1016/j.jclinepi.2013.08.003

Galbraith, R. (1988). Graphical display of estimates having differing standard errors. *Technometrics*, *30*(3), 271-281.

Garfield, E. (2006). The history and meaning of the journal impact factor. *Jama*, *295*(1), 90-93. https://doi.org/10.1001/jama.295.1.90

Garner, P., Hopewell, S., Chandler, J., MacLehose, H., Schünemann, H. J., Akl, E. A., Beyene, J., Chang, S., Churchill, R., Dearness, K., Guyatt, G., Lefebvre, C., Liles, B., Marshall, R., Martínez García, L., Mavergames, C., Nasser, M., Qaseem, A., Sampson, M., . . . Schünemann, H. J. (2016). When and how to update systematic reviews: consensus and checklist. *Bmj*, *354*, i3507. https://doi.org/10.1136/bmj.i3507

Garritty, C., Gartlehner, G., Nussbaumer-Streit, B., King, V. J., Hamel, C., Kamel, C., Affengruber, L., & Stevens, A. (2021). Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. *J Clin Epidemiol*, *130*, 13-22. https://doi.org/10.1016/j.jclinepi.2020.10.007

Garritty, C., Tsertsvadze, A., Tricco, A. C., Sampson, M., & Moher, D. (2010). Updating systematic reviews: an international survey. *PLoS One*, *5*(4), e9914. https://doi.org/10.1371/journal.pone.0009914

Gates, M., Gates, A., Pieper, D., Fernandes, R. M., Tricco, A. C., Moher, D., Brennan, S. E., Li, T., Pollock, M., & Lunny, C. (2022). Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. *Bmj*, *378*.

Goel, H., Raheja, D., & Nadar, S. K. (2024). Evidence-based medicine or statistically manipulated medicine? Are we slaves to the P-value? *Postgrad Med J*. https://doi.org/10.1093/postmj/qgae012

Goldkuhle, M., Bender, R., Akl, E. A., van Dalen, E. C., Nevitt, S., Mustafa, R. A., Guyatt, G. H., Trivella, M., Djulbegovic, B., Schünemann, H., Cinquini, M., Kreuzberger, N., & Skoetz, N. (2021). GRADE Guidelines: 29. Rating the certainty in time-to-event outcomes-Study limitations due to censoring of participants with missing data in intervention studies. *J Clin Epidemiol*, *129*, 126-137. https://doi.org/10.1016/j.jclinepi.2020.09.017

Gómez-Polo, C., Montero, J., Gómez-Polo, M., de Parga, J. A., & Celemin-Viñuela, A. (2017). Natural Tooth Color Estimation Based on Age and Gender. *J Prosthodont*, *26*(2), 107-114. https://doi.org/10.1111/jopr.12345

© Álvaro Limones Burgos, 2024

Grimes, D. A., & Schulz, K. F. (2002). An overview of clinical research: the lay of the land. *Lancet*, *359*(9300), 57-61. https://doi.org/10.1016/s0140-6736(02)07283-5

Guyatt, G., Ebrahim, S., Alonso-Coello, P., Johnston, B. C., Mathioudakis, A. G., Briel, M., Mustafa, R. A., Sun, X., Walter, S. D., Heels-Ansdell, D., Neumann, I., Kahale, L. A., Iorio, A., Meerpohl, J., Schünemann, H. J., & Akl, E. A. (2017). GRADE guidelines 17: assessing the risk of bias associated with missing participant outcome data in a body of evidence. *J Clin Epidemiol*, *87*, 14-22. https://doi.org/10.1016/j.jclinepi.2017.05.005

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., DeBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., & Schünemann, H. J. (2011). GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*, *64*(4), 383-394. https://doi.org/10.1016/j.jclinepi.2010.04.026

Guyatt, G., Oxman, A. D., Kunz, R., Atkins, D., Brozek, J., Vist, G., Alderson, P., Glasziou, P., Falck-Ytter, Y., & Schünemann, H. J. (2011). GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*, *64*(4), 395-400. https://doi.org/10.1016/j.jclinepi.2010.09.012

Guyatt, G., Oxman, A. D., Kunz, R., Brozek, J., Alonso-Coello, P., Rind, D., Devereaux, P. J., Montori, V. M., Freyschuss, B., Vist, G., Jaeschke, R., Williams, J. W., Jr., Murad, M. H., Sinclair, D., Falck-Ytter, Y., Meerpohl, J., Whittington, C., Thorlund, K., Andrews, J., & Schünemann, H. J. (2011). GRADE guidelines 6. Rating the quality of evidence--imprecision. *J Clin Epidemiol*, *64*(12), 1283-1293. https://doi.org/10.1016/j.jclinepi.2011.01.012

Guyatt, G., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Falck-Ytter, Y., Jaeschke, R., Vist, G., Akl, E. A., Post, P. N., Norris, S., Meerpohl, J., Shukla, V. K., Nasser, M., & Schünemann, H. J. (2011). GRADE guidelines: 8. Rating the quality of evidence--indirectness. *J Clin Epidemiol*, *64*(12), 1303-1310. https://doi.org/10.1016/j.jclinepi.2011.04.014

Guyatt, G., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., Alonso-Coello, P., Glasziou, P., Jaeschke, R., Akl, E. A., Norris, S., Vist, G., Dahm, P., Shukla, V. K., Higgins, J., Falck-Ytter, Y., & Schünemann, H. J. (2011). GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *J Clin Epidemiol*, *64*(12), 1294-1302. https://doi.org/10.1016/j.jclinepi.2011.03.017

Guyatt, G., Oxman, A. D., Montori, V., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., Djulbegovic, B., Atkins, D., Falck-Ytter, Y., Williams, J. W., Jr., Meerpohl, J., Norris, S. L., Akl, E. A., & Schünemann, H. J. (2011). GRADE guidelines: 5. Rating the quality of evidence--publication bias. *J Clin Epidemiol*, *64*(12), 1277-1282. https://doi.org/10.1016/j.jclinepi.2011.01.011

Guyatt, G., Oxman, A. D., Santesso, N., Helfand, M., Vist, G., Kunz, R., Brozek, J., Norris, S., Meerpohl, J., Djulbegovic, B., Alonso-Coello, P., Post, P. N., Busse, J. W., Glasziou, P., Christensen, R., & Schünemann, H. J. (2013). GRADE guidelines: 12. Preparing summary

© Álvaro Limones Burgos, 2024

of findings tables-binary outcomes. *J Clin Epidemiol*, *66*(2), 158-172. https://doi.org/10.1016/j.jclinepi.2012.01.012

Guyatt, G., Oxman, A. D., Sultan, S., Brozek, J., Glasziou, P., Alonso-Coello, P., Atkins, D., Kunz, R., Montori, V., Jaeschke, R., Rind, D., Dahm, P., Akl, E. A., Meerpohl, J., Vist, G., Berliner, E., Norris, S., Falck-Ytter, Y., & Schünemann, H. J. (2013). GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *J Clin Epidemiol*, *66*(2), 151-157. https://doi.org/10.1016/j.jclinepi.2012.01.006

Guyatt, G., Oxman, A. D., Sultan, S., Glasziou, P., Akl, E. A., Alonso-Coello, P., Atkins, D., Kunz, R., Brozek, J., Montori, V., Jaeschke, R., Rind, D., Dahm, P., Meerpohl, J., Vist, G., Berliner, E., Norris, S., Falck-Ytter, Y., Murad, M. H., & Schünemann, H. J. (2011). GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol*, *64*(12), 1311-1316. https://doi.org/10.1016/j.jclinepi.2011.06.004

Guyatt, G., Oxman, A. D., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., Montori, V., Akl, E. A., Djulbegovic, B., Falck-Ytter, Y., Norris, S. L., Williams, J. W., Jr., Atkins, D., Meerpohl, J., & Schünemann, H. J. (2011). GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *J Clin Epidemiol*, *64*(4), 407-415. https://doi.org/10.1016/j.jclinepi.2010.07.017

Guyatt, G., Thorlund, K., Oxman, A. D., Walter, S. D., Patrick, D., Furukawa, T. A., Johnston, B. C., Karanicolas, P., Akl, E. A., Vist, G., Kunz, R., Brozek, J., Kupper, L. L., Martin, S. L., Meerpohl, J. J., Alonso-Coello, P., Christensen, R., & Schunemann, H. J. (2013). GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. *J Clin Epidemiol*, *66*(2), 173-183. https://doi.org/10.1016/j.jclinepi.2012.08.001

Guyatt, G., Zhao, Y., Mayer, M., Briel, M., Mustafa, R., Izcovich, A., Hultcrantz, M., Iorio, A., Alba, A. C., Foroutan, F., Sun, X., Schunemann, H., DeBeer, H., Akl, E. A., Christensen, R., & Schandelmaier, S. (2023). GRADE guidance 36: updates to GRADE's approach to addressing inconsistency. *J Clin Epidemiol*, *158*, 70-83. https://doi.org/10.1016/j.jclinepi.2023.03.003

Hakoum, M. B., Jouni, N., Abou-Jaoude, E. A., Hasbani, D. J., Abou-Jaoude, E. A., Lopes, L. C., Khaldieh, M., Hammoud, M. Z., Al-Gibbawi, M., Anouti, S., Guyatt, G., & Akl, E. A. (2017). Characteristics of funding of clinical trials: cross-sectional survey and proposed guidance. *BMJ Open*, *7*(10), e015997. https://doi.org/10.1136/bmjopen-2017-015997

Hamel, C., Michaud, A., Thuku, M., Skidmore, B., Stevens, A., Nussbaumer-Streit, B., & Garritty, C. (2021). Defining Rapid Reviews: a systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews. *J Clin Epidemiol*, *129*, 74-85. https://doi.org/10.1016/j.jclinepi.2020.09.041

Hao, Q., Gao, Y., Zhao, Y., Murad, M. H., Mustafa, R., Ansari, M. T., Schünemann, H. J., Rind, D. M., Brignardello-Petersen, R., & Guyatt, G. (2023). GRADE concept 6: a novel application of external indirect evidence into GRADE ratings of evidence certainty in

© Álvaro Limones Burgos, 2024

network meta-analysis. *J Clin Epidemiol*, *163*, 95-101. https://doi.org/10.1016/j.jclinepi.2023.09.006

Harbord, R. M., Egger, M., & Sterne, J. A. (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*, *25*(20), 3443-3457. https://doi.org/10.1002/sim.2380

Harrison, S., Jones, H. E., Martin, R. M., Lewis, S. J., & Higgins, J. P. T. (2017). The albatross plot: A novel graphical tool for presenting results of diversely reported studies in a systematic review. *Res Synth Methods*, *8*(3), 281-289. https://doi.org/10.1002/jrsm.1239

Hattle, M., Burke, D. L., Trikalinos, T., Schmid, C. H., Chen, Y., Jackson, D., & Riley, R. D. (2022). Multivariate meta-analysis of multiple outcomes: characteristics and predictors of borrowing of strength from Cochrane reviews. *Syst Rev*, *11*(1), 149. https://doi.org/10.1186/s13643-022-01999-0

Hayden, J. A., van der Windt, D. A., Cartwright, J. L., Côté, P., & Bombardier, C. (2013). Assessing bias in studies of prognostic factors. *Ann Intern Med*, *158*(4), 280-286. https://doi.org/10.7326/0003-4819-158-4-201302190-00009

Herrera, D., Berglundh, T., Schwarz, F., Chapple, I., Jepsen, S., Sculean, A., Kebschull, M., Papapanou, P. N., Tonetti, M. S., & Sanz, M. (2023). Prevention and treatment of peri-implant diseases-The EFP S3 level clinical practice guideline. *J Clin Periodontol*, *50 Suppl 26*, 4-76. https://doi.org/10.1111/jcpe.13823

Hickel, R., Peschke, A., Tyas, M., Mjör, I., Bayne, S., Peters, M., Hiller, K. A., Randall, R., Vanherle, G., & Heintze, S. D. (2010). FDI World Dental Federation - clinical criteria for the evaluation of direct and indirect restorations. Update and clinical examples. *J Adhes Dent*, *12*(4), 259-272. https://doi.org/10.3290/j.jad.a19262

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & VA, W. (2022). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). *Cochrane*, Available from www.training.cochrane.org/handbook.

Hooijmans, C. R., Rovers, M. M., de Vries, R. B., Leenaars, M., Ritskes-Hoitinga, M., & Langendam, M. W. (2014). SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol*, *14*, 43. https://doi.org/10.1186/1471-2288-14-43

Hopewell, S., Clarke, M., Moher, D., Wager, E., Middleton, P., Altman, D. G., & Schulz, K. F. (2008). CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med*, *5*(1), e20. https://doi.org/10.1371/journal.pmed.0050020

Hopewell, S., Loudon, K., Clarke, M. J., Oxman, A. D., & Dickersin, K. (2009). Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev*, *2009*(1), Mr000006. https://doi.org/10.1002/14651858.MR000006.pub3

© Álvaro Limones Burgos, 2024

Hosseini, M., & Gotfredsen, K. (2012). A feasible, aesthetic quality evaluation of implant-supported single crowns: an analysis of validity and reliability. *Clin Oral Implants Res*, *23*(4), 453-458. https://doi.org/10.1111/j.1600-0501.2011.02162.x

Hultcrantz, M., Rind, D., Akl, E. A., Treweek, S., Mustafa, R. A., Iorio, A., Alper, B. S., Meerpohl, J. J., Murad, M. H., Ansari, M. T., Katikireddi, S. V., Östlund, P., Tranæus, S., Christensen, R., Gartlehner, G., Brozek, J., Izcovich, A., Schünemann, H., & Guyatt, G. (2017). The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol*, *87*, 4-13. https://doi.org/10.1016/j.jclinepi.2017.05.006

Husereau, D., Drummond, M., Augustovski, F., de Bekker-Grob, E., Briggs, A. H., Carswell, C., Caulley, L., Chaiyakunapruk, N., Greenberg, D., Loder, E., Mauskopf, J., Mullins, C. D., Petrou, S., Pwu, R. F., Staniszewska, S., & Force, C. I. G. R. P. T. (2022). Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *Bmj*, *376*, e067975. https://doi.org/10.1136/bmj-2021-067975

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., Ioannidis, J. P., Straus, S., Thorlund, K., Jansen, J. P., Mulrow, C., Catalá-López, F., Gøtzsche, P. C., Dickersin, K., Boutron, I., Altman, D. G., & Moher, D. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*, *162*(11), 777-784. https://doi.org/10.7326/m14-2385

Jarde, A., Losilla, J.-M., Vives, J., & F. Rodrigo, M. (2013). Q-Coh: A tool to screen the methodological quality of cohort studies in systematic reviews and meta-analyses. *International Journal of Clinical and Health Psychology*, *13*(2), 138-146. https://doi.org/https://doi.org/10.1016/S1697-2600(13)70017-6

Juszczak, E., Altman, D. G., Hopewell, S., & Schulz, K. (2019). Reporting of Multi-Arm Parallel-Group Randomized Trials: Extension of the CONSORT 2010 Statement. *Jama*, *321*(16), 1610-1620. https://doi.org/10.1001/jama.2019.3087

Katz, D. L., Williams, A. L., Girard, C., Goodman, J., Comerford, B., Behrman, A., & Bracken, M. B. (2003). The evidence base for complementary and alternative medicine: methods of Evidence Mapping with application to CAM. *Altern Ther Health Med*, *9*(4), 22-30.

Kicinski, M., Springate, D. A., & Kontopantelis, E. (2015). Publication bias in meta-analyses from the Cochrane Database of Systematic Reviews. *Stat Med*, *34*(20), 2781-2793. https://doi.org/10.1002/sim.6525

L'Abbé, K. A., Detsky, A. S., & O'Rourke, K. (1987). Meta-analysis in clinical research. *Ann Intern Med*, *107*(2), 224-233. https://doi.org/10.7326/0003-4819-107-2-224

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation Science*, *5*(1), 69. https://doi.org/10.1186/1748-5908-5-69

© Álvaro Limones Burgos, 2024

Limones, A., Celemín-Viñuela, A., Romeo-Rubio, M., Castillo-Oyagüe, R., Gómez-Polo, M., & Martínez Vázquez de Parga, J. A. (2022). Outcome measurements and quality of randomized controlled clinical trials of tooth-supported fixed dental prostheses: A systematic review and qualitative analysis. *J Prosthet Dent*. https://doi.org/10.1016/j.prosdent.2022.04.022

Limones, A., Molinero-Mourelle, P., Azevedo, L., Romeo-Rubio, M., Correia, A., & Gómez-Polo, M. (2020). Zirconia-ceramic versus metal-ceramic posterior multiunit tooth-supported fixed dental prostheses: A systematic review and meta-analysis of randomized controlled trials. *J Am Dent Assoc*, *151*(4), 230-238.e237. https://doi.org/10.1016/j.adaj.2019.12.013

Limones, A., Saez-Alcaide, L. M., Diaz-Parreno, S. A., Helm, A., Bornstein, M. M., & Molinero-Mourelle, P. (2020). Medication-related osteonecrosis of the jaws (MRONJ) in cancer patients treated with denosumab VS. zoledronic acid: A systematic review and meta-analysis. *Med Oral Patol Oral Cir Bucal*, *25*(3), e326-e336. https://doi.org/10.4317/medoral.23324

Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J., & Denniston, A. K. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med*, *26*(9), 1364-1374. https://doi.org/10.1038/s41591-020-1034-x

Liu, X., Rivera, S. C., Moher, D., Calvert, M. J., & Denniston, A. K. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *Bmj*, *370*, m3164. https://doi.org/10.1136/bmj.m3164

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*, *22*(4), 719-748.

Martin, C., Littlewood, S. J., Millett, D. T., Doubleday, B., Bearn, D., Worthington, H. V., & Limones, A. (2023). Retention procedures for stabilising tooth position after treatment with orthodontic braces. *Cochrane Database Syst Rev*, *5*(5), Cd002283. https://doi.org/10.1002/14651858.CD002283.pub5

Meng, Z., Wang, J., Lin, L., & Wu, C. (2024). Sensitivity analysis with iterative outlier detection for systematic reviews and meta-analyses. *Stat Med*, *43*(8), 1549-1563. https://doi.org/10.1002/sim.10008

Millard, L. A., Flach, P. A., & Higgins, J. P. (2016). Machine learning to assist risk-of-bias assessments in systematic reviews. *Int J Epidemiol*, *45*(1), 266-277. https://doi.org/10.1093/ije/dyv306

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet*, *354*(9193), 1896-1900. https://doi.org/10.1016/s0140-6736(99)04149-5

© Álvaro Limones Burgos, 2024

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*, *4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research*, *19*, 539-549.

Montaner, S. (1887). *Los Alfabetos: Los antiguos signos usados por el hombre en la construcción del mundo civilizado*.

Montero, J., López, J. F., Vicente, M. P., Galindo, M. P., Albaladejo, A., & Bravo, M. (2011). Comparative validity of the OIDP and OHIP-14 in describing the impact of oral health on quality of life in a cross-sectional study performed in Spanish adults. *Med Oral Patol Oral Cir Bucal*, *16*(6), e816-821. https://doi.org/10.4317/medoral.16851

Moons, K. G., de Groot, J. A., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., Reitsma, J. B., & Collins, G. S. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*, *11*(10), e1001744. https://doi.org/10.1371/journal.pmed.1001744

Morgan, R. L., Yepes-Nuñez, J. J., Ewusie, J., Mbuagbaw, L., Chang, S., Baldeh, T., Hempel, S., Helfand, M., Shekelle, P., Wilt, T. J., & Schünemann, H. J. (2023). Improving grading of recommendations assessment, development, and evaluation evidence tables part 4: a three-arm noninferiority randomized trial demonstrates improved understanding of content in summary of findings tables with a new format. *J Clin Epidemiol*, *154*, 125-135. https://doi.org/10.1016/j.jclinepi.2022.12.001

Munn, Z., Barker, T. H., Moola, S., Tufanaru, C., Stern, C., McArthur, A., Stephenson, M., & Aromataris, E. (2020). Methodological quality of case series studies: an introduction to the JBI critical appraisal tool. *JBI Evid Synth*, *18*(10), 2127-2133. https://doi.org/10.11124/jbisrir-d-19-00099

Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *Evid Based Med*, *21*(4), 125-127. https://doi.org/10.1136/ebmed-2016-110401

Murad, M. H., Verbeek, J., Schwingshackl, L., Filippini, T., Vinceti, M., Akl, E. A., Morgan, R. L., Mustafa, R. A., Zeraatkar, D., Senerth, E., Street, R., Lin, L., Falck-Ytter, Y., Guyatt, G., & Schünemann, H. J. (2023). GRADE GUIDANCE 38: Updated guidance for rating up certainty of evidence due to a dose-response gradient. *J Clin Epidemiol*, *164*, 45-53. https://doi.org/10.1016/j.jclinepi.2023.09.011

Mustafa, R. A., El Mikati, I. K., Murad, M. H., Hultcrantz, M., Steingart, K. R., Yang, B., Leeflang, M. M. G., Akl, E. A., Dahm, P., & Schünemann, H. J. (2024). GRADE guidance 37: rating imprecision in a body of evidence on test accuracy. *J Clin Epidemiol*, *165*, 111189. https://doi.org/10.1016/j.jclinepi.2023.10.005

© Álvaro Limones Burgos, 2024

Neumann, I., Quiñelen, E., Nahuelhual, P., Burdiles, P., Celedón, N., Cerda, K., Herrera-Omegna, P., Kraemer, P., Cancino, K. D., Valenzuela, J. P., Sepúlveda, D., Morgano, G. P., Akl, E. A., & Schünemann, H. J. (2022). Using Explicit Thresholds were valuable for judging Benefits and Harms in partially contextualized GRADE Guidelines. *J Clin Epidemiol*, *147*, 69-75. https://doi.org/10.1016/j.jclinepi.2022.03.017

O'Loughlin, K. T. (2020). Current Policies: Policy Statement on Evidence-Based Dentistry. *American Dental Association* 105.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology*, *134*, 178-189. https://doi.org/10.1016/j.jclinepi.2021.03.001

Pandis, N., Fleming, P. S., Hopewell, S., & Altman, D. G. (2015). The CONSORT Statement: Application within and adaptations for orthodontic trials. *Am J Orthod Dentofacial Orthop*, *147*(6), 663-679. https://doi.org/10.1016/j.ajodo.2015.03.014

Pandor, A., Kaltenthaler, E., Martyn-St James, M., Wong, R., Cooper, K., Dimairo, M., O'Cathain, A., Campbell, F., & Booth, A. (2019). Delphi consensus reached to produce a decision tool for SelecTing Approaches for Rapid Reviews (STARR). *J Clin Epidemiol*, *114*, 22-29. https://doi.org/10.1016/j.jclinepi.2019.06.005

Patel, D. R., O'Brien, T., Petrie, A., & Petridis, H. (2014). A systematic review of outcome measurements and quality of studies evaluating fixed tooth-supported restorations. *J Prosthodont*, *23*(6), 421-433. https://doi.org/10.1111/jopr.12160

Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., . . . Würbel, H. (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLoS Biol*, *18*(7), e3000410. https://doi.org/10.1371/journal.pbio.3000410

Perry, R., Whitmarsh, A., Leach, V., & Davies, P. (2021). A comparison of two assessment tools used in overviews of systematic reviews: ROBIS versus AMSTAR-2. *Systematic Reviews*, *10*(1), 273. https://doi.org/10.1186/s13643-021-01819-x

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc*, *13*(3), 141-146. https://doi.org/10.1097/xeb.0000000000000050

Peters, M. D. J., Godfrey, C., McInerney, P., Munn, Z., Tricco, A. C., & Khalil, H. (2020). Chapter 11: Scoping Reviews (2020 version). *JBI Manual for Evidence Synthesis*, *Available from https://synthesismanual.jbi.global*. https://doi.org/https://doi.org/10.46658/JBIMES-20-12

© Álvaro Limones Burgos, 2024

Piaggio, G., Elbourne, D. R., Pocock, S. J., Evans, S. J., & Altman, D. G. (2012). Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. *Jama*, *308*(24), 2594-2604. https://doi.org/10.1001/jama.2012.87802

Pogue, J. M., & Yusuf, S. (1997). Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Control Clin Trials*, *18*(6), 580-593; discussion 661-586. https://doi.org/10.1016/s0197-2456(97)00051-2

Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., & Koffel, J. B. (2021). PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews(). *J Med Libr Assoc*, *109*(2), 174-200. https://doi.org/10.5195/jmla.2021.962

RevMan 5. (2020). *Review Manager 5. Version 5.4.* In The Cochrane Collaboration.

Richards, D., & Lawrence, A. (1995). Evidence based dentistry. *Br Dent J*, *179*(7), 270-273. https://doi.org/10.1038/sj.bdj.4808896

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *Bmj*, *342*, d549. https://doi.org/10.1136/bmj.d549

Ruggiero, S. L., Dodson, T. B., Aghaloo, T., Carlson, E. R., Ward, B. B., & Kademani, D. (2022). American Association of Oral and Maxillofacial Surgeons' Position Paper on Medication-Related Osteonecrosis of the Jaws-2022 Update. *J Oral Maxillofac Surg*, *80*(5), 920-943. https://doi.org/10.1016/j.joms.2022.02.008

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & W.S., R. (1996). Evidence based medicine: what it is and what it isn't. *Bmj*, *312*, 71.

Salameh, J. P., Bossuyt, P. M., McGrath, T. A., Thombs, B. D., Hyde, C. J., Macaskill, P., Deeks, J. J., Leeflang, M., Korevaar, D. A., Whiting, P., Takwoingi, Y., Reitsma, J. B., Cohen, J. F., Frank, R. A., Hunt, H. A., Hooft, L., Rutjes, A. W. S., Willis, B. H., Gatsonis, C., . . . McInnes, M. D. F. (2020). Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. *Bmj*, *370*, m2632. https://doi.org/10.1136/bmj.m2632

Santesso, N., Glenton, C., Dahm, P., Garner, P., Akl, E. A., Alper, B., Brignardello-Petersen, R., Carrasco-Labra, A., De Beer, H., Hultcrantz, M., Kuijpers, T., Meerpohl, J., Morgan, R., Mustafa, R., Skoetz, N., Sultan, S., Wiysonge, C., Guyatt, G., & Schünemann, H. J. (2020). GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol*, *119*, 126-135. https://doi.org/10.1016/j.jclinepi.2019.10.014

Saran, A., & White, H. (2018). Evidence and gap maps: a comparison of different approaches. *Campbell Systematic Reviews*, *14*(1), 1-38. https://doi.org/10.4073/cmdp.2018.2

Sarkis-Onofre, R., Poletto-Neto, V., Cenci, M. S., Pereira-Cenci, T., & Moher, D. (2017). Impact of the CONSORT Statement endorsement in the completeness of reporting of randomized

© Álvaro Limones Burgos, 2024

clinical trials in restorative dentistry. *J Dent*, *58*, 54-59. https://doi.org/10.1016/j.jdent.2017.01.009

Savovic, J., Turner, R. M., Mawdsley, D., Jones, H. E., Beynon, R., Higgins, J. P. T., & Sterne, J. A. C. (2018). Association Between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-Epidemiologic Study. *Am J Epidemiol*, *187*(5), 1113-1122. https://doi.org/10.1093/aje/kwx344

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*, *63*(8), 834-840. https://doi.org/10.1016/j.jclinepi.2010.02.005

Schünemann, H., Brożek, J., Guyatt, G., & Oxman, A. (2013). Manual GRADE para calificar la calidad de la evidencia y la fuerza de la recomendación (1ª Ed. en español). P.A Orrego & M.X. Rojas (Trans.) Mar 2017. *Publicación Original: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html*.

Schünemann, H. J., Cuello, C., Akl, E. A., Mustafa, R. A., Meerpohl, J. J., Thayer, K., Morgan, R. L., Gartlehner, G., Kunz, R., Katikireddi, S. V., Sterne, J., Higgins, J. P., & Guyatt, G. (2019). GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol*, *111*, 105-114. https://doi.org/10.1016/j.jclinepi.2018.01.012

Schünemann, H. J., Mustafa, R., Brozek, J., Santesso, N., Alonso-Coello, P., Guyatt, G., Scholten, R., Langendam, M., Leeflang, M. M., Akl, E. A., Singh, J. A., Meerpohl, J., Hultcrantz, M., Bossuyt, P., & Oxman, A. D. (2016). GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *J Clin Epidemiol*, *76*, 89-98. https://doi.org/10.1016/j.jclinepi.2016.01.032

Schünemann, H. J., Mustafa, R. A., Brozek, J., Santesso, N., Bossuyt, P. M., Steingart, K. R., Leeflang, M., Lange, S., Trenti, T., Langendam, M., Scholten, R., Hooft, L., Murad, M. H., Jaeschke, R., Rutjes, A., Singh, J., Helfand, M., Glasziou, P., Arevalo-Rodriguez, I., . . . Guyatt, G. H. (2019). GRADE guidelines: 22. The GRADE approach for tests and strategies-from test accuracy to patient-important outcomes and recommendations. *J Clin Epidemiol*, *111*, 69-82. https://doi.org/10.1016/j.jclinepi.2019.02.003

Schünemann, H. J., Mustafa, R. A., Brozek, J., Steingart, K. R., Leeflang, M., Murad, M. H., Bossuyt, P., Glasziou, P., Jaeschke, R., Lange, S., Meerpohl, J., Langendam, M., Hultcrantz, M., Vist, G. E., Akl, E. A., Helfand, M., Santesso, N., Hooft, L., Scholten, R., . . . Guyatt, G. H. (2020a). GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol*, *122*, 129-141. https://doi.org/10.1016/j.jclinepi.2019.12.020

Schünemann, H. J., Mustafa, R. A., Brozek, J., Steingart, K. R., Leeflang, M., Murad, M. H., Bossuyt, P., Glasziou, P., Jaeschke, R., Lange, S., Meerpohl, J., Langendam, M., Hultcrantz, M., Vist, G. E., Akl, E. A., Helfand, M., Santesso, N., Hooft, L., Scholten, R., . . . Guyatt, G. H. (2020b). GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and

© Álvaro Limones Burgos, 2024

presenting it in evidence profiles and summary of findings tables. *J Clin Epidemiol*, *122*, 142-152. https://doi.org/10.1016/j.jclinepi.2019.12.021

Schünemann, H. J., Neumann, I., Hultcrantz, M., Brignardello-Petersen, R., Zeng, L., Murad, M. H., Izcovich, A., Morgano, G. P., Baldeh, T., Santesso, N., Cuello, C. G., Mbuagbaw, L., Guyatt, G., Wiercioch, W., Piggott, T., De Beer, H., Vinceti, M., Mathioudakis, A. G., Mayer, M. G., . . . Akl, E. A. (2022). GRADE guidance 35: update on rating imprecision for assessing contextualized certainty of evidence and making decisions. *J Clin Epidemiol*, *150*, 225-242. https://doi.org/10.1016/j.jclinepi.2022.07.015

Shan, T., Tay, F. R., & Gu, L. (2021). Application of Artificial Intelligence in Dentistry. *J Dent Res*, *100*(3), 232-244. https://doi.org/10.1177/0022034520969115

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*, *7*, 10. https://doi.org/10.1186/1471-2288-7-10

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *Bmj*, *358*, j4008. https://doi.org/10.1136/bmj.j4008

Sheth, V. H., Shah, N. P., Jain, R., Bhanushali, N., & Bhatnagar, V. (2022). Development and validation of a risk-of-bias tool for assessing in vitro studies conducted in dentistry: The QUIN. *J Prosthet Dent*. https://doi.org/10.1016/j.prosdent.2022.05.019

Skoetz, N., Goldkuhle, M., van Dalen, E. C., Akl, E. A., Trivella, M., Mustafa, R. A., Nowak, A., Dahm, P., Schünemann, H., & Bender, R. (2020). GRADE guidelines 27: how to calculate absolute effects for time-to-event outcomes in summary of findings tables and Evidence Profiles. *J Clin Epidemiol*, *118*, 124-131. https://doi.org/10.1016/j.jclinepi.2019.10.015

Slade, G. D. (1997). Derivation and validation of a short-form oral health impact profile. *Community Dent Oral Epidemiol*, *25*(4), 284-290. https://doi.org/10.1111/j.1600-0528.1997.tb00941.x

Song, Y., Darzi, A., Ballesteros, M., García, L. M., Alonso-Coello, P., Arayssi, T., Bhaumik, S., Chen, Y., Cluzeau, F., & Ghersi, D. (2019). Extending the RIGHT statement for reporting adapted practice guidelines in healthcare: the RIGHT-Ad@ pt Checklist protocol. *BMJ Open*, *9*(9), e031767.

Sponchiado-Júnior, E. C., Vieira, W. A., Frozoni, M., Herkrath, F. J., & de-Jesus-Soares, A. (2021). CONSORT Compliance in Randomized Clinical Trials of Regenerative Endodontic Treatments of Necrotic Immature Teeth: A Scoping Review. *J Endod*, *47*(11), 1751-1766. https://doi.org/10.1016/j.joen.2021.07.017

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A. W., Churchill, R.,

© Álvaro Limones Burgos, 2024

Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., . . . Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Bmj*, *355*, i4919. https://doi.org/10.1136/bmj.i4919

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H. Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., . . . Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj*, *366*, l4898. https://doi.org/10.1136/bmj.l4898

Stevens, A., Garritty, C., Hersi, M., & Moher, D. (2018). Developing PRISMA-RR, a reporting guideline for rapid reviews of primary studies (Protocol). Disponible en: http://www.equator-network.org/wp-content/uploads/2018/2002/PRISMA-RR-protocol.pdf.

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., Moher, D., Becker, B. J., Sipe, T. A., & Thacker, S. B. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*, *283*(15), 2008-2012. https://doi.org/10.1001/jama.283.15.2008

Sweeting, M. J., Sutton, A. J., & Lambert, P. C. (2004). What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Stat Med*, *23*(9), 1351-1375. https://doi.org/10.1002/sim.1761

Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol*, *58*(9), 894-901. https://doi.org/10.1016/j.jclinepi.2005.01.006

Tipton, E., Pustejovsky, J. E., & Ahmadi, H. (2019). A history of meta-regression: Technical, conceptual, and practical developments between 1974 and 2018. *Res Synth Methods*, *10*(2), 161-179. https://doi.org/10.1002/jrsm.1338

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467-473. https://doi.org/10.7326/m18-0850

Uttley, L., Quintana, D. S., Montgomery, P., Carroll, C., Page, M. J., Falzon, L., Sutton, A., & Moher, D. (2023). The problems with systematic reviews: a living systematic review. *J Clin Epidemiol*, *156*, 30-41. https://doi.org/10.1016/j.jclinepi.2023.01.011

Vallejo, I. (2023). *El infinito en un junco: la invención de los libros del mundo antiguo*. Siruela.

Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Res Synth Methods*, *1*(2), 112-125. https://doi.org/10.1002/jrsm.11

© Álvaro Limones Burgos, 2024

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*, *370*(9596), 1453-1457. https://doi.org/10.1016/s0140-6736(07)61602-x

Wallach, J. D. (2019). Meta-analysis Metastasis. *JAMA Intern Med*, *179*(11), 1594-1595. https://doi.org/10.1001/jamainternmed.2019.2999

Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.

Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., & Kleijnen, J. (2003). The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*, *3*, 25. https://doi.org/10.1186/1471-2288-3-25

Whiting, P., Savovic, J., Higgins, J. P., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J., Churchill, R., & group, R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*, *69*, 225-234. https://doi.org/10.1016/j.jclinepi.2015.06.005

Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M., Sterne, J. A., & Bossuyt, P. M. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, *155*(8), 529-536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

Wilensky, A., Shapira, L., Limones, A., & Martin, C. (2023). The efficacy of implant surface decontamination using chemicals during surgical treatment of peri-implantitis: A systematic review and meta-analysis. *J Clin Periodontol*, *50 Suppl 26*, 336-358. https://doi.org/10.1111/jcpe.13794

Williamson, P. R., Altman, D. G., Bagley, H., Barnes, K. L., Blazeby, J. M., Brookes, S. T., Clarke, M., Gargon, E., Gorst, S., Harman, N., Kirkham, J. J., McNair, A., Prinsen, C. A. C., Schmitt, J., Terwee, C. B., & Young, B. (2017). The COMET Handbook: version 1.0. *Trials*, *18*(Suppl 3), 280. https://doi.org/10.1186/s13063-017-1978-4

Wittneben, J. G., Wismeijer, D., Brägger, U., Joda, T., & Abou-Ayash, S. (2018). Patient-reported outcome measures focusing on aesthetics of implant- and tooth-supported fixed dental prostheses: A systematic review and meta-analysis. *Clin Oral Implants Res*, *29 Suppl 16*, 224-240. https://doi.org/10.1111/clr.13295

Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. (2019). PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med*, *170*(1), 51-58. https://doi.org/10.7326/m18-1376

© Álvaro Limones Burgos, 2024

Xie, F., Shemilt, I., Vale, L., Ruiz, F., Drummond, M. F., Lord, J., Herrmann, K. H., Rojas, M. X., Zhang, Y., Canelo-Aybar, C., Alonso-Coello, P., Shamliyan, T., & Schünemann, H. J. (2023). GRADE guidance 23: considering cost-effectiveness evidence in moving from evidence to health-related recommendations. *J Clin Epidemiol*, *162*, 135-144. https://doi.org/10.1016/j.jclinepi.2023.08.001

Yan, R., Jiang, R., Hu, L., Deng, Y., Wen, J., & Jiang, X. (2022). Establishment and assessment of rodent models of medication-related osteonecrosis of the jaw (MRONJ). *Int J Oral Sci*, *14*(1), 41. https://doi.org/10.1038/s41368-022-00182-4

Yang, B., Mustafa, R. A., Bossuyt, P. M., Brozek, J., Hultcrantz, M., Leeflang, M. M. G., Schünemann, H. J., & Langendam, M. W. (2021). GRADE Guidance: 31. Assessing the certainty across a body of evidence for comparative test accuracy. *J Clin Epidemiol*, *136*, 146-156. https://doi.org/10.1016/j.jclinepi.2021.04.001

Yao, X., Kumar, M. V., Su, E., Flores Miranda, A., Saha, A., & Sussman, J. (2024). Evaluating the efficacy of artificial intelligence tools for the automation of systematic reviews in cancer research: A systematic review. *Cancer Epidemiol*, *88*, 102511. https://doi.org/10.1016/j.canep.2023.102511

Yeves, J. A. (2019). *Documentos con pinturas: Diplomática, historia y arte*. Museo Lázaro Galdiano.

Zavalis, E. A., Rameau, A., Saraswathula, A., Vist, J., Schuit, E., & Ioannidis, J. P. (2024). Availability of evidence and comparative effectiveness for surgical versus drug interventions: an overview of systematic reviews and meta-analyses. *BMJ Open*, *14*(1), e076675. https://doi.org/10.1136/bmjopen-2023-076675

Zeng, L., Brignardello-Petersen, R., Hultcrantz, M., Mustafa, R. A., Murad, M. H., Iorio, A., Traversy, G., Akl, E. A., Mayer, M., Schünemann, H. J., & Guyatt, G. H. (2022). GRADE Guidance 34: update on rating imprecision using a minimally contextualized approach. *J Clin Epidemiol*, *150*, 216-224. https://doi.org/10.1016/j.jclinepi.2022.07.014

Zeng, L., Brignardello-Petersen, R., Hultcrantz, M., Siemieniuk, R. A. C., Santesso, N., Traversy, G., Izcovich, A., Sadeghirad, B., Alexander, P. E., Devji, T., Rochwerg, B., Murad, M. H., Morgan, R., Christensen, R., Schünemann, H. J., & Guyatt, G. H. (2021). GRADE guidelines 32: GRADE offers guidance on choosing targets of GRADE certainty of evidence ratings. *J Clin Epidemiol*, *137*, 163-175. https://doi.org/10.1016/j.jclinepi.2021.03.026

Zhang, Y., Alonso-Coello, P., Guyatt, G. H., Yepes-Nuñez, J. J., Akl, E. A., Hazlewood, G., Pardo-Hernandez, H., Etxeandia-Ikobaltzeta, I., Qaseem, A., Williams, J. W., Jr., Tugwell, P., Flottorp, S., Chang, Y., Zhang, Y., Mustafa, R. A., Rojas, M. X., & Schünemann, H. J. (2019). GRADE Guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. *J Clin Epidemiol*, *111*, 94-104. https://doi.org/10.1016/j.jclinepi.2018.01.013

© Álvaro Limones Burgos, 2024

Zhang, Y., Coello, P. A., Guyatt, G. H., Yepes-Nuñez, J. J., Akl, E. A., Hazlewood, G., Pardo-Hernandez, H., Etxeandia-Ikobaltzeta, I., Qaseem, A., Williams, J. W., Jr., Tugwell, P., Flottorp, S., Chang, Y., Zhang, Y., Mustafa, R. A., Rojas, M. X., Xie, F., & Schünemann, H. J. (2019). GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences-inconsistency, imprecision, and other domains. *J Clin Epidemiol*, *111*, 83-93. https://doi.org/10.1016/j.jclinepi.2018.05.011

Zhang, Y., Huang, L., Wang, D., Ren, P., Hong, Q., & Kang, D. (2021). The ROBINS-I and the NOS had similar reliability but differed in applicability: A random sampling observational studies of systematic reviews/meta-analysis. *J Evid Based Med*, *14*(2), 112-122. https://doi.org/10.1111/jebm.12427

© Álvaro Limones Burgos, 2024