



**FACULTAD DE FARMACIA**

**UNIVERSIDAD COMPLUTENSE**

**TRABAJO FIN DE GRADO**

**ANÁLISIS DESCRIPTIVO DE LOS ENSAYOS  
CLÍNICOS INICIADOS EN ESPAÑA EN 2014  
PUBLICADOS EN CLINICALTRIALS.GOV**

Autor: Luis Castro González

D.N.I.: 76132493-D

Tutor: José Manuel Massó Córdoba

Convocatoria: Febrero 2016

# Índice

1. RESUMEN.	2
2. ABSTRACT.	2
3. INTRODUCCIÓN Y ANTECEDENTES.	3
4. OBJETIVOS.	4
5. METODOLOGÍA.	4
6. RESULTADOS Y DISCUSIÓN.	5
7. DESARROLLO DE UNA APLICACIÓN WEB.	13
8. CONCLUSIONES.	13
Referencias	15
A. ANEXO.	16

## 1. RESUMEN.

*Clinicaltrials.gov* es una base datos de estudios clínicos en la que se deben registrar todos los ensayos clínicos de EE. UU. Es de carácter internacional por lo que estudiamos su relevancia analizando los ensayos clínicos iniciados en España en 2014 aquí publicados. Utilizando el lenguaje de programación estadística *R* tratamos los datos para adaptarlos a nuestro cometido. Hay publicados 712 ensayos clínicos de los 714 aprobados por la AEMPS. Las características de los ensayos españoles publicados en *Clinicaltrials.gov* no son completamente exactas a las descritas por la AEMPS debido a la no obligación de suministrar todos los datos si el ensayo no es de EE. UU. España es la 7<sup>a</sup> potencia en investigación por número de estudios publicados en *clinicaltrials.gov*. Tras el análisis podemos concluir con que los ensayos clínicos españoles tienen un perfil medio de asignación aleatoria, intervención paralela, enmascaramiento abierto, de fase 3, cuyo objetivo es determinar la eficacia y seguridad de algún tratamiento y en el que participan hombres y mujeres, adultos y ancianos. Adicionalmente se crea una aplicación web con el paquete *Shiny* de *R* que permite de forma dinámica e interactiva reproducir este análisis y ampliarlo según necesidades, disponible en <http://ensayosclinicos.infarmon.es/>. Ahora se conoce mejor la investigación clínica en España y se dispone de un análisis descriptivo de la misma.

## 2. ABSTRACT.

*Clinicaltrials.gov* is a database of clinical studies in which all interventional studies performed in USA must be registered. Because of its international nature, we assess its relevance with an analysis of the clinical trials started in Spain during 2014 which are published here. All the data are processed using the *R* statistical language, to adapt them to our problem. 712 out of 714 Spanish clinical trials approved by AEMPS are available, but their features are not completely accurate since the publication is not mandatory for the studies outside the USA. Spain is the 7th country by number of publications in *clinicaltrials.gov*. After the analysis we can conclude that the Spanish trials have the following predominant profile: randomized allocation, parallel assignment, doubled blind masking, phase 3, with the aim of determining the efficacy and safety of the treatments, and involving male, female, adults, and senior patients. Additionally, a web application has been developed using the *Shiny* package of *R*, which allows a dynamic and interactive replication of this analysis, deeper if desired, and available on the website: <http://ensayosclinicos.infarmon.es/>. As a result of this work, the situation of clinical research in Spain is better known, and a descriptive analysis is available.

### 3. INTRODUCCIÓN Y ANTECEDENTES.

El Real Decreto 1090/2015, de 4 de diciembre, por el que se regulan los ensayos clínicos con medicamentos, los Comités de Ética de la Investigación con medicamentos y el Registro Español de Estudios Clínicos, publicado en el Boletín Oficial del Estado [1] define ensayo clínico como «un estudio clínico que cumpla cualquiera de las siguientes condiciones: 1.<sup>o</sup> Se asigna de antemano al sujeto de ensayo a una estrategia terapéutica determinada, que no forma parte de la práctica clínica habitual del Estado miembro implicado. 2.<sup>o</sup> La decisión de prescribir los medicamentos en investigación se toma junto con la de incluir al sujeto en el estudio clínico. 3.<sup>o</sup> Se aplican procedimientos de diagnóstico o seguimiento a los sujetos de ensayo que van más allá de la práctica clínica habitual».

La Declaración de Helsinki [2], un manifiesto similar al de los Derechos humanos pero dirigido a las investigaciones médicas en humanos, establece que «se debe registrar cada ensayo clínico en una base de datos de acceso público antes de reclutar el primer sujeto».

Existen muchas bases de datos sobre ensayos clínicos, España cuenta con el Registro Español de estudios clínicos (REec) [3] que proporciona información a pacientes y profesionales de la salud sobre investigaciones clínicas permitidas en España para informar sobre los estudios que se están realizando en relación a una enfermedad o con uno o varios medicamentos determinados. La Biblioteca Nacional de Medicina del Instituto Nacional de Salud de Estados Unidos de América junto con la Agencia de Alimentos y Drogas (FDA) cuenta con la base de datos *Clinicaltrials.gov* [4]. La diferencias más llamativas son el carácter internacional de *clinicaltrials.gov* y la facilidad que alcanza para el acceso masivo a la información de estudios clínicos. La información que se recopila aquí es suministrada y actualizada por el promotor o por el investigador principal del estudio clínico y comprende ensayos clínicos, estudios observacionales y estudios de medicamentos de uso compasivo. Participan estudios de 191 países siendo obligatorio que se registren todos los ensayos clínicos estadounidenses en los que intervenga un fármaco o producto sanitario. Recibe 65.000 visitas diarias. De este modo *clinicaltrials.gov* es una de las bases de datos de estudios clínicos más importantes del mundo.

Al consultar la base de datos es posible filtrar estudios por parámetros como el tipo de estudio, el estado, el tipo de diseño, el promotor, el objetivo, el título, los grupos de edad y género que intervienen, la fase en la que se encuentra, así como la nacionalidad, fechas de interés, etc. Sin embargo no permite filtrar por las características del diseño como son la asignación, el modelo de intervención, enmascaramiento, el propósito principal, etc. Los resultados de la búsqueda se muestran por estudios individuales, no obteniéndose un análisis descriptivo o resumen de las características de los estudios filtrados, lo que proporcionaría una visión general de los ensayos.

Debido a la fuerte relevancia de *clinicaltrials.gov*, se ha decidido realizar un análisis descriptivo de los ensayos clínicos iniciados en España en 2014 publicados en esta base de datos. Este análisis está centrado en las características de los diseños de los estudios además de otros datos de interés como es la financiación. Gracias a este estudio comprobamos la fiabilidad de los registros de *clinicaltrials.gov* en función de las aprobaciones de la Agencia Española de Medicamentos y Productos Sanitarios (AEMPS), que en su memoria anual de actividades [5] tiene una sección dedicada a ensayos clínicos. Igualmente vamos a conocer mejor la situación de la investigación clínica en España y como son los ensayos clínicos en España.

## 4. OBJETIVOS.

Los principales objetivos de este trabajo son:

- Realizar un análisis descriptivo en el que se muestren como están diseñados los ensayos clínicos en España registrados en *clinicaltrials.gov*.
- Comprobar la fiabilidad de los datos de *clinicaltrials.gov* comparando nuestro análisis con el de la Agencia Española de Medicamentos y Productos Sanitarios.
- Crear una herramienta para la consulta del análisis descriptivo.

## 5. METODOLOGÍA.

La diana de estudio fueron los ensayos clínicos iniciados en España en 2014 y publicados en la base de datos *clinicaltrials.gov*. Escogimos 2014 y no 2015 porque a fecha de comienzo y realización de este trabajo no se había publicado la memoria de actividades de la AEMPS de 2015 donde se recoge un análisis de los ensayos clínicos aprobados ese año.

Se filtró la búsqueda por el parámetro país (Country) España (Spain). Se descargaron los 22 campos disponibles para todos los estudios realizados en España y registrados en esta base de datos en formato .csv (valores separados por coma). Para adaptar los datos al análisis, realizar el propio análisis descriptivo y exportar los resultados en formato gráfico se utilizó el lenguaje de programación *R* [6], versión 3.2.3. *R* es un lenguaje y entorno para estadística informática y es software libre.

Se creó una secuencia de comandos o *script* en *R* para adaptar los datos al análisis, analizar y representar los resultados en gráficas. Este *script* se puede encontrar en el anexo a este

documento. De esta forma todo el trabajo realizado es completamente reproducible y revisable. A grandes rasgos el *script* realiza las siguientes operaciones:

1. Carga en memoria el archivo de datos descargado.
2. Filtra por año de inicio (*Start.Date*) para seleccionar solamente los ensayos clínicos iniciados en 2014, ya que *clinicaltrials.gov* no proporciona este criterio de búsqueda.
3. Procesamiento de la variable diseños del estudio (*Study.Designs*) para separar las distintas características de cada ensayo clínico y poder trabajar con ellas.
4. Traducción a español de los campos a mostrar en las gráficas.
5. Generación de las gráficas.

Adicionalmente se creó una aplicación web que permite realizar un análisis personalizado de los datos de forma interactiva y dinámica, de modo tal que este trabajo pueda ser reproducido sin la necesidad de conocer el lenguaje de programación *R*. Para esto se utilizó el paquete *Shiny* [7] versión 0.13, que «proporciona una elegante y potente infraestructura web para construir aplicaciones usando *R*».

La aplicación creada se puede consultar en la dirección <http://ensayosclinicos.infarmon.es/> [8].

## 6. RESULTADOS Y DISCUSIÓN.

En primer lugar vemos la cantidad de ensayos clínicos y estudios observacionales iniciados en España en 2014 y publicados en *clinicaltrials.gov* que como se puede observar en la figura 6.1 se han iniciado 712 ensayos clínicos. El número de ensayos clínicos aprobados por la AEMPS en España en 2014 fueron 714. Este dato confirma la relevancia de *clinicaltrials.gov* ya que el 99,72% de los ensayos clínicos aprobados por la AEMPS fueron publicados en *clinicaltrials.gov*.

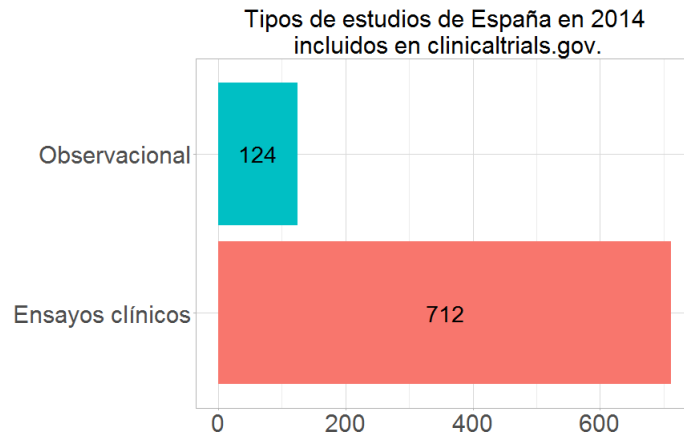


Figura 6.1: Ensayos clínicos y estudios observacionales.

En la figura 6.2 se muestran los grupos de edades que intervienen en los distintos ensayos clínicos. Destaca la participación conjunta de adultos y ancianos mientras que solo en el 14.33 % de los ensayos clínicos participan niños. En la tabla 6.1 se muestran los datos acumulados que comparándolos con los de la AEMPS se comprueba que son muy próximos. En el informe de la AEMPS solo se muestra la cantidad de estudios que incluyen niños, ancianos y adultos.

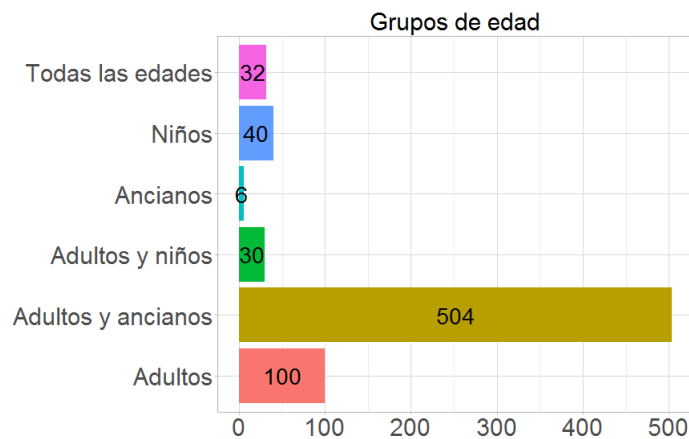


Figura 6.2: Grupos de edades.

En la figura 6.3 se muestra la distribución por sexos en los ensayos clínicos. Según la AEMPS en el 93 % de los ensayos clínicos participan hombres y mujeres, según *clinicaltrials.gov* en el 86 %.

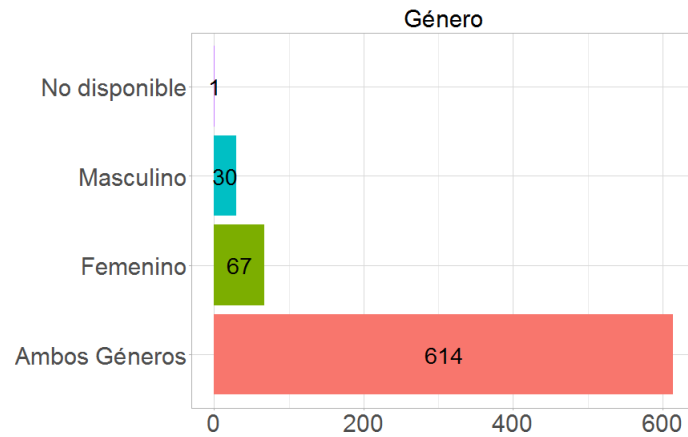


Figura 6.3: Géneros.

En la figura 6.4 se muestran las fases en las que se desarrollan los ensayos clínicos. Hay ensayos clínicos registrados con dos fases, esto quiere decir que se desarrollan en una fase que es una combinación de dos fases. En la figura 6.4 están incluidos dentro de la fase más avanzada a la que pertenecen, es decir, aquellos estudios de fase 1 y 2 están incluidos en fase 2 y aquellos de fase 2 y 3 están incluidos en fase 3. *Clinicaltrials.gov* trabaja con 5 fases que se definen a continuación:

- Fase 0: estudios exploratorios que implican una exposición muy limitada al fármaco, sin un objetivo terapéutico o diagnóstico (por ejemplo, screening y estudios de microdosis).
- Fase 1: estudios en los que normalmente participan voluntarios sanos y enfatizan en la seguridad. El objetivo es encontrar los efectos adversos más frecuentes y a menudo conocer la farmacocinética del medicamento.
- Fase 2: estudios que recogen información sobre la efectividad y farmacodinámica, es decir, si el medicamento tiene efecto en personas con la enfermedad y a que dosis. Por ejemplo, unos participantes reciben el medicamento de estudio y se compara con otros participantes que reciben otro producto, generalmente una sustancia inactiva llamada placebo. Se sigue comprobando la seguridad y los efectos adversos a corto plazo.
- Fase 3: estudios que recogen más información sobre seguridad y eficacia estudiando diferentes poblaciones, diferentes dosis y administrando el medicamento en condiciones más próximas a la clínica diaria de los pacientes.
- Fase 4: estudios postautorización. Recogen información adicional sobre la seguridad, eficacia y uso óptimo del medicamento.

Comparando con la AEMPS en la tabla 6.1 encontramos diferencias debido a la existencia de una fase 0 y a que no siempre se especifica la fase en *clinicaltrials.gov*.

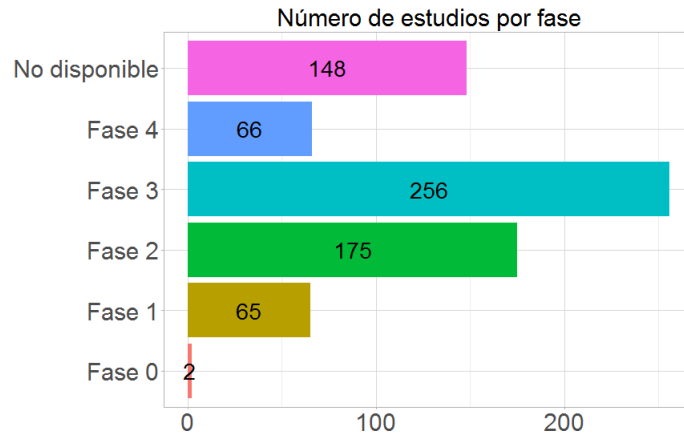


Figura 6.4: Fases de los ensayos.

En la figura 6.5 podemos observar el tipo principal de promotor de los ensayos clínicos. *Clinicaltrials.gov* incluye los siguientes posibles tipos de promotor:

- Institutos Nacionales de Salud (INS).
- Otras agencias federales de Estados Unidos de América como por ejemplo la Administración de Alimentos y Drogas, centros para el control y prevención de enfermedades y el Departamento de Asuntos Geriátricos.
- Industria: compañías farmacéuticas.
- Otros: individuales, universidades, organizaciones, etc.

La AEMPS solamente clasifica al promotor como laboratorio farmacéutico y no comercial. El promotor principal de los ensayos clínicos es la industria farmacéutica seguida muy por debajo por otros organismos privados como universidades y fundaciones.

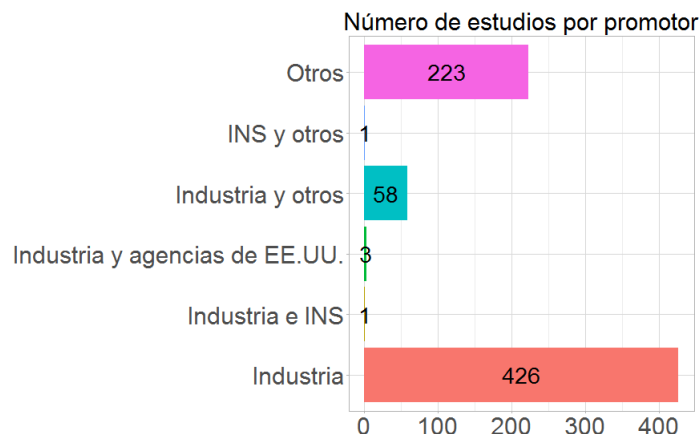


Figura 6.5: Tipo de promotor.

La tabla 6.1 resume el análisis comparativo llevado a cabo hasta el momento.

Dato	Clinicaltrials.gov	AEMPS
Ensayos clínicos	712	714
Ensayos clínicos en fase 1	65	135
Ensayos clínicos en fase 2	175	207
Ensayos clínicos en fase 3	256	286
Ensayos clínicos en fase 4	66	86
Ensayos clínicos que incluyen adultos	94 %	91 %
Ensayos clínicos que incluyen ancianos	72 %	72 %
Ensayos clínicos que incluyen niños	14 %	13 %
Ensayos clínicos que incluyen hombres y mujeres	86 %	93 %
Ensayos clínicos de laboratorios farmacéuticos	68 %	72 %
Ensayos clínicos de no comerciales	31 %	28 %

Tabla 6.1: Resumen comparativo.

A continuación se analizan las características de diseño de los ensayos clínicos:

- **Asignación:** hace referencia a como se ha procedido a la hora de asignar el objeto de estudio a los participantes. Puede ser asignación aleatoria o no aleatoria. Podemos comprobar en la tabla 6.2 que la mayoría de los ensayos clínicos son de asignación aleatoria. La asignación aleatoria es preferible a la no aleatoria porque se evitan posibles sesgos provocados por la subjetividad de las personas encargadas de la asignación.

Asignación	Número	Porcentaje	Enmascaramiento	Número	Porcentaje
Aleatoria	508	71.35 %	Abierto	371	52.11 %
No aleatoria	74	10.39 %	Simple ciego	69	9.69 %
No disponible	130	18.26 %	Doble ciego	272	38.20 %

Tabla 6.2: Asignación y enmascaramiento.

- **Enmascaramiento:** hace referencia al conocimiento que tiene el participante y/o el investigador sobre qué se está administrando al propio participante, si el medicamento de estudio, un medicamento comparador o un placebo. El enmascaramiento puede ser abierto cuando todos conocen el tratamiento que los pacientes están recibiendo, simple ciego cuando el participante no sabe lo que está recibiendo y doble ciego cuando ni el participante ni el investigador conocen el tratamiento que recibe el paciente. Comprobamos en la tabla 6.2 que el enmascaramiento es normalmente abierto o doble ciego. Un enmascaramiento doble ciego es la elección ideal ya que se

evitan posibles sesgos provocados por la subjetividad de las personas. Se ha comprobado con el uso de la aplicación que aquellos estudios que son no aleatorizados son abiertos, salvo cinco excepciones y que por tanto los aleatorizados son los únicos doble ciego, salvo una excepción.

- **Intervención:** hace referencia a como intervienen los participantes en el ensayo clínico. Puede ser que haya un solo grupo, o que haya dos grupos donde uno toma el medicamento a estudio y otro toma otro medicamento o un placebo. Cuando participan dos grupos, si los grupos son fijos en el tiempo se trata de intervención paralela, si por el contrario en un momento del ensayo clínico los participantes del grupo control pasan a tomar el medicamento de estudio y viceversa se trata de intervención cruzada y si van cambiando progresivamente al grupo de estudio se trata de intervención factorial. La tabla 6.3 indica que la mayoría de los ensayos clínicos tuvieron una intervención paralela.

Intervención	Número	Porcentaje
Un solo grupo	182	25.56 %
Paralela	495	69.52 %
Cruzada	31	4.35 %
Factorial	3	0.42 %
No disponible	1	0.14 %

Tabla 6.3: Intervención.

- **Propósito:** hace referencia al tipo de objeto de estudio. El propósito puede ser el tratamiento terapéutico, el tratamiento paliativo, investigación en servicios sanitarios, la prevención, el diagnóstico, ciencia básica, etc. En la tabla 6.4 se muestra que el propósito más común es el tratamiento. Se ha comprobado con el uso de la aplicación que los estudios de fase 2 y 3 tienen en su amplia mayoría como propósito el tratamiento, lo que concuerda con la propia definición de estas fases.

Propósito	Número	Porcentaje
Tratamiento	586	82.30 %
Tratamiento paliativo	16	2.25 %
Diagnóstico	15	2.11 %
Prevención	50	7.02 %
Ciencia básica	6	0.84 %
Investigación en servicios sanitarios	2	0.28 %
No disponible	37	5.20 %

Tabla 6.4: Propósito.

- Objetivo: hace referencia a la característica en estudio. El objetivo puede ser la seguridad, eficacia, farmacodinámica, farmacocinética, bioequivalencia y biodisponibilidad. En el cuadro 6.5 se muestra que establecer la seguridad y eficacia de un tratamiento es el objetivo principal. Se ha comprobado con el uso de la aplicación que los estudios de fase 2 y 3 tienen como objetivo la seguridad y eficacia, salvo cinco excepciones. Los ensayos en fase 1 son los ocupan la gran mayoría de objetivos de bioequivalencia, biodisponibilidad, farmacocinética y farmacodinámica. Lo que concuerda con la propia definición de estas fases.

Objetivo	Número	Porcentaje
Seguridad y eficacia	396	55.62 %
Seguridad	50	7.02 %
Eficacia	174	24.44 %
Farmacocinética y farmacodinámica	5	0.70 %
Farmacocinética	9	1.26 %
Farmacodinámica	1	0.14 %
Bioequivalencia	4	0.56 %
Biodisponibilidad	1	0.14 %
No disponible	72	10.11 %

Tabla 6.5: Objetivo.

Puesto	País	Número de ensayos clínicos
1 <sup>o</sup>	EE. UU.	7056
2 <sup>o</sup>	Canadá	1276
3 <sup>o</sup>	Francia	1271
4 <sup>o</sup>	Reino Unido	1131
5 <sup>o</sup>	Alemania	1003
6 <sup>o</sup>	China	915
7 <sup>o</sup>	ESPAÑA	787
8 <sup>o</sup>	Italia	703
9 <sup>o</sup>	Holanda	607
10 <sup>o</sup>	Bélgica	563
11 <sup>o</sup>	Brasil	500
12 <sup>o</sup>	Dinamarca	460
13 <sup>o</sup>	Israel	399
14 <sup>o</sup>	Polonia	399
15 <sup>o</sup>	Australia	393
16 <sup>o</sup>	Japón	375
17 <sup>o</sup>	Suiza	362
18 <sup>o</sup>	Suecia	350
19 <sup>o</sup>	República Checa	255
20 <sup>o</sup>	Austria	252

Tabla 6.6: Ensayos clínicos por país.

En la tabla 6.6 se comparan los datos de España en 2014 con los datos publicados en *clinicaltrials.gov* de otros países europeos y otras potencias internacionales de la investigación clínica. Vemos que España se sitúa en el 7<sup>o</sup> puesto de ensayos clínicos registrados en *clinicaltrials.gov* en 2014. Como ya dijimos anteriormente, el registro de un ensayo clínico en *clinicaltrials.gov* solo es obligatorio en EE.UU. por lo que este dato es meramente orientativo. Sin embargo podemos decir que en España, en lo referente a ensayos clínicos, se está trabajando de forma adecuada para garantizar las transparencias de dichas investigaciones.

Una vez comprobada la fiabilidad de *clinicaltrials.gov* hemos hecho otro análisis que se resume en la tabla 6.7 y predice el próximo análisis de la AEMPS para el año 2015. Como podemos observar ha habido menos ensayos clínicos, aumentando la participación de ancianos y disminuyendo la de niños, sucede lo mismo los promotores comerciales y no comerciales. Esta prospección se ha llevado a cabo gracias a la secuencia de comandos o *script* creado a lo largo del trabajo y que se encuentra en el anexo. Solo hubo que modificar en el *script*, en la línea siguiente al comentario «# Filtramos por año de inicio.», el año 2014 por 2015.

Dato	Análisis 2015	Análisis 2014
Ensayos clínicos	584	712
Ensayos clínicos en fase 1	47	65
Ensayos clínicos en fase 2	146	175
Ensayos clínicos en fase 3	197	256
Ensayos clínicos en fase 4	59	66
Ensayos clínicos que incluyen adultos	554 (94.86 %)	94 %
Ensayos clínicos que incluyen ancianos	457 (78.25 %)	72 %
Ensayos clínicos que incluyen niños	58 (9.93 %)	14 %
Ensayos clínicos que incluyen hombres y mujeres	511 (87.5 %)	86 %
Ensayos clínicos de laboratorios farmacéuticos	383 (65.58 %)	68 %
Ensayos clínicos de no comerciales	201 (34.42 %)	31 %

Tabla 6.7: Análisis de los ensayos clínicos de España en 2015.

Finalmente se ha desarrollado una aplicación web que permite obtener todos estos resultados según los criterios de consulta que se deseen. Hemos decidido crearla para darle valor a la reproducibilidad de este trabajo y para disponer de una herramienta de consulta de información de ensayos clínicos que resuelve ciertos problemas que presenta *clinicaltrials.gov*. A fecha de realización de este trabajo solo se pueden hacer consultas a los ensayos clínicos de España en 2014. Como trabajo futuro se ampliará la base de datos a más países y todas la fechas disponibles. La aplicación se puede consultar en <http://ensayosclnicos.infarmon.es/> [8].

## 7. DESARROLLO DE UNA APLICACIÓN WEB.

Como ya se ha comentado, se ha creado una aplicación web que permite reproducir este trabajo. Además permite ampliar el análisis y o realizar otro centrado en los criterios deseados. Se puede acceder a la aplicación a través de <http://ensayosclinicos.infarmon.es/> [8].

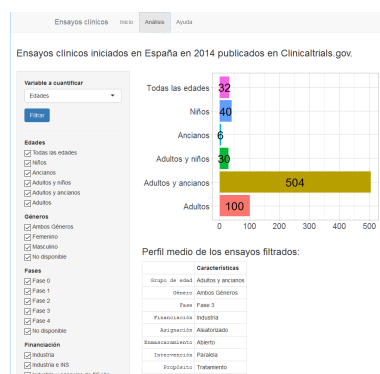


Figura 7.1: Aplicación web.

Como se puede ver en la figura 7.1 la aplicación es sencilla de utilizar. Una vez que se accede a ella pulsamos en la pestaña de «Análisis» y ya obtenemos los primeros resultados. Por defecto están establecidos los filtros necesarios para mostrar los mismos resultados que aquí se han mostrado. Para modificar la consulta solo hace falta marcar o desmarcar los criterios deseados. Siempre habrá que tener marcada al menos una característica de cada variable para que se desarrolle el análisis.

## 8. CONCLUSIONES.

*Clinicaltrials.gov* tiene una relevancia muy fuerte como base de datos de ensayos clínicos como hemos podido comprobar al comparar los estudios aquí registrados con los autorizados por la AEMPS. Además, los datos proporcionados son muy próximos a los de la AEMPS lo que otorga una alta fiabilidad a *clinicaltrials.gov* como base de datos de ensayos clínicos.

El perfil medio del ensayo clínico de España en 2014 incluye participantes adultos y ancianos de ambos géneros, es de fase 3, financiado y desarrollado por la industria farmacéutica, la asignación del objeto de estudio es aleatoria, el enmascaramiento del ensayo es abierto, la intervención de los participantes es en grupos paralelos y la meta del estudio es determinar la eficacia y seguridad de algún tratamiento.

Finalmente en <http://ensayosclinicos.infarmon.es/> [8] se puede reproducir el análisis llevado a cabo en este trabajo además de poder calcular nuevos datos de interés.

## Referencias

- [1] España, “Real Decreto 1090/2015, de 4 de diciembre, por el que se regulan los ensayos clínicos con medicamentos, los Comités de Ética de la Investigación con medicamentos y el Registro Español de Estudios Clínicos.,” *BOE*, vol. 307, pp. 121923–121964, Diciembre 2015.
- [2] A. M. Mundial, “Declaración de Helsinki de la AMM - Principios Éticos para las investigaciones médicas en seres humanos.,” Séptima revisión, 2013.
- [3] “Registro Español de Estudios Clínicos.” <https://reec.aemps.es/>. Visitado 25/01/2016.
- [4] “Clinicaltrials.” <https://clinicaltrials.gov/>. Visitado: 24/01/2015.
- [5] A. E. de Medicamentos y Productos Sanitarios, *Memoria de Actividades de la AEMPS 2014*. 2015.
- [6] R. C. Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.
- [7] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson, *shiny: Web Application Framework for R*, 2016. R package version 0.13.0.
- [8] “Ensayos Clínicos Infarmon.” <http://ensayosclinicos.infarmon.es>. Visitado: 25/01/2016.

## A. ANEXO.

```

# Cargamos librerías.
library(dplyr)
library(ggplot2)

# Cargamos los datos.
datos <- read.csv("TFG/Datos/DatosSpain.csv", stringsAsFactors = FALSE)

# Preparamos las gráficas.
tema <- theme_light() + theme(axis.text = element_text(size = 30), axis.title.y = element_
  _blank(), axis.title.x = element_blank(), legend.position = "None", plot.title =
  element_text(size = 30))
texto <- geom_text(aes(label = ..count.., y = ..count../2), stat = "count", size = 10)

# Traducimos los nombres de las columnas.
datos <- datos %>% rename(Género = Gender, Edades = Age.Groups, Fases = Phases,
  Financiación = Funded.Bys, Tipo = Study.Types, Diseño = Study.Designs, Inicio = Start
  .Date, NCT = NCT.Number, Título = Title, Enfermedades = Conditions)

# Filtramos por año de inicio.
datos <- datos %>% filter(grepl("2014$", Inicio))

# Traducimos los nombres de las variables.
datos$Tipo <- datos$Tipo %>% replace(datos$Tipo == "Observational", "Observacional")
datos$Tipo <- datos$Tipo %>% replace(datos$Tipo == "Interventional", "Ensayos_clínicos")

# Hacemos la primera gráfica antes de aislar los ensayos clínicos.
gdeestudios <- ggplot(data = datos, aes(x = Tipo, fill = Tipo)) + geom_bar(aes(y = ..
  count..)) + texto + coord_flip() + labs(x = "Tipos_de_estudio", y = "Número_de
  _estudios", title = "Tipos_de_estudios_de_España_en_2014\\n_incluidos_en_
  clinicaltrials.gov.")

# Aislamos los ensayos clínicos.
datos <- datos %>% filter(grepl("Ensayos_clínicos", Tipo))

# Creamos las variables de las características de los Diseños.
Asignación <- rep("No_disponible", length(datos$Tipo))
Objetivo <- rep("No_disponible", length(datos$Tipo))
Intervención <- rep("No_disponible", length(datos$Tipo))
Enmascaramiento <- rep("No_disponible", length(datos$Tipo))
Propósito <- rep("No_disponible", length(datos$Tipo))

# Separamos el contenido de Diseño por el símbolo "|".
datos$Diseño <- strsplit(as.character(datos$Diseño), "|", fixed = TRUE)

# Separamos el contenido de Diseño por el símbolo ":".
for (i in 1:length(datos$Diseño)) {
  for (j in 1:length(datos$Diseño[[i]])) {
    datos$Diseño[[i]][j] <- strsplit(as.character(datos$Diseño[[i]][j]), ":", fixed =
      TRUE)
  }
}

# Llenamos las variables de las características del Diseño.
for (i in 1:length(datos$Diseño)) {
  for (j in 1:length(datos$Diseño[[i]])) {
    if(!is.na(datos$Diseño[[i]][[j]][1]) && datos$Diseño[[i]][[j]][1]!="Allocation")

```

```

    Asignación[[i]] <- datos$Diseño[[i]][[j]][2]
  if(!is.na(datos$Diseño[[i]][[j]][1]) && datos$Diseño[[i]][[j]][1]!="Endpoint_
    Classification") Objetivo[[i]] <- datos$Diseño[[i]][[j]][2]
  if(!is.na(datos$Diseño[[i]][[j]][1]) && datos$Diseño[[i]][[j]][1]!="Intervention_
    Model") Intervención[[i]] <- datos$Diseño[[i]][[j]][2]
  if(!is.na(datos$Diseño[[i]][[j]][1]) && datos$Diseño[[i]][[j]][1]!="Masking")
    Enmascaramiento[[i]] <- datos$Diseño[[i]][[j]][2]
  if(!is.na(datos$Diseño[[i]][[j]][1]) && datos$Diseño[[i]][[j]][1]!="Primary_Purpose")
    Propósito[[i]] <- datos$Diseño[[i]][[j]][2]
}
}

# Traducimos y simplificamos algunas variables.
Asignación <- replace(Asignación, Asignación == "_Randomized", "Aleatorizado")
Asignación <- replace(Asignación, Asignación == "_Non-Randomized", "No_aleatorizado")
Objetivo <- replace(Objetivo, Objetivo == "_Safety/Efficacy_Study", "Seguridad/Eficacia")
Objetivo <- replace(Objetivo, Objetivo == "_Safety_Study", "Seguridad")
Objetivo <- replace(Objetivo, Objetivo == "_Pharmacokinetics/Dynamics_Study", "
  Farmacocinética/Farmacodinámica")
Objetivo <- replace(Objetivo, Objetivo == "_Pharmacokinetics_Study", "Farmacocinética")
Objetivo <- replace(Objetivo, Objetivo == "_Pharmacodynamics_Study", "Farmacodinámica")
Objetivo <- replace(Objetivo, Objetivo == "_Efficacy_Study", "Eficacia")
Objetivo <- replace(Objetivo, Objetivo == "_Bio-equivalence_Study", "Bioequivalencia")
Objetivo <- replace(Objetivo, Objetivo == "_Bio-availability_Study", "Biodisponibilidad")
Intervención <- replace(Intervención, Intervención == "_Single_Group_Assignment", "Un_
  solo_grupo")
Intervención <- replace(Intervención, Intervención == "_Parallel_Assignment", "Paralela")
Intervención <- replace(Intervención, Intervención == "_Factorial_Assignment", "Factorial
  ")
Intervención <- replace(Intervención, Intervención == "_Crossover_Assignment", "Cruzada")
Enmascaramiento <- replace(Enmascaramiento, grepl("_Double*", Enmascaramiento), "Doble_
  ciego")
Enmascaramiento <- replace(Enmascaramiento, grepl("_Single*", Enmascaramiento), "Simple_
  ciego")
Enmascaramiento <- replace(Enmascaramiento, grepl("_Open*", Enmascaramiento), "Abierto")
Propósito <- replace(Propósito, Propósito == "_Treatment", "Tratamiento")
Propósito <- replace(Propósito, Propósito == "_Supportive_Care", "Tratamiento_paliativo")
Propósito <- replace(Propósito, Propósito == "_Prevention", "Prevención")
Propósito <- replace(Propósito, Propósito == "_Health_Services_Research", "Servicios_
  sanitarios")
Propósito <- replace(Propósito, Propósito == "_Diagnostic", "Diagnóstico")
Propósito <- replace(Propósito, Propósito == "_Basic_Science", "Ciencia_básica")

# Ahora convertimos variables en dataframes y las añadimos a la tabla de datos.
datos <- cbind(datos, as.data.frame(Asignación), as.data.frame(Objetivo), as.data.frame(
  Intervención), as.data.frame(Enmascaramiento), as.data.frame(Propósito))

# Seguimos traduciendo y simplificando variables.
datos$Edades <- replace(datos$Edades, datos$Edades == "Senior", "Ancianos")
datos$Edades <- replace(datos$Edades, datos$Edades == "Child|Adult|Senior", "Todas_las_
  edades")
datos$Edades <- replace(datos$Edades, datos$Edades == "Child|Adult", "Adultos_y_niños")
datos$Edades <- replace(datos$Edades, datos$Edades == "Child", "Niños")
datos$Edades <- replace(datos$Edades, datos$Edades == "Adult|Senior", "Adultos_y_ancianos
  ")
datos$Edades <- replace(datos$Edades, datos$Edades == "Adult", "Adultos")
datos$Fases <- replace(datos$Fases, datos$Fases == "", "No_disponible")
datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_1|Phase_2", "Fase_2")
datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_2|Phase_3", "Fase_3")

```

```

datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_0", "Fase_0")
datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_1", "Fase_1")
datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_2", "Fase_2")
datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_3", "Fase_3")
datos$Fases <- replace(datos$Fases, datos$Fases == "Phase_4", "Fase_4")
datos$Género <- replace(datos$Género, datos$Género == "null", "No_disponible")
datos$Género <- replace(datos$Género, datos$Género == "Both", "Ambos_Géneros")
datos$Género <- replace(datos$Género, datos$Género == "Female", "Femenino")
datos$Género <- replace(datos$Género, datos$Género == "Male", "Masculino")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Industry", "
  Industria")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Industry|NIH", "
  Industria_e_INS")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Industry|Other", "
  Industria_y_otros")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Industry|U.S.
  Fed", "Industria_y_agencias_de_EE.UU.")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Other", "Otros")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Other|Industry", "
  Industria_y_otros")
datos$Financiación <- replace(datos$Financiación, datos$Financiación == "Other|NIH", "INS_
  y_otros")

# Generamos variables para ver datos acumulativos.
ancianos <- 0
adultos <- 0
niños <- 0
hombresymujeres <- 0
industria <- 0
nocomercial <- 0
for (i in 1:length(datos$Edades)) {
  if (datos$Edades[[i]] == "Ancianos" | datos$Edades[[i]] == "Adultos_y_ancianos" | datos
    $Edades[[i]] == "Todas_las_edades") ancianos <- ancianos + 1
  if (datos$Edades[[i]] == "Adultos" | datos$Edades[[i]] == "Adultos_y_ancianos" | datos$
    Edades[[i]] == "Adultos_y_niños" | datos$Edades[[i]] == "Todas_las_edades") adultos
    <- adultos + 1
  if (datos$Edades[[i]] == "Niños" | datos$Edades[[i]] == "Adultos_y_niños" | datos$
    Edades[[i]] == "Todas_las_edades") niños <- niños + 1
  if (datos$Género[[i]] == "Ambos_Géneros") hombresymujeres <- hombresymujeres + 1
  if (datos$Financiación[[i]] == "Industria" | datos$Financiación[[i]] == "Industria_y_
    agencias_de_EE.UU." | datos$Financiación[[i]] == "Industria_e_INS" | datos$
    Financiación[[i]] == "Industria_y_otros") industria <- industria + 1
  if (datos$Financiación[[i]] == "Otros" | datos$Financiación[[i]] == "INS_y_otros")
    nocomercial <- nocomercial + 1
}
pAncianos <- round(ancianos*100/length(datos$Edades), 2)
padultos <- round(adultos*100/length(datos$Edades), 2)
pniños <- round(niños*100/length(datos$Edades), 2)
phombresymujeres <- round(hombresymujeres*100/length(datos$Género), 2)
pindustria <- round(industria*100/length(datos$Financiación), 2)
pnocomercial <- round(nocomercial*100/length(datos$Financiación), 2)
pAsignación <- round(summary(datos$Asignación)*100/length(datos$Asignación), 2)
pIntervención <- round(summary(datos$Intervención)*100/length(datos$Intervención), 2)
penmascaramiento <- round(summary(datos$Enmascaramiento)*100/length(datos$Enmascaramiento
), 2)
pPropósito <- round(summary(datos$Propósito)*100/length(datos$Propósito), 2)
pobjetivo <- round(summary(datos$Objetivo)*100/length(datos$Objetivo), 2)

# Generamos las gráficas.

```

```

gdeEdad <- ggplot(data = datos, aes(x = Edades, fill = Edades)) + geom_bar(aes(y = ..
  count..)) + texto + labs(x = "Grupos_de_edad", y = "Número_de_personas", title = "
  Grupos_de_edad") + tema + coord_flip()
gdeGénero <- ggplot(data = datos, aes(x = Género, fill = Género)) + geom_bar(aes(y = ..
  count..)) + texto + labs(x = "Género", y = "Número_de_personas", title = "Género") +
  tema + coord_flip()
gdeFases <- ggplot(data = datos, aes(x = Fases, fill = Fases)) + geom_bar(aes(y = ..count
  ..)) + texto + labs(x = "Fases", y = "Número_de_estudios", title = "Número_de_
  estudios_por_fase") + tema + coord_flip()
gdeFinanciación <- ggplot(data = datos, aes(x = Financiación, fill = Financiación)) +
  geom_bar(aes(y = ..count..)) + texto + labs(x = "Promotor", y = "Número_de_estudios",
  title = "Número_de_estudios_por_promotor") + tema + coord_flip()
gdeAsignación <- ggplot(data = datos, aes(x = Asignación, fill = Asignación)) + geom_bar(
  aes(y = ..count..)) + texto + labs(x = "Asignación", y = "Número_de_estudios", title
  = "Número_de_estudios_por_tipo_de_Asignación") + tema + coord_flip()
gdeIntervención <- ggplot(data = datos, aes(x = Intervención, fill = Intervención)) +
  geom_bar(aes(y = ..count..)) + texto + labs(x = "Intervención", y = "Número_de_
  estudios", title = "Número_de_estudios_por_tipo_de_Intervención") + tema + coord_flip
  ()
gdeEnmascaramiento <- ggplot(data = datos, aes(x = Enmascaramiento, fill =
  Enmascaramiento)) + geom_bar(aes(y = ..count..)) + texto + labs(x = "Enmascaramiento"
  , y = "Número_de_estudios", title = "Número_de_estudios_por_tipo_de_enmascaramiento")
  + tema + coord_flip()
gdePropósito <- ggplot(data = datos, aes(x = Propósito, fill = Propósito)) + geom_bar(aes
  (y = ..count..)) + texto + labs(x = "Propósito", y = "Número_de_estudios", title = "
  Número_de_estudios_por_Propósito") + tema + coord_flip()
gdeObjetivo <- ggplot(data = datos, aes(x = Objetivo, fill = Objetivo)) + geom_bar(aes(y
  = ..count..)) + texto + labs(x = "Objetivo", y = "Número_de_estudios", title = "
  Número_de_estudios_por_objetivo") + tema + coord_flip()

# Exportamos los datos para mantener actualizada la aplicación.
datos$Diseño <- as.character(datos$Diseño)
write.csv(datos, "TFG/Shiny/datos/datos.csv", fileEncoding = "UTF-8", row.names = FALSE)

```