Received: 17 January 2025 | Accepted: 28 March 2025

DOI: 10.1111/opo.13507

ORIGINAL ARTICLE

OPO THE COLLEGE OF OPTOMETRISTS

Comparison of pupil size measurement: Repeatability and agreement across four devices with different measurement principles under varied lighting conditions

Irene Martínez-Alberquilla[1,2] | **Laura Rico-del-Viejo**[1,2] | **Abinaya Priya Venkataraman**[3] | **Alberto Dominguez-Vicent**[3]

[1]Clinical and Experimental Eye Research Group, Faculty of Optics and Optometry, Universidad Complutense de Madrid, Madrid, Spain

[2]Department of Optometry and Vision, Faculty of Optics and Optometry, Universidad Complutense de Madrid, Madrid, Spain

[3]Unit of Optometry, Division of Eye and Vision, Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden

Correspondence
Irene Martínez-Alberquilla, Faculty of Optics and Optometry, Universidad Complutense de Madrid, Madrid, Spain.
Email: irenem07@ucm.es

Abstract

Aim: To evaluate the repeatability of pupil size measurements obtained with four different devices under varying lighting conditions and to assess the agreement between them.

Methods: This prospective study included 80 healthy participants with a mean age of 27.5 ± 8.8 years (range 19–58). Pupil size was measured under mesopic (4 lux) and low photopic (50 lux) conditions using two open-field devices (tabletop and handheld PowerRefractors) and two closed-field devices (MS-39 and NIDEK Tonoref III). Three consecutive measurements were taken per eye for each device. Repeatability was analysed using the repeatability limit (R Limit), while the agreement was assessed with Bland–Altman analysis.

Results: All devices demonstrated R Limits lower than 1 mm under both low photopic and mesopic conditions. The NIDEK Tonoref III had the best repeatability under low photopic conditions (R Limit = 0.52 mm), while the tabletop PowerRefractor had the best R Limit under mesopic conditions (0.68 mm). The agreement analysis under mesopic conditions revealed smaller mean differences and narrower limits of agreement (LoA) (approximately 2 mm) among devices with similar designs (tabletop with handheld PowerRefractor and MS-39 with NIDEK Tonoref III), while combinations of designs exhibited wider variability. Mesopic conditions yielded more consistent LoA across all device pairs.

Conclusions: All devices showed excellent repeatability. Instruments with similar measurement designs demonstrated better agreement, particularly under low photopic conditions. Significant variability existed when comparing open- and closed-field devices. These findings underscore the importance of selecting appropriate instruments for clinical and research applications, and highlight the need for further validation of pupillometry devices.

KEYWORDS
agreement, mesopic, photopic, pupil size, pupillometry, repeatability

INTRODUCTION

Precise measurement of the pupil diameter is fundamental in various medical and research fields, particularly in ophthalmology and optometry,[1] as it can affect the visual outcomes in procedures such as corneal refractive surgery, intraocular lens implantation and contact lens fitting.[2–6] Currently, a wide variety of pupillometers are used in clinical practice to assess the pupil diameter as either a primary or secondary objective.[7–10] These devices have different designs, and each of them has its own advantages and limitations.

Factors such as lighting conditions, visual stimuli, emotional state or social interactions will affect pupil size during the measurement.[11,12] Moreover, the type of device used to measure pupil size, whether open- or closed-field instruments, will likely influence the results.[13–15] Open-field pupillometers measure pupil size in an environment where the subject's view is unrestricted, whereas closed-field pupillometers enclose the subject's eyes, lighting and stimuli in a controlled environment. The scientific literature has highlighted the role of accommodation in pupil measurements with concepts such as proximal and instrument accommodation,[16] which are likely to influence the accommodative response and, therefore, affect the accuracy of pupil size measurement.

Given the wide variety of devices available for measuring pupil size, it is crucial to gain an understanding of their measurement capabilities. To allow comparison between studies using different devices and to understand the accuracy of each instrument, agreement and repeatability studies are required. Such validation is essential for enhancing clinical accuracy and patient outcomes.

Several open[14,17] and closed-field devices[18,19] have been evaluated in the literature for pupil diameter measurement across diverse populations, light conditions and clinical contexts. However, there is limited scientific evidence regarding agreement between open- and closed-field instruments, as well as their reliability in various illumination settings.

Therefore, this study aimed to evaluate the repeatability of four different pupillometers under different lighting conditions and to assess their agreement, in order to determine which instrument design is most appropriate for accurately measuring pupil size, providing essential insight into their applicability in both research and clinical settings.

METHODS

This prospective study was conducted on healthy subjects at the Optometry Clinic (Eye Centre of Excellence, Karolinska Institute, Stockholm, Sweden). The study conformed to the principles of the Declaration of Helsinki. Regional Ethics Committee approval was obtained from the Swedish Ethical Review Authority, and signed informed consent was obtained from all subjects. Participants were between 18 and 70 years of age. Individuals with dry eye symptoms, a history of ocular disorders, trauma, pharmacological treatment, surgery, pupillary abnormalities or any other condition that could affect the pupil shape or response were excluded. This was ensured through a detailed clinical history and an ophthalmological examination conducted prior to the measurements being obtained. The demographic data collected included age, gender and refraction.

Instrumentation

Four instruments were used in this study to measure pupil sizes under different lighting conditions.

Key points

- Precise measurement of the pupil diameter is fundamental in research as it can affect the visual outcome in several procedures.
- All of the devices included here showed excellent repeatability. Instruments with similar measurement designs (open-field or closed-field) demonstrated better agreement, particularly under low photopic conditions.
- These findings provide valuable information regarding pupil size measurement under different light conditions, enabling clinicians to determine the most appropriate instrument according to its characteristics and the target population for assessment.

Tabletop (TT) power refractor (TT PowerRef)

The PowerRef 3 (Plusoptix GmbH, plusoptix.com) is an infra-red video-based eccentric tabletop photorefractor that measures gaze position, pupil size and meridional refraction binocularly at 50 Hz. The PowerRef was placed outside the participant's line of sight and an infra-red (IR)-reflecting, transparent hot mirror was used to reflect the IR light from the PowerRef3 into the eyes. Hence, the instrument did not obstruct the participant's view, thereby allowing open-field viewing during the measurements. Participants fixated on a stimulus consisting of a series of random words composed of five letters, which changed every 5 s to ensure stable fixation. A tablet computer (8th generation Apple iPad, resolution 2160 × 1620 pixels, apple.com/in/ipad/) was used to present the fixation target at 4 m. The pupil size measurements ranged from 4.0 to 8.0 mm in 0.1 mm steps.

Handheld (HH) autorefractor (AR)

The Plusoptix A12C is an HH AR (Plusoptix GmbH, plusoptix.com) that measures the gaze, refraction and pupil size binocularly. During the measurement, the examiner holds the instrument at a distance of 1 m from the participant, who fixates on the nose of an internal 'smiley face' target. The pupil size measurement range is from 3.0 to 8.0 mm in 0.1 mm increments. This instrument is designed so that the measurement is taken while the participant is seated without using a chin rest.

MS-39

The MS-39 (CSO, software version v.4.1.0.7, csoitalia.it/) is an anterior segment optical coherence tomographer

(AS-OCT) that combines Placido disc corneal topography with high-resolution OCT-based anterior segment tomography. This instrument uses a 950 nm light-emitting diode source for pupil size measurements, which can be obtained at different lighting conditions, namely, scotopic (0.04 lux), mesopic (4 lux) and photopic (50 lux).

AR

The NIDEK Tonoref III (NIDEK Co., nidek-intl.com/) is an autorefractometer, autokeratometer, non-contact tonometer and pachymeter. This is a tabletop device that combines super luminescent diode light with a highly sensitive charge-coupled device (CCD) camera. This AR has an integrated pupillography tool to measure the pupil size in a range from 1.0 to 10.0 mm in 0.1 mm increments.

Procedure

The instruments were calibrated before the measurements, and the order of testing was randomised. One eye per participant was randomly selected for the study, and three consecutive measurements were performed with each device. Measurements were taken under mesopic (4 lux) and low photopic (50 lux) conditions. The illumination levels were established based on the illumination settings available in the MS-39 in order to ensure the same illumination among all instruments. The lighting conditions were controlled with the EC1 luxometer (Hagner AB, hagner.se). Photopic measurements were taken after the mesopic measurements, and participants were adapted to the light condition for at least 2 min before each measurement.[9,10]

Statistical analysis

Descriptive analysis was performed including the mean and standard deviation (SD) of the pupil size. Repeatability and agreement were assessed according to British and International Standards.[20,21] The repeatability limit (R Limit) was used to assess repeatability, which was calculated as $1.96\sqrt{2} \times Sw$ (where Sw is the within-subject SD). The R Limit represents the likely limits within which 95% of the measurements should lie.[20,21] Differences in pupil size between each pair of instruments were analysed with repeated measures ANOVA.

When significant differences were identified, Bonferroni post-hoc test was used. For statistical significance, a *p*-value of 0.05 was set as the threshold.

A Bland–Altman analysis for repeated measures was used to describe the agreement between the instruments.[20] These analyses were performed for both lighting conditions. Correlations between age and device measurement with different illumination conditions were also calculated using Pearson's correlation coefficient. Correlations were considered strong, moderately strong, fair and poor if they were >0.80, between 0.50 and 0.80, between 0.30 and 0.50 and <0.30, respectively.[22]

A sample size (*n*) calculation was performed considering both repeatability and agreement analysis. For the repeatability analysis, *n* was calculated considering the number of repeated measurements (*m*) and confidence level (CL)[21] for the estimated Sw as $CL = \frac{1.96}{\sqrt{2 \cdot n(m-1)}}$. For a CL of 0.125 mm and three repeated measurements, 62 eyes were required. For agreement, the following formula was used: $\text{desired CI of LoA} = 1.96 \times \sqrt{\frac{3s^2}{n}}$, where *s* is the SD of the differences, CI the confidence interval and LoA the limits of agreement. We considered the desired CI for the LoA in the current study to be 0.125 mm for pupil size. Using this value and the s value (the SD of the differences) obtained in a subset of the first 20 eyes, the minimum sample required was 80 eyes. Then, considering both n values, the minimum sample size should be at least 80 eyes.

RESULTS

Eighty eyes of 80 participants (58 women and 22 men) were included in the study. The mean age was 27.5 ± 8.8 years (range 19–58 years) and the mean spherical equivalent was −1.07 ± 2.23 D (range −13.0–2.63 D).

Table 1 summarises the descriptive metrics and repeatability outcomes for the pupil diameter measurements obtained with the different devices under low photopic and mesopic lighting conditions. Under photopic conditions, the AR exhibited the largest mean pupil diameter (5.85 mm), whereas the MS-39 reported the smallest mean diameter (4.66 mm). For mesopic conditions, the MS-39 also measured the smallest diameter (5.37 mm), followed by the TT (6.30 mm) and the HH (6.84 mm) PowerRefs.

TABLE 1 Pupil size results and repeatability metrics.

	Mean ± SD (mm)		R Limit (mm)	
Instrument	Low photopic	Mesopic	Low photopic	Mesopic
TT power refractor	5.16 ± 1.02	6.30 ± 0.95	0.65	0.68
HH power refractor	5.35 ± 1.17	6.84 ± 1.01	0.91	0.78
MS-39 AS-OCT	4.66 ± 0.83	5.37 ± 0.86	0.61	0.71
AR	5.85 ± 0.89	N/A	0.52	N/A

Abbreviations: AR, autorefractor; AS-OCT, anterior segment optical coherence tomographer; HH, handheld; N/A, not applicable; R Limit, repeatability limit; TT, tabletop.

Regarding the repeatability results, all R Limits were <1.00 mm. Under low photopic conditions, the lowest R Limit was obtained with the AR (0.52 mm), while the highest R Limit was obtained with the HH PowerRef (0.91 mm). Under mesopic conditions, the repeatability values were similar among the devices, the highest value being 0.78 mm for the HH PowerRef.

Table 2 summarises the agreement findings across instruments under low photopic conditions. This table shows the mean differences between devices and the limits of agreement (LoA) with their intervals. The mean difference was <1 mm for all device comparisons except the MS-39 versus AR, where the mean difference was almost 2 mm. The TT PowerRef versus HH PowerRef comparison showed the smallest mean difference (−0.20 mm). Regarding the LoA intervals, the narrowest interval was obtained for the comparison between MS-39 and AR (2.08 mm) and the widest interval was obtained for the comparison between the HH PowerRef and the MS-39 (4.88 mm).

Table 3 displays the agreement results for pupil size measurements obtained under mesopic conditions. The smallest mean difference was observed between the two PowerRefs (−0.52 mm) and the largest mean difference was observed between the HH PowerRef and the MS-39 (1.47 mm). In terms of LoA, all device comparisons showed similar LoA intervals (about 2.5 mm).

Significant negative correlations were found between age and the pupil size measured with the four devices under different lighting conditions (Table 4). The correlation coefficients ranged from 0.20 to 0.30 (moderate correlation), and in most cases, they were statistically significant ($p < 0.05$). For the TT PowerRef and HH PowerRef, the correlations under mesopic conditions (−0.39 and −0.34, respectively) were higher than those obtained under low photopic conditions.

DISCUSSION

Pupil size measurement plays a crucial role within the fields of ophthalmology and optometry.[2–4,6,23] Considering the existence of a variety of pupillometers with different designs, it becomes essential to assess the repeatability of each device and their agreement. This evaluation provides valuable insights that can be applied in clinical practice to improve outcomes and enhance the accuracy of pupillary assessments. In this study, four different pupillometers were compared, two open field and two closed field. Open-field devices are more representative of natural viewing conditions, whereas closed-field devices are less reflective of such conditions.

Repeatability analysis showed that both closed-field instruments resulted in the lowest R Limit (about 0.6 mm) under photopic conditions. However, it should be considered that the repeatability values in all cases were <1 mm. Similarly, the R limit values were also <1 mm under mesopic conditions for all instruments. These results show that the instrument design does not affect the instrument repeatability under low photopic or mesopic conditions.

Considering the lack of studies evaluating the repeatability and agreement of the devices used in the present study, comparisons will be attempted using results from other devices reported in the existing literature, although direct comparisons may not always be feasible due to

TABLE 2 Agreement between instruments for the pupil sizes measured under low photopic conditions.

Instruments	Low photopic condition (mm)				
	Mean difference	LoA interval	LoA+	LoA−	*p*-Value
TT vs. HH	−0.20	2.31	0.96	−1.36	>0.99
TT vs. MS-39 AS-OCT	0.49	3.89	2.43	−1.46	0.007
TT vs. AR	−0.70	3.59	1.09	−2.50	<0.001
HH vs. MS-39 AS-OCT	0.68	4.88	3.10	−1.77	<0.001
HH vs. AR	0.57	4.54	2.84	−1.71	0.01
MS-39 AS-OCT vs. AR	1.92	2.08	2.23	0.15	<0.001

Abbreviations: AR, autorefractor; AS-OCT, anterior segment optical coherence tomographer; HH, handheld PowerRef; LoA, limit of agreement; TT, tabletop PowerRef.

TABLE 3 Agreement between instruments for the pupil sizes measured under mesopic conditions.

Instruments	Mesopic condition (mm)				
	Mean difference	LoA interval	LoA+	LoA−	*p*-Value
TT vs. HH	−0.52	2.52	0.74	−1.77	0.001
TT vs. MS-39 AS-OCT	0.93	2.15	2.00	−0.15	<0.001
HH vs. MS-39 AS-OCT	1.47	2.70	2.82	0.12	<0.001

Abbreviations: AS-OCT, anterior segment optical coherence tomographer; HH, handheld PowerRef; LoA, limits of agreement; TT, tabletop PowerRef.

TABLE 4 Correlation between age and device measurement.

Instruments	Lighting conditions	Age: Pearson correlation	Age: *p*-Value
TT power refractor	Low photopic	−0.27	0.05
	Mesopic	−0.39	0.004
HH power refractor	Low photopic	−0.24	0.11
	Mesopic	−0.34	0.02
MS-39 AS-OCT	Low photopic	−0.27	0.045
	Mesopic	−0.27	0.048
AR	Low photopic	−0.27	0.046

Abbreviations: AR, autorefractor; AS-OCT, anterior segment optical coherence tomographer; HH, handheld; R Limit, repeatability limit; TT, tabletop.

differences in methodologies and metrics. For example, Otero et al.[17] studied the agreement and repeatability of pupil measurements obtained with different open-field autorefractors and pupillometers, including the PowerRef II (plusoptix.com), which showed the largest variability under mid-mesopic conditions. The Sw was 0.17, 0.25 and 0.19 mm for low (0.05 lux), mid (0.80 lux) and high (20.00 lux) mesopic light levels, respectively. Using these Sw values and the formula described in the statistical analysis section, the R Limits could be calculated as 1.14, 1.40 and 1.21 mm, respectively. These R Limits were greater than the current results, which were all <1 mm. These differences may be explained by variations in the measurement method (binocular vs. monocular), the fixation target (distance and type of target) and the use of an older version of the device (PowerRef II).

Additionally, Satou et al.[14] compared the repeatability of two open-field devices, the Spot Vision Screener (hillrom.com), which is an HH vision screening device and the Grand Seiko WAM 5500 (grandseiko.com), a binocular open-field tabletop autorefractor with an illuminance of 200 lux (photopic). In that investigation, measurements were taken under binocular conditions, and the repeatability was described in terms of the intraclass correlation coefficient (ICC). The values were 0.87 for the Spot Vision and 0.97 for the WAM 5500, with a mean difference of 0.00 ± 0.44 mm. Unfortunately, comparisons with results obtained in the present study cannot be performed as different metrics were used.

Considering closed-field devices, the MS-39 has been used previously to measure anterior segment variables, such as central corneal and epithelial thickness, corneal diameter, aqueous depth, corneal power or higher-order aberrations in healthy patients,[24] as well as those with keratoconus,[25] cataract[26] and previously undergone a small-incision lenticule extraction procedure.[27] Similarly, the NIDEK Tonoref III (nidek-intl.com) has been used to measure central corneal thickness[28] and the amplitude of accommodation.[29] However, there is a lack of information regarding the pupil size measurement repeatability.

Additionally, Xu et al.[18] observed excellent repeatability with the Scansys Scheimpflug-based corneal topographer (Scansys TA517 3D Anterior Segment Analyser, mediworks.biz/en) for photopic pupil diameters (ICC = 0.956) and good repeatability for scotopic pupil diameters (ICC = 0.767). Kim et al.[19] also reported excellent repeatability with the Oculus Pentacam HR (pentacam.com) and the IOLMaster 700 (zeiss.com), with ICCs of 0.907 and 0.938, respectively, being obtained in a dark room.

Overall, the present results showed good repeatability, with pupil size fluctuations being <1 mm in 95% of the measurements. Unfortunately, direct comparisons cannot be made with the previous literature as they have used ICC metrics. According to McAlinden et al.,[20] the repeatability coefficient offers a more practical and clinically meaningful measure of precision compared with the ICC, which should be avoided.

In addition, agreement among devices was assessed using the Bland–Altman test for repeated measurements. A noteworthy finding was that under low photopic conditions, the smallest LoA intervals were observed between devices that shared similar designs. Comparisons between both binocular open-field devices (TT and HH PowerRefs) or between monocular closed-field devices resulted in the narrowest LoA intervals (about 2 mm). These results show that devices operating on similar measurement designs tended to produce more consistent results compared with devices using different principles. Furthermore, Guler Alis et al.[30] reported significant differences in pupil size obtained with the Plusoptix device when comparing binocular and monocular viewing (4.5 and 5.0 mm, respectively). This may help explain the discrepancies observed between binocular open-field and monocular closed-field devices. All measurements in the present study were performed binocularly without occluding the contralateral eye. However, in closed-field devices (AR and MS-39), the contralateral eye lacks a fixation target and was covered by the device itself. Therefore, this could be considered a "monocular" condition, potentially contributing to differences in pupil sizes between the types of devices.

There is limited literature regarding the agreement between the devices used in this study regarding pupil size measurement. However, other investigators have assessed this metric with different instruments. For example, Pérez-Bartolomé et al.[31] compared the Anterion AS-OCT (business-lounge.heidelbergengineering.com) and Pentacam HR (both closed-field devices) for measuring pupil diameter in healthy eyes under "standard light conditions", although the authors do not specify whether these conditions corresponded to photopic, mesopic or scotopic lighting. They showed that the LoA ranged from −0.35 to 4.20 mm with a mean difference of 1.92 mm, which is similar to the mean difference obtained in the current study between the MS-39 and AR instruments (1.919 mm). However, Pérez-Bartolomé et al. concluded that the devices were not interchangeable for pupil measurement. Kim et al.[19] also evaluated the repeatability and agreement of pupil diameter values obtained with the Pentacam HR and IOLMaster 700, observing significantly different pupil diameters and a low degree of agreement (LoA interval of 2.46 mm). It should be noted that the subjects in that study were older than the sample in the current investigation (67.2 years vs. 27.5 years), which might have impacted the outcome.

Under mesopic conditions, agreement analysis revealed notable differences compared with low photopic conditions. While the TT and HH PowerRefs continued to exhibit the smallest mean difference (−0.516 mm), the LoA intervals were more consistent across all device pairs. This contrasts with the findings under photopic conditions, where the LoA intervals were significantly wider for mixed comparisons (e.g., open-field vs. closed-field devices). This reduced variability suggests that mesopic conditions may reduce the influence of device-specific factors, which appear to play a larger role under photopic conditions. However, the mean differences between devices were significantly higher, ranging from −0.516 to 1.466 mm. These results agree with the findings of Kanclerz et al.,[32] who studied the agreement between the Nidek ARK-1 autorefractor (nidek-intl.com) and the Lenstar LS-900 biometer (haag-streit.com) for the measurement of pupil size of healthy eyes in dimly lit conditions (the authors did not specify the exact illuminance). The mean difference between the devices was 0.90 mm, and the 95% LoA ranged between −0.80 and 2.60 mm, showing poor agreement between devices, which cannot be considered interchangeable for this parameter.

Considering that pupil size is an ocular parameter that plays a major role in several ocular treatments,[2–6] clinicians should be mindful of the advantages and limitations of the chosen instrument for pupil size assessment, focusing on the patient's characteristics. In this regard, devices such as the Plusoptix A12C and TT PowerRef are not recommended for older patients with smaller pupil diameters,[33] particularly under photopic conditions. The rationale behind this observation is the floor effect observed at around 3 and 4 mm pupil diameters, which not only limits the accuracy of these devices in this population but also potentially reduces repeatability and agreement with other devices that have a wider measurement range. As closed-field pupillometers, such as the AS-OCT MS-39 and the NIDEK Tonoref III, may be able to capture a wider range of pupil sizes with greater precision, agreement between these devices could be affected. Given these considerations, it is essential to select carefully the appropriate device based on the population and testing conditions to ensure reliable and accurate measurements.

On the other hand, the decrease in pupil diameter associated with age has been reported extensively.[34,35] For this reason, correlations between age and pupil size under different light conditions were assessed. Under photopic conditions, previous studies found that the difference in pupil size between younger and older subjects was minimised. As expected, the results showed negative, moderate correlations for both mesopic and low photopic conditions. Moreover, these correlations were stronger under mesopic conditions, as previously found by Guillon et al.[36]

This study has several limitations that should be addressed. First, measurements were taken under controlled low photopic and mesopic conditions, which may not fully replicate real-world scenarios. Second, confounding factors such as age, refractive status or pupil dynamics were not taken into account, which could have affected the outcomes. Additionally, the diversity of participants was limited, restricting the generalisability of the results to broader populations. Finally, although good repeatability of pupil size has been reported in the literature,[14,18,19] it is important to note that obtaining an exact determination of the pupil size is challenging due to constant variations during the measurement process.[37] Therefore, future studies should address these factors to provide a more comprehensive assessment of the devices.

In conclusion, this study highlights the varying degrees of repeatability and agreement depending on the lighting environment and the type of device used. While open- and closed-field pupillometers can provide reliable and repeatable pupil size measurements, their agreement depends on the ambient lighting. Under low photopic conditions, devices having the same measurement principle showed narrower LoA intervals, suggesting closer agreement, whereas different types of devices exhibited greater variability, limiting their interchangeability. Under mesopic conditions, LoA intervals were more consistent across all pairings, although the mean differences observed between devices were relatively high. These results emphasise the need for caution when comparing or interpreting measurements across instruments and encourage further investigation into the factors influencing measurement discrepancies.

AUTHOR CONTRIBUTIONS

Irene Martínez-Alberquilla: Data curation (lead); formal analysis (equal); investigation (lead); methodology (equal); writing – original draft (lead); writing – review and editing (equal). **Laura Rico-del-Viejo:** Formal analysis (equal); investigation (equal); writing – original draft (equal); writing – review and editing (equal). **Abinaya Priya Venkataraman:** Conceptualization (equal); project administration (equal); resources (equal); supervision (equal); validation (equal); writing – review and editing (equal). **Alberto Dominguez-Vicent:** Conceptualization (equal); investigation (equal); project administration (equal); resources (equal); software (equal); supervision (equal); writing – review and editing (lead).

FUNDING INFORMATION

None.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest related to this study.

ORCID

Irene Martínez-Alberquilla https://orcid.org/0000-0003-4870-5862

REFERENCES

1. Wilhelm H, Wilhelm B. Clinical applications of pupillography. *J Neuroophthalmol*. 2003;23:42–9.
2. Myung D, Schallhorn S, Manche EE. Pupil size and LASIK: a review. *J Refract Surg*. 2013;29:734–41.
3. Teshigawara T, Meguro A, Mizuki N. The effect of age, postoperative refraction, and pre- and postoperative pupil size on halo size and intensity in eyes implanted with a trifocal or extended depth-of-focus lens. *Clin Ophthalmol*. 2021;15:4141–52.
4. Du H, Zhang B, Wang Z, Xiong L. Quality of vision after myopic refractive surgeries: SMILE, FS-LASIK, and ICL. *BMC Ophthalmol*. 2023;23:291. https://doi.org/10.1186/s12886-023-03045-6
5. Fernández J, Rodríguez-Vallejo M, Martínez J, Burguera N, Piñero DP. Pupil dependence assessment with multifocal intraocular lenses through visual acuity and contrast sensitivity defocus curves. *Eur J Ophthalmol*. 2021;31:2989–96.
6. Monsálvez-Romín D, González-Méijome JM, Esteve-Taboada JJ, García-Lázaro S, Cerviño A. Light distortion of soft multifocal contact lenses with different pupil size and shape. *Cont Lens Anterior Eye*. 2020;43:130–6.
7. Fernández J, Burguera N, Rocha-de-Lossada C, Rodríguez-Calvo-de-Mora M, Rodríguez-Vallejo M. Agreement between two devices for measuring pupil diameter in patients implanted with multifocal intraocular lenses. *Vision*. 2023;7:40. https://doi.org/10.3390/vision7020040
8. Schallenberg M, Bangre V, Steuhl KP, Kremmer S, Selbach JM. Comparison of the Colvard, Procyon, and Neuroptics pupillometers for measuring pupil diameter under low ambient illumination. *J Refract Surg*. 2010;26:134–43.
9. Scheffel M, Kuehne C, Kohnen T. Comparison of monocular and binocular infrared pupillometers under mesopic lighting conditions. *J Cataract Refract Surg*. 2010;36:625–30.
10. Bradley JC, Bentley KC, Mughal AI, Bodhireddy H, Brown SM. Dark-adapted pupil diameter as a function of age measured with the NeurOptics pupillometer. *J Refract Surg*. 2011;27:202–7.
11. Mathôt S. Pupillometry: psychology, physiology, and function. *J Cogn*. 2018;1:16. https://doi.org/10.5334/joc.18
12. Ferencová N, Višňovcová Z, Bona Olexová L, Tonhajzerová I. Eye pupil – A window into central autonomic regulation via emotional/cognitive processing. *Physiol Res*. 2021;70(Suppl4):S669–S682.
13. Bradley JC, Cohn CD, Wu PW, Brown SM. Comparison of a monocular pupillometer and the pupillometry function of a binocular free-viewing autorefractor. *J Cataract Refract Surg*. 2011;37:1257–62.
14. Satou T, Takahashi Y, Niida T. Comparison of refractive value and pupil size under monocular and binocular conditions between the spot vision screener and binocular open-field autorefractor. *Strabismus*. 2020;28:186–93.
15. Rosen ES, Gore CL, Taylor D, Chitkara D, Howes F, Kowalewski E. Use of a digital infrared pupillometer to assess patient suitability for refractive surgery. *J Cataract Refract Surg*. 2002;28:1433–8.
16. Lee EC, Lee JW, Park KR. Experimental investigations of pupil accommodation factors. *Invest Ophthalmol Vis Sci*. 2011;52:6478–85.
17. Otero C, Aldaba M, Ferrer O, Gascón A, Ondategui-Parra JC, Pujol J. Suitability of open-field autorefractors as pupillometers and instrument design effects. *Int J Ophthalmol*. 2017;10:567–72.
18. Xu W, Zhai C, Yusufu M, Gan J, Li S, Zheng Y, et al. Repeatability and agreement between a reference Scheimpflug tomographer and a low-cost Scheimpflug system. *J Cataract Refract Surg*. 2023;49:614–9.
19. Kim NH, Kim HJ, Cho SC, Han KE. Repeatability and agreement of chord mu between Scheimpflug tomography and swept-source optical coherence tomography. *Korean J Ophthalmol*. 2023;37:510–7.
20. McAlinden C, Khadka J, Pesudovs K. Statistical methods for conducting agreement (comparison of clinical tests) and precision (repeatability or reproducibility) studies in optometry and ophthalmology. *Ophthalmic Physiol Opt*. 2011;31:330–8.
21. McAlinden C, Khadka J, Pesudovs K. Precision (repeatability and reproducibility) studies and sample-size calculation. *J Cataract Refract Surg*. 2015;41:2598–604.
22. Chan YH. Biostatistics 104: correlational analysis. *Singapore Med J*. 2003;44:614–9.
23. McKelvie J, McArdle B, McGhee C. The influence of tilt, decentration, and pupil size on the higher-order aberration profile of aspheric intraocular lenses. *Ophthalmology*. 2011;118:1724–31.
24. Venkataraman AP, Domínguez-Vicent A, Selin P, Brautaset R, Montés-Micó R. Precision of a new SS-OCT biometer to measure anterior segment parameters and agreement with 3 instruments with different measurement principles. *J Cataract Refract Surg*. 2023;50:486–91.
25. Schiano-Lomoriello D, Bono V, Abicca I, Savini G. Repeatability of anterior segment measurements by optical coherence tomography combined with Placido disk corneal topography in eyes with keratoconus. *Sci Rep*. 2020;10:1124. https://doi.org/10.1038/s41598-020-57926-7
26. Wang Q, Chen M, Ning R, Savini G, Wang Y, Zhang T, et al. The precision of a new anterior segment optical coherence Tomographer and its comparison with a swept-source OCT-based optical biometer in patients with cataract. *J Refract Surg*. 2021;37:616–22.
27. Ning R, Huang X, Jin Y, Lei CS, Ma X, Xu S, et al. Corneal higher-order aberrations measurements: precision of SD-OCT/Placido topography and comparison with a Scheimpflug/Placido topography in eyes after small-incision lenticule extraction. *Ophthalmol Therapy*. 2023;12:1595–610.
28. Desmond T, Arthur P, Watt K. Comparison of central corneal thickness measurements by ultrasound pachymetry and 2 new devices, Tonoref III and RS-3000. *Int Ophthalmol*. 2019;39:917–23.
29. Weng CC, Hwang DK, Liu CJL. Repeatability of the amplitude of accommodation measured by a new generation autorefractor. *PLoS One*. 2020;15:e0224733. https://doi.org/10.1371/journal.pone.0224733
30. Guler Alis M, Alis A. Compatibility of pupil size measured with Nidek ARK-1a table top autorefractometer and Plusoptix A12C photoscreener. *J Binocul Vis Ocul Motil*. 2021;71:161–6.
31. Pérez-Bartolomé F, Rocha-De-Lossada C, Sánchez-González JM, Feu-Basilio S, Torras-Sanvicens J, Peraza-Nieves J. Anterior-segment swept-source ocular coherence tomography and Scheimpflug imaging agreement for keratometry and pupil measurements in healthy eyes. *J Clin Med*. 2021;10:5789. https://doi.org/10.3390/jcm10245789
32. Kanclerz P, Przewłócka K, Wang X. Inter-device measurement variability of vital data parameters for keratorefractive and cataract refractive surgery. *Ther Adv Ophthalmol*. 2021;13:25158414211045750. https://doi.org/10.1177/25158414211045750
33. Fernández J, Rodríguez-Vallejo M, Martínez J, Burguera N, Piñero DP. Pupil diameter in patients with multifocal intraocular lenses. *J Refract Surg*. 2020;36:750–6.
34. Winn B, Whitaker D, Elliott DB, Phillips NJ. Factors affecting light-adapted pupil size in normal human subjects. *Invest Ophthalmol Vis Sci*. 1994;35:1132–7.
35. Nakamura K, Bissen-Miyajima H, Oki S, Onuma K. Pupil sizes in different Japanese age groups and the implications for intraocular lens choice. *J Cataract Refract Surg*. 2009;35:134–8.
36. Guillon M, Dumbleton K, Theodoratos P, Gobbe M, Wooley CB, Moody K. The effects of age, refractive status, and luminance on pupil size. *Optom Vis Sci*. 2016;93:1093–100.
37. Donnenfeld E. The pupil is a moving target: centration, repeatability, and registration. *J Refract Surg*. 2004;20:S593–S596.

How to cite this article: Martínez-Alberquilla I, Rico-del-Viejo L, Venkataraman AP, Dominguez-Vicent A. Comparison of pupil size measurement: Repeatability and agreement across four devices with different measurement principles under varied lighting conditions. *Ophthalmic Physiol Opt*. 2025;00:1–7. https://doi.org/10.1111/opo.13507