

Device-dependent white-to-white discrepancy between ARGOS and ALADDIN optical biometers: a prospective agreement study with decision- level lens sizing simulation

Alper Güneş

alpergunes@gmail.com

Tokat Gaziosmanpaşa University Faculty of Medicine

Mehmet Yusuf Tahaoğlu¹

Tokat Gaziosmanpaşa University Faculty of Medicine

Ender Şener¹

Tokat Gaziosmanpaşa University Faculty of Medicine

Raşit Kılıç¹

Tokat Gaziosmanpaşa University Faculty of Medicine

Helin Deniz Demir

Tokat Gaziosmanpaşa University Faculty of Medicine

Article

Keywords:

Posted Date: August 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9932132/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Device-dependent white-to-white discrepancy between ARGOS and ALADDIN optical biometers: a prospective agreement study with decision-level lens sizing simulation

Alper Güneş¹, Mehmet Yusuf Tahaoğlu¹, Ender Şener¹, Raşit Kılıç¹, Helin Deniz Demir¹

¹Department of Ophthalmology, Faculty of Medicine, Tokat Gaziosmanpaşa University, Tokat, Türkiye

Correspondence to: Alper Güneş, Department of Ophthalmology, Faculty of Medicine, Tokat Gaziosmanpaşa University, 60100 Tokat, Türkiye. alpergunes@gmail.com; ORCID: 0000-0001-9213-8805.

Abstract

Accurate ocular biometry underpins intraocular lens (IOL) selection. However, clinical interchangeability rather than mere statistical agreement between swept-source optical coherence tomography (SS-OCT) and optical low-coherence interferometry (OLCI) biometers remains underexplored. In this prospective, single-center, eye-based agreement study, 100 eyes of 100 patients with age-related cataract underwent biometry with ARGOS (SS-OCT) and ALADDIN (OLCI) by a single examiner. Axial length (AL), anterior chamber depth (ACD), lens thickness (LT), mean keratometry (Km), and white-to-white (WTW) corneal diameter were compared using intraclass (ICC) and Lin's concordance (CCC) correlation coefficients, Bland-Altman limits of agreement, Passing-Bablok regression, a priori clinical tolerance thresholds, and a WTW-based phakic IOL sizing simulation. Inter-device agreement was excellent for AL, ACD, LT, and Km (all ICCs ≥ 0.977), with at least 96% of eyes within the predefined tolerance thresholds. In contrast, WTW differed systematically (bias -0.180 mm; ICC=0.793; Cohen's $d=0.88$; $P<0.001$), with only 34% of eyes within ± 0.10 mm. In the sizing simulation, 40% of eyes were assigned a different phakic IOL size depending on the device providing the WTW, the ALADDIN-derived size being larger in 90% of discordant eyes. ARGOS and ALADDIN are interchangeable for routine cataract biometry but not for WTW-dependent phakic IOL sizing, where a single device should be used.

Introduction

Among the optical biometry technologies currently in clinical use, optical low-coherence interferometry (OLCI) and swept-source optical coherence tomography (SS-OCT) are the most widely adopted [3,4]. The ALADDIN biometer (Topcon Corp., Tokyo, Japan) employs OLCI at 830 nm and combines a 24-ring Placido-disc topographer with interferometric axial measurement. The ARGOS biometer (Movu Inc., Komaki, Japan) uses SS-OCT at 1060 nm, providing two-dimensional cross-sectional imaging of the entire eye with segmental refractive indices for each ocular compartment, a feature that theoretically improves AL accuracy in eyes with extreme biometrics[5,6].

Comparative studies have generally reported excellent agreement among modern biometers for AL and ACD [7,8]. However, WTW has consistently emerged as the most device-dependent

parameter[9,10]. In a phakic-IOL-oriented comparison of four anterior-segment imaging devices, Calvo-Sanz et al specifically identified WTW as the parameter with the poorest inter-device correlation among the biometric inputs used for posterior-chamber phakic IOL sizing[11]. This may carry direct clinical consequences, because incorrect WTW-driven sizing can produce an abnormal vault, angle closure, anterior subcapsular cataract, or endothelial cell loss [12].

Despite a growing body of literature on biometer comparisons, studies directly evaluating the clinical interchangeability rather than mere statistical agreement of ARGOS and ALADDIN remain limited. A recent prospective comparison reported excellent ICC values for AL and ACD between these two devices, but poor agreement for WTW (ICC=0.338)[13]. Independent cohorts also reported significant inter-device differences in WTW and keratometry between ARGOS and ALADDIN[14,15]. Importantly, prior reports have relied predominantly on statistical metrics (ICC, Pearson correlation, paired differences) and have not quantified how often inter-device WTW disagreement translates into a discordant phakic IOL size assignment in routine clinical practice.

We hypothesized that, although ARGOS and ALADDIN would show excellent statistical agreement for most parameters, WTW measurements would fall outside a clinically meaningful tolerance threshold, and that this disagreement would translate into a clinically relevant rate of phakic IOL size discordance. To our knowledge, this is among the first studies to translate the ARGOS–ALADDIN WTW disagreement into a decision-level phakic IOL sizing simulation using predefined clinical tolerance thresholds, alongside effect-size estimation, Lin's concordance correlation coefficient and Passing-Bablok regression.

Methods

Ethical Approval

The study adhered to the tenets of the Declaration of Helsinki (as revised in 2013) and was approved by Tokat Gaziosmanpaşa University Faculty of Medicine Clinical Research Ethics Committee. (approval number 23-KAEK-194; date of approval 14 March 2023). Written informed consent was obtained from all participants prior to inclusion. Reporting follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement and the Guidelines for Reporting Reliability and Agreement Studies (GRRAS)^[16,17].

Study Design and Participants

This was a prospective, single-center, eye-based agreement study conducted at a tertiary university ophthalmology department between April 2023 and March 2024. Inclusion criteria were phakic eyes with age-related cataract and sufficient media clarity for optical biometry. Exclusion criteria were prior intraocular or refractive surgery, corneal opacities or scarring, dense cataracts precluding optical measurement, retinal disease, glaucoma, severe dry eye and poor fixation. When both eyes were eligible, one eye was selected randomly using a computer-generated random number sequence to avoid inter-eye correlation. A flow diagram of participant selection is provided as Supplementary Fig. S1.

Biometric Devices

ARGOS (Movu Inc., Komaki, Japan): an SS-OCT biometer operating at 1060 nm that provides two-dimensional cross-sectional imaging of the entire eye. Biometric parameters are computed using segmental refractive indices (cornea 1.376; aqueous and vitreous 1.336; lens 1.410). WTW is derived from infrared en-face OCT images by automated detection of the limbal boundary.

ALADDIN (Topcon Corp., Tokyo, Japan): an OLCI biometer operating at 830 nm that combines a 24-ring Placido-disc corneal topographer with interferometric axial measurement. WTW is estimated by an edge-detection algorithm applied to anterior-segment images derived from the visible-spectrum Placido reflection pattern. Parameters analyzed for both devices were AL, ACD (epithelium to anterior lens), LT, K1, K2, mean keratometry ($K_m = (K1+K2)/2$) and WTW.

Measurement Protocol

All measurements were performed in a single session by the same examiner (>5 years of experience with both devices) under standardized mesopic lighting without pharmacological pupil dilation. The order of devices was randomized for each patient using a computer-generated 1:1 allocation list to minimize order and tear-film effects, and a brief interval of approximately 3 minutes separated the two examinations to allow tear-film recovery. For each device, three consecutive measurements were obtained per eye; only readings meeting the manufacturer's internal quality criteria (signal-to-noise ratio thresholds and fixation indicators) were accepted, and the mean of valid readings was used for analysis. If quality criteria were not met, the measurement was repeated.

Sample Size

For an ICC agreement study with two raters, an expected ICC of 0.90, a null ICC of 0.70, $\alpha=0.05$ and power=0.80, the minimum required sample is approximately 46 subjects^[18]. The 100 eyes included here provide additional power to detect smaller systematic differences, to estimate Bland-Altman LoA with narrower confidence intervals, and to enable AL-stratified subgroup exploration.

Statistical Analysis

Analyses were performed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA) and Python 3 (packages pingouin, SciPy and statsmodels) for non-standard analyses (Lin's CCC, Passing-Bablok regression, and confidence intervals for the LoA). Normality was assessed using the Shapiro–Wilk test. Descriptive statistics are reported as mean $\pm$ standard deviation (SD) with range. Paired-samples t-tests compared measurements between devices, with $P<0.05$ considered statistically significant; Wilcoxon signed-rank tests were used to confirm results for non-normally distributed differences. Given the a priori hypothesis-driven design and the limited number of comparisons (five parameters), no formal multiplicity correction was applied, but P-values are reported for each parameter.

Agreement was assessed with ICC (two-way mixed-effects model, absolute agreement, single rater) reported with 95% confidence intervals (CI); values were interpreted as poor (<0.50), moderate ($0.50\text{--}0.75$), good ($0.75\text{--}0.90$) and excellent (>0.90). Lin's concordance correlation coefficient (CCC) was computed to jointly capture precision and accuracy. Bland-Altman analysis was used to derive systematic bias and 95% LoA (bias $\pm 1.96 \times$ SD of differences), with 95% CIs for both the bias and the LoA according to Bland's method^[19]. Proportional bias was assessed by linear regression of the differences on the means. Method-comparison regression was performed using Passing-Bablok regression with bootstrap-based confidence intervals^[20]. Effect sizes were quantified using Cohen's d ($d = \text{mean difference} / \text{SD of differences}$); $d \geq 0.80$ was considered a large effect.

Clinical interchangeability was evaluated using a priori tolerance thresholds: ± 0.10 mm for AL, ACD, LT and WTW, and ± 0.25 D for Km. These thresholds were selected on the basis of established clinical relevance: a 0.10 mm AL error corresponds to approximately 0.27 D of refractive prediction error^[1], and 0.25 D represents the smallest clinically meaningful keratometric difference for monofocal IOL calculation. For WTW, the ± 0.10 mm threshold was selected

because WTW is used as a sizing input in phakic IOL and sulcus-related lens selection, where small systematic inter-device differences may shift borderline eyes across discrete sizing categories; the threshold was therefore intended to reflect clinical-decision sensitivity rather than measurement repeatability alone. The proportion of eyes within each threshold was reported with 95% binomial (Clopper–Pearson) confidence intervals.

Two pre-specified secondary analyses were performed. First, an AL-stratified subgroup analysis grouped eyes as short ($AL < 22.5$ mm), normal ($22.5 \leq AL \leq 24.5$ mm) or long ($AL > 24.5$ mm). Second, a phakic IOL sizing simulation independently used the WTW of each device to assign an implantable collamer lens (ICL) size according to a representative WTW-based nomogram (12.1 mm for $WTW < 11.0$ mm; 12.6 mm for $11.0 \leq WTW < 11.5$ mm; 13.2 mm for $11.5 \leq WTW < 12.1$ mm; 13.7 mm for $WTW \geq 12.1$ mm). The number of eyes for which the assigned size differed between devices was reported with the corresponding 95% Clopper–Pearson confidence interval, together with the direction of disagreement. Lin's CCC and Passing-Bablok regression were reported in line with concordance and method-comparison reporting standards^[20,21].

Results

Participants and Demographics

One hundred eyes of 100 patients with a mean age of 64.1 ± 7.4 years (range 46–78 years) were enrolled; 58 patients (58%) were female and 42 (42%) male, and the right eye was studied in 50 patients (50%). All eyes were phakic with mild-to-moderate age-related cataract. The cohort spanned short to long axial lengths (mean ARGOS AL 23.52 ± 1.21 mm, range 20.45–26.17 mm) and a wide keratometric range (mean ARGOS Km 43.65 ± 1.55 D, range 39.09–47.21 D), confirming a representative cataract population. Biometric measurements were successfully obtained with both devices in all 100 patients, with no acquisition failures. Demographic and baseline biometric characteristics are summarized in Table 1.

Variable	Value
Age (years), mean $\pm$ SD (range)	64.1 ± 7.4 (46–78)
Sex, female / male, n (%)	58 / 42 (58 / 42)
Eye studied, OD / OS, n (%)	50 / 50 (50 / 50)
Axial length, mean $\pm$ SD (range), mm	23.52 ± 1.21 (20.45–26.17)
Anterior chamber depth, mean $\pm$ SD, mm	3.20 ± 0.30
Lens thickness, mean $\pm$ SD, mm	4.45 ± 0.41

Variable	Value
Mean keratometry, mean $\pm$ SD (range), D	43.65 $\pm$ 1.55 (39.09–47.21)
White-to-white, mean $\pm$ SD (range), mm	11.67 $\pm$ 0.41 (10.73–12.66)

Table 1 Demographic and baseline biometric characteristics of the study cohort (n=100 eyes of 100 patients). Continuous variables are presented as mean $\pm$ SD with range where informative; categorical variables as n (%). All biometric values derive from the ARGOS device.

Comparative biometric measurements with Bland-Altman bias and 95% LoA, ICC, CCC, P-values, Cohen's d effect sizes and Passing-Bablok regression coefficients are summarized in Table 2. The proportion of eyes within clinical tolerance thresholds is presented in Table 3, and the simulated phakic IOL size cross-tabulation in Table 4. Bland-Altman plots are presented in Figs. 1 and 2, and the phakic IOL sizing simulation in Fig. 3.

Parameter	ARGOS	ALADDIN	Bias (95% CI)	95% LoA	ICC (95% CI)	CCC	d	P	PB slope (95% CI)
AL (mm)	23.52 $\pm$ 1.21	23.51 $\pm$ 1.20	+0.009 (−0.005, +0.023)	−0.130, +0.148	0.998 (0.997, 0.999)	0.998	0.13	0.206	1.000 (0.992, 1.004)
ACD (mm)	3.20 $\pm$ 0.30	3.19 $\pm$ 0.31	+0.010 (−0.002, +0.021)	−0.106, +0.125	0.981 (0.972, 0.987)	0.981	0.16	0.107	1.034 (1.000, 1.077)
LT (mm)	4.45 $\pm$ 0.41	4.44 $\pm$ 0.39	+0.006 (−0.011, +0.023)	−0.161, +0.173	0.977 (0.966, 0.985)	0.977	0.07	0.482	0.989 (0.973, 1.000)
Km (D)	43.65 $\pm$ 1.55	43.71 $\pm$ 1.54	−0.056 (−0.077, −0.034)	−0.268, +0.156	0.997 (0.996, 0.998)	0.997	0.52	<0.001	0.996 (0.987, 1.003)
WTW (mm)	11.67 $\pm$ 0.41	11.85 $\pm$ 0.40	−0.180 (−0.220, −0.139)	−0.579, +0.220	0.793 (0.697, 0.858)	0.794	0.88	<0.001	0.966 (0.868, 1.080)

Table 2 Inter-device biometric agreement with Bland-Altman, ICC, CCC, effect size and Passing-Bablok regression (n=100 eyes). Bias = mean difference (ARGOS – ALADDIN); 95% CI of the bias by t-distribution. 95% LoA = bias $\pm$ 1.96 $\times$ SD of differences (95% CIs for each LoA in Supplementary Table S1). ICC = intraclass correlation coefficient (two-way mixed-effects, absolute agreement, single rater). CCC = Lin's concordance correlation coefficient (95% CI: AL 0.998–0.998; ACD 0.979–0.982; LT 0.975–0.979; Km 0.997–0.997; WTW 0.745–0.834). d = Cohen's d (|mean difference|/SD of differences); d $\geq$ 0.80 = large effect. PB = Passing-Bablok regression; slope=1 and intercept=0 indicate no systematic difference; intercepts (original units): AL −0.010; ACD −0.116; LT +0.045; Km +0.253; WTW +0.590. The WTW row is highlighted as the parameter of clinical concern. AL = axial length; ACD = anterior chamber depth; LT = lens thickness; Km = mean keratometry; WTW = white-to-white corneal diameter.

Axial Length, Anterior Chamber Depth and Lens Thickness

Inter-device agreement was excellent for AL (ICC=0.998; CCC=0.998; d=0.13), ACD (ICC=0.981; CCC=0.981; d=0.16) and LT (ICC=0.977; CCC=0.977; d=0.07). Within the $\pm$ 0.10 mm tolerance threshold, 98%, 98% and 96% of eyes met clinical interchangeability for AL, ACD and LT, respectively. Passing-Bablok slopes were close to unity for all three parameters and intercepts were close to zero. A small but statistically significant proportional bias was detected for LT (slope of differences on means = 0.050, P=0.018); however, its absolute magnitude is well below the predefined clinical tolerance and is not clinically relevant. The trivial effect sizes

($d < 0.20$) further support that the observed systematic differences are statistically and clinically negligible.

Keratometry

For Km, the bias was -0.056 D, with narrow 95% LoA of -0.268 to $+0.156$ D and excellent agreement (ICC=0.997; CCC=0.997); 98% of eyes were within the ± 0.25 D tolerance threshold. Although the inter-device difference reached statistical significance ($P < 0.001$), the magnitude of the bias was small and well within a clinically acceptable range for routine monofocal IOL power calculation.

White-to-White Corneal Diameter

WTW measurements showed a statistically and clinically significant systematic discrepancy. ALADDIN consistently yielded larger WTW values than ARGOS (mean 11.85 ± 0.40 mm vs 11.67 ± 0.41 mm; bias -0.180 mm; 95% CI of bias -0.220 to -0.139 mm; 95% LoA -0.579 to $+0.220$ mm; ICC=0.793, 95% CI 0.697–0.858; CCC=0.794; $P < 0.001$). The Cohen's d of 0.88 represents a large effect size. Only 34% of eyes (34/100; 95% Clopper–Pearson CI 24.8–44.2) fell within the ± 0.10 mm tolerance threshold, indicating that WTW measurements from the two devices should not be considered clinically interchangeable. Passing-Bablok regression gave a slope close to unity (0.966; 95% CI 0.868–1.080) but an intercept of $+0.590$ mm, consistent with a constant systematic offset between devices rather than a proportional bias; the linear regression of differences on means confirmed no significant proportional bias (slope=0.026; $P=0.62$). On a tolerance sensitivity analysis, the proportion of eyes within agreement for WTW rose only modestly with more permissive thresholds 34% at ± 0.10 mm, 51% at ± 0.20 mm and 69% at ± 0.30 mm whereas AL, ACD and LT were already within tolerance in $\geq 96\%$ of eyes at ± 0.10 mm; WTW thus remained the least interchangeable parameter across all thresholds tested (Supplementary Table S3).

Parameter	Tolerance	Eyes within (n/N)	% (95% CI)
AL	± 0.10 mm	98/100	98.0 (93.0–99.8)
ACD	± 0.10 mm	98/100	98.0 (93.0–99.8)
LT	± 0.10 mm	96/100	96.0 (90.1–98.9)
Km	± 0.25 D	98/100	98.0 (93.0–99.8)
WTW	± 0.10 mm	34/100	34.0 (24.8–44.2)

Table 3 Proportion of eyes within a priori clinical tolerance thresholds (n=100 eyes). Tolerance thresholds defined a priori on the basis of clinical relevance. 95% CIs by the exact binomial (Clopper–Pearson) method.

Phakic IOL Sizing Simulation

When each device's WTW was used independently to assign a phakic IOL size according to the representative WTW-based nomogram, 60 of 100 eyes (60%) received the same lens size from both devices and 40 of 100 (40%; 95% CI 30.3–50.3) received a different lens size. The disagreement was strongly directional: ALADDIN-derived sizing was larger than ARGOS-derived sizing in 36 of these 40 discordant eyes (90%) and smaller in only 4 (10%). The cross-tabulation of size assignments is presented in Table 4 and Fig. 3. These findings translate the observed WTW measurement bias into an explicit clinical-decision consequence, illustrating that a systematic inter-device WTW difference of approximately 0.18 mm shifts a substantial fraction of eyes across one or more sizing categories.

ARGOS-derived size	ALADDIN 12.1	ALADDIN 12.6	ALADDIN 13.2	ALADDIN 13.7	Row total
12.1 mm	0	5	1	0	6
12.6 mm	0	12	16	0	28
13.2 mm	0	3	35	14	52
13.7 mm	0	0	1	13	14
Column total	0	20	53	27	100

Table 4 Cross-tabulation of simulated phakic IOL size assignments derived independently from ARGOS WTW and ALADDIN WTW (n=100 eyes). Diagonal cells indicate concordant assignments (60 eyes); off-diagonal cells indicate discordant assignments (40 eyes), of which 36 (90%) reflect ALADDIN assigning a larger lens size than ARGOS. Nomogram thresholds: 12.1 mm for WTW<11.0 mm; 12.6 mm for 11.0≤WTW<11.5 mm; 13.2 mm for 11.5≤WTW<12.1 mm; 13.7 mm for WTW≥12.1 mm.

Axial-Length Subgroup Analysis

The cohort included 20 short eyes (AL<22.5 mm), 57 normal eyes (22.5≤AL≤24.5 mm) and 23 long eyes (AL>24.5 mm). Inter-device agreement for AL remained excellent in all three groups (ICC 0.986, 0.998 and 0.956 in short, normal and long eyes, respectively), although point agreement was numerically lowest in long eyes. The WTW discrepancy persisted across all three AL subgroups, with the proportion of eyes within the ±0.10 mm tolerance threshold reaching only 45% (9/20), 32% (18/57) and 30% (7/23) in short, normal and long eyes, respectively, and similar bias magnitudes (−0.16, −0.19 and −0.17 mm). The systematic WTW discrepancy was therefore not driven by axial length (Supplementary Fig. S2).

Comparison with Prior ARGOS-vs-ALADDIN Literature

Supplementary Table S2 places the present findings in the context of recently published ARGOS-vs-ALADDIN comparisons and other contemporary SS-OCT/OLCI agreement studies. Across studies, agreement for AL and ACD is uniformly excellent, agreement for Km is consistently

good-to-excellent, and WTW agreement is the least consistent parameter. Our WTW ICC of 0.793 is higher than the 0.338 reported by Munteanu et al^[13], most plausibly explained by differences in patient demographics, lighting conditions, software/firmware versions and the spread of WTW values in each sample. Nevertheless, our LoA range of approximately 0.80 mm and the failure of two-thirds of eyes to meet the ± 0.10 mm tolerance threshold confirm the central finding that WTW is the parameter where ARGOS and ALADDIN diverge clinically.

Discussion

This prospective study evaluated the clinical interchangeability of biometric measurements obtained with two widely used optical biometers, ARGOS (SS-OCT) and ALADDIN (OLCI), in 100 cataract patients. Both devices yielded highly interchangeable measurements for AL, ACD, LT, and Km, with excellent ICC and CCC values (>0.97), trivial-to-small effect sizes ($d \leq 0.52$), and $\geq 96\%$ of eyes within the a priori tolerance thresholds. In contrast, WTW showed a systematic bias of -0.180 mm with a large effect size ($d=0.88$), wide 95% LoA, and only 34% of eyes within the ± 0.10 mm threshold; most importantly, this WTW discrepancy translated directly into a discordant WTW-based phakic IOL size assignment in 40% of eyes, with a strongly asymmetric direction of disagreement.

The excellent agreement for axial and corneal-power parameters is consistent with the established literature on modern optical biometers[7,8,22] and is here supported jointly by ICC, CCC, Passing-Bablok slopes close to unity, and the proportion of eyes within clinically defined tolerance thresholds. The concordance of these complementary metrics provides stronger evidence for interchangeability than any single statistic, and it indicates that, for routine monofocal cataract surgery planning, ARGOS and ALADDIN can be used interchangeably for AL, ACD, LT, and Km.

Mechanism of the WTW Discrepancy

Three methodological differences plausibly explain the systematic WTW bias. First, wavelength and tissue penetration: ARGOS uses 1060 nm infrared light, which penetrates beyond the visible scleral edge and may identify a deeper anatomical limbal boundary, whereas ALADDIN relies on a visible-spectrum Placido reflection that highlights the more anteriorly located optical limbus. Second, imaging modality and edge-detection algorithm: the two devices acquire anterior-segment

images using different light sources and detectors and apply different automated limbal boundary-detection algorithms, so the same anatomical structure can yield systematically different chord measurements. Third, anatomical reference point: the cornea-iris junction, scleral spur, and pigmented limbal boundary lie within a 0.1–0.3 mm transition zone, and small algorithmic preferences for one landmark over another produce systematic, device-specific offsets[11,15]. The Passing-Bablok intercept of +0.590 mm together with a slope close to unity (0.966) and the absence of significant proportional bias ($P=0.62$) support a constant additive offset rather than a magnitude-dependent disagreement, the pattern expected from a systematic anatomical-landmark difference rather than a calibration error.

Notably, the direction of the WTW discrepancy in our cohort ALADDIN is larger than that in ARGOS and opposite to that reported in one prior comparison[15], reinforcing that the WTW offset is not only device-dependent but also context-dependent (software version, illumination, examiner). The present findings do not establish which device measures the true anatomical limbus-to-limbus distance; rather, they indicate that WTW values are device-dependent and should not be used interchangeably. Published normative WTW values may vary according to population and measurement platform[23]; therefore, the present findings should be interpreted as evidence of device-dependence rather than anatomical superiority of either device. Resolving which value is anatomically correct would require a reference standard such as gonioscopy or anterior-segment OCT with manual landmark identification, which was beyond the scope of this study.

Clinical Relevance: From Millimeters to Lens Sizes

This cohort comprised cataract surgery candidates rather than phakic-IOL candidates, who are typically younger and more myopic with different anterior-segment anatomy; the present exercise is therefore a decision-level simulation designed to expose device-driven decision variability, not an outcome-validated phakic-IOL sizing study. The principal clinical novelty of this study is the translation of the measurement-level WTW discrepancy into a decision-level outcome. Phakic IOL platforms use WTW as one of the sizing inputs and are available only in discrete size steps; for the STAAR EVO ICL, available sizes are separated by approximately 0.5–0.6 mm. In the present cohort, 40% of eyes received a different lens-size assignment based solely on the device that produced the WTW, and the ALADDIN-derived size was larger than the ARGOS-derived size in

90% of discordant cases. This finding cannot be derived from ICC or LoA alone; a clinical-tolerance threshold combined with a decision-anchored simulation is required to expose it.

Two caveats are essential. First, the simulation was not validated against postoperative vault or sulcus-to-sulcus measurements and therefore represents a decision-level discordance rather than an outcome-validated difference; it does not establish which device would produce the better surgical outcome. Second, the result should not be interpreted as evidence that either device is unsafe. The appropriate conclusion is narrower and more robust. When WTW informs millimetric surgical decisions, the two devices should not be mixed, because doing so introduces a non-trivial probability of a different lens-size decision.

From a practical perspective, these findings suggest the following. For routine monofocal cataract surgery, ARGOS and ALADDIN appear clinically interchangeable for AL, ACD, LT, and Km. For toric IOL planning, Km agreement was clinically acceptable; however, astigmatism magnitude and axis should still be interpreted together with corneal topography when clinically relevant. For phakic IOL or ICL sizing, WTW measurements should not be used interchangeably between devices, and the same device should be used consistently throughout the preoperative assessment and any subsequent revision. For sulcus-supported IOL or capsular tension ring sizing, WTW should be interpreted cautiously because it is only an indirect proxy for internal anterior-segment dimensions, and direct sulcus-to-sulcus imaging is preferable when available.

Axial-Length Subgroup Observations

AL-stratified analysis confirmed that both the excellent agreement for axial measurements and the WTW discrepancy are robust across short, normal, and long eyes. The numerically lower (but still excellent) AL ICC in long eyes (0.956) is consistent with prior reports that segmental refractive indices may produce modest but systematic differences relative to traditional group-refractive-index methods in highly myopic eyes[5,6]; this signal should be interpreted cautiously in our subgroup of 23 long eyes and warrants confirmation in a larger high-myopia cohort. Critically, WTW disagreement was uniform across all three AL subgroups ($\geq 55\%$ of eyes

outside the ± 0.10 mm tolerance in each group), indicating that it is not eye-length-dependent and therefore unlikely to be mitigated by restricting comparisons to particular biometric phenotypes.

Strengths and Limitations

The principal strength of this study is its explicit conceptual separation of statistical agreement from clinical interchangeability, operationalized through a priori tolerance thresholds, effect-size estimation, Bland-Altman analysis with confidence intervals for the limits of agreement, Passing-Bablok regression, Lin's CCC, AL-stratified subgroup analysis, and a decision-anchored phakic IOL sizing simulation. Several limitations should be acknowledged. First, this was a single-center study conducted in a routine cataract population, which may limit generalizability to eyes with irregular corneas, prior refractive surgery, dense media opacity[24], or extreme biometric characteristics. Second, only one eye per patient was included, which eliminated inter-eye correlation but reduced the available sample size relative to bilateral designs. Third, although three measurements per device were used internally for the device-averaged readings, a formal coefficient-of-repeatability (intra-device) analysis was not separately reported in the present cohort. Fourth, pupil diameter was not recorded systematically, which may be relevant for WTW assessment because visible-spectrum anterior-segment edge detection can be influenced by limbal contrast, illumination, and pupil-related image characteristics. Fifth, the clinical consequences of the WTW discrepancy were evaluated by simulation rather than by direct outcome measurement. Postoperative vault, sulcus-to-sulcus correlation, and refractive outcomes were not assessed, so the simulation represents a decision-level discordance rather than an outcome-validated difference. Sixth, the simulation used a representative WTW-only nomogram and did not incorporate other inputs (such as ACD, sulcus-to-sulcus, or manufacturer-specific calculators) used in current clinical practice. Future studies should combine paired device measurements with manufacturer-specific calculators, sulcus-to-sulcus imaging, and postoperative vault analysis to establish outcome-level rather than decision-level disagreement.

In conclusion, ARGOS and ALADDIN provide highly interchangeable measurements for AL, ACD, LT, and Km in routine cataract surgery planning. In contrast, WTW corneal diameter shows a clinically meaningful systematic bias, a large effect size, wide limits of agreement, and a low proportion of eyes within the predefined ± 0.10 mm tolerance threshold; in a decision-anchored simulation, this measurement-level discrepancy translates into a discordant WTW-

based phakic IOL size assignment in approximately 40% of eyes. A clear distinction should therefore be drawn between routine cataract biometry, for which the devices are largely interchangeable, and WTW-dependent surgical decisions such as phakic IOL, ICL, or sulcus-supported lens sizing, for which the devices should not be mixed. WTW values should not be used interchangeably when they inform millimetric surgical decisions; a single biometric device should be used consistently in that setting, and the device used should be documented in the surgical record.

References

1. Hoffer, K. J. & Savini, G. Update on intraocular lens power calculation study protocols: the better way to design and report clinical trials. *Ophthalmology* **128**, e115–e120 (2021).
2. Olsen, T. Calculation of intraocular lens power: a review. *Acta Ophthalmol. Scand.* **85**, 472–485 (2007).
3. Srivannaboon, S., Chirapapaisan, C., Chonpimai, P. & Locket, S. Clinical comparison of a new swept-source optical coherence tomography-based optical biometer and a time-domain optical coherence tomography-based optical biometer. *J. Cataract Refract. Surg.* **41**, 2224–2232 (2015).
4. Shammas, H. J., Ortiz, S., Shammas, M. C., Kim, S. H. & Chong, C. Biometry measurements using a new large-coherence-length swept-source optical coherence tomographer. *J. Cataract Refract. Surg.* **42**, 50–61 (2016).
5. Kurian, M. et al. Biometry with a new swept-source optical coherence tomography biometer: repeatability and agreement with an optical low-coherence reflectometry device. *J. Cataract Refract. Surg.* **42**, 577–581 (2016).
6. Akman, A., Asena, L. & Güngör, S. G. Evaluation and comparison of the new swept source OCT-based IOLMaster 700 with the IOLMaster 500. *Br. J. Ophthalmol.* **100**, 1201–1205 (2016).
7. Huang, J. et al. Comprehensive comparison of axial length measurement with three swept-source OCT-based biometers and partial coherence interferometry. *J. Refract. Surg.* **35**, 115–120 (2019).

8. Gao, R. et al. Comparison of ocular biometric measurements between a new swept-source optical coherence tomography and a common optical low coherence reflectometry. *Sci. Rep.* **7**, 2484 (2017).
9. Yang, C. M., Lim, D. H., Kim, H. J. & Chung, T. Y. Comparison of two swept-source optical coherence tomography biometers and a partial coherence interferometer. *PLoS One* **14**, e0223114 (2019).
10. Nemeth, G. & Modis, L. Jr. Ocular measurements of a swept-source biometer: repeatability data and comparison with an optical low-coherence interferometry biometer. *J. Cataract Refract. Surg.* **45**, 789–797 (2019).
11. Calvo-Sanz, J. A., Poyales, F., Zhou, Y., Arias-Puente, A. & Garzón, N. Agreement between the biometric measurements used to calculate the size of the implantable collamer lenses measured with four different technologies. *Indian J. Ophthalmol.* **70**, 1586–1592 (2022).
12. Chen, X. et al. Effect of the difference between the white-to-white and sulcus-to-sulcus on vault and the related factors after ICL implantation. *Ophthalmol. Ther.* **10**, 947–955 (2021).
13. Munteanu, M. et al. Agreement in biometric parameters between swept-source optical coherence tomography and optical low-coherence interferometry: insights into clinical precision. *J. Clin. Med.* **14**, 1407 (2025).
14. Al Barri, L., Mercea, N., Yasar, I. I., Munteanu, M. & Stanca, H. T. Evaluation of refractive predictive accuracy in intraocular lens power calculations: a comparative study of swept-source optical coherence tomography and optical low-coherence interferometry. *J. Clin. Med.* **14**, 1201 (2025).
15. Al Barri, L. et al. Interdevice agreement of keratometry, astigmatism vectors, and ocular biometry in cataract candidates: SS-OCT (Argos) vs. OLCI (Aladdin) vs. Scheimpflug-Placido (Sirius). *Bioengineering (Basel)* **13**, 296 (2026).
16. von Elm, E. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* **370**, 1453–1457 (2007).
17. Kottner, J. et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J. Clin. Epidemiol.* **64**, 96–106 (2011).

18. Bonett, D. G. Sample size requirements for estimating intraclass correlations with desired precision. *Stat. Med.* **21**, 1331–1335 (2002).
19. Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–160 (1999).
20. Passing, H. & Bablok, W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. *J. Clin. Chem. Clin. Biochem.* **21**, 709–720 (1983).
21. Lin, L. I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**, 255–268 (1989).
22. Arriola-Villalobos, P. et al. Agreement and clinical comparison between a new swept-source optical coherence tomography-based optical biometer and an optical low-coherence reflectometry biometer. *Eye (Lond.)* **31**, 437–442 (2017).
23. Natung, T. et al. Ocular biometry characteristics and corneal astigmatism in cataract surgery candidates at a tertiary care center in North-East India. *Indian J. Ophthalmol.* **67**, 1417–1423 (2019).
24. González-Godínez, S. et al. Ocular biometry in dense cataracts: comparison of partial-coherence interferometry, swept-source optical coherence tomography and immersion ultrasound. *Indian J. Ophthalmol.* **70**, 107–111 (2022).

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

A.G. conceived and designed the study, performed the measurements, analyzed and interpreted the data, and wrote the main manuscript text. M.Y.T., E.Ş., and R.K. contributed to data acquisition, data curation, and critical revision of the manuscript. H.D.D. supervised the study and critically revised the manuscript. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Data availability

The datasets generated and analysed during the current study are available from the corresponding author (A.G.) on reasonable request.

Additional information

An earlier version of this manuscript has been deposited as a preprint on Research Square (<https://doi.org/10.21203/rs.3.rs-7035462/v1>).

Figure legends

Figure 1 Bland-Altman plot for white-to-white (WTW) corneal diameter comparing ARGOS (SS-OCT) and ALADDIN (OLCI) biometers (n=100 eyes). The solid red line represents the mean difference (bias = -0.180 mm); dashed blue lines indicate the 95% limits of agreement (-0.579 to $+0.220$ mm), with shaded bands indicating their 95% confidence intervals; the shaded green band represents the predefined clinical tolerance threshold (± 0.10 mm). Only 34% of eyes fell within the tolerance band, indicating that WTW measurements are not clinically interchangeable between devices.

Figure 2 Bland-Altman plots for axial length (AL), anterior chamber depth (ACD), lens thickness (LT) and mean keratometry (Km) comparing ARGOS (SS-OCT) and ALADDIN (OLCI) biometers (n=100 eyes). Solid red lines represent mean differences (bias); dashed blue lines indicate 95% limits of agreement; shaded green bands represent predefined clinical tolerance thresholds. For all four parameters, the majority of eyes fell within the tolerance band, supporting clinical interchangeability.

Figure 3 Simulated phakic IOL size assignment based on WTW from each device. (A) Cross-tabulation heatmap of simulated ICL sizes derived independently from ARGOS-WTW (rows) and ALADDIN-WTW (columns); concordant assignments (diagonal) account for 60% of eyes, and off-diagonal cells indicate discordant assignments. (B) Direction of discordance: in 40 eyes with discordant assignment, ALADDIN-derived sizing was larger than ARGOS-derived sizing in 36 (90%) and smaller in 4 (10%).

Figure 1

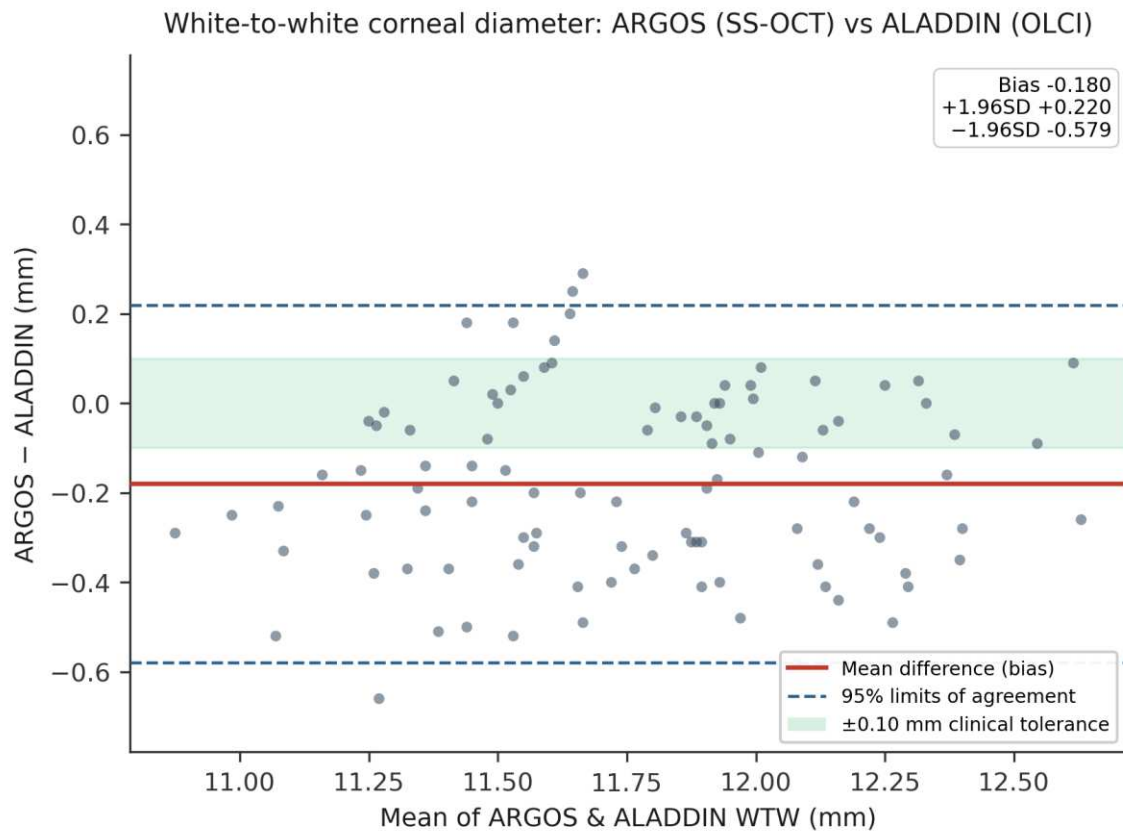


Figure 2

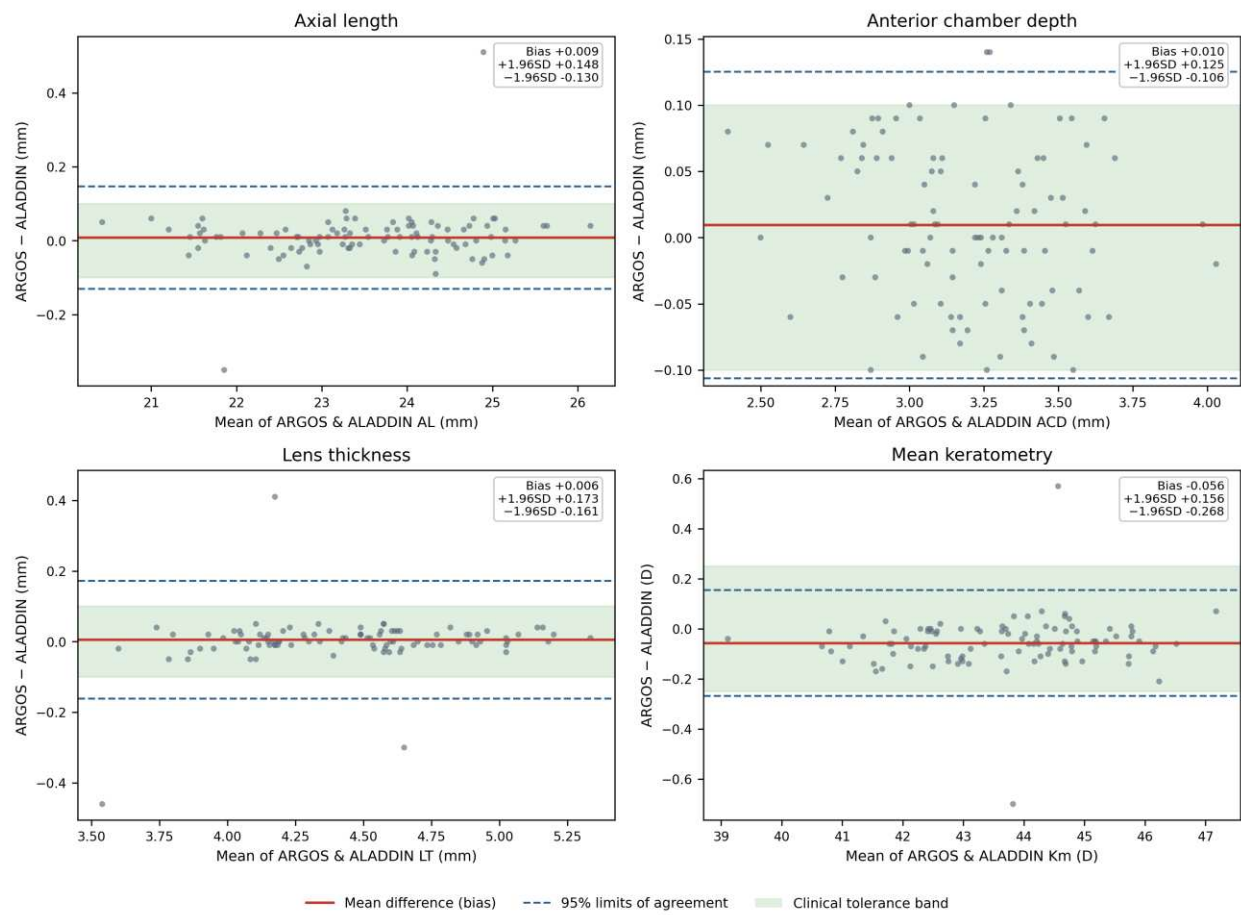
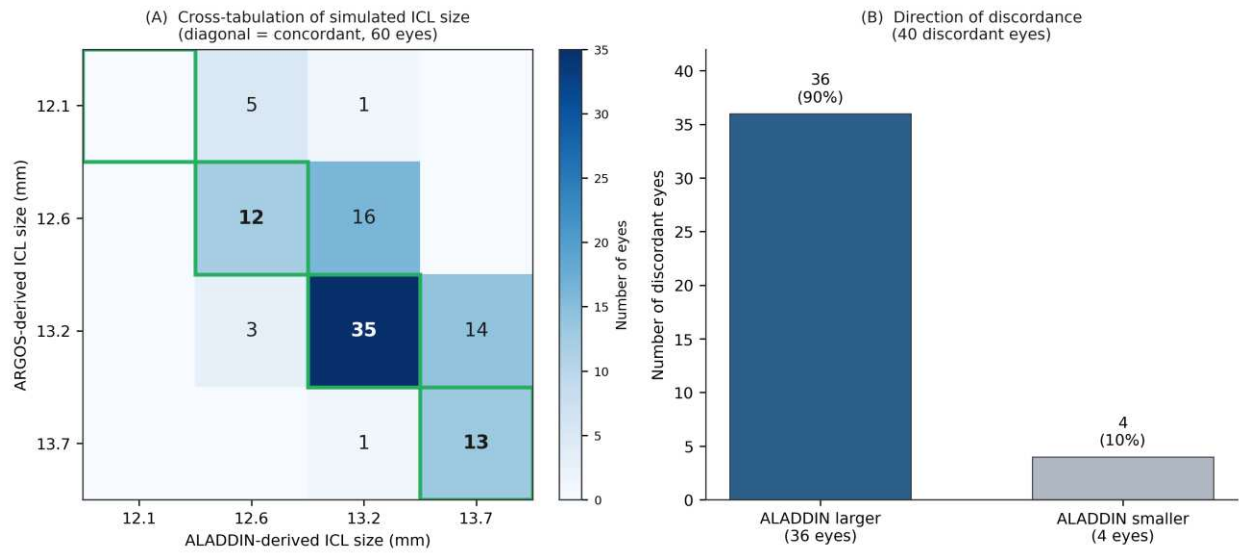
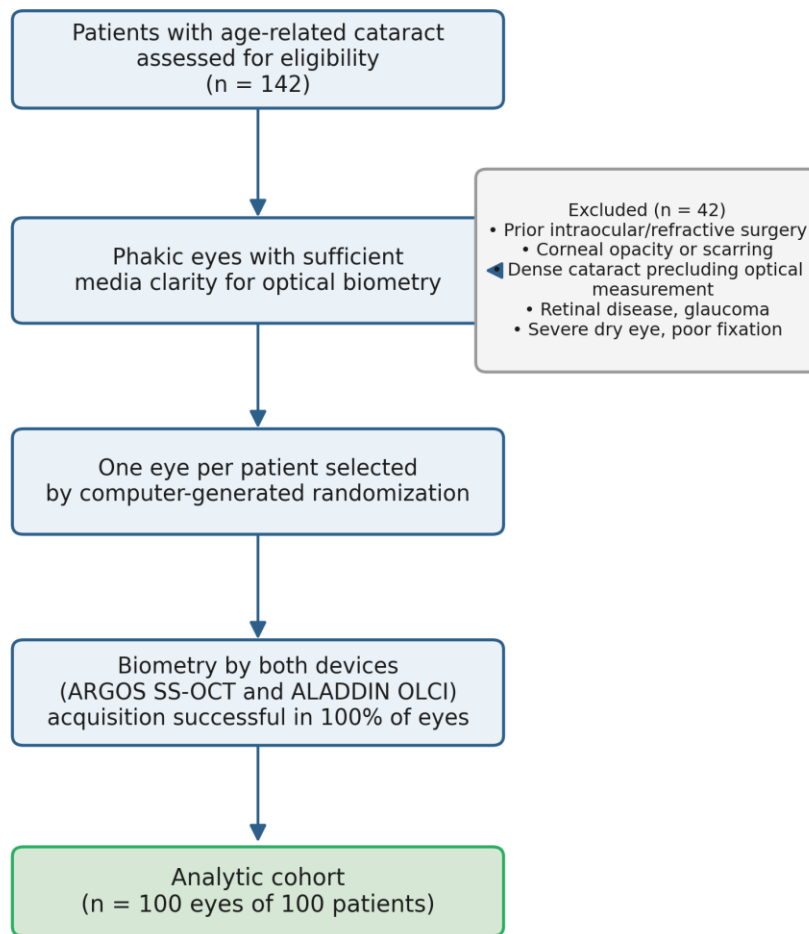


Figure 3

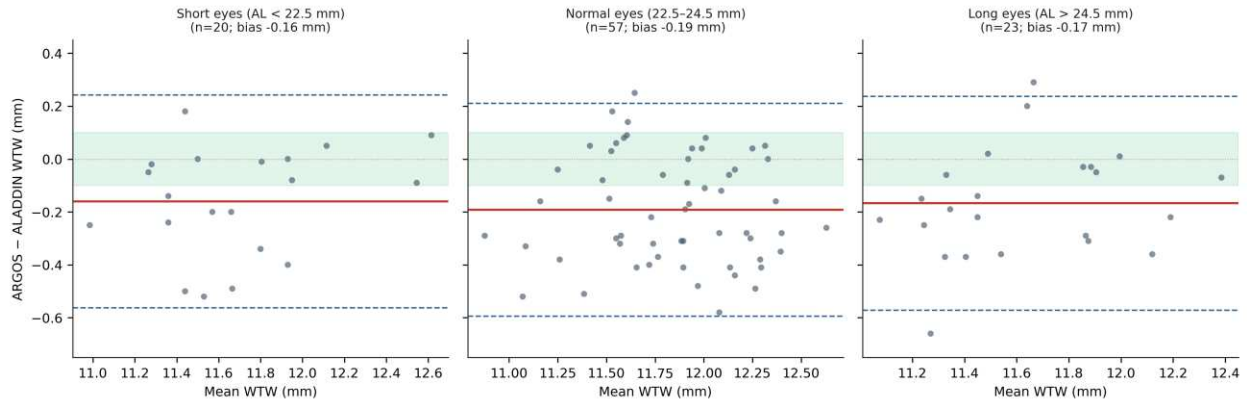


Supplementary information

Supplementary Fig. S1. STROBE flow diagram of participant selection. Of 142 patients screened during the study period, 42 were excluded and 100 were enrolled in the final analytic cohort; biometric data acquisition was successful in 100% of enrolled eyes.



Supplementary Fig. S2. Bland-Altman plots for WTW stratified by axial-length subgroup (short, normal and long eyes). The systematic WTW discrepancy persisted in all three subgroups, with similar bias magnitudes (-0.16 , -0.19 and -0.17 mm) and similarly low proportions of eyes within the ± 0.10 mm tolerance threshold (45%, 32% and 30%).



Supplementary Table S1. 95% confidence intervals for the limits of agreement and Passing-Bablok intercepts for all five biometric parameters.

Parameter	Bias	Lower LoA (95% CI)	Upper LoA (95% CI)	PB intercept
AL (mm)	+0.009	-0.130 (-0.154, -0.105)	+0.148 (+0.123, +0.172)	-0.010
ACD (mm)	+0.010	-0.106 (-0.126, -0.086)	+0.125 (+0.105, +0.146)	-0.116
LT (mm)	+0.006	-0.161 (-0.190, -0.131)	+0.173 (+0.143, +0.202)	+0.045
Km (D)	-0.056	-0.268 (-0.305, -0.231)	+0.156 (+0.119, +0.193)	+0.253
WTW (mm)	-0.180	-0.579 (-0.650, -0.509)	+0.220 (+0.150, +0.290)	+0.590

95% LoA = bias $\pm 1.96 \times$ SD of differences; 95% CIs for each LoA calculated using $SE = \sqrt{(3s^2/n)}$ and the t-distribution with $n-1$ degrees of freedom (Bland & Altman, 1999). PB intercept = Passing-Bablok intercept (slopes and their CIs are shown in Table 2). For WTW, the intercept of +0.590 mm with a slope near unity (0.966) is consistent with a constant additive offset rather than a proportional bias.

Supplementary Table S2. Summary of published ARGOS–ALADDIN and related SS-OCT/OLCI biometry comparisons.

Study	Devices	Eyes (n)	AL	ACD	Km	WTW
Munteanu et al. 2025[13]	ARGOS vs ALADDIN	170	0.975	0.960	0.924*	0.338
Al Barri et al. 2025[14,15]	ARGOS vs ALADDIN	Not uniformly reported	Excellent	Excellent	Sig. diff.	Lower
Yang et al. 2019[9]	ARGOS vs IOLMaster 700 vs PCI	146	>0.99	>0.95	>0.95	Device-dep.
Present study	ARGOS vs ALADDIN	100	0.998	0.981	0.997	0.793

Values are presented as reported in the original articles. Where exact ICC values were not available for all parameters, the direction and clinical interpretation of agreement are summarized descriptively. *Average of K1 and K2 ICCs reported by Munteanu et al. AL = axial length; ACD = anterior chamber depth; Km = mean keratometry; WTW = white-to-white corneal diameter; PCI = partial coherence interferometry; Sig. diff. = significant difference; Device-dep. = device-dependent.

Supplementary Table S3. Tolerance sensitivity analysis: proportion of eyes within ± 0.10 , ± 0.20 and ± 0.30 mm for the four length-based biometric parameters.

Parameter	Within ± 0.10 mm	Within ± 0.20 mm	Within ± 0.30 mm
AL	98 (98%)	98 (98%)	98 (98%)

Parameter	Within ± 0.10 mm	Within ± 0.20 mm	Within ± 0.30 mm
ACD	98 (98%)	100 (100%)	100 (100%)
LT	96 (96%)	96 (96%)	97 (97%)
WTW	34 (34%)	51 (51%)	69 (69%)

Values are n (%) of 100 eyes; the same $\pm 0.10/\pm 0.20/\pm 0.30$ mm thresholds are applied to all four parameters to allow direct comparison. AL, ACD and LT reach near-complete agreement at the tightest threshold, whereas WTW remains below 70% even at ± 0.30 mm, confirming that the inter-device WTW discrepancy is not an artefact of an arbitrarily strict tolerance. For AL, the two eyes outside ± 0.10 mm are gross outliers (0.35 and 0.51 mm) that remain outside even the ± 0.30 mm threshold. AL = axial length; ACD = anterior chamber depth; LT = lens thickness; WTW = white-to-white corneal diameter.