



LETTER TO THE EDITOR

Agreement analysis between swept-source optical coherence tomography and Laguna ONhE: statistical considerations

Enes Toklu¹ ¹. Department of Ophthalmology, Bakircay University Faculty of Medicine, Izmir Cigli Training and Research Hospital, Izmir, Turkey.

Dear Editor,

We read with great interest the article entitled “Agreement between structural parameters provided by the swept-source OCT and Laguna ONhE”, which evaluated the agreement between swept-source optical coherence tomography (OCT) and the Laguna ONhE software for optic nerve head measurements⁽¹⁾. The authors should be commended for addressing an important clinical issue, as low-cost, color fundus photography-based tools may improve access to structural glaucoma assessment in settings where OCT is not readily available. The use of Bland–Altman analysis, random selection of one eye per participant, and an internal calibration approach represents notable methodological strengths.

Nevertheless, several analytical aspects may affect the interpretation of interchangeability between the two methods. First, the Bland–Altman results require careful presentation. The Methods section states that 95% limits of agreement (LoA) were calculated; however, the tables appear to report narrow confidence intervals around the mean bias rather than the actual LoA. This distinction is clinically relevant. Confidence intervals around the mean difference estimate the precision of the average bias, whereas LoA describe the expected range of disagreement for individual eyes^(2,3). For a method proposed as a potential alternative to OCT, individual-level agreement is more informative than mean bias alone.

Second, agreement should not be inferred primarily from *p* values. After calibration, the mean disc-area bias was no longer statistically significant, which was interpreted as evidence of correction of systematic measurement error. However, the absence of a statistically significant paired difference does not demonstrate that two methods agree sufficiently for clinical use. Two devices may exhibit minimal average bias while still showing substantial disagreement in individual patients. Therefore, reporting postcalibration LoA, absolute prediction error, and the proportion of eyes falling within clinically acceptable error ranges would provide a more clinically meaningful assessment.

Third, the Bland–Altman plots suggest possible proportional bias for several parameters, particularly the area cup-to-disc ratio, vertical cup-to-disc ratio, and rim volume. When measurement differences vary across the measurement range, a single mean bias may not adequately characterize agreement. This issue is particularly relevant because the large-disc subgroup exhibited more pronounced bias. Because large optic discs are precisely those in which distinguishing physiological cupping from glaucomatous damage can be challenging, disc-size-dependent disagreement warrants cautious interpretation.

The calibration equation for disc area is promising, but its validation could be described more comprehensively. Because the calibration and validation samples

<http://dx.doi.org/10.5935/0004-2749.2026-0216>**Submitted for publication:**

July 16, 2026

Accepted for publication:

September 9, 2026

Funding:

This study received no specific financial support.

Disclosure of potential conflicts of interest:

The author declares no potential conflicts of interest.

Corresponding author:

Enes Toklu

E-mail: enestoklu@gmail.com

Data Availability Statement:

The datasets generated and/or analyzed during the current study are already available.

Edited by**Editor-in-Chief:** Newton Kara-Júnior**Associate Editor:** Carolina Pelegrini
Barbosa Gracitelli

were drawn from the same retrospective cohort, the analysis represents internal validation. External validation in independent populations, using different image-acquisition settings, disc phenotypes, and distributions of glaucoma severity, would be necessary before broader clinical implementation. In addition, the equivalence analysis would be more transparent if the confidence intervals for the two one-sided tests and the clinical rationale for the $\pm 0.10 \text{ mm}^2$ and $\pm 0.20 \text{ mm}^2$ margins were reported⁽⁴⁾.

Finally, the exclusion of eyes for which Laguna ONhE could not perform automated analysis should be considered when interpreting real-world applicability. Agreement among analyzable images may overestimate practical performance if excluded eyes had poorer image quality, atypical discs, media opacity, or more advanced disease. A comparison between included and excluded eyes would help clarify the generalizability of the findings.

In conclusion, this study provides a valuable contribution by evaluating an accessible tool for optic nerve head assessment. Our comments are intended to emphasize that Laguna ONhE and OCT should not yet be considered fully interchangeable for all structural parameters. The evidence appears strongest for calibrated disc-area estimation, whereas the cup-to-disc ratios and rim volume require more cautious interpretation.

AUTHOR CONTRIBUTIONS:

Significant contribution to conception and design: Enes Toklu.

Data Acquisition: not applicable.

Data Analysis and Interpretation: Enes Toklu.

Manuscript Drafting: Enes Toklu.

Significant intellectual content revision of the manuscript: Enes Toklu.

Final approval of the submitted manuscript: Enes Toklu.

Statistical analysis: Enes Toklu.

Obtaining funding: not applicable.

Supervision of administrative, technical, or material support: not applicable.

Research group leadership: not applicable.

REFERENCES

1. Silva Filho GV, Caiado GC, Ricardo OM, Teixeira SH, Prata TS, Gracitelli CP, et al. Agreement between structural parameters provided by the swept-source OCT and Laguna ONhE. *Arq Bras Oftalmol*. 2026;89(5):e20250377.
2. Gerke O. Reporting standards for a Bland-Altman agreement analysis: a review of methodological reviews. *Diagnostics (Basel)*. 2020;10(5):334.
3. Giavarina D. Understanding Bland Altman analysis. *Biochem Med (Zagreb)*. 2015;25(2):141-51.
4. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci*. 2017;8(4):355-62.



LETTER TO THE EDITOR

Reply to “Agreement analysis between swept-source optical coherence tomography and Laguna ONhE: statistical considerations”

Gilvan Vilarinho da Silva Filho¹ , Gustavo Coelho Caiado¹ , Olívia Moura de Paula Ricardo² ,
Sergio Henrique Teixeira¹ , Tiago dos Santos Prata¹ ,
Carolina Pelegrini Barbosa Gracitelli¹ , Augusto Paranhos Jr¹

1. Glaucoma Sector, Department of Ophthalmology and Visual Sciences, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil.

2. Glaucoma Instituto, Belo Horizonte, MG, Brazil.

Dear Editor,

We thank Dr. Toklu for his careful reading of our article⁽¹⁾ and for his constructive comments, which we address below.

Limits of agreement. We acknowledge that Tables 2 and 3 presented the 95% confidence intervals (95% CIs) of the mean bias rather than the 95% limits of agreement (LoA), which were presented only graphically in Figures 2 and 3. Among the 349 eyes with valid measurements on both devices, the limits of agreement (LoA) for optical coherence tomography (OCT) minus Laguna ONhE were -0.45 to $+0.58$ mm² for disc area, -0.23 to $+0.40$ for the area cup-to-disc ratio, -0.18 to $+0.36$ for the vertical cup-to-disc ratio, and -0.23 to $+0.19$ mm³ for rim volume⁽²⁾.

Equivalence and individual agreement. We agree that the absence of a statistically significant difference does not demonstrate agreement. Accordingly, our interpretation of disc area was based on the Two One-Sided Tests (TOST) procedure⁽³⁾. Before calibration, the 90% CI of the mean difference ($+0.042$ to $+0.088$ mm²) fell within the ± 0.10 mm² equivalence margin (TOST, $p=0.006$). The mean absolute difference was 0.18 mm², and 41% and 70% of eyes exhibited differences within ± 0.10 and ± 0.20 mm², respectively.

Upon reexamining the calibration analysis in response to these comments, we found that the postcalibration results reported in the article (bias, 0.025 mm²; $p=0.3$; Figure 2B and Table 5) had inadvertently been calculated using the uncalibrated Laguna ONhE values from the validation set. The equation derived from the calibration set ($n=244$) was as follows: $\text{OCT disc area} = 0.277 + 0.902 \times \text{Laguna disc area}$. Applying this equation to the validation set ($n=105$), the bias was -0.049 mm² (95% CI, -0.097 to -0.001 ; $p=0.047$), the LoA were -0.54 to $+0.44$ mm², and the 90% CI (-0.089 to -0.009 mm²) remained within the ± 0.10 mm² equivalence margin (TOST, $p=0.019$). The mean absolute difference was 0.15 mm² both before and after calibration, whereas 50% and 74% of calibrated estimates fell within ± 0.10 and ± 0.20 mm², respectively. Thus, our conclusion regarding group-level equivalence for disc area remains unchanged, whereas calibration does not appear to improve individual-level agreement.

Proportional bias and large discs. We agree. Regression of the differences on the means demonstrated proportional bias for the area cup-to-disc ratio (slope, 0.54), vertical cup-to-disc ratio (0.52), and rim volume (-0.25 ; all $p<0.001$) as well as a small

<http://dx.doi.org/10.5935/0004-2749.2026-0216>

Corresponding author:

Augusto Paranhos Jr

E-mail: augusto.paranhos@gmail.com



effect for disc area (0.07; $p=0.03$). Because the large-disc subgroup was defined according to OCT disc area, part of the observed bias may reflect regression to the mean. When the subgroup was instead defined according to the mean disc area measured by both methods ($n=50$), the disc-area bias was 0.11 mm^2 (95% CI, $0.02-0.20 \text{ mm}^2$). Thus, disagreement in large discs persists and warrants caution.

Validation and equivalence margins. We agree that the calibration was validated only internally. External validation in independent populations is the next step, and the present findings provide a rationale for such a study. No established threshold exists for clinically relevant differences in disc area. The ± 0.10 and $\pm 0.20 \text{ mm}^2$ margins were defined by the authors and corresponded to approximately 5% and 10%, respectively, of the mean disc area in our sample. Our equivalence findings therefore apply within these prespecified margins.

Excluded eyes. Automated analysis of ophthalmic images has inherent limitations. Laguna ONhE was unable to analyze 86 eyes, and an additional 13 eyes were excluded because

of low OCT signal strength. Therefore, our findings apply to images yielding valid measurements on both devices. The exclusion of eyes that could not be analyzed by Laguna ONhE was acknowledged as a limitation of our study.

In conclusion, we agree with Dr. Toklu that Laguna ONhE disc-area estimates are equivalent to those obtained with OCT at the group level; however, the two methods should not be considered interchangeable for individual eyes, particularly with respect to the cup-to-disc ratios and rim volume.

REFERENCES

1. Silva Filho GV, Caiado GC, Ricardo OM, Teixeira SH, Prata TS, Gracitelli CP, et al. Agreement between structural parameters provided by the swept-source OCT and Laguna ONhE. *Arq Bras Oftalmol.* 2026;89(5):e20250377.
2. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-10. Comment in: *Lancet.* 2004;363(9403):164; author reply 154-5.
3. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. *Soc Psychol Personal Sci.* 2017;8(4):355-62.