Journal of Karnali Academy of Health Sciences
An Official Journal of Karnali Academy of Health Sciences

Reproducibility and reliability of magnetic resonance imaging and B-mode ultrasound in ocular axial length measurement

Prasanna Ghimire[1], **Avinash Singh**[1], **Hari Sapkota**[1]

[1]Department of Radiodiagnosis and Imaging, Nepalgunj Medical College Teaching Hospital, Kohalpur, Nepal

ABSTRACT

Introduction: Accurate ocular axial length (AL) measurement is critical for intraocular lens power calculation and refractive outcomes. While optical biometry is the clinical standard, dense media opacities limit its utility. This study evaluated B-mode ultrasound (B-USG) and MRI reproducibility and reliability for ocular AL measurement and assess inter-method agreement as alternative AL measurement methods.

Methods: This prospective study included 102 adults (mean age 31.5±6.5 years) with normal ocular anatomy and emmetropic refraction. AL measurement used B-USG (9–12 MHz) and 1.5 Tesla MRI. Two blinded observers performed independent measurements; one repeated measurement after one day. Reliability was assessed using Intraclass Correlation Coefficient (ICC), Standard Error of Measurement (SEM), and Minimum Detectable Change (MDC). The agreement was evaluated using Bland–Altman analysis and Pearson correlation.

Results: Mean AL was 23.618±0.726 mm for B-USG and 24.033±0.727 mm for MRI. Both methods showed excellent intra-observer (ICC=0.92 for ultrasound; ICC=0.95 for MRI) and inter-observer reliability (ICC=0.90 for ultrasound; ICC=0.93 for MRI). Strong linear correlation was observed ($r=0.95$, $p<0.001$). Bland-Altman analysis revealed significant systematic bias, with MRI measuring systematically longer values (+0.42 mm, 95% CI: 0.38–0.46), with wide limits of agreement (−0.15 to +0.99 mm), indicating non-interchangeable methods.

Conclusion: B-USG and MRI are highly reliable and reproducible for AL measurement. However, consistent systematic bias and wide agreement limits restrict interchangeable use for high absolute accuracy applications. These modalities are valuable optical biometry alternatives, provided single method consistency.

Keywords: Biometry, B-mode ultrasound, MRI, Reproducibility, Reliability

INTRODUCTION

Accurate measurement of ocular axial length (AL), the distance from the anterior corneal surface to the retinal pigment epithelium, is fundamental in ophthalmology for intraocular lens (IOL) power calculation, where a 0.1 mm error can cause a refractive error of 0.25–0.35 diopters. [1] It is also used for monitoring pathologies such as myopia progression, glaucoma risk, and nanophthalmos.

Optical biometry, based on partial coherence interferometry, is the gold standard due to its non-contact nature and high precision, but it is limited by dense media opacities. [2] Alternative methods are needed when optical biometry is unfeasible. [3] Conventional A-scan ultrasound relies on operator skill and probe alignment, whereas B-mode ultrasound (B-USG) provides real-time imaging ensuring alignment. [4] B-mode ultrasound further offers additional advantages, including easy availability, wide accessibility in routine radiology and ophthalmology practice, and relatively low cost, making it a practical option in resource-limited settings.

Copyright © 2025 by the author(s), wherein the author(s) are the only owners of the copyright of the published content
Licensing: This published content is distributed under the terms of the Creative Commons Attribution International License (CC BY 4.0) license, and is free to access on the Journal's website. The author(s) retain ownership of the copyrights and publishing rights without limitations for their content, and they grant others permission to copy, use, print, share, modify, and distribute the article's content even for commercial purposes. Disclaimer: This publication's claims, opinions, and information are the sole creations of the specific author(s) and contributor(s). Errors in the contents and any repercussions resulting from the use of the information included within are not the responsibility of the publisher, editor, or reviewers. Regarding any jurisdictional assertions in any published articles, their contents, and the author's institutional affiliations, the Journal and its publisher maintain their objectivity.

Corresponding Author:
Dr. Prasanna Ghimire
Email: drprasannaghimire@gmail.com

Date of submission: September 23, 2025
Date of Acceptance: December 29, 2025
Date of Publication: January 10, 2026

DOI: https://doi.org/10.61814/jkahs.v8i3.1048

MRI, unaffected by media opacities, offers high-resolution images of the eye and orbit, making it a potential alternative. This study evaluates intra- and inter-observer reliability of B-USG and MRI and assesses their agreement using Bland-Altman analysis. [5]

METHODS

This prospective, observational study was conducted at Nepalgunj Medical College Teaching Hospital from October 2023 to March 2025. The study was approved by the Institutional Ethics Committee (Ref: NGMC-IRC-04/080-81) and written informed consent was obtained from all participants.

A total of 102 adults (51 male, 51 females; mean age 31.48 ± 6.52 years) referred for MRI of the brain and orbit were enrolled.

Inclusion criteria were: i) age 21–45 years to represent a stable adult population where ocular axial length remains relatively constant minimizing from ongoing ocular growth in younger individuals and agerelated degenerative changes (e.g., cataract, presbyopia, or retinal pathology) in older adults thereby ensuring measurement reliability; ii) a normal ophthalmological examination, spherical equivalent refraction within ±0.5 diopters, and iii) normal orbital and brain anatomy on MRI.

Exclusion criteria were prior ocular surgery or pathology.

Based on pilot data and established methodologies for reliability and agreement studies, 102 participants provided 80% power to detect clinically meaningful differences with $\alpha=0.05$.

Ocular biometry with Bmode ultrasound was performed using a GE P6 ultrasound system with a linear 9–12 MHz transducer, following established Bmode protocols. MRI axial length measurements were adapted from prior validated methodologies. [6]

Ocular biometry was performed using a GE P6 ultrasound system with a linear 9-12 MHz transducer, following established Bmode ultrasound protocols for ocular axial length measurement. [7] With the patient supine and eyes gently closed, coupling gel was applied over the upper eyelid. The probe was positioned to acquire real-time sagittal images that bisected the pupil and optic nerve head, thus ensuring alignment along the visual axis. The axial length was measured from the anterior corneal surface to the posterior retinal echo. The average of three measurements was recorded per session.

Imaging was performed on a 1.5 Tesla GE Signa Creator scanner with a dedicated orbital coil. The protocol included high-resolution 3D T2-weighted sequences (slice thickness 0.8–1.0 mm, FOV 180 mm, matrix 512×512) optimized for globe assessment, consistent with prior validated methodology. [2] Additional standard T1, T2, and STIR sequences were acquired for orbital evaluation. AL was measured on sagittal reformations from the anterior corneal surface to the posterior scleral wall using the scanner's measurement tools.

Two blinded observers independently performed all measurement. To evaluate intra-observer reliability, Observer A repeated measurements one day later.

Data were analyzed using IBM SPSS Statistics v28.0 (IBM Corp., Armonk, NY). Reliability was assessed with the Intraclass Correlation Coefficient (ICC) using a two-way mixed-effects model, interpreted as: <0.50 = poor, 0.50–0.75 = moderate, 0.75–0.90 = good, >0.90 = excellent [5]. Measurement precision was further evaluated with the Standard Error of Measurement (SEM) and Minimum Detectable Change (MDC).

Agreement between B-USG and MRI was examined with Bland–Altman analysis to determine mean difference (bias) and 95% limits of agreement. Pearson correlation was used to assess the strength of linear association. Subgroup analyses (gender, age, axial length, refractive status) were conducted with one-way ANOVA. Statistical significance was set at $p < 0.05$.

RESULTS

All 102 participants (51 males, 51 females; mean age 31.48 ± 6.52 years) completed both B-USG and MRI examinations without complications. All had normal ocular anatomy and emmetropic refraction (spherical equivalent ±0.50 D). All MRI examinations were reported as normal.

Axial length (AL) measurements were normally distributed for both modalities. Mean AL was 23.618 ± 0.726 mm (range 22.41–25.25 mm) with B-USG and 24.033 ± 0.727 mm (range 22.82–25.66 mm) with MRI (Table 1).

Table 1: Comprehensive descriptive statistics of axial length measurements

Parameter	B-mode Ultrasound	MRI
Combined Mean ± SD (mm)	23.618 ± 0.726	24.033 ± 0.727
Range (mm)	22.410-25.253	22.820-25.662
95% CI	23.47-23.77	23.89-24.18
Median (IQR)	23.52 (23.1-24.2)	23.94 (23.5-24.6)

Individual measurement precision was assessed by calculating the standard deviation of the three consecutive measurements for each examination. Both modalities demonstrated excellent within-session precision. The mean intra-examination SD was 0.048 ± 0.021 mm for B-USG and 0.039 ± 0.018 mm for MRI, with >98% of all measurements showing SD <0.1 mm.

Both modalities achieved excellent intra-observer, inter-observer, and test–retest reliability (ICC ≥ 0.90) (Table 2). MRI consistently showed slightly higher ICC values than B-USG. SEM and MDC values were lower for MRI, indicating higher measurement stability.

Table 2: Reliability of B-mode ultrasound and MRI for ocular axial length measurement

Reliability Metric	B-mode Ultrasound	MRI	Clinical Interpretation
Intra-observer ICC (95% CI)	0.92 (0.88–0.95)	0.95 (0.93–0.97)	Excellent
Inter-observer ICC (95% CI)	0.90 (0.86–0.93)	0.93 (0.89–0.96)	Excellent
Test-retest ICC (95% CI)	0.91 (0.87–0.94)	0.94 (0.91–0.96)	Excellent
Standard Error of Measurement (SEM)	0.036	0.030	-
Minimum Detectable Change (MDC_{95})	0.100	0.083	-

Although strongly correlated ($r = 0.95$, $p < 0.001$), Bland–Altman analysis demonstrated a systematic bias, with MRI measurements averaging 0.42 mm longer than B-USG (95% CI: 0.38–0.46 mm) (Figure 1). Limits of agreement ranged from −0.15 to +0.99 mm (Table 3). This magnitude of difference corresponds to a refractive error of >1.0 diopter, precluding interchangeable use.

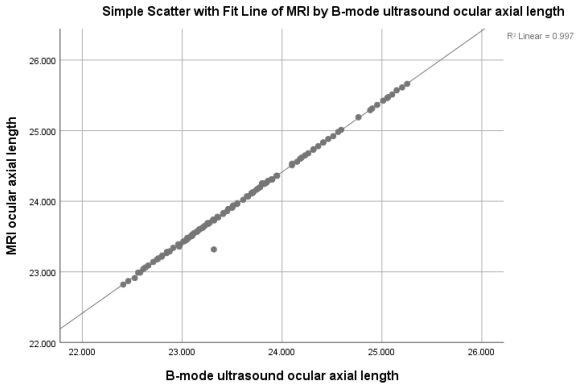

Figure 1: Scatter plot illustrating the relationship between B-mode USG and MRI average ocular axial length.

Table 3: Detailed inter-method agreement analysis

Agreement Parameter	Statistical Value	95% Confidence Interval
Correlation Analysis		
Pearson Correlation (r)	0.95	0.93-0.97
R^2 (variance explained)	90.3%	-
Systematic Bias Assessment		
Mean Difference (MRI - B-mode)	+0.42 mm	0.38-0.46 mm
Standard Deviation of Differences	0.29 mm	-
Statistical Significance		
Paired t-test	t = 14.6	-
p-value	<0.001	Highly significant
Limits of Agreement		
Lower Limit (Mean - 1.96×SD)	-0.15 mm	-0.21 to -0.09
Upper Limit (Mean + 1.96×SD)	+0.99 mm	0.93 to 1.05

Most participants (57.9 %) showed MRI–B-USG differences between 0.20 and 0.60 mm, with a small proportion exceeding 0.80 mm (Table 4). Only 6.8% showed minimal or reversed differences.

Table 4: Distribution of measurement differences between MRI and B-mode ultrasound

Difference Range (mm)	Frequency (%)	Cumulative %	Clinical Impact*
< -0.10	3(2.9%)	2.9%	B-mode > MRI (rare)
-0.10 to 0.00	4(3.9%)	6.9%	Minimal difference
0.00 to 0.20	18(17.6%)	24.5%	Small MRI advantage
0.20 to 0.40	31(30.4%)	54.9%	Moderate difference
0.40 to 0.60	28(27.5%)	82.4%	Clinically significant
0.60 to 0.80	15(14.7%)	97.1%	Large difference
0.80 to 1.00	2(2.0%)	99.0%	Very large difference
> 1.00	1(1.0%)	100.0%	Extreme difference

Clinical impact estimated for IOL power calculation (0.1 mm AL error ≈ 0.25-0.35 D refractive error)

One-way ANOVA showed no statistically significant differences in the mean inter-method AL difference across gender (p=0.847), age groups (p=0.623), axial length categories (p=0.412), or refractive status (p=0.789) (Table 5). The systematic bias was consistent across the study population.

Table 5: Subgroup analysis of inter-method agreement

Subgroup	N	Mean Difference ± SD (mm)	95% LoA (mm)	p-value*
Gender Analysis				
Male	51	0.42 ± 0.29	-0.15 to +0.99	0.847
Female	51	0.42 ± 0.29	-0.15 to +0.99	
Age Group Analysis				
21-30 years	52	0.41 ± 0.28	-0.14 to +0.96	0.623
31-40 years	38	0.43 ± 0.30	-0.16 to +1.02	
41-45 years	12	0.42 ± 0.31	-0.19 to +1.03	
Axial Length Categories				
Short eyes (<22.5 mm)	15	0.40 ± 0.27	-0.13 to +0.93	0.412
Normal eyes (22.5-24.5 mm)	72	0.42 ± 0.29	-0.15 to +0.99	
Long eyes (>24.5 mm)	15	0.44 ± 0.31	-0.17 to +1.05	
Refractive Status Analysis				
Emmetropic (±0.25 D)	82	0.42 ± 0.28	-0.13 to +0.97	0.789
Low hyperopia (+0.26 to +0.50 D)	10	0.43 ± 0.32	-0.20 to +1.06	
Low myopia (-0.26 to -0.50 D)	10	0.41 ± 0.27	-0.12 to +0.94	

***p-values from one-way ANOVA comparing mean differences between subgroups**

DISCUSSION

The present study systematically evaluated the reproducibility and reliability of B-mode ultrasound (B-USG) and magnetic resonance imaging (MRI) for ocular axial length (AL) measurement in a prospective cohort of healthy adults. Both modalities demonstrated excellent intra- and inter-observer reliability, with intraclass correlation coefficients (ICCs) exceeding 0.90 across all assessments. [8] These results are concordant with established benchmarks for clinical measurement tools and are broadly consistent with prior reliability studies of ultrasound biometry. [1,2,8] The findings confirm that both techniques are capable of generating stable and internally consistent measurements when applied under standardized conditions. [9,10]

Beyond reliability, the central finding of this investigation was the demonstration of a systematic inter-modality bias. MRI consistently produced longer AL estimates than B-USG, with a mean difference of +0.42 mm (95% CI: 0.38–0.46 mm) and wide limits of agreement (−0.15 to +0.99 mm). Although the two modalities were strongly correlated (r = 0.95), this magnitude of bias is clinically consequential. Prior optical biometry literature has established that even a 0.1 mm discrepancy in AL measurement can translate into a postoperative refractive shift of approximately 0.25–0.35 diopters. [1] Accordingly, the observed bias corresponds to a refractive deviation exceeding 1.0 diopter, a level of error that would be considered unacceptable in contemporary cataract surgery, where precision targets often fall within ±0.50 diopters.

The mechanistic basis of this discrepancy warrants careful consideration. B-USG derives its posterior reference point from acoustic reflections at the vitreoretinal interface, whereas MRI, owing to its superior soft-tissue contrast, may resolve boundaries at the choroid–sclera junction or even the external scleral contour. [11, 12] This discrepancy in anatomical endpoints, rather than stochastic error, provides a plausible explanation for the systematic nature of the bias. Comparable findings have been documented in smaller series, suggesting that this is an intrinsic limitation of cross-modality comparison rather than a methodological artefact. [11,12]

The clinical implications are twofold. First, methodological consistency is imperative. Although both B-USG and MRI demonstrated excellent internal reproducibility, their lack of interchangeability precludes alternating between modalities within the same patient. In longitudinal contexts such as monitoring myopia progression interchanging modalities may generate artefactual impressions of axial elongation. [11] In surgical contexts, particularly intraocular lens (IOL) power calculation, the use of non-standardized modalities could compromise refractive precision and ultimately visual outcomes. [13] Second, while both modalities remain viable alternatives to optical biometry when media opacities preclude light-based techniques, their deployment must be carefully contextualized. Selection should be guided by availability, cost considerations, and institutional expertise, rather than assumptions of equivalence. [4,14,15]

The methodological strengths of the present study include its prospective design, a well-characterized and demographically balanced cohort, and rigorous statistical analysis incorporating both reliability indices and agreement metrics. The use of two blinded observers, with repeated measurements, minimized observer bias and strengthened the robustness of findings. Importantly, all participants had normal orbital anatomy on MRI, reducing the likelihood of confounding structural abnormalities.

Nevertheless, several limitations must be acknowledged. The single-center setting may restrict external validity, and the exclusive inclusion of emmetropic individuals limits extrapolation to populations with high refractive error, who may represent a more clinically relevant group. Furthermore, optical biometry, as the gold standard for AL measurement, was not included in the comparative framework due to pragmatic constraints. [16] Finally, the study cohort comprised healthy participants rather than individuals with dense media opacities, the very population in which MRI or B-USG would serve as the principal alternatives to optical techniques.

Future research should therefore extend these findings through multicenter designs, larger and more heterogeneous populations, and inclusion of pathological eyes. Direct three-way comparisons among optical biometry, B-USG, and MRI would allow not only validation but also the development of cross-modality calibration algorithms. Such work is essential to determine whether MRI can be integrated as a

reliable adjunct or substitute within the broader clinical armamentarium of ocular biometry, particularly in surgical planning and refractive research. [6]

CONCLUSION

This prospective study demonstrates that both B-mode ultrasound and magnetic resonance imaging yield excellent intra- and inter-observer reliability for ocular axial length measurement in adults with normal refraction and orbital anatomy. Despite this reproducibility, the modalities are not interchangeable, as MRI consistently produces longer measurements than B-USG, with a systematic bias of approximately 0.42 mm and wide limits of agreement. Given that this discrepancy exceeds clinically acceptable thresholds for intraocular lens power calculation, modality consistency is imperative in both surgical and longitudinal applications.

Accordingly, while B-USG and MRI represent valuable alternatives when optical biometry is precluded, their use must be restricted to internally consistent protocols rather than substituted interchangeably. Further research incorporating direct comparison with optical biometry and evaluation in pathological eyes is required to delineate the precise clinical role of MRI within ocular biometry.

DECLARATION

Acknowledgements

None

Ethical Approval

This research was approved by IRC/of Nepalgunj Medical College and Teaching Hospital with the reference number of 04/080-81 on 4th August, 2023.

Consent/Assent

Informed written consent was obtained from the all the participants before data collection.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

PG reviewed the literature, conceptualized and designed the research, contributed to data acquisition, analysis, statistical work, and manuscript preparation and review; AS contributed to research concepts, design, data acquisition, manuscript preparation, and review; HS contributed to research concepts, design, data acquisition, manuscript preparation, and review. All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the research work.

Conflicts of Interest

There are no conflict of interest.

Funding

The author(s) received no external fund for this research.

REFERENCES

1. Melles RB, Holladay JT, Chang WJ. Accuracy of intraocular lens calculation formulas. Ophthalmology. 2018;125(2):16978. | DOI |
2. Findl O, Kriechbaum K, Sacu S, Kiss B, Polak K, Nepp J, et al. Influence of operator experience on the performance of ultrasound biometry compared to optical biometry before cataract surgery. J Cataract Refract Surg. 2003;29(10):19505. | DOI |
3. Khorrami-Nejad M, Khodair AM, Khodaparast M, Babapour Mofrad F, Dehghanian Nasrabadi F. Comparison of the ocular ultrasonic and optical biometry devices in the different quality measurements. J Optom. 2023;16(4):28495. | DOI |
4. Wong TY, Foster PJ, Ng TP, Tielsch JM, Johnson GJ, Seah SK. Variations in ocular biometry in an adult Chinese population in Singapore: the Tanjong Pagar Survey. Invest Ophthalmol Vis Sci. 2001;42(1):7380. | PubMed |
5. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):30710. | PubMed |
6. Gutiérrez-Vázquez J, Pérez-Sánchez LI, Satrústegui-Lapetra M, Arevalo-Manso JJ, Gómez-Herrera JJ, Criado-Álvarez JJ. Measurement of the anterior chamber depth and ocular axial length: transpalpebral Bmode ultrasound using an 18MHz linear probe compared with the IOL Master 500. Med Ultrason. 2021;23(1):4854. | DOI |
7. Akduman EI, Nacke RE, Leiva PM, Akduman L. Accuracy of ocular axial length measurement with MRI. Ophthalmologica. 2008;222(6):3979. | DOI |
8. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):15563. | DOI |
9. Sabih R, Saad M, Arshad H, Sohail S, Faheem R. Comparative analysis of refractive outcomes using optical and ultrasound biometry in phacoemulsification cataract surgery. J Health Rehabil Res. 2024;4(3):15. | DOI |
10. Cass K, Thompson CM, Tromans C, Wood IC. Evaluation of the validity and reliability of Ascan ultrasound biometry with a single use disposable cover. Br J Ophthalmol. 2002;86(3):3449. | DOI |
11. Park UC, Lee EK, Kim BH, Oh BL. Decreased choroidal and scleral thicknesses in highly myopic eyes with posterior staphyloma. Sci Rep. 2021;11(1):7987. | DOI |
12. Patel S, Marshall J, Fitzke FW 3rd. Refractive index of the human corneal epithelium and stroma. J Refract Surg. 1995;11(2):1005. | DOI |
13. De Bernardo M, Cione F, Capasso L, Coppola A, Rosa N. A formula to improve the reliability of optical axial length measurement in IOL power calculation. Sci Rep. 2022;12(1):18845. | DOI |
14. Walter KA, Gagnon MR, Hoopes PC Jr, Dickinson PJ. Accurate intraocular lens power calculation after myopic laser in situ keratomileusis, bypassing corneal power. J Cataract Refract Surg. 2006;32(3):4259. | DOI |
15. Wiseman SJ, Tatham AJ, Meijboom R, Terrera GM, Hamid C, Doubal FN, et al. Measuring axial length of the eye from magnetic resonance brain imaging. BMC Ophthalmol. 2022;22(1):54. | DOI |
16. Ganekar A, Mandlik H, Joshi A, Tripathi A, Divya N. Study of accuracy of axial length measurements by two biometric methods with reference to postoperative spherical equivalents in patients undergoing phacoemulsification. Indian J Clin Exp Ophthalmol. 2020;6(4):5959. | DOI |