

Probabilistic Reasoning Layers for Intraocular Lens Power Prediction: Toward Explainable AI in Ophthalmology

A. Tsimpouras

October 2025

Abstract

Accurate intraocular lens (IOL) power calculation remains one of the most critical yet uncertain steps in cataract surgery. Current formulas-ranging from empirical regression-based equations to modern machine-learning models-produce deterministic outputs that fail to represent patient-level variability or measurement uncertainty.

This paper introduces a probabilistic reasoning layer, built on SHAP-informed feature attribution and ensemble-based variance propagation, that augments existing prediction systems with interpretable confidence estimation. The framework aims to improve transparency, mitigate outlier risks, and strengthen human-AI collaboration in ophthalmic decision-making. This paper is presented as a conceptual proposal and methodological note.

1. Introduction

Predicting postoperative refractive outcomes continues to pose challenges despite advances in biometric precision and modeling. Although contemporary IOL formulas increasingly incorporate machine learning, many function as opaque predictors lacking interpretability and explicit uncertainty representation. Such opacity limits clinical trust and adoption-particularly for atypical biometric profiles, where prediction errors have the greatest impact on patient outcomes.

This work proposes an auxiliary probabilistic reasoning layer designed to complement any existing model. By translating complex feature interactions into transparent, clinically meaningful explanations, the approach provides both predictive performance and interpretive clarity, aligning algorithmic inference with established surgical heuristics.

2. Proposed Framework

The proposed reasoning layer operates atop a pre-trained IOL power prediction model and introduces three synergistic components:

1. Feature Attribution

Shapley Additive Explanations (SHAP) quantify the individual contribution of each biometric feature (e.g., axial length, keratometry, anterior chamber depth) to the final prediction.

2. Uncertainty Propagation

A variance-weighted ensemble aggregates the outputs of multiple regressors to produce confidence intervals reflecting both data-driven and model-driven uncertainty.

3. Human-Readable Mapping

Model insights are expressed through interpretable metrics that connect measurable input shifts to refractive prediction variance, enabling clinicians to visualize confidence bounds in diopter space.

Together, these components establish a bridge between classical clinical heuristics and AI-driven inference, allowing for more transparent and auditable decision support in cataract surgery.

3. Prototype Implementation

A proof-of-concept system can be implemented using standard open-source ML components while maintaining clinical interpretability:

1. Base Modeling Layer

Employ ensemble regressors such as Gradient Boosting, Random Forest, or hybrid stacking architectures within scikit-learn. Each model predicts postoperative refraction from biometric inputs (axial length, keratometry, anterior chamber depth, lens thickness).

2. Explainability Layer

Integrate SHAP for per-feature contribution analysis. TreeExplainer or KernelExplainer can be used depending on the base model. The resulting SHAP values form the weighting scheme for subsequent uncertainty modeling.

3. Uncertainty and Calibration Module

Apply variance-weighted ensembling or Bayesian posterior reweighting across multiple regressors to estimate prediction intervals. Optional Monte-Carlo sampling at inference time quantifies epistemic uncertainty.

4. Personalization Adapter

Introduce surgeon- or site-specific embeddings that fine-tune model priors using local outcome data, enabling adaptive calibration without retraining the core ensemble.

This modular structure allows the reasoning layer to be appended to existing IOL formulas or learning systems with minimal integration cost while preserving interpretability for clinical review.

4. Discussion

The proposed reasoning layer extends IOL prediction systems beyond point estimates, offering interpretable uncertainty quantification that mirrors clinical reasoning. By decomposing feature contributions and propagating variance through ensemble aggregation, the framework allows clinicians to visualize why and how a prediction is made. This directly addresses one of the primary obstacles in AI-assisted ophthalmology: limited trust due to model opacity.

In practical terms, the system outputs diopter confidence intervals and explanatory mappings aligned with surgeon heuristics—for instance, showing how small keratometry or axial-length deviations shift expected refractive outcomes. Such interpretability enhances both clinical transparency and liability readiness, two essential factors for regulatory and real-world adoption.

Unlike deep learning black boxes, this modular approach can be audited, versioned, and calibrated to site-specific data without sacrificing traceability. It thus supports a gradual transition toward human-AI collaborative decision systems in cataract surgery—where machine predictions complement rather than replace clinician judgment.

5. Future Work

Further research will focus on three main directions:

1. Clinical Validation and Benchmarking

Evaluate the reasoning layer on large, multi-center cataract surgery datasets to assess its reliability across device types, demographics, and surgical techniques. Comparative studies against current state-of-the-art formulas (Barrett, Kane, Holladay 2, etc.) will determine measurable gains in prediction stability and interpretability.

2. Integration with Clinical Workflows

Develop a lightweight application programming interface (API) or plug-in to embed the reasoning layer into existing electronic medical record (EMR) or biometry systems. This will test usability, latency, and user-experience factors critical for surgeon adoption.

3. Regulatory and Human-Factors Evaluation

Explore the framework’s compliance with medical-device standards for explainable AI (e.g., EU AI Act, FDA Good Machine Learning Practice) and conduct human-in-the-loop studies measuring trust calibration, cognitive load, and decision consistency.

Longer-term extensions include incorporating multimodal data (OCT, topography) and refining the Bayesian component for hierarchical learning across institutions. Collectively, these steps aim to translate the reasoning-layer paradigm from a research prototype into a clinically deployable, auditable decision-support tool.

References

1. Lundberg, S. M., & Lee, S. I. (2017). *A Unified Approach to Interpreting Model Predictions*. Advances in Neural Information Processing Systems (NeurIPS).
2. Kane, J. X., Van Heerden, A., Atik, A., & Petsoglou, C. (2020). *Accuracy of 13 Intraocular Lens Power Formulas Across Axial Lengths*. Journal of Cataract & Refractive Surgery, 46(1), 8-14.
3. Yamauchi, T., Tabuchi, H., et al. (2021). *Prediction of Postoperative Refraction Using Machine Learning Ensembles in Cataract Surgery*. Scientific Reports, 11(1), 20413.
4. Doshi-Velez, F., & Kim, B. (2017). *Towards a Rigorous Science of Interpretable Machine Learning*. arXiv preprint arXiv:1702.08608.
5. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
6. Holzinger, A., Biemann, C., Pattichis, C. S., & Kell, D. B. (2017). *What Do We Need to Build Explainable AI Systems for the Medical Domain?* Reviews in the Journal of Biomedical Informatics, 94, 103133.

Acknowledgments

The author thanks colleagues and mentors for discussions that informed the early stages of this work. Special appreciation is extended to the open-source and academic communities developing tools for explainable AI and medical ML interpretability.

Author Note

This paper is presented as a **conceptual and methodological proposal**. No clinical data were used or analyzed. All opinions and views expressed are solely those of the author and do not reflect those of any institution or organization.

License

This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

You are free to share (copy and redistribute the material in any medium or format) and adapt (remix, transform, and build upon the material) for any purpose, even commercially, provided that you give appropriate credit to the original author.

© 2025 A. Tsimpouras. To view a copy of this license, visit:

<https://creativecommons.org/licenses/by/4.0/>.