Title Page:

Title: A Deep Learning Model to Predict Postoperative Refraction in Cataract Surgery

Authors: Aaron Hao Tan PhD(C) MASc[1], Luqmaan Moolla MD(C) MEng[2], Matt Schlenker MD FRCSC[3,4]

Affiliations:

1. University of Toronto Faculty of Applied Science and Engineering, Toronto, Ontario, Canada
2. University of Toronto Temerty School of Medicine, Toronto, Ontario, Canada
3. Department of Ophthalmology and Vision Sciences, University of Toronto, Toronto, Ontario, Canada
4. Kensington Eye Institute, University of Toronto, Toronto, Ontario, Canada

Corresponding Author: Dr. Matthew B. Schlenker MD, FRCSC
Kensington Vision & Research Centre
Suite 501, 340 College Street
Toronto, Ontario, Canada
M5T 3A9
Phone: 416-928-1335
Fax: 416-928-5075
Email: matt.schlenker@gmail.com

Word Count: 1198

Key Points:

New Instrument: Deep learning artificial intelligence algorithm that can predict postoperative refraction for cataract surgery with the inputs of axial length, keratometry 1, keratometry 2, A-constant, anterior chamber depth, age, sex, and lens power.

Finding: This deep learning algorithm was able to predict postoperative refraction with an overall accuracy of 82.5% for cases within ±0.50D of refractive error, which beat the commonly used Barrett Universal II (80.4%), and two conventional machine learning algorithms, gradient boosting regressor (73.1%) and random forest regressor (66.8%) in a dataset of 2490 eyes.

Conclusion: Our novel deep learning model demonstrated a higher accuracy and precision than both a modern IOL power formula (Barrett Universal II) and two classical ML approaches in predicting postoperative refraction in our data set of 2490 eyes.

Introduction

Cataracts are the leading cause of blindness worldwide, and the treatment of this ocular condition is the most common surgical procedure in ophthalmology.[1] A crucial factor in surgical planning is determining the appropriate intraocular lens (IOL) power to ensure a postoperative refractive outcome which yields acceptable visual results. Preoperative IOL power formulas consider several biometric factors when making an IOL power prediction. Currently available IOL equations, such as Barrett Universal II, Olsen, Haigis, and Holladay 1 can achieve 80.8%, 78.7%, 77.1%, and 76.6% of cases within ±0.50D of refractive error, respectively.[2] These accuracies, which may seem encouraging, still mean that approximately 20% of cases fall outside of an acceptable range of refractive success. In addition, many of these equations cannot accurately and consistently predict postoperative refraction for eyes with atypical biometric parameters, such as short or long axial lengths (ALs).[2] This is likely due to the non-linear relationship between common eyes and those with extreme biometric parameters.[3]

To overcome the limitations of the available IOL equations, researchers have explored artificial intelligence (AI) algorithms, such as machine learning (ML), to identify complex relationships within datasets to increase the accuracy and precision of postoperative refractive predictions.[4] However, many of the proposed ML approaches are based on classical techniques, such as support vector regression and random forest, with only a few having explored deep learning models.[5-7] Of these deep learning approaches, the proposed models either had a limited number of input variables [5-6] or a multi-stage pipeline,[7] which restricts their predictive performance because of the sparse feature space and cascading errors, respectively. Thus, the objective of this project is to develop a novel software instrument based on an end-to-end deep learning model to better predict postoperative refraction during cataract surgery for all eyes, including those with extreme biometric parameters.

Approach

Deep learning is a branch of ML that utilizes an artificial neural network architecture and labelled dataset to perform classification and regression tasks.[8] Unlike classical ML algorithms, deep learning employs multiple neural layers, in combination with non-linear activation functions, to approximate high-dimensional data with low-dimensional features.[8]

In this study, our dataset consisted of 2490 eyes with biometric data that included AL, keratometry 1, keratometry 2, A-constant, anterior chamber depth, age, sex, lens power used, and the observed 1-month postoperative refraction collected from an IOLMaster 500 machine. Our dataset underwent a 70-30 split to create a training and testing set of 1743 and 747 eyes, respectively.

In terms of the proposed deep learning model, we have designed an architecture with 5 neural layers (1 input layer, 3 hidden layers, 1 output layer). The input layer dimension is 8 to include all the biometric features available in the dataset, and the output layer dimension is 1 to predict the postoperative refraction. The hidden layer dimensions are tuned empirically to be 64, 128, and 64. To create a universal function approximator, a Leaky Rectified Linear Unit (L-ReLU) function was incorporated to include non-linearity. During training, the network's parameters were updated with stochastic gradient descent using a mean squared error function and a batch size of 64. To prevent the model from overfitting to the training data, early stopping was adopted. The proposed model was developed with PyTorch[9] and converged in approximately 130 epochs.

To benchmark the performance of the proposed deep learning model, two classical ML methods that have been used by previous AI studies for postoperative refraction prediction were implemented using the Sci-kit Learn library.[10] These methods include random forest and gradient boosting regressor. Each model was fitted on the same training set as the proposed deep learning model to predict postoperative refraction. The Barrett Universal II formula was also included as a benchmark as it consistently performs as one of the best non-learning-based IOL equations available.[2] To quantify the results, the test set was further categorized into short (AL < 22.5mm), intermediate (22.5 < AL < 25.5mm), and long eyes (AL > 25.5mm) to illustrate each method's accuracy for different subsets of the population. The primary outcome with the percentage of cases within ±0.50D of refractive error when using our novel neural network, random forest, gradient boosting regressor or Barrett Universal II formula.

Results and Discussion

Our novel deep learning model achieved an overall accuracy of 82.5% for cases within ±0.50D of refractive error, followed by Barrett Universal II (80.4%), gradient boosting regressor (73.1%), and random forest regressor (66.8%) on our test set. When analyzing the accuracies across a range of axial lengths, the proposed deep learning model outperformed both gradient boosting and random forest as seen in Figure 1A. In comparison with the Barrett Universal II formula, the deep learning approach was able to predict more accurately for short and long axial length cases (Figure 1A). For eyes with axial lengths between 22.5mm and 25.5mm, the prediction accuracies between the proposed deep learning model and Barrett II's formula were comparable (83.5% and 83.8% respectively). This trend is further illustrated in Figure 1B, where the proposed deep learning model was able to predict with better accuracy than all benchmarking algorithms, and matching that of the Barrett Universal II formula, for eyes with intermediate axial lengths.

In terms of precision, Figure 1C presents the standard deviation (SD), and the coefficient of variation (CV) for the accuracies of each method across all axial lengths. In particular, the proposed deep learning model had the lowest value for both SD and CV at (6.92, 8.56), respectively, compared to compared to Barrett II (11.81, 15.94), gradient boosting regressor (15.97, 24.21) and random forest regressor (16.55, 27.62). An F-test was conducted to further validate the precisions, where the results showed that they are statistically different with 90% probability. As a result, both the prediction accuracy and precision of our proposed deep learning model outperformed the classical machine learning methods, as well as the Barrett Universal II IOL power formula, which further validates its efficacy and generalizability across eyes with a range of axial lengths.

Conclusion

In conclusion, our novel deep learning model demonstrated a higher accuracy and precision than both a modern IOL power formula (Barrett Universal II) and two classical ML approaches in predicting postoperative refraction in our data set of 2490 eyes. Notably, our deep learning model was able to predict accurately and consistently across a range of ALs. This is a challenging area for modern IOL equations, as they have difficulty in accurately determining postoperative refraction for high myopes or hyperopes.[3] With that said, even though our dataset consists of a respectable 2490 eyes, more cases in the training set would be beneficial in increasing the performance of the deep learning model.[8] Furthermore, many of the modern IOL equations are starting to include additional parameters such as patient age, white-to white measurements, and lens thickness to increase their accuracy.[3] We are currently in the process of collecting a larger dataset that includes these preoperative ocular measurements. Once this is available, the next iteration of our deep learning model can be developed, and an extensive study can be conducted in a comparative analysis against alternative IOL power equations. Our hope is this new instrument can revolutionize preoperative cataract surgery planning and be developed further to address the 20% of cases with poor refractive accuracy.

Fig 1.

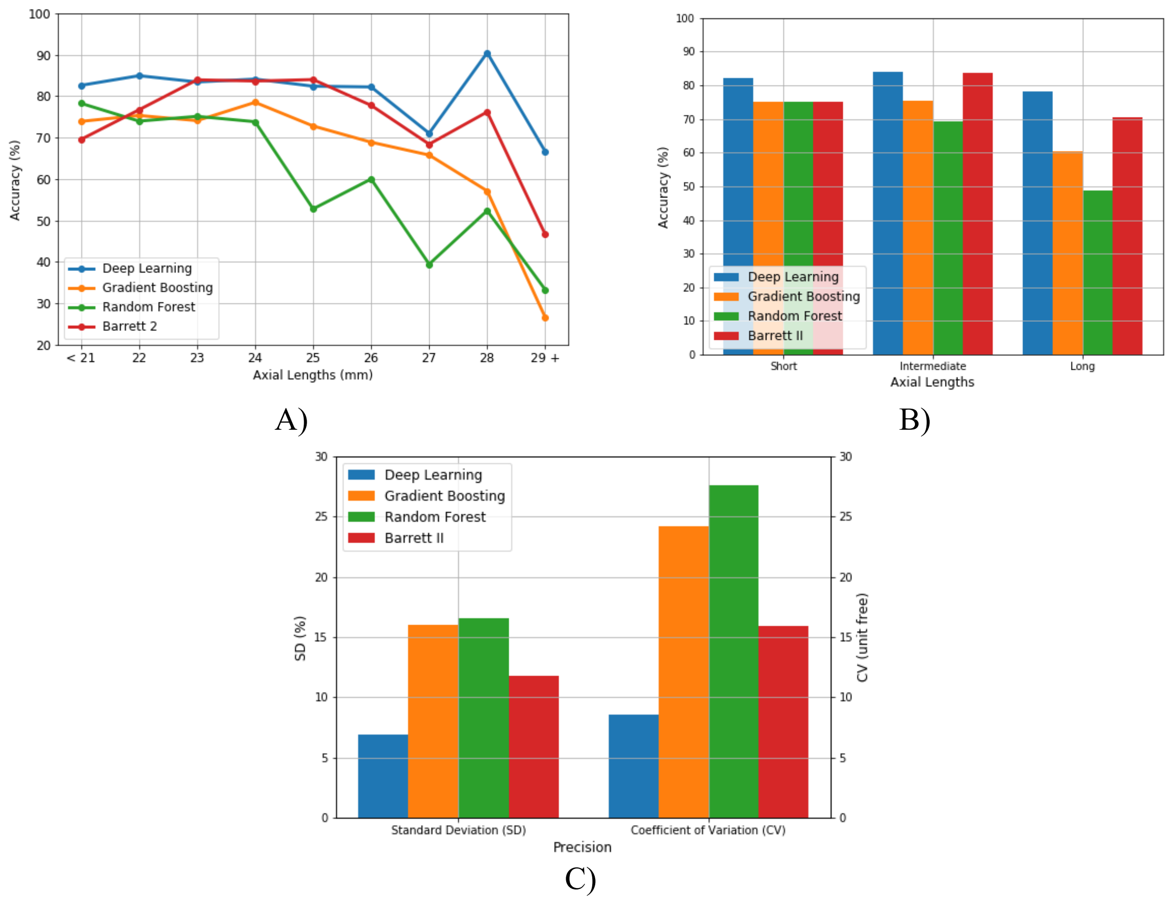

Figure Legend(s)

Fig 1. A) Accuracy at +/-0.50D across different axial lengths, B) Accuracy at +/-0.5D in short, intermediate, and long axial length groups C) Precision represented through standard deviation, and the coefficient of variation for the accuracies of each method across all axial lengths

Acknowledgements:

Conflicts of interest/Competing Interests: The authors of this paper declare that they are no conflicts of interest.

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

Authors' Contributions: All authors contributed to the manuscript and have reviewed/approved its contents.

Other Substantial Contributors: We thank Austin Pereira MD MEng, who is affiliated with the Department of Ophthalmology and Vision Sciences at the University of Toronto for his contributions to this project and for his help in writing this paper.

Data Accessibility: The data that support the findings of this study are available on request from the corresponding author.

References

1. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*. 2011;96(5):614-618. doi:10.1136/bjophthalmol-2011-300539
2. Melles RB, Holladay JT, Chang WJ. Accuracy of Intraocular Lens Calculation Formulas. *Ophthalmology*. 2018;125(2):169-178. doi:10.1016/j.ophtha.2017.08.027
3. Siddiqui AA, Devgan U. Mastering Lens Calculations: New Formulas and Comparisons. *Current Ophthalmology Reports*. 2018;6(4):233-236. doi:10.1007/s40135-018-0186-z
4. Kapoor R, Walters SP, Al-Aswad LA. The current state of artificial intelligence in ophthalmology. *Survey of Ophthalmology*. 2019;64(2):233-240. doi:10.1016/j.survophthal.2018.09.002
5. Yamauchi T, Tabuchi H, Takase K, Masumoto H. Use of a Machine Learning Method in Predicting Refraction after Cataract Surgery. *Journal of Clinical Medicine*. 2021;10(5):1103. doi:10.3390/jcm10051103
6. Fernández-Álvarez JC, Hernández-López I, Cruz-Cobas PP, Cárdenas-Díaz T, Batista-Leyva AJ. Using a multilayer perceptron in intraocular lens power calculation. *Journal of Cataract & Refractive Surgery*. 2019;45(12):1753-1761. doi:10.1016/j.jcrs.2019.07.035
7. Ladas J, Ladas D, Lin SR, Devgan U, Siddiqui AA, Jun AS. Improvement of Multiple Generations of Intraocular Lens Calculation Formulae with a Novel Approach Using Artificial Intelligence. *Translational Vision Science & Technology*. 2021;10(3):7. doi:10.1167/tvst.10.3.7
8. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. *WIREs Data Mining and Knowledge Discovery*. 2019;9(4). doi:10.1002/widm.1312

9. PyTorch. https://pytorch.org/. Accessed May 26, 2021.

10. Scikit-learn. scikit. https://scikit-learn.org/stable/. Accessed May 26, 2021.