AI-powered effective lens position prediction improves the accuracy of existing lens formulas

Tingyang Li,[1] Joshua Stein,[2,3,4] Nambi Nallasamy [1,2]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bjophthalmol-2020-318321).

[1]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA
[2]Kellogg Eye Center, Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, Michigan, USA
[3]Department of Health Management and Policy, University of Michigan School of Public Health, Ann Arbor, Michigan, USA
[4]Center for Eye Policy and Innovation, University of Michigan, Ann Arbor, Michigan, USA

Correspondence to
Dr. Nambi Nallasamy, Kellogg Eye Center, University of Michigan, 1000 Wall St, Ann Arbor, MI 48105, USA; nnallasa@med.umich.edu

Received 12 November 2020
Revised 21 February 2021
Accepted 12 March 2021

To cite: Li T, Stein J, Nallasamy N. *Br J Ophthalmol* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjophthalmol-2020-318321

ABSTRACT

Aims To assess whether incorporating a machine learning (ML) method for accurate prediction of postoperative anterior chamber depth (ACD) improves the refraction prediction performance of existing intraocular lens (IOL) calculation formulas.

Methods A dataset of 4806 patients with cataract was gathered at the Kellogg Eye Center, University of Michigan, and split into a training set (80% of patients, 5761 eyes) and a testing set (20% of patients, 961 eyes). A previously developed ML-based method was used to predict the postoperative ACD based on preoperative biometry. This ML-based postoperative ACD was integrated into new effective lens position (ELP) predictions using regression models to rescale the ML output for each of four existing formulas (Haigis, Hoffer Q, Holladay and SRK/T). The performance of the formulas with ML-modified ELP was compared using a testing dataset. Performance was measured by the mean absolute error (MAE) in refraction prediction.

Results When the ELP was replaced with a linear combination of the original ELP and the ML-predicted ELP, the MAEs±SD (in Diopters) in the testing set were: 0.356±0.329 for Haigis, 0.352±0.319 for Hoffer Q, 0.371±0.336 for Holladay, and 0.361±0.331 for SRK/T which were significantly lower ($p<0.05$) than those of the original formulas: 0.373±0.328 for Haigis, 0.408±0.337 for Hoffer Q, 0.384±0.341 for Holladay and 0.394±0.351 for SRK/T.

Conclusion Using a more accurately predicted postoperative ACD significantly improves the prediction accuracy of four existing IOL power formulas.

INTRODUCTION

The estimation of postoperative intraocular lens (IOL) position is essential to IOL power calculations for cataract surgery. Norrby and Olsen have reported that inaccuracy in the prediction of the postoperative anterior chamber depth (ACD) is the number one source of error for postoperative refraction prediction.[1 2] In addition to its vital role in IOL formulas, the postoperative ACD is also a critical variable in ray tracing, where the uncertainty in the postoperative ACD directly affects the accuracy of the results. Methods to improve the accuracy of the prediction of postoperative ACD have been studied for decades. In first-generation formulas, the lens position was represented by a constant. Later, more and more preoperative biometric variables such as the axial length (AL) and the corneal power were added to calculate the postoperative IOL position. In 1993, Holladay first proposed the term 'expected lens position' or ELP to indicate the location of the lens as it relates to a given optical model of the eye.[3] The ELP estimates in SRK/T, Holladay1 and Hoffer Q are derived based on theoretical formulas. The ELP estimate in the Haigis formula is a simple linear combination of the AL and the preoperative ACD. Although ELP was initially intended to estimate the position of the IOL, ELPs in the aforementioned formulas were developed to account for different formula-specific assumptions and regression results.[1 4] In order to reflect the use of ELP to account for these formula-specific assumptions and regression results, the term ELP today refers to 'effective lens position' rather than 'expected lens position'. In view of the limitations of the ELP in existing formulas, recently, more efforts have been devoted to constructing ELPs that better reflect the true location of the IOL.[5–9] New IOL power prediction methods have also been developed based on the new-generation ELP prediction methods, and they have shown that using a more accurately predicted IOL position helps to improve the IOL power prediction accuracy.[5]

It is so far largely unexplored whether inserting a more accurately predicted ELP into existing formulas improves refraction prediction accuracy. This is an important question because: (1) it provides a fast and efficient way to modify and improve on existing IOL formulas whose reliability has been tested extensively; (2) such research can provide supports for translating the continued improvements in accuracy in postoperative ACD prediction into better refraction predictions in published formulas. Several previous studies had modified the ELPs in existing formulas in order to achieve better refraction prediction results in certain cataract cases. Modification of ELP calculation in the Haigis formula for sulcus-implanted IOLs was reported to improve performance.[10] Kim *et al* adjusted the ELP estimation in SRK/T formulas with the corneal height in postrefractive patients and achieved satisfactory accuracy.[11] It remains to be explored whether improvement of ELP estimates for in-the-bag IOL placement can improve IOL power calculations of existing formulas for general cataract patients.

Since most recently published IOL formulas (eg, Barrett Universal II,[12 13] Holladay 2, Olsen formula[14]) are either not disclosed to the public or do not have the option to customise the value of ELP during the prediction of postoperative refraction, here we applied our previously developed postoperative ACD prediction methods to a dataset of 4806 cataract surgery patients and replaced the ELP estimates in 4 existing IOL formulas: Haigis, Hoffer Q, Holladay and SRK/T. We combined our machine learning (ML) prediction of true postoperative ACD with the original ELP estimated by each formula and substituted this updated ELP prediction for each formula. We then compared the refraction prediction performance of each formula using its original and enhanced ELP estimates. The findings reported here demonstrate that existing formulas can benefit from improved methods for predicting true postoperative ACD.

MATERIALS AND METHODS

Postoperative ACD prediction ML model

In previous work,[15] we developed an ML-based postoperative ACD prediction model, which predicts the postoperative ACD (in mm) based on preoperative biometry. Here, in the presented study, an ACD prediction ML model was trained using the method and dataset (847 patients, 1205 eyes, 4137 records) described in the previous research. The dataset was composed of the preoperative and postoperative biometry measured by the Lenstar LS900 optical biometers (Haag-Streit USA, EyeSuite software V.i9.1.0.0) at the University of Michigan's Kellogg Eye Center. The postoperative ACD was defined as the distance from the front surface of the cornea to the front surface of the IOL. The postoperative ACD predicted by the ML model is referred to as ELP_{ML} in this manuscript.

Data collection

In this study, biometry records were collected using the same approach as for the development of the ML postoperative ACD prediction model at University of Michigan's Kellogg Eye Center.[15]

The inclusion criteria were: (1) patients who had cataract surgery (Current Procedural Terminology (CPT) code=66 984 or 66 982) but no prior refractive surgery and no additional surgical procedures at the time of cataract surgery. (2) The implanted lens was an Alcon SN60WF single-piece acrylic monofocal lens (Alcon, USA). Each case in the dataset corresponds to one operation of a single eye with preoperative and postoperative information. The preoperative information includes the measurements of the AL, lens thickness (LT), ACD, flat keratometry (K1), steep keratometry (K2), and the average keratometry which was calculated as $K = \frac{K1+K2}{2}$. The postoperative information includes the postoperative refraction (spherical component SC and cylindrical component CC) where the time when it was recorded was closest to 1 month (30 days) after surgery. Since the patients were measured in a lane of 10 feet long (3.048 m), which was shorter than the standard length of 20 feet (6 m), the SC was adjusted for the vergence distance by adding $\frac{1}{6} - \frac{1}{test\ distance\ in\ meters} = \frac{1}{6} - \frac{1}{3.048} = -0.1614$ according to Simpson and Charman's recommendation.[16] The spherical equivalent (SE) refraction was therefore calculated as $SE\ refraction = (SC - 0.1614) + 0.5CC$. Samples that were used to train the postoperative ACD prediction ML model were excluded from the dataset so that the dataset better simulates unseen samples.

The dataset in total consisted of 4806 patients (figure 1). The dataset was split into a training dataset used for the development of the methods and a testing dataset used for performance comparison. Eighty per cent of the patients were randomly assigned to the training set, and the rest of the patients (20%) were assigned to the testing set. For patients who had more than one associated case in the testing set (ie, patients who had both eyes operated on), one case was randomly selected to ensure each patient had the same weight when the prediction performance was evaluated. At the end of this process, the training set had 3845 patients (5761 eyes), and the testing set had 961 patients (961 eyes).

Linear regression model

We implemented four existing formulas (Haigis, Hoffer Q, Holladay, and SRK/T) in Python based on their publications.[17–24] The existing formulas calculated the ELP (ELP_F) as a function of the preoperative biometry (figure 1): $ELP_F = f_0(biometry)$. The predicted ELP (ELP_F) was then used to predict the postoperative refraction: $refraction = f_1(ELP_F, biometry)$. Here, the goal was to reduce the refraction prediction error by replacing ELP_F with a different value, ELP_F'. Our approach involves two steps: (1) finding the theoretically most optimal ELP values, (2) modelling the most optimal ELP with ELP_F and the ML-predicted postoperative ACD, denoted ELP_{ML}.

In the first step, the most optimal ELP (denoted ELP_{BC}) was found by the standard method of back-calculating the ELP when the predicted refraction was set to equal the true refraction (ie,

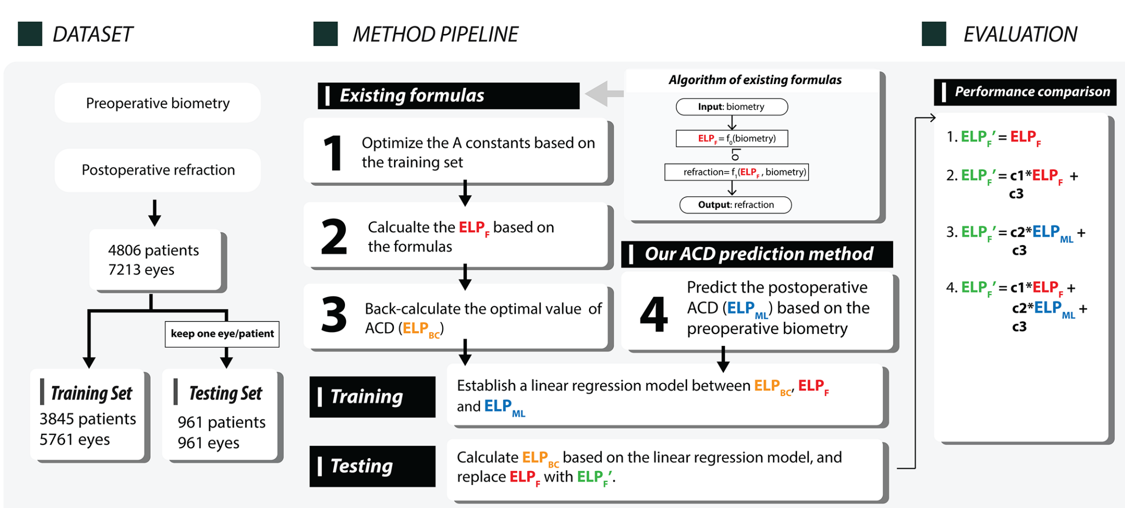

Figure 1 The analysis pipeline of the presented study. ELP_F = the effective lens position (ELP) estimated by the existing formulas. ELP_{ML} = the postoperative anterior chamber depth (ACD) predicted by the machine learning method. ELP_{BC} = the back-calculated ELP (see main text). ELP_F' is a term that refers to a new ELP that is used to replace the ELP_F in the existing formulas.

Table 1 The summary statistics for the patient demographics for the training and testing dataset

Characteristic	Training set	Testing set
Gender	Male: 2514 eyes (43.6%), Female: 3247 eyes (56.4%)	Male: 425 eyes (44.2%), Female: 536 eyes (55.8%)
Age at surgery (years)	70.99±9.61	70.10±10.24
Preoperative K (D)	43.85±1.64	43.90±1.66
Preoperative AL (mm)	24.19±1.40	24.20±1.41
Preoperative LT (mm)	4.54±0.45	4.53±0.45
Preoperative ACD (mm)	3.24±0.41	3.26±0.41
Postoperative refraction (D)	−0.53±0.96	−0.57±0.90

For the age at surgery, preoperative biometry, and postoperative refraction, the mean±standard deviation (SD) is shown in the table.
ACD, anterior chamber depth; AL, axial length; D, Diopter; K, keratometry; LT, lens thickness.

$f_1\left(ELP_{BC}, biometry\right) = true\ refraction$). In other words, when $ELP'_F = ELP_{BC}$, the refraction prediction errors of all patients equal zero. More details on the computation of ELP_{BC} can be found in online supplemental materials.

After the computation of ELP_F, ELP_{ML}, and ELP_{BC}, we modelled ELP_{BC} using a linear function of ELP_F and/or ELP_{ML} so as to obtain an approximation of the most optimal ELP using available variables. We compared four different approaches of approximating ELP_{BC} : (1) original, $ELP'_F = ELP_F$: using the original ELP_F, (2) Formula LR, $ELP'_F = c_1 \cdot ELP_F + c_3$: using linearly adjusted ELP_F, (3) ML LR, $ELP'_F = c_2 \cdot ELP_{ML} + c_3$: using linearly adjusted ELP_{ML}, (4) Formula & ML LR, $ELP'_F = c_1 \cdot ELP_F + c_2 \cdot ELP_{ML} + c_3$: using a linear combination of ELP_F and ELP_{ML}. Here, c_1, c_2, and c_3 are constants. Outliers with large refraction errors (ie, $error \geq mean\ error + 2 \cdot standard\ deviation$ or $error \leq mean\ error - 2 \cdot standard\ deviation$) were excluded for each formula before establishing the linear regression model, in order to obtain better modelling results. The refraction prediction errors were calculated as $error =\ predicted\ refraction - true\ refraction$. The linear regression was performed using scikit-learn 0.20.3.

On the testing set, ELP'_F was calculated based on the values of c_1, c_2, and c_3 obtained through linear regression. The predicted refraction was calculated as $refraction = f_1\left(ELP'_F, biometry\right)$. The mean absolute error (MAE), median absolute error (MedAE) and mean error (ME) were calculated for performance comparison.

A-constant optimisation

The A-constants for the formulas were optimised based on the training dataset so that the ME in refraction prediction was closest to zero. The A-constants were optimised separately for the unmodified formulas and formulas with a modified ELP estimate (see additional details in the A-constant optimisation section and online supplemental figure S1). The optimised A-constants for the original formulas were: a0=−0.733, a1=−0.234, a2=0.217 for Haigis, ACD constant=5.724 for Hoffer Q, surgeon factor=1.864 for Holladay, and A=119.089 for SRK/T (online supplemental table S1).

Statistical analysis

Linear regression analysis was used to assess the significance of the correlation between ELP_F, ELP_{ML}, and ELP_{BC}. To test

Table 2 The Pearson correlation coefficients (R) between ELP_F, ELP_{ML}, and ELP_{BC}

Index	Variable pairs	Haigis	Hoffer Q	Holladay1	SRK/T
1	ELP_F vs. ELP_{ML}	0.751	0.676	0.698	0.636
2	ELP_{BC} vs. ELP_F	0.621	0.730	0.622	0.633
3	ELP_{BC} vs. ELP_{ML}	0.532	0.544	0.534	0.524

The ELP_{BC} and ELP_F were calculated using the A constants optimised based on the original formulas. P values of all correlations were <0.05. All R were rounded to three decimal places.

whether the MAE and ME of different methods were significantly different, a Friedman test followed by a post hoc paired Wilcoxon signed-rank test with Bonferroni correction was used. Statistical significance was defined as the p value<0.05. All the above analyses were performed with Python V.3.7.3.

RESULTS

Dataset overview

The cases in the training and testing datasets had a similar distribution according to the summary statistics shown in table 1. As elaborated in he Materials and methods section, we calculated ELP_F, ELP_{ML} and ELP_{BC} based on the formulas and their optimised A-constants. The mean and SD of the ELPs calculated based on the original formulas were summarised in online supplemental table S2. ELP_{BC} and ELP_F had similar mean values in contrast to ELP_{ML}.

The Pearson correlation coefficients (R) between ELP_F, ELP_{ML}, and ELP_{BC} were shown in table 2. Three ELP-related variables were positively intercorrelated with each other. The correlation coefficients, R, between ELP_{BC} and ELP_{ML} were the weakest among the three pairs of variables across all formulas.

Linear regression results on the training set

Linear regression models were established based on the training set and the R^2 of alternative linear models were shown in table 3. The coefficients of the fitted linear regression line are shown in online supplemental table S3. The mean and SD of the ELP'_F resulting from different models are shown in online supplemental table S4. For 'Formula LR', the R^2 was larger than that of 'ML LR' for all four formulas. For 'Formula & ML LR', the R^2 was larger than that when one of ELP_F or ELP_{ML} was excluded from the linear combination for all four formulas.

Refraction prediction performance comparison on the testing set

We tested the performance of four scenarios on the testing set and summarised the MAE and SD in table 4. The ME and MedAE were shown in online supplemental tables S5 and S6. Statistical tests were used to compare the difference in the MAEs of different models (see the Materials and methods section).

Table 3 The R^2 of alternative least-squares linear regression models in the training set

Index	Methods	Haigis	Hoffer Q	Holladay1	SRK/T
1	Formula LR	0.377	0.541	0.579	0.394
2	ML LR	0.376	0.442	0.426	0.378
3	Formula & ML LR	**0.425**	**0.622**	**0.605**	**0.482**

The outlier cases were removed before calculating the above values. The largest R^2 among three methods is marked in bold for each formula. P-values of all correlations were < 0.05.

Table 4 Performance in the testing set

Index	Methods	Haigis	Hoffer Q	Holladay1	SRK/T
1	Original	0.373±0.328	0.408±0.337	0.384±0.341	0.394±0.351
2	Formula LR	0.373±0.328 (0.0%)	0.374±0.321 (8.3%)	0.388±0.342 (−1.1%)	0.391±0.345 (0.8%)
3	ML LR	0.391±0.346 (−4.8%)	0.454±0.375 (-21.4%)	0.434±0.364 (−13.0%)	0.397±0.344 (−1.5%)
4	Formula & ML LR	**0.356±0.329 (9.0%)**	**0.352±0.319 (22.5%)**	**0.371±0.336 (3.4%)**	**0.361±0.331 (9.1%)**

The MAE ±SD and the percentage reduction in MAE compared with 'Original' for alternative linear models in the testing set. All MAE and SD were rounded to three decimal places. The percentage reduction was calculated as $\frac{\text{MAE of a given method} - \text{MAE of original}}{\text{MAE of original}} \cdot 100\%$. All percentage reduction values were rounded to one decimal place. The method with the smallest MAE among four alternative methods is marked in bold for each formula.

Using a linear combination of ELP_F and ELP_{ML}, the refraction prediction results of four existing formulas were significantly improved compared with original ELP_F (statistical test results shown in online supplemental tables S7 and S8).

We further compared the MAEs of 'Original' and 'Formula & ML LR' among patients with short, medium and long AL (online supplemental table S9). It was observed that the short and medium AL groups had a higher percentage decrease in MAE than the long AL group for Hoffer Q and SRK/T. For Haigis, the medium AL group achieved higher decrease than the other two groups. And for Holladay, the long AL group achieved more decrease in MAE than the other two groups.

DISCUSSION

In this study, we applied a previously developed ML method for postoperative ACD prediction to an unseen dataset of 4806 cataract surgery patients to assess whether it was possible to improve the performance of existing IOL formulas (Haigis, Hoffer Q, Holladay, and SRK/T) by replacing each formula's ELP estimate.

We computed three ELP-related quantities: the ML-predicted postoperative ACD (ELP_{ML}), formula-predicted ELP (ELP_F), and a back-calculated ELP (ELP_{BC}) that minimised the refraction error for each eye in the dataset. They are strongly correlated with each other (table 2), which indicates that (1) ELP_F and ELP_{ML} are both predictive of the most optimal ELP ELP_{BC}, (2) ELP_F and ELP_{ML} contain partially overlapping information, which is consistent with our expectation. ELP_{ML} is an estimation of the value of the true postoperative ACD. On the other hand, ELP_F was designed by the originators of each formula to serve a similar purpose but was based on the theoretical assumptions in each formula. Our findings are consistent with observations of previous studies that the ELP estimates made by IOL formulas were numerically different from the true postoperative ACD.[9]

Using a training dataset of 3845 patients, we sought to evaluate whether the machine-predicted postoperative ACD, ELP_{ML}, was able to provide information that could be used to refine each formula's predicted ELP, ELP_F. We established regression models between the ELP_{ML}, ELP_F, and ELP_{BC} to evaluate whether a linear combination of ELP_{ML} and ELP_F used in place of the original ELP_F could lower the refraction prediction error. Using the modified ELPs, we obtained significantly lower MAEs in refraction prediction compared with the formulas with the original ELPs on the unseen testing set (table 4). Notably, the accurately predicted postoperative ACD (ELP_{ML}) alone did not outperform the original ELP (ELP_F) when it was inserted into the formulas (table 4, row 3 compared with row 1). This is likely because the original method of calculating ELP in each formula compensates for its particular model of the eye and its associated assumptions. Our ELP_{ML}, however, does not have any components that compensate for the assumptions and constants in the formulas. On the other hand, ELP_{ML} has information about the true postoperative ACD, which it appears can beneficially alter the original ELP estimate.

In this study, the A-constants were optimised separately when ELP_F was replaced with different ELP_F'. The means of ELP_F', as shown in online supplemental table S4, were numerically close to those of ELP_F as shown in online supplemental table S2. However, in our method, the similarity between ELP_F' and ELP_F was not among the restrictions and goals of the optimisation. The reason that ELP_F' and the original ELP_F had similar means might be that the other parts of each formula put restrictions on the values of ELP in order to obtain reasonable results. This could also be the reason why ELP_{BC} and ELP_F had similar means as shown in online supplemental table S2.

Previous studies involving replacement of ELP in existing formulas have focused on special cases, such as sulcus implantation and postrefractive surgery eyes, where ELP estimates of traditional formulas would be expected to be inapplicable.[10,11] However, the method for replacing ELP estimates presented here provides a simple way of improving the refraction prediction performance of existing formulas for the general cataract surgery population. While it would be ideal to evaluate this method on modern formulas such as Barrett Universal II or Holladay 2, the absence of published equations for these formulas prevents such a study. As such, we studied the application of the ML predicted postoperative ACD in four existing formulas whose mathematical equations were published. Although it awaits to be further validated, similar results can likely be transferred to other refraction prediction methods, since many modern IOL power formulas use predicted postoperative ACD as an intermediate step for predicting postoperative refraction. A limitation of the study was the absence of an external validation set, despite the use of a large unseen testing dataset (961 eyes). Accordingly, evaluation of the method at additional institutions and the extension to additional formulas will be future directions of this work.

In summary, the results of this study demonstrate that an ML method for postoperative ACD prediction based on postoperative optical biometry can be incorporated into a variety of existing IOL power formulas to improve their accuracy in refraction prediction.

Contributors TL: data analysis, programming and writing of the manuscript. JS: data collection. NN: data collection, guidance on method development, and writing of the manuscript.

Funding This work was supported by the Lighthouse Guild, New York, NY (JDS) and National Eye Institute, Bethesda, MD, 1R01EY026641-01A1 (JDS).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Institutional review board approval was obtained for the study, and it was determined that informed consent was not required because of its retrospective nature and the anonymised data used in this study. The study was carried out in accordance with the tenets of the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from the Sight Outcomes Research Collaborative (SOURCE) repository for participating institutions and are not publicly available.

ORCID iD
Nambi Nallasamy http://orcid.org/0000-0001-7501-7198

REFERENCES

1 Norrby S. Sources of error in intraocular lens power calculation. *J Cataract Refract Surg* 2008;34:368–76.
2 Olsen T. Sources of error in intraocular lens power calculation. *J Cataract Refract Surg* 1992;18:125–9.
3 Holladay JT. Refractive power calculations for intraocular lenses in the phakic eye. *Am J Ophthalmol* 1993;116:63–6.
4 Norrby S, Bergman R, Hirnschall N, *et al*. Prediction of the true IOL position. *Br J Ophthalmol* 2017;101:1440–6.
5 Martinez-Enriquez E, Pérez-Merino P, Durán-Poveda S, *et al*. Estimation of intraocular lens position from full crystalline lens geometry: towards a new generation of intraocular lens power calculation formulas. *Sci Rep* 2018;8:9829.
6 Hirnschall N, Amir-Asgari S, Maedel S, *et al*. Predicting the postoperative intraocular lens position using continuous intraoperative optical coherence tomography measurements. *Invest Ophthalmol Vis Sci* 2013;54:5196–203.
7 Chang Y-C, Cabot F, Williams S. Pre-Operative prediction of post-cataract surgery IOL position using anterior chamber depth and lens thickness determined with Extended-depth OCT. *Invest Ophthalmol Vis Sci* 2017;58:2717.
8 Satou T, Shimizu K, Tsunehiro S, *et al*. Development of a new intraocular lens power calculation method based on lens position estimated with optical coherence tomography. *Sci Rep* 2020;10:6501.
9 Tamaoki A, Kojima T, Tanaka Y, *et al*. Prediction of effective lens position using multiobjective evolutionary algorithm. *Transl Vis Sci Technol* 2019;8:64.
10 Eom Y, Song JS, Kim HM. Modified Haigis formula effective lens position equation for ciliary Sulcus-Implanted intraocular lenses. *Am J Ophthalmol* 2016;161:142–9.
11 Kim DH, Kim MK, Wee WR. Estimation of intraocular lens power calculation after myopic corneal refractive surgery: using corneal height in anterior segment optical coherence tomography. *Korean J Ophthalmol* 2015;29:195–202.
12 Barrett GD. Intraocular lens calculation formulas for new intraocular lens implants. *J Cataract Refract Surg* 1987;13:389–96.
13 Barrett GD. An improved universal theoretical formula for intraocular lens power prediction. *J Cataract Refract Surg* 1993;19:713–20.
14 Olsen T. The Olsen formula. In: *Intraocular lens power calculations*, 2004: 27–40.
15 Li T, Yang K, Stein JD, *et al*. Gradient boosting decision tree algorithm for the prediction of postoperative intraocular lens position in cataract surgery. *Transl Vis Sci Technol* 2020;9:38.
16 Simpson MJ, Charman WN. The effect of testing distance on intraocular lens power calculation. *J Refract Surg* 2014;30:726.
17 Retzlaff JA, Sanders DR, Kraff MC. Development of the SRK/T intraocular lens implant power calculation formula. *J Cataract Refract Surg* 1990;16:333–40.
18 Hoffer KJ. The Hoffer Q formula: a comparison of theoretic and regression formulas. *J Cataract Refract Surg* 1993;19:700–12.
19 Haigis W, Lege B, Miller N, *et al*. Comparison of immersion ultrasound biometry and partial coherence interferometry for intraocular lens calculation according to Haigis. *Graefes Arch Clin Exp Ophthalmol* 2000;238:765–73.
20 Holladay JT, Musgrove KH, Prager TC. A three-part system for refining intraocular lens power calculations. J Cataract Refract Surg 1988;14:17–24..
21 Correction. *J Cataract Refract Surg* 1994;20:677.
22 Erratum. *J Cataract Refract Surg* 1990;16:528.
23 Hoffer KJ. Reply: Errata in printed Hoffer Q formula. *J Cataract Refract Surg* 2007;33:2–3.
24 Zuberbuhler B, Morrell AJ. Errata in printed Hoffer Q formula. *J Cataract Refract Surg* 2007;33:2.

Formula	Constant	Optimized Constants	Mean Error
Haigis	a0 (a1 = 0.234, a2 = 0.217)	-0.736	0.000319
Hoffer Q	ACD	5.726	0.000157
Holladay1	Surgeon factor	1.867	0.000121
SRK/T	A constant	119.091	-0.000204

Table S1 The optimized A constants and the corresponding mean error for the original formulas in the training dataset. The A constants were optimized so that the absolute value of the mean error was minimized. The mean errors were calculated after excluding the outliers (see main text). The mean errors in the training set were rounded to three significant figures.

Dataset	Variables	Holladay1	SRK/T	Hoffer Q	Haigis
Training dataset	ELP_F	4.00 ± 0.41	5.82 ± 0.50	5.86 ± 0.38	5.28 ± 0.36
	ELP_{BC}	4.08 ± 0.92	5.85 ± 0.83	5.97 ± 1.00	5.32 ± 0.76
	ELP_{ML}	4.67 ± 0.27			
Testing dataset	ELP_F	4.00 ± 0.39	5.83 ± 0.47	5.87 ± 0.38	5.28 ± 0.36
	ELP_{BC}	4.06 ± 0.72	5.84 ± 0.65	5.95 ± 0.80	5.31 ± 0.61
	ELP_{ML}	4.68 ± 0.26			

Table S2 The mean ± standard deviation (SD) for ELP_F, ELP_{ML}, and ELP_{BC} in the training and testing dataset when $ELP_F' = ELP_F$. The ELP_{BC} and ELP_F were calculated using the corresponding formula with the optimized A-constants therefore their values vary with different formulas. The values of ELP_{ML} only depend on the values of the preoperative biometry. The outliers were not removed when the above summary statistics were calculated. All values were rounded to two decimal places.

Methods	Holladay1	SRK/T	Hoffer Q	Haigis
Formula LR	c_1 = 1.27 c_3 = -1.01	c_1 = 0.76 c_3 = -1.67	c_1 = 1.44 c_3 = -26.65	c_1 = 1.00 c_3 = -5.88
ML LR	c_2 = 1.65 c_3 = -3.80	c_2 = 1.31 c_3 = -0.33	c_2 = 1.62 c_3 = -1.70	c_2 = 1.24 c_3 = -0.53
Formula & ML LR	c_1 = 0.98 c_2 = 0.61 c_3 = -2.77	c_1 = 0.47 c_2 = 0.79 c_3 = -2.55	c_1 = 1.09 c_2 = 0.65 c_3 = -6.90	c_1 = 0.58 c_2 = 0.68 c_3 = -4.44

Table S3 The coefficients (c_1 and c_2) and the intercept c_3 for the linear regression model established based on the training dataset. All values were rounded to two decimal places.

Dataset	Method	Holladay1	SRK/T	Hoffer Q	Haigis
Training dataset	Formula LR	4.06 ± 0.52	5.80 ± 0.38	5.90 ± 0.55	5.28 ± 0.36
	ML LR	3.92 ± 0.44	5.78 ± 0.35	5.86 ± 0.43	5.26 ± 0.33
	Formula & ML LR	3.98 ± 0.53	5.80 ± 0.40	5.89 ± 0.55	5.27 ± 0.37
Testing dataset	Formula LR	4.04 ± 0.46	5.80 ± 0.31	5.91 ± 0.53	5.28 ± 0.34
	ML LR	3.90 ± 0.39	5.79 ± 0.32	5.87 ± 0.39	5.27 ± 0.30
	Formula & ML LR	3.97 ± 0.46	5.79 ± 0.34	5.89 ± 0.49	5.28 ± 0.34

Table S4 The mean ± standard deviation (SD) for ELP_F' in the training and testing dataset. The ELP_{BC} and ELP_F were calculated using the corresponding formula with the optimized A-constants therefore their values vary with different formulas. The values of ELP_{ML} only depend on the values of the preoperative biometry. The outliers were not removed when the above summary statistics were calculated.

Methods	Holladay1	SRK/T	Hoffer Q	Haigis
Original	-0.020 ± 0.513	-0.008 ± 0.528	-0.020 ± 0.529	-0.025 ± 0.496
Formula LR	0.008 ± 0.517	-0.003 ± 0.522	-0.017 ± 0.492	-0.018 ± 0.496
ML LR	0.057 ± 0.563	0.001 ± 0.525	0.007 ± 0.589	-0.001 ± 0.523
Formula & ML LR	0.009 ± 0.500	-0.008 ± 0.490	-0.016 ± 0.475	-0.014 ± 0.484

Table S5 The mean error (ME) ± standard deviation (SD) of alternative linear models in the testing set. All values were rounded to three decimal places.

Methods	Holladay1	SRK/T	Hoffer Q	Haigis
Original	0.299	0.307	0.330	0.283
Formula LR	0.305	0.310	0.293	0.283
ML LR	0.351	0.308	0.366	0.304
Formula & ML LR	0.290	0.273	0.268	0.263

Table S6 The median absolute error (MedAE) of alternative linear models in the testing set. All values were rounded to three decimal places.

Statistic	Holladay1	SRK/T	Hoffer Q	Haigis
Friedman chi-square test statistic	37.29	39.13	117.42	37.25
p-value	4.00e-08	1.63e-08	2.78e-25	4.07e-08

Table S7 The Friedman test statistic and the p-values for comparing the testing set results of different methods. All Friedman statistics were rounded to two decimal places. All p-values were rounded to three significant figures.

Formula	Methods	ML LR	Formula LR	Formula & ML LR
Haigis	Formula LR	**1.7E-01**	**/**	**/**
	Formula & ML LR	1.8E-11	2.6E-03	/
	Original	**1.7E-01**	**1.0E+00**	2.7E-03
Hoffer Q	Formula LR	1.5E-10	/	/
	Formula & ML LR	1.7E-26	3.6E-05	/
	Original	4.1E-05	1.5E-10	5.1E-17
Holladay 1	Formula LR	1.5E-04	/	/
	Formula & ML LR	4.4E-12	3.0E-04	/
	Original	1.4E-05	**1.0E+00**	9.9E-03
SRK/T	Formula LR	**1.0E+00**	/	/
	Formula & ML LR	1.7E-12	7.0E-06	/
	Original	**1.0E+00**	**1.0E+00**	1.1E-05

Table S8 The post hoc test results of four existing formulas for comparing the testing set performance of different methods. The insignificant p-values (p≥0.05) were highlighted in bold.

Method	Formulas	**Short AL** ($AL < 22mm$) **n=28**	**Medium AL** ($22mm \leq AL \leq 26mm$) **n=832**	**Long AL** ($AL > 26mm$) **n=100**
Original	Haigis	0.321 ± 0.234	0.373 ± 0.332	0.383 ± 0.315
	Hoffer Q	0.524 ± 0.295	0.396 ± 0.335	0.480 ± 0.350
	Holladay1	0.397 ± 0.224	0.364 ± 0.322	0.541 ± 0.464
	SRK/T	0.438 ± 0.236	0.386 ± 0.337	0.452 ± 0.465
Formula & ML LR	Haigis	0.330 ± 0.285 (-2.8%)	0.353 ± 0.331 (5.5%)	0.394 ± 0.319 (-2.8%)
	Hoffer Q	0.338 ± 0.264 (35.5%)	0.344 ± 0.318 (13.1%)	0.420 ± 0.338 (12.5%)
	Holladay1	0.392 ± 0.257 (1.3%)	0.356 ± 0.320 (2.2%)	0.486 ± 0.445 (10.3%)
	SRK/T	0.375 ± 0.284 (14.3%)	0.351 ± 0.324 (9.0%)	0.438 ± 0.391 (3.2%)

Table S9 The mean absolute error (MAE) ± standard deviation in the testing set for patients with short, medium, and long axial length (AL). All MAE and SD were rounded to three decimal places. For "Formula & ML LR", the percentage reduction in MAE compared to "Original" were shown. The percentage reduction was calculated as

$\frac{MAE\ of\ Original - MAE\ of\ Formula\ \&\ ML\ LR}{MAE\ of\ Original} \cdot 100\%$ and was rounded to one decimal place. The letter "n" is the number of cases in each AL group.

Calculation of ELP_{BC}

As described in the main text, the postoperative refraction was predicted using a function of ELP_F and preoperative biometry: $predicted\ refraction = f_1(ELP_F, biometry)$. Here we define ELP_{BC} as follows: when $ELP_F = ELP_{BC}$, $f_1(ELP_{BC}, biometry) - true\ refraction = 0$ holds for all cases. In other words, when the ELP estimation equals ELP_{BC}, the refraction prediction error equals zero for all cases. Based on the above definition, the value of ELP_{BC} can be found by solving for the x in the equation $f_1(x, biometry) - true\ refraction = 0$, where $biometry$and $true\ refraction$ are known. For a given case, there were always no more than two roots for the above function because of the quadratic nature of the formulas. When there were two roots, the smaller root was taken as ELP_{BC} because of two main reasons: (1) the greater root was usually >50, which was not within a physiologically meaningful range for ELP; (2) practically when the larger roots were used as ELP_{BC}, the R^2 in the training set was significantly lower than that obtained with the smaller root (data are not shown). The function $f_1(x, biometry) - true\ refraction = 0$ was solved programmatically using scipy.optimize.fsolve (scipy 1.2.1) in Python 3.7.3.

A-Constant Optimization

When ELP_F was not replaced with a modified value ELP'_F (Figure S1, upper part), the A-constants of the formulas were optimized in the standard way: first, compute the mean refraction prediction error when the A-constant takes different values, then, the A-constant that gives the smallest absolute mean error is the most optimal A-constant.

When $ELP_F = ELP'_F$ (Figure S1, lower part), the A-constants were optimized based on the same concept, the value of ELP'_F changes with the values of the A-constant. The pseudo-code for the A-constant optimization process is shown below. The value of ELP_{ML} does not change with the A-constant. The value of ELP_F and ELP_{BC} changes with the A-constants.

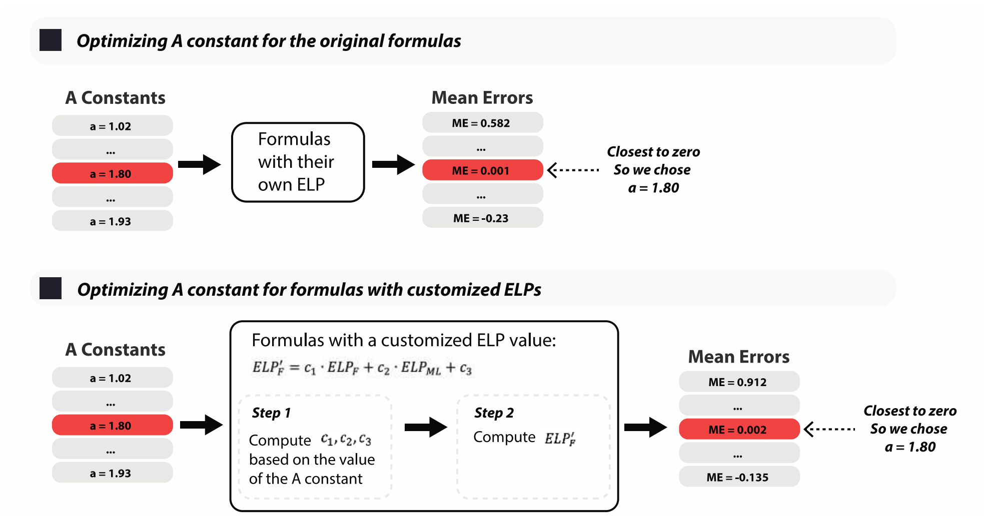

Figure S1 The pipeline of the A-constant optimization procedure. The numbers in the figure are not real data.

Algorithm 1: A-constant optimization when $ELP_F = ELP'_F$

ELP_{ML} ← compute ELP_{ML} using the machine learning model
FOR a **IN** A-constant search space

    ELP_F ← compute ELP_F based on the formula with a as the A constant
    ELP_{BC} ← compute ELP_{BC} based on the formula with a as the A constant

    coefficients c_1, c_2, and c_3 ← model ELP_{BC} as a linear function of ELP_{ML} and/or ELP_F.
    $ELP'_F \leftarrow c_1 \cdot ELP_F + c_2 \cdot ELP_{ML} + c_3$
    predicted refraction ← compute the predicted refraction based on a and ELP'_F
    mean error ←compute the mean error based on the predicted refraction and the true refraction

END FOR

The most optimal A-constant← the A-constant that gives the smallest absolute mean error

Li T, *et al. Br J Ophthalmol* 2021;0:1–5. doi: 10.1136/bjophthalmol-2020-318321