Ray tracing intraocular lens calculation performance improved by AI-powered postoperative lens position prediction

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

[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]Center for Eye Policy and Innovation, University of Michigan, Ann Arbor, Michigan, USA
[4]Department of Health Management and Policy, University of Michigan School of Public Health, Ann Arbor, Michigan, USA

Correspondence to
Dr Nambi Nallasamy, University of Michigan Department of Ophthalmology and Visual Sciences, Ann Arbor, Michigan, USA; nnallasa@med.umich.edu

Received 18 August 2021
Accepted 31 October 2021

To cite: Li T, Reddy A, Stein JD, *et al*. *Br J Ophthalmol* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/bjophthalmol-2021-320283

ABSTRACT

Aims To assess whether incorporating a machine learning (ML) method for accurate prediction of postoperative anterior chamber depth (ACD) improves cataract surgery refraction prediction performance of a commonly used ray tracing power calculation suite (OKULIX).

Methods and analysis A dataset of 4357 eyes of 4357 patients with cataract was gathered at the Kellogg Eye Center, University of Michigan. A previously developed machine learning (ML)–based method was used to predict the postoperative ACD based on preoperative biometry measured with the Lenstar LS900 optical biometer. Refraction predictions were computed with standard OKULIX postoperative ACD predictions and ML-based predictions of postoperative ACD. The performance of the ray tracing approach with and without ML-based ACD prediction was evaluated using mean absolute error (MAE) and median absolute error (MedAE) in refraction prediction as metrics.

Results Replacing the standard OKULIX postoperative ACD with the ML-predicted ACD resulted in statistically significant reductions in both MAE (1.7% after zeroing mean error) and MedAE (2.1% after zeroing mean error). ML-predicted ACD substantially improved performance in eyes with short and long axial lengths ($p<0.01$).

Conclusions Using an ML-powered postoperative ACD prediction method improves the prediction accuracy of the OKULIX ray tracing suite by a clinically small but statistically significant amount, with the greatest effect seen in long eyes.

INTRODUCTION

Postoperative intraocular lens (IOL) position estimation is essential to IOL power calculations for cataract surgery. Inaccuracy in prediction of the postoperative anterior chamber depth (ACD) has been reported to be the primary remaining source of error in IOL power calculations.[1 2] Methods for predicting postoperative ACD have evolved over the past several decades. First-generation lens calculation formulas represented postoperative ACD by a constant. As more biometric variables have become available, additional preoperative measurements such as the axial length and corneal power have been added to methods for estimating the postoperative IOL position.

Most modern IOL calculation formulas involve computation of postoperative refraction using Gaussian optics, which relies on the assumption that incoming rays are paraxial, in addition to empirically determined adjustment factors. The primary empirical adjustments for these modern formulas (such as Barrett Universal II, Holladay 2 and SRK/T) are made through the use of effective lens position (ELP) as an intermediate quantity to indicate the location of the lens as it relates to a given optical model of the eye.[3] First introduced by Holladay in 1993,[3] ELP was initially intended to estimate the position of the IOL in the postoperative eye. In practice, however, the postoperative ACD and the optimal location of the principal plane of the IOL in a given formula's optical model of the eye are not numerically equal.[1 4] The optimal ELP in a given eye can be back-calculated with knowledge of the eye's postoperative refraction. Achieving an accurate prediction of the optimal ELP for all patients is a more challenging task, however, and represents an ongoing limitation for modern formulas.

Numerical ray tracing represents an alternative to Gaussian optics for the purpose of IOL power calculation. Ray tracing involves the direct calculation of refraction of rays of light at each medium change within the eye using Snell's law. Studies have demonstrated that ray tracing performance is comparable with that of state-of-the-art IOL calculation formulas in normal eyes and may provide improved IOL calculation accuracy in certain populations.[5 6]

The more data available regarding the index of refraction of each medium, the curvature of each surface (including the anterior and posterior surfaces of the cornea and the intraocular lens implant) and position of each of these refracting surfaces relative to one another, the more accurate a ray tracing calculation is. As such, ray tracing calculations are likely to benefit from improved methods for predicting the actual anatomical postoperative IOL position.

In previous work, our group has demonstrated that (1) it is possible to improve on estimates of the true anatomical postoperative IOL position through the use of a gradient-boosting machine learning (ML) approach, and (2) incorporation of this ML-predicted postoperative IOL position can be used to refine ELP estimates in existing IOL formulas and improve their accuracy.[7 8]

We investigate here whether our previously described ML method for prediction of postoperative IOL position is able to improve the accuracy of the commonly used OKULIX ray tracing suite for intraocular lens power calculation. The standard approach to postoperative IOL position prediction in the OKULIX suite employs a regression model based on axial length and thickness of the crystalline lens. In the work presented here, we sought to determine whether, holding all else equal, replacement of the standard OKULIX postoperative IOL position prediction method with that of a highly accurate ML-based predictor had the potential to improve the accuracy of ray tracing calculations.

MATERIALS AND METHODS

Data collection

Preoperative and postoperative biometry records of patients with cataract were exported from the Lenstar LS900 optical biometers (Haag-Streit USA Inc, EyeSuite software V.i9.1.0.0) at the University of Michigan's Kellogg Eye Center. Patients who had cataract surgery but no prior corneal surgery and no additional surgical procedures at the time of cataract surgery were included. Only surgery cases involving the implantation of Alcon SN60WF single-piece acrylic monofocal lenses (Alcon Inc., USA) were included in the study because it is the most commonly implanted lens at the Kellogg Eye Center. Cases that were used to train the postoperative ACD prediction ML model were excluded from the dataset so that the dataset involved only unseen samples. One eye was selected at random for each patient who had undergone surgery in both eyes, so that all cases in the final dataset were independent of each other. The preoperative information gathered included the measurements of the axial length (AL), crystalline lens thickness (LT), anterior chamber depth (ACD), the radius of curvature in the flat meridian (R1), the radius of curvature in the steep meridian (R2), patient gender and selected IOL power. As defined by the Lenstar LS900 optical biometer, the postoperative ACD represents the distance from the anterior surface of the cornea to the anterior surface of the IOL. The postoperative refraction (including the spherical component (SC) and cylindrical component (CC)) records were obtained. The spherical equivalent (SE) refraction was calculated as $SE\ refraction = (SC - 0.1614) + 0.5CC$. The constant 0.1614 was used to account for the length (10 feet, 3.048 m) of the examination lane according to Simpson and Charman's recommendation.[9] For each patient, the postoperative record that was generated closest to 1 month (ie, 30 days) after surgery was included. Details about the collection and processing of the dataset can be found in our previous publications.[8 9]

Performance comparison between OKULIX and ML-based approach

The dataset in total consisted of the aforementioned preoperative and postoperative data for 4357 eyes of 4357 patients (figure 1). Our postoperative ACD prediction ML model (referred to as the 'Base' model in our prior work) was used to compute predictions of postoperative ACD (in mm) for each eye in the dataset based on the preoperative data (AL, CCT, preoperative ACD, LT, R1, R2 and WTW) and patient gender. Details about the ML method can be found in our previous publication.[10] The OKULIX standalone PC software suite (OKULIX V.9.20; Panopsis GmbH, Mainz, Germany) was used to compute refraction predictions based on the available preoperative biometry (AL, LT, R1, R2 and preoperative ACD) and laterality of the case. In addition to the postoperative refraction, OKULIX also predicts the postoperative aqueous depth (AD) as an intermediate value which in downstream pipelines serves as one of the input variables for the prediction of the refraction. By design, the OKULIX software allows the users to replace its predicted postoperative AD with

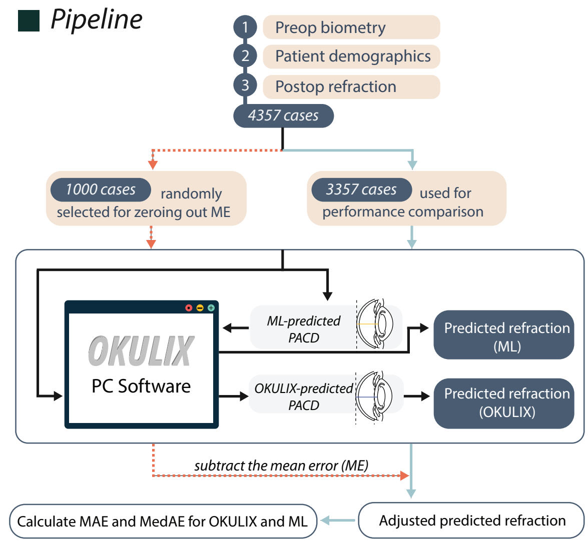

Figure 1 Analysis pipeline of the presented study. MAE, mean absolute error; MedAE, median absolute error; PACD, postoperative anterior chamber depth. The 'OKULIX-predicted PACD' was not a direct output from OKULIX, but calculated as $PACD = PAD + CCT$, where PAD is the OKULIX-predicted postoperative aqueous depth (AD). The eye illustration was created by TL.

a custom value. We therefore compared the refraction prediction accuracy of OKULIX when different postoperative anterior chamber depths (PACDs) were used: (1) its standard internal prediction of postoperative ACD and (2) the postoperative ACD prediction from our ML model. OKULIX by default outputs a predicted AD defined as the distance from the posterior surface of the cornea to the anterior surface of the IOL, instead of a predicted PACD (defined as distance from the front surface of the cornea to the anterior surface of the IOL). In order to transform the ML-predicted postoperative ACD to a predicted postoperative AD, the preoperative central corneal thickness (CCT) was subtracted from the ML-predicted postoperative ACD. Similarly, to compare the OKULIX and ML-predicted PACD, we transformed the OKULIX-predicted postoperative AD to postoperative ACD prediction by adding the preoperative CCT. The equations used to transform predicted postoperative ACD to AD and AD to ACD are shown as follows, where the OKULIX-predicted postoperative AD is the direct output from OKULIX and the ML-predicted postoperative ACD is the direct output from the ML method.

$$OKULIX\ predicted\ postop.\ ACD = OKULIX\ predicted\ postop.\ AD + preop.\ CCT$$

$$ML\ predicted\ postop.\ AD = ML\ predicted\ postop.\ ACD - preop.\ CCT$$

The dataset was split into an optimisation dataset and a performance comparison dataset, where the former was used for zeroing out the mean error and the latter was used for prediction performance comparison. A total of 1000 cases were randomly set aside as the optimisation dataset. Increasing the size of the optimisation dataset did not significantly change the results (results not shown). The remaining 3357 cases were used to assess performance of the ray tracing calculations with and without the ML-predicted PACD. Mean absolute error (MAE), median absolute error (MedAE) and mean error (ME) of refraction predictions were computed for performance comparison. The prediction error was defined as follows. A negative error corresponds to a more myopic prediction, and a positive error corresponds to a more hyperopic prediction.

$$prediction\ error = predicted\ refraction - actual\ refraction$$

We further bootstrapped the performance comparison dataset to obtain the confidence intervals for the MAE, MedAE, and differences in MAE and MedAE. Specifically, a random sampling of 3357 cases with replacement was applied 10 000 times to the performance comparison dataset, and four metrics were calculated for those 10 000 bootstrap samples as follows. Let p_i be the absolute error of the OKULIX-based approach for the i-th case and q_i be the absolute error of the ML-based approach for the i-th case in the bootstrap dataset. For each bootstrap dataset, we calculated (1) the mean absolute error as $mean\left\{p_1, p_2, \ldots, p_{3357}\right\}$ and $mean\left\{q_1, q_2, \ldots, q_{3357}\right\}$, (2) the median absolute error as $median\left\{p_1, p_2, \ldots, p_{3357}\right\}$ and $median\left\{q_1, q_2, \ldots, q_{3357}\right\}$, (3) the mean of the differences between the absolute error as $mean\left\{p_1 - q_1, p_2 - q_2, \ldots, p_{3357} - q_{3357}\right\}$ and (4) the median of the differences between the absolute error as $median\left\{p_1 - q_1, p_2 - q_2, \ldots, p_{3357} - q_{3357}\right\}$.

Zeroing of mean error

To account for possible systematic differences in our patient population, we zeroed out the mean errors of the OKULIX predictions individually for the OKULIX and ML approaches. This was done by computing the mean error on the optimisation dataset and subtracting it from the predictions on the performance comparison dataset (figure 1). The aforementioned scoring metrics (eg, MAE) were computed after the zeroing of mean error.[11]

Statistical analysis

A paired Wilcoxon test was performed to evaluate the significance of the difference between the OKULIX and ML-predicted PACD. The same test (Wilcoxon test) was performed to test whether the prediction errors of the OKULIX and ML-based approach were significantly different. Statistical significance was defined as the p value <0.05. All the aforementioned analyses were performed with Python V.3.7.3.

RESULTS

Dataset overview

The dataset in total consisted of 4357 cataract surgery cases of 4357 patients. Among those patients, 1919 (44.04%) were males and 2438 (55.96%) were females. A summary of the dataset is shown in table 1. The OKULIX-predicted postoperative ACD had a mean of 5.13 mm and was significantly longer than the ML-predicted postoperative ACD (mean=4.68 mm) (Wilcoxon-test p value<0.01).

Table 1 Baseline characteristics of the dataset

Characteristic	Mean±SD	Median	Range Min	Range Max
Age at surgery (years)	70.66±9.53	71.24	13.23	89.45
Preoperative K (D)	43.85±1.61	43.84	33.42	50.39
Preoperative AL (mm)	24.21±1.38	24.01	20.44	31.22
Preoperative LT (mm)	4.52±0.45	4.51	2.50	5.99
Preoperative ACD (mm)	3.26±0.41	3.27	2.08	5.15
OKULIX-predicted postoperative ACD (mm)	5.13±0.29	5.12	4.25	6.24
ML-predicted postoperative ACD (mm)	4.68±0.26	4.69	3.85	5.59
Postoperative refraction (D)	−0.56±0.96	−0.41	−12.16	3.34

The highly myopic postoperative refraction value ('−12.16' in the last row in the table) was from a patient with high myopia and unilateral cataract who chose a highly myopic target refraction to avoid anisometropia.
ACD, anterior chamber depth; AL, axial length; D, dioptre; LT, lens thickness; ML, machine learning.

Refraction prediction performance comparison

The OKULIX and ML-based approaches were compared using the performance comparison dataset, and the results are shown in table 2. The refraction prediction errors in the optimisation dataset were significantly different for the OKULIX and ML-based approaches (table 2, column 1) (Wilcoxon-test p value <0.01). The unadjusted predictions in the optimisation dataset from OKULIX tended to be more myopic compared with the true refraction (ME=−0.329 D), while the unadjusted predictions from the ML-based method tended to be more hyperopic (ME=0.211 D) (although to a lesser extent). After subtracting the mean errors from the optimisation dataset, the mean errors of the OKULIX and ML-based approaches became much closer to zero in the performance comparison dataset (table 2, column 2). The mean absolute errors (MAEs) in refraction prediction were significantly lower for the ML-based approach compared with the OKULIX-based approach (table 2, column 3) (Wilcoxon-test p value <0.01).

The bootstrap distributions of MAE, MedAE and the mean/median difference in absolute errors are depicted in figure 2. According to the bootstrap results, the 95% CI for MAEs was [0.3459, 0.3677] for OKULIX and [0.3404, 0.3615] for ML. The 95% CI for MedAE was [0.2679, 0.2901] for OKULIX and [0.2638, 0.2858] for ML. The 95% CI for the mean of the differences in absolute errors was [0.0017, 0.0101]. The 95% CI for the median of differences in absolute errors was [0.0003, 0.0083].

Refraction prediction performance comparison in different axial length groups

We further summarised the performance of the OKULIX and ML-based approaches in different axial length (AL) groups in table 3. There were significant differences in the short and long AL groups and insignificant difference in the medium length group.

Table 2 Performance comparison set results

	Unadjusted ME in optimisation dataset ±SD	ME in performance comparison dataset ±SD	MAE in performance comparison dataset ±SD	MedAE in performance comparison dataset
OKULIX	−0.329±0.471	0.018±0.478	0.357±0.318	0.280
ML	0.211±0.470	0.019±0.470	0.351±0.313	0.274
% improvement	/	/	1.7%	2.1%

The first column shows the ME ±SD in the optimisation dataset, where the ME was calculated from the unadjusted predictions from OKULIX and ML. The second column shows the ME ±SD in the performance comparison dataset, after subtracting the corresponding ME in the optimisation dataset from the original prediction from OKULIX and ML. The per cent improvement was calculated as $\frac{OKULIX\ performance - ML\ performance}{OKULIX\ performance}$.
MAE, mean absolute error; ME, mean error; MedAE, median absolute error; ML, machine learning.

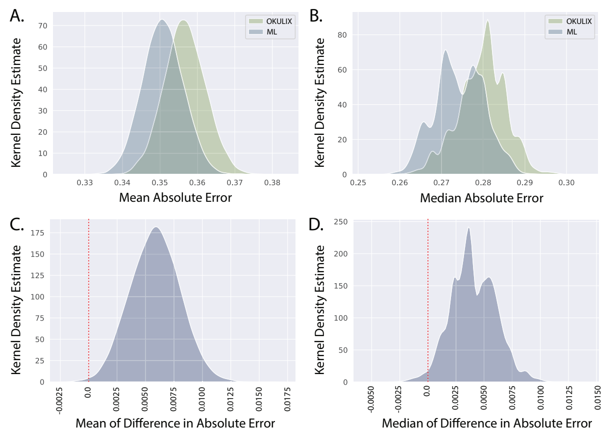

Figure 2 Bootstrap results in the performance comparison dataset. (A) The distribution of the mean absolute error. (B) The distribution of the median absolute error. (C) The distribution of the means of the difference in the absolute errors of OKULIX and the machine learning (ML) method. (D) The distribution of the medians of the difference in the absolute errors of OKULIX and the ML method. The red dashed line marks the location of 0.0.

DISCUSSION

In this study, we sought to determine whether using a ML-based approach to prediction of postoperative IOL position could improve the performance of a ray tracing approach to IOL power calculations.

Ray tracing offers several advantages over traditional geometrical optics approaches to IOL power calculation. In particular, through the calculation of refraction of light rays at each refracting surface in the eye, factors such as irregular corneal curvature and varying pupil sizes can be accounted for with a ray tracing approach. The incorporation of more detailed surface information, including corneal tomography with Scheimpflug imaging or optical coherence tomography, offers the potential to maintain accuracy for a broader range of eyes than traditional methods. Since ray tracing accuracy depends primarily on the accuracy of anatomical measurements, rather than on theoretical quantities such as ELP, ray tracing may have greater long-term potential in comparison with traditional IOL calculation methods as techniques for measuring the size, shape and anatomical location of components of the eye's optical system continue to evolve.

Table 3 Results comparison in different axial length groups

AL group	Number of cases (%)	Wilcoxon-test p value	Method	±SD	MedAE
Short (AL≤23 mm)	589 (17.5)	<0.01	OKULIX	0.394±0.333	0.313
			ML	0.373±0.328	0.290
			% improvement	5.3%	7.3%
Medium (23 mm<AL≤26 mm)	2429 (72.4)	0.22	OKULIX	0.340±0.306	0.265
			ML	0.343±0.307	0.270
			% improvement	−0.9%	−0.5%
Long (AL >26 mm)	339 (10.1)	<0.01	OKULIX	0.409±0.361	0.322
			ML	0.365±0.328	0.271
			% improvement	10.8%	15.8%

This table shows the performance of OKULIX and ML-based approaches for patients in the short, medium and long axial length group. The per cent improvement was calculated as $\frac{\text{OKULIX performance} - \text{ML performance}}{\text{OKULIX performance}}$. The Wilcoxon test was performed to compare the difference in the absolute errors between OKULIX and ML-based approaches in three different axial length groups.
AL, axial length; MAE, mean absolute error; MedAE, median absolute error; ML, machine learning.

The physical location of the IOL (postoperative ACD) is one component of the post-cataract surgery optical system that is not directly measurable preoperatively. However, our group has previously developed and validated a ML method for accurate postoperative ACD prediction using preoperative optical biometry, gender and intended IOL power.

In previous work, our group demonstrated that incorporating this ML method for postoperative ACD prediction into the ELP calculations of traditional IOL calculation formulas significantly improved the refraction prediction performance of these IOL formulas.[8] Since ray tracing methods in theory require only the true postoperative ACD, rather than an ELP, it would logically follow that utilisation of our validated ML method for postoperative ACD prediction could improve the accuracy of ray tracing IOL power calculations. In the study presented here, we evaluated whether replacing the standard model for postoperative ACD in the commonly used OKULIX ray tracing suite with our ML-based postoperative ACD prediction could improve the refraction prediction performance of the OKULIX suite.

We found that replacing the standard OKULIX postoperative ACD model (based on a linear regression of axial length and lens thickness) with our ML-based postoperative ACD model resulted in statistically significant improvements in the MAE and MedAE of ray tracing refraction predictions. It appears likely that the internal ACD predictions of OKULIX were sufficiently accurate in normal length eyes (as would be expected for a linear regression approach) such that no significant difference in refraction prediction was seen in the normal length eye population. However, utilisation of our ML approach significantly improved refraction prediction MAE and MedAE in patients with short eyes (5.3% reduction in MAE, 7.3% reduction in MedAE) and those with long eyes (10.8% reduction in MAE, 15.8% reduction in MedAE). Although the overall improvement and the improvement in the short eyes were clinically small, we believe the results are of clinical significance for the long eyes.

The results highlighted the impact of the accuracy of postoperative ACD prediction on the accuracy of the refraction prediction for ray tracing methods. In relation to this study, our previous research has shown that substituting the ELP with a more accurately predicted postoperative ACD improved the accuracy of existing IOL formulas.[8] Since more and more studies on ELP (or postoperative ACD) prediction have been published, it is important to synchronise the progress in ELP prediction with efforts on refraction prediction. Our research provided a support for the vital role of postoperative ACD prediction in refraction prediction, suggesting that new generation ELP prediction methods should be considered when developing new IOL formulas and should be considered for incorporation with existing IOL formulas as an easily achievable refinement.

Since our ML model for postoperative ACD prediction was demonstrated to outperform linear regression in its prior validation study, it would be expected that incorporation of this ML model into the OKULIX suite would improve refraction prediction performance. While the means of the ACD predictions and refraction predictions were different between the standard and ML-based approaches, subtracting off the mean error in refraction prediction on an optimisation subset represents a straightforward (and previously described) method for addressing these systematic differences. The postoperative ACD prediction of OKULIX appeared to be longer than the ones predicted by the ML method (table 1), and naturally the unadjusted refraction predictions of OKULIX were more myopic (table 2). We corrected this systematic error by subtracting the mean error for the performance comparison dataset.

While our study was intentionally limited to eyes with normal corneas (in order to test the standard version of the ML model described in our group's prior work), a clear future direction is to apply the keratometry-independent version of our ML model for predicting postoperative ACD to the ray tracing analysis of eyes with abnormal corneas, such as those with ectasia, prior refractive surgery or prior keratoplasty. This subset of patients with abnormal corneas is one group for which ray tracing has been demonstrated to have clear advantages over traditional methods for IOL power calculation. However, the accurate prediction of postoperative ACD in this population is more challenging due to the absence of reliable keratometry data.

In addition, our study was limited to a retrospective sample and further investigation in a prospective manner would be of value.

In summary, this study demonstrates that incorporation of a validated ML method for postoperative ACD prediction can significantly improve ray tracing IOL calculation performance. Further investigation into the efficacy of this approach in eyes with ectatic and post-refractive surgery corneas is warranted.

Contributors TL: data analysis, programming and writing of the manuscript. AR: data collection, programming. JDS: data collection. NN: data analysis, programming, data collection, guidance on method development and writing of the manuscript. NN is responsible for the overall content as guarantor. The guarantor accepts full responsibility for the finished work and the conduct of the study, had access to the data, and controlled the decision to publish.

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 applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iDs
Tingyang Li http://orcid.org/0000-0003-2559-184X
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 Einighammer J, Oltrup T, Bende T, *et al*. Calculating intraocular lens geometry by real ray tracing. *J Refract Surg* 2007;23:393–404.
6 Hoffmann PC, Lindemann CR. Intraocular lens calculation for aspheric intraocular lenses. *J Cataract Refract Surg* 2013;39:867–72.
7 Tamaoki A, Kojima T, Tanaka Y, *et al*. Prediction of effective lens position using multiobjective evolutionary algorithm. *Transl Vis Sci Technol* 2019;8:64.
8 Li T, Stein J, Nallasamy N. AI-powered effective lens position prediction improves the accuracy of existing lens formulas. *Br J Ophthalmol* 2021. doi:10.1136/bjophthalmol-2020-318321. [Epub ahead of print: 09 Apr 2021].
9 Simpson MJ, Charman WN. The effect of testing distance on intraocular lens power calculation. *J Refract Surg* 2014;30:726.
10 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.
11 Preußner P-R, Hoffmann P, Petermeier K. Vergleich zwischen Raytracing und IOL-Formeln der 3. Generation. *Klin Monatsbl Augenheilkd* 2009;226:83–9.