IOL Formula Constants: Strategies for Optimization and Defining Standards for Presenting Data

Achim Langenbucher a, Nóra Szentmáry b, c, Alan Cayless d, Michael Müller e, Timo Eppig a, Simon Schröder a, Ekkehart Fabian e

Department of Experimental Ophthalmology, Saarland University, Homburg/Saar, Germany; Dr. Rolf M. Schwiete Center for Limbal Stem Cell and Aniridia Research, Saarland University, Homburg/Saar, Germany; Department of Ophthalmology, Semmelweis-University, Budapest, Hungary; School of Physical Sciences, The Open University, Milton Keynes, UK; Augenzentrum Rosenheim, Rosenheim, Germany

Keywords
Lens power calculation formulae · Formula constant optimization · Customized constants · Non-linear optimization · Mean absolute prediction error

Abstract
Purpose: The aim of this study is to present strategies for optimization of lens power (IOLP) formula constants and to show options how to present the results adequately. Methods: A dataset of N = 1,601 preoperative biometric values, IOLP data and postoperative refraction data was split into a training set and a test set using a random sequence. Based on the training set, we calculated the formula constants for established lens calculation formulae with different methods. Based on the test set, we derived the formula prediction error (PE) as difference of the achieved refraction from the formula predicted refraction. Results: For formulae with 1 constant, it is possible to back-calculate the individual constant for each case using formula inversion. However, this is not possible for formulae with >1 constant. In these cases, more advanced concepts such as non-linear optimization strategies are necessary to derive the formula constants. During cross-validation, measures such as the mean absolute or the root mean squared PE or the ratio of cases within mean absolute PE (MAE) limits could be used as quality measures. Conclusions: Different constant optimization concepts yield different results. To test the performance of optimized formula constants, a cross-validation strategy is mandatory. We recommend performance curves, where the ratio of cases within absolute PE limits is plotted against the MAE.

Background
In cataract surgery, most intraocular lens (IOL) power calculations worldwide are performed with theoretical-optical formulae [1–3]. These formulae require biometric measurements of the patient eye before cataract surgery derived with a biometer. They also involve formula constants, which adapt the general formula to the characteristics of the specific IOL [4]. These characteristics include the optics and haptic design as well as the material properties. Depending on the philosophy of the IOL calculation formula, the constants in most of the classical formulae directly interact with the effective lens position or the lens power (IOLP) [5].

© 2021 The Author(s).
Published by S. Karger AG, Basel
When an IOL company launches a new type of IOL to the market, the formula constants provided to surgeons for IOLP calculation are mostly estimated from the respective constants of similar lens types already on the market [4, 6], or they are derived in some pre-market studies based on a very limited number of cases. Once the lens has been on the market for a while, more clinical data are available including preoperative biometric data, the IOLP implanted, and the postoperative refractive outcome in terms of subjective sphere and cylinder or spherical equivalent. This allows the formula constants to be successively refined for better refractive outcomes in the future. This is a typical forward prediction process, where the data of previous cataract surgeries are used to predict the proper formula constant for cataract surgeries in the future.

However, there are no standards established in the literature specifying how to calculate formula constants for a specific dataset [4, 7, 8]. There is also no consensus on how many data points are required for reliable prediction of formula constants. There is also no standard for presentation of results, which would facilitate a comparison of different studies [9].

From the theory, it is clear that in all processes of forward prediction the dataset used for calculation of the formula constant should not be used for validation or verification. This implies that the dataset should be split into 2 non-overlapping sets: a training set, used for derivation of the formula constants, and a test or validation set used for verifying the performance of the formula constant in terms of a cross-validation. If validation is not performed on an independent dataset, there might be some overestimation of the performance of the formula constants [9]. There are several options for cross-validation: with only one dataset available, established techniques include “holdout,” “K fold cross-validation,” or “repeated random subsampling.” In contrast to holdout, where the dataset is simply split into training and test sets using a binary random sequence, K fold cross-validation splits the data into K partitions and the formula constant is calculated based on K-1 partitions by excluding partition I from the calculation and then testing with the excluded partition I. This process is repeated until each partition has been excluded and used as the test partition once. With repeated random subsampling, a random partition is separated out before calculation of the formula constant and this separated dataset then used for validation of the formula constant. However, this carries the risk that some data points may never be included in the training or validation data sets [10].

Furthermore, the number of data points required for formula constant optimization is unclear. In general, with increasing number, we expect a more reliable constant if the data pool is consistent, but the number depends also on the quality of the dataset and the composition of the study population [4].

Furthermore, there is no consensus or standard by which the target parameter should be optimized, or on the appropriate norm for optimization. In simple terms, we could back-calculate an appropriate individual formula constant for each data point within a dataset, and from these individual results, we could consider, for example, the mean or median value as an optimized constant. However, taking the mean or median as the optimized formula, constant does not guarantee that the mean or median of the refraction error is zeroed. On the other hand, we could solve the formula for the predicted refraction; calculate the deviation of the achieved refraction from the formula predicted refraction considered as “prediction error” (PE) and use non-linear optimization algorithms for minimizing, for example, the mean absolute PE (MAE) or the mean root squared PE (RMSE). Different target parameters and optimization criteria yield different results for the optimized formula constant, and there is no consensus on which technique should be used [9].

There are even various options for presentation of the results, making direct comparisons difficult. In some studies, the mean and standard deviation of the PE are documented, whereas in other studies, we find results on the MAE or the MRSE. Again, in other studies, we find distribution measures of the individual formula constants such as mean, median, or standard error, or the authors document the performance curves for MAE or RMSE or the portion of eyes which are within limits of a quarter, a half, or 1 dioptre of PE.

The purpose of this study is to show, using a large dataset of a cataract population with preoperative biometric data, the IOLP of the implanted lens, and postoperative refraction data:

- How cross-validation works in a clinical setting, and how the results compare with training and validation both carried out on the entire dataset,
- What the differences are in optimizing for different target parameters or using different target criteria or norms,
- To explain the meaning of different representation formats for our results, and
- To make recommendations on which presentation format shows the results most appropriately.
**Table 1. Descriptive statistics of the entire dataset with mean, SD, median, minimum and maximum, 5, and 95% quantiles (90% confidence intervals)**

<table>
<thead>
<tr>
<th>N = 1,452</th>
<th>AL in mm</th>
<th>ACD in mm</th>
<th>LT in mm</th>
<th>R1 in mm</th>
<th>R2 in mm</th>
<th>IOLP in D</th>
<th>SEQ in D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>23.497</td>
<td>3.099</td>
<td>4.608</td>
<td>7.741</td>
<td>7.594</td>
<td>21.519</td>
<td>−0.509</td>
</tr>
<tr>
<td>SD</td>
<td>1.082</td>
<td>0.389</td>
<td>0.422</td>
<td>0.271</td>
<td>0.271</td>
<td>2.849</td>
<td>0.859</td>
</tr>
<tr>
<td>Median</td>
<td>23.419</td>
<td>3.101</td>
<td>4.606</td>
<td>7.723</td>
<td>7.588</td>
<td>21.5</td>
<td>−0.375</td>
</tr>
<tr>
<td>Minimum</td>
<td>20.635</td>
<td>2.000</td>
<td>3.345</td>
<td>6.907</td>
<td>6.434</td>
<td>8.0</td>
<td>−4.0</td>
</tr>
<tr>
<td>Maximum</td>
<td>29.041</td>
<td>4.354</td>
<td>5.794</td>
<td>8.771</td>
<td>8.581</td>
<td>30.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Quantile 5%</td>
<td>22.016</td>
<td>2.461</td>
<td>3.912</td>
<td>7.236</td>
<td>7.175</td>
<td>16.5</td>
<td>−2.5</td>
</tr>
<tr>
<td>Quantile 95%</td>
<td>25.529</td>
<td>3.753</td>
<td>5.311</td>
<td>8.212</td>
<td>8.056</td>
<td>26.0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

AL refers to the axial length, ACD to the phakic anterior chamber depth measured from the corneal front apex to the anterior apex of the crystalline lens, LT to the central thickness of the crystalline lens, R1 and R2 to the corneal front surface radius in the flat and steep meridian, IOLP to the refractive power of the implanted lens, and SE to the spherical equivalent of postoperative refraction. IOLP, lens power; AL, axial length; LT, lens thickness; SD, standard deviation.

**Methods**

**Measurement Data**

In this retrospective study, we analysed a dataset with 1,601 clinical measurements of a cataract population from Augenzentrum Rosenheim, which was transferred to us. The anonymized data contained preoperative biometric data derived with the IOL-Master 700 (Carl-Zeiss-Meditec, Jena, Germany) including axial length (AL), external phakic anterior chamber depth measured from the corneal front apex to the anterior apex of the crystalline lens (ACD), lens thickness, corneal front surface radius measured in the flat (R1), and in the steep meridian (R2). In all cases, a Sen-sar 1 piece IOL (Johnson & Johnson, Brunswick, NJ, USA) was implanted. Beside refractive power of the implanted lens (IOLP), the postoperative refraction (sphere and cylinder) 6–8 weeks after cataract surgery was recorded. From the total of 1,601 measurements, N = 1,452 complete measurements with a postoperative visual acuity of 0.6 or higher were used for formula constant optimization. Eyes with missing data (mostly anterior chamber depth) were excluded. The spherical equivalent of postoperative refraction (SEQ) was calculated as sphere + ½·cylinder, and the mean corneal front surface radius was calculated as R = ½ (R1 + R2). The descriptive data on pre-cataract biometry, IOLP, and postoperative refraction are summarized in Table 1.

**Calculation Strategy**

The anonymized Excel data (.xlsx-format) were imported into MATLAB (Matlab version 2019b, The Math Works, Natick, MA, USA) for further processing. For all eyes, an AL correction according to the Cooke formula [11, 12] was performed. The individual formula constant was back-calculated for each case for the SRKT formula [13, 14], the Holladay1 formula [17], and the simplified Haigis with 1 formula constant a0 and standard values for a1 = 0.4 and a2 = 0.1 [2, 4]. For the SRKT formula 2 different strategies were applied, one according to the concept described in the original paper [13, 14] (SRK2 style) and one with an inversion of the SRKT formula [4].

For all calculations, the deviation of the achieved postoperative SEQ from the formula predicted SEQ was quoted as the PE. AE in this context refers to the absolute value of PE and SE to the squared value of PE.

For testing the formula performance and cross-validation, the N = 1,452 data points were split randomly into a training set (70%, 1,017 cases) and a test set (30%, 435 cases) [10]. For the training set, the mean and median A constant (Amean and Amedian for formula inversion and A0mean and A0median for SRK2 style calculation) for the SRKT formula, the mean and median pACD constant (pACDmean and pACDmedian) for the Hoffer-Q formula, the mean and median SF constant (SFmean and SFmedian) for the Holladay1 formula, and the mean and median a0 constant (a0mean and a0median) for the simplified Haigis formula were calculated.

In addition, a non-linear optimization algorithm (Levenberg-Marquardt algorithm, [18, 19]) was implemented and applied to derive formula constants from the training set for the SRKT (Amean, A0mean), the Hoffer-Q (pACDmean, pACDmedian), the Holladay1 (SFmean, SFmedian), the simplified Haigis (a0mean, a0median) and the simplified Haigis formula were calculated. In the next step, we calculated the MAE, the median absolute error, and the RMSE (RMSE = square-root of SE) for each lens calculation formula and each optimized constant using the test dataset. Then for all formulae, we derived the portion of cases with AE ≤0.25 dpt, AE ≤0.5 dpt, AE ≤1.0 dpt, and AE ≤2.0 dpt. In a final step, we calculated the MAE based on the training set (used to optimize the constants) to evaluate the robustness of the optimized constants to variation of the formula constant. For the formulae with 1 constant (SRKT, Hoffer-Q, Holladay1, simplified Haigis), the formula constant optimized for the lowest RMSE of...
Accuracies of various prediction error models:

- **SRKT (optimized for RMSE):**
  - n = 435
  - A constant: 119.1076 mm
  - MAE: 0.41606 d
  - RMSE: 0.54553 d

- **Hoffer-Q (optimized for RMSE):**
  - n = 435
  - pACD constant: 5.6764 mm
  - MAE: 0.43983 d
  - RMSE: 0.57681 d

- **Holladay1 (optimized for RMSE):**
  - n = 435
  - SF constant: 1.8892 mm
  - MAE: 0.42118 d
  - RMSE: 0.56237 d

- **Haigis (optimized for RMSE):**
  - n = 435
  - a0 constant: 1.5222 mm
  - a0 constant: 5.3966
  - a2: 0.62474
  - a3: -0.097119
  - MAE: 0.34959 d
  - RMSE: 0.47149 d

(For legend see next page.)
refraction was varied by ±1 in steps of 0.01 and the portion of cases with AE ≤0.25 dpt, AE ≤0.5 dpt, and AE ≤1.0 dpt (in the training set) was recorded. For the Haigis formula with 3 constants (a0/a1/a2), we varied a0 by ±1, a1 by ±0.2, and a2 by ±0.1 each in 100 equi-distant steps starting from a0/a1/a2RMSE and calculated for each constant triplet the MAE to test the robustness of the constant triplet to variations in a0, in a1, and in a2.

**Results**

The formula constants derived from the training set for all formulae under test are summarized in Table 2. For formulae with 1 formula constant (SRKT, Hoffer-Q, Holladay1, simplified Haigis with optimized a0), the mean and median of the individual formula constant are listed. For the SRKT formula, we have also provided the mean and median of the individual formula constant calculated according to the concept proposed in the original paper (SRK2 style, A0 mean and A0 median, Lit…). In addition, for all formulae, we have listed the respective formula constants calculated from a non-linear optimization process in terms of minimizing the mean absolute and the root mean squared refraction error.

The distribution of the PE for all formulae with constant optimization for RMSE is displayed in Figure 1. Optimization was performed on the training set (N = 1,017), and cross-validation was done on the test set (N = 435).

The respective formula constants listed in Table 2. Figure 1a refers to the SRKT formula (ARMSE), Figure 1b to the Hoffer-Q formula (pACD RMSE), Figure 1c to the Holladay1 formula (SF RMSE), Figure 1d to the simplified Haigis formula with customized a0 and preset values for a1 = 0.4 and a2 = 0.1 (a0 RMSE), and Figure 1e to the Haigis formula with constant triplet (a0/a1/a2 RMSE).

The distributions of the absolute PE AE for all different constant optimization strategies and each formula are shown in Figure 2. Again, optimization was performed on the training set, and cross-validation was done on the test set. The respective formula constants listed in Table 2. Figure 2a refer to the AE of the SRKT formula (A0 mean, A0 median, A mean, A median, A MAE, and A RMSE), Figure 2b to the Hoffer-Q formula (pACD mean, pACD median, pACD MAE, and pACD RMSE), Figure 2c to the Holladay1 formula (SF mean, SF median, SF MAE, and SF RMSE), Figure 2d to the simplified Haigis formula with customized a0 and preset values for a1 = 0.4 and a2 = 0.1 (a0 mean, a0 median, a0 MAE, and a0 RMSE), and Figure 2e to the Haigis formula with constant triplet (a0/a1/a2 MAE and a0/a1/a2 RMSE).

The robustness of formula constants is shown in Figure 3, in terms of the ratio of cases within PE limits (Fig. 3a–d) or MAE (Fig. 3e). For the formulae with one constant, this constant was varied in limits of ±1 (Fig. 3a: SRKT formula; Fig. 3b: Hoffer-Q formula; Fig. 3c: Holladay1 formula; and Fig. 3d: simplified Haigis formula)

### Table 2. Optimized constants for the SRKT, Hoffer-Q, Holladay1, simplified Haigis formula with optimized a0, and Haigis formula with 3 optimized constants (a0/a1/a2) based on the training set

<table>
<thead>
<tr>
<th>N = 1,017 training data</th>
<th>SRKT formula</th>
<th>Hoffer-Q formula</th>
<th>Holladay formula</th>
<th>Haigis formula a0 constant</th>
<th>Haigis a0/a1/a2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean of individual constant</td>
<td>A mean = 118.981</td>
<td>pACD mean = 5.636</td>
<td>SF mean = 1.833</td>
<td>a0 mean = 1.468</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A0 mean = 119.305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median of individual constant</td>
<td>A median = 118.993</td>
<td>pACD median = 5.611</td>
<td>SF median = 1.820</td>
<td>a0 median = 1.454</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A0 median = 119.322</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimized for MAE</td>
<td>A MAE = 119.016</td>
<td>pACD MAE = 5.596</td>
<td>SF MAE = 1.819</td>
<td>a0 MAE = 1.450</td>
<td>1.720/0.495/0.076</td>
</tr>
<tr>
<td>Optimized for RMSE</td>
<td>A RMSE = 119.016</td>
<td>pACD RMSE = 5.606</td>
<td>SF RMSE = 1.824</td>
<td>a0 RMSE = 1.446</td>
<td>1.565/0.443/0.089</td>
</tr>
</tbody>
</table>

For the formulae with 1 constant, the optimized constant was derived from the mean and the median of the individually back-calculated constant. The individual constants for the SRKT formula were back-calculated using formula inversion (A mean and A median) and using the SRK2 style strategy (A0 mean and A0 median). For all formulae, the optimized constant was also derived using a non-linear optimization process for lowest MAE and RMSE. MAE, mean absolute error; RMSE, root mean squared error.

**Figure 1.** PE as the deviation of achieved spherical equivalent – formula predicted refraction. Formula constants have been optimized on the training set (N = 1,017) and are cross-validated on the test set (N = 435). In this example, PE is shown for a formula constant optimization for minimizing the RMSE. Refers to the SRKT (a), to the Hoffer–Q (b), to the Holladay1 (c), to the simplified Haigis with optimized a0 and preset values for a1 = 0.4 and a2 = 0.1 (d), and to the Haigis formula with triplet constant (a0/a1/a2) optimization (e). PE, prediction error.
Distribution of absolute prediction error AE

Ratio of cases within prediction error limits, %

Absolute prediction error AE ≤ ...

Optimized for RMSE
Optimized for MAE
Mean (individual a0 constant)
Median (individual a0 constant)
Mean (individual pACD constant)
Median (individual pACD constant)
Mean (individual SF constant)
Median (individual SF constant)
Mean (individual A constant)
Median (individual A constant) SRK2 style

(For legend see next page.)
with optimized $a_0$ and $a_1 = 0.4$ and $a_2 = 0.1$. For the Haigis formula with 3 formula constants ($a_0/a_1/a_2$), $a_0$ was varied in limits of $\pm 1$, $a_1$ in limits of $\pm 0.2$, and $a_2$ in limits of $\pm 0.1$ from the constant triplet optimized for RMSE (Fig. 3e). MAE data coded in colour are clipped to $\leq 2.0$ dpt for display. All constant triplets ($a_0/a_1/a_2$) located on the blue hyperplane yield good results in terms of a low MAE.

In Figure 4, the ratio of cases within limits of absolute PE $AE \leq 0.25 \text{ dpt}$, $\leq 0.5 \text{ dpt}$, and $\leq 1.0 \text{ dpt}$ is displayed. Formula constants were optimized on the training set ($N = 1,017$) for minimum RMSE and cross-validated on the test set ($N = 435$).

A comparison of all formulae under test is shown in Figure 5. The formula performance in terms of the absolute PE MAE is displayed for the SRKT, the Hoffer-Q, the Holladay1, the simplified Haigis with optimized $a_0$ and standard values for $a_1 = 0.4$ and $a_2 = 0.1$, and for the Haigis formula with constant triplet optimization $a_0/a_1/a_2$. Optimization for minimum RMSE was performed on the training set and cross-validated on the test set.

**Discussion**

In the last 20 years, patient expectations for an excellent visual performance after cataract surgery have increased greatly, as cataract surgery becomes more and more standardized and complications very rare. However, there is still controversial discussion over which formula should be used for the general cases [1–3] or special situations such as long or short eyes or uncommon anterior segment geometries [7, 20, 21]. General rules for selecting the “best formula” are difficult to define. There are currently many different competing calculation concepts and most of the formula authors are self-opinionated that their own philosophy of calculating the power of an IOL implant outperforms other concepts.

The key to success in formula-based IOLP calculation is the use of appropriate formula constants [2, 7, 22–25]. These constants adapt a generic formula, which is a more general formulation of a mathematical concept, to a specific IOL type, special surgery conditions, patient’s ethnicities, or measurement equipment. The formula constants provided by the manufacturer can be used as a good estimate or starting point for further optimization. From the basic idea, a constant optimization can be performed post hoc if results of a sufficient number of representative clinical data with a lens type have already been collected [24]. The result of the optimization process is then applied to subsequent cataract procedures in terms of a forward prediction.

Such a constant optimization process requires all biometric data which feed into the formula for IOLP calculation. In addition to the biometric data, the power of the lens and the postoperative refraction in terms of sphere and cylinder or spherical equivalent are required. For formulae with one constant, there is a straightforward option for calculating the formula constant. The formula can be re-organized and solved for the constant, and for each clinical case, we could back-calculate which formula constant is required for the biometric data together with the power of the implanted lens to yield the refraction actually achieved after cataract surgery. For each clinical case, we obtain an individual constant, and the mean or median of all individual constants could be quoted as an optimized constant for a large dataset. However, this strategy does not optimize the dataset for the refractive outcome, but rather for any of the measures in the distribution function of the individual constants.

For formulae with $>1$ constant, this simple concept of back-calculation the individual constant for each clinical case fails. For example, for the Haigis formula with 3 constants, the effective lens position $d$ can be back-calculated, and in most constant optimization strategies, a multivariate linear regression is used to calculate the constant triplet $a_0/a_1/a_2$ as intercept ($a_0$) and 2 weighting factors ($a_1$ and $a_2$) from the linear regression ($d = a_0 + a_1\cdot \text{ACD} + a_2\cdot \text{AL}$) [7, 9, 24]. Most of the modern IOLP calculation formulae are not published, but the classical formulae such as SRKT, Hoffer-Q, Holladay1, or Haigis formula were published $>20$ years ago. Only one of these formulae (SRKT) gives some hint in the original paper as to how the lens constant should be optimized, which was considered in this paper as “SRK2 style” optimization (and the optimized constants $A_0\text{mean}$ and $A_0\text{median}$) in addition to $A_0\text{mean}$, $A_0\text{median}$, $A_0\text{MAE}$, and $A_0\text{RMSE}$) (c), to the simplified Haigis formula with customized $a_0$ and preset values for $a_1 = 0.4$ and $a_2 = 0.1$ ($A_0\text{mean}$, $A_0\text{median}$, $A_0\text{MAE}$, and $A_0\text{RMSE}$) (d), and to the Haigis formula with constant triplet ($a_0/a_1/a_2\text{MAE}$ and $a_0/a_1/a_2\text{RMSE}$) (e). PE, prediction error; RMSE, root mean squared error.
Robustness of SRKT A constant, $n = 1,017$ training data

Robustness of Hoffer-Q pACD constant, $n = 1,017$ training data

Robustness of holladay1 SF constant, $n = 1,017$ training data

Robustness of haigis a0 constant, $n = 1,017$ training data

MAE in D, robustness of haigis a0/a1/a2 triple constant optimization

(For legend see next page.)
the classical formula inversion (optimized constants $A_{\text{mean}}$ and $A_{\text{median}}$). There are no general rules for formula constant optimization. With steadily increasing computing capacity and speed, we are no longer bound to straightforward calculation concepts with direct backcalculation of the formula constant. Instead, non-linear optimization algorithms have been developed with a very high performance which could optimize any target parameter with any optimization criterion. This means, that instead of extracting any measure from the distribution of all individual constants, we could, for example, optimize for the mean, the mean absolute, the median, or the RMSE in terms of deviation of the achieved refraction after cataract surgery from the formula predicted refraction. In other words, we replace any measure from the statistical distribution of individual constants by a measure which has high relevance for the patient and her/his refractive outcome. In the present paper, we have implemented the Levenberg-Marquardt algorithm [18, 19] as a non-linear gradient descent method to search for an optimized constant which yields the lowest mean absolute or RMSE in addition to the straightforward calculation of the constants using formula inversion.

In reality, in most of the datasets, the difference between all the optimization strategies is expected to be low. We found only slight differences between the constants derived with formula inversion and the constants based on non-linear optimizations. However, in the case of datasets with outliers or skewed distributions of parameters instead of a “clean” dataset where data are acquired under strict study conditions (as was the case here with our data set), the difference between both strategies might be higher.

In most of the papers concerning formula constant optimization, the dataset is not split into training and test or validation data [4, 7, 9, 23, 24]. That means that the dataset is first used to derive the optimized formula constants, and later, the same dataset is used to test for the perfor-

![Fig. 3. Robustness of formula constants in terms of ratio of cases within PE limits (a-d) or MAE (e). Formula constants were optimized on the training set (N = 1,017) for minimum RMSE and cross-validated on the test set (N = 435). PE, prediction error; RMSE, root mean squared error.](image)

![Fig. 4. Ratio of cases within limits of absolute PE AE ≤0.25 dpt, ≤0.5 dpt, and ≤1.0 dpt. Formula constants were optimized on the training set (N = 1,017) for minimum RMSE and cross-validated on the test set (N = 435). PE, prediction error; RMSE, root mean squared error.](image)

![Fig. 5. Comparison of the performance of all formulae under test in terms of MAE. Formula constants were optimized on the training set (N = 1,017) for minimum RMSE and cross-validated on the test set (N = 435). PE, prediction error; MAE, mean absolute prediction error; RMSE, root mean squared error.](image)
mance of the formula with the optimized constants. From all concepts in artificial intelligence or machine learning, we know that a strict separation of the dataset into training and test set is mandatory in terms of a cross-validation [10]. The entire dataset is split into training data for constant optimization and test set for validation. In this paper, we used a randomly selected 70% subset of our large dataset for training and the remaining 30% for validation. Only when testing the robustness of our constants [25] were the training data also used for testing the performance for variation of the constant; this was done to ensure that the robustness graphs have a direct clinical impact.

In this paper, we have attempted to describe different concepts of constant optimization based on a large study population of cataract patients treated with 1 intraocular lens type. All the data were collected at 1 clinical centre, and the clinical settings appear very homogeneous for this dataset as shown in Table 1. In Table 2, we have listed all optimized constants derived with different optimization strategies. For all formulae with 1 formula constant (SRKT, Hoffer-Q, Holladay1, and simplified Haigis formula), we back-calculated the individual constant by formula inversion (A\text{mean} and A\text{median}) and the SRK2 like style (A0\text{mean} and A0\text{median}). RMSE, root mean squared error; MEDAE, median absolute error.

<table>
<thead>
<tr>
<th>N = 435 test data</th>
<th>Mean absolute refraction error MAE in dpt</th>
<th>Median absolute refraction error MEDAE in dpt</th>
<th>Root mean squared refraction error RMSE in dpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRKT formula</td>
<td>A\text{mean} 0.391</td>
<td>A\text{median} 0.390</td>
<td>A0\text{mean} 0.428</td>
</tr>
<tr>
<td>Hoffer-Q formula</td>
<td>pACD\text{mean} 0.417</td>
<td>pACD\text{median} 0.416</td>
<td>pACD\text{MAE} 0.417</td>
</tr>
<tr>
<td>Holladay1 formula</td>
<td>SF\text{mean} 0.390</td>
<td>SF\text{median} 0.390</td>
<td>SF\text{MAE} 0.390</td>
</tr>
<tr>
<td>Simplified Haigis formula with a0</td>
<td>a0\text{mean} 0.369</td>
<td>a0\text{median} 0.368</td>
<td>a0\text{MAE} 0.368</td>
</tr>
<tr>
<td>Haigis formula with a0/a1/a2</td>
<td>a0/a1/a2\text{MAE} 0.366</td>
<td>a0/a1/a2\text{RMSE} 0.366</td>
<td></td>
</tr>
</tbody>
</table>

(A\text{mean} and A\text{median} refer to the mean and median of the individual back-calculated constant, and (A\text{MAE} and A\text{RMSE}) to the formula constant optimization for mean absolute and root mean squared refraction error. For the SRKT formula, there are 2 types of back-calculation for the individual formula constant: direct formula inversion (A\text{mean} and A\text{median}) and the SRK2 like style (A0\text{mean} and A0\text{median}). RMSE, root mean squared error; MEDAE, median absolute error.)
Strategies for Formula Constant Optimization

The results for the PE using the constants optimized for the lowest RMSE are shown in the histograms of Figure 1 for all formulae. The mean and median of the PE are not necessarily zero as we did not optimize for a mean PE, but for the mean absolute or root mean squared PE. In the graphs, we have included the best-fit normal distribution and the mean absolute and RMSE. The respective values for the MAE, the median absolute error, together with the RMSE for all optimization strategies for all formulae are listed in Table 3. The MAE ranged between 0.366 and 0.433 dpt, the median absolute PE between 0.288 and 0.341 dpt, and the RMSE between 0.488 and 0.572 dpt.

To our knowledge, the robustness to variation of formula constants has not been investigated so far. For this analysis, we used the training set (previously used for constant optimization) and varied the constants starting from the optimized constant in terms of a minimum RMSE. The respective graphs for the formulae using 1 constant are shown in Figure 3a–d. The 3 curves refer to the ratio of cases within absolute PE limits of AE ≤0.25, ≤0.5, and ≤1.0 dpt. The vertical lines indicate the various optimized constants. From the graphs, we see that the tolerance in the SRKT formula is larger than the Hoffer-Q formula in our test set in terms of absolute PE AE [25].

Further, data splitting is always performed with a random sequence, and depending on the selection of the training and test sets the results of cross-validation could somehow diverge. This means that repeating the entire formula constant optimization process and cross-validation using another random sequence for data splitting could produce different results. In a large dataset as in our study, the differences in the cross-validated data are not expected to depend too much on the random sequence for data splitting, but in smaller datasets, this could be a problem. However, data splitting and cross-validation exactly reflect the situation in clinical life, where we optimize constants on an existing dataset and use the constant for predicting the IOLP in a new cataract case. For our dataset, we tested the performance of formula constants optimized in the training set for RMSE if applied on the test set, the training set, and the entire dataset. The performance in terms of ratio of cases within AE limits of ≤0.25, ≤0.5, and ≤1.0 dpt is listed in Table 4.

What we see is that in most cases the optimized constants perform better in terms of a higher ratio of cases within
limits on the training set compared to the test set. This result seems obvious. The higher the coherence between the training and test set the closer the results for the training and test set. From a mathematical standpoint, the more individual a IOLP calculation scheme is in terms of degrees of freedom, the better the formula would be expected to reproduce the training set with the optimized constants. However, cross-validation with a mutually exclusive test set might lead to an overfitting with the consequence that the performance might be degraded significantly. As we deal with a large dataset in this study and with formulae with 1 or 3 constants, the effect of overfitting is low.

In conclusion, this study shows that:

1. For IOL calculation formulae which are disclosed and public domain, there are different strategies of formula constant optimizations. For formulae with 1 constant, we could back-calculate an individual formula constant for each case and derive the optimized constant from the mean or median of the individual constants. For formulae with >1 constant, the constant optimization is more difficult as we cannot back-calculate the appropriate constant for each case.

2. In addition to classical constant optimization strategies, we could always implement a non-linear optimization strategy, where any target parameter is optimized for any target criterion. An optimization for the minimal PE seems to be closest to the needs of the surgeon and patients as the predictability of the refractive outcome is the most important quality criterion in modern cataract surgery. Typical target criteria in most cases are the minimization of the RMSE or the MAE.

3. Evaluating the results of formula constant optimization always requires cross-validation. For such a cross-validation, the dataset has to be split into a training set used for calculating the constants, and a test set used for validation of the constants. Without cross-validation, in most cases, the performance of the constants would be overestimated.

4. There are different options for presenting the results of formula constant optimization. We feel that performance curves which show the ratio of cases within limits of the absolute PE are the most appropriate tools and yield more information than the MAE or the ratio of cases within selected absolute PE limits.

### Table 4. Ratio of cases within absolute PE AE limits of ≤0.25 dpt, ≤0.5 dpt, and ≤1.0 dpt for the test set, the training set, and the entire dataset

<table>
<thead>
<tr>
<th>Ratio of cases within AE limits in %</th>
<th>AE ≤0.25 dpt</th>
<th>AE ≤0.5 dpt</th>
<th>AE ≤1.0 dpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRKT formula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test set, N = 435</td>
<td>40.69</td>
<td>69.89</td>
<td>95.40</td>
</tr>
<tr>
<td>Training set, N = 1,017</td>
<td>41.69</td>
<td>71.49</td>
<td>95.28</td>
</tr>
<tr>
<td>All, N = 1,452</td>
<td>41.39</td>
<td>71.01</td>
<td>95.32</td>
</tr>
<tr>
<td>Hoffer-Q formula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test set, N = 435</td>
<td>40.00</td>
<td>68.97</td>
<td>93.33</td>
</tr>
<tr>
<td>Training set, N = 1,017</td>
<td>39.43</td>
<td>69.42</td>
<td>95.97</td>
</tr>
<tr>
<td>All, N = 1,452</td>
<td>39.60</td>
<td>69.28</td>
<td>95.18</td>
</tr>
<tr>
<td>Holladay1 formula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test set, N = 435</td>
<td>42.75</td>
<td>71.49</td>
<td>93.33</td>
</tr>
<tr>
<td>Training set, N = 1,017</td>
<td>43.26</td>
<td>70.70</td>
<td>96.56</td>
</tr>
<tr>
<td>All, N = 1,452</td>
<td>43.11</td>
<td>70.94</td>
<td>95.59</td>
</tr>
<tr>
<td>Simplified Haigis formula with a0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test set, N = 435</td>
<td>44.14</td>
<td>71.26</td>
<td>95.86</td>
</tr>
<tr>
<td>Training set, N = 1,017</td>
<td>44.84</td>
<td>74.83</td>
<td>96.85</td>
</tr>
<tr>
<td>All, N = 1,452</td>
<td>44.63</td>
<td>73.76</td>
<td>96.56</td>
</tr>
<tr>
<td>Haigis formula with a0/a1/a2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test set, N = 435</td>
<td>44.14</td>
<td>72.87</td>
<td>96.32</td>
</tr>
<tr>
<td>Training set, N = 1,017</td>
<td>45.03</td>
<td>74.53</td>
<td>96.76</td>
</tr>
<tr>
<td>All, N = 1,452</td>
<td>44.63</td>
<td>73.76</td>
<td>96.56</td>
</tr>
</tbody>
</table>

Constants were optimized for minimum RMSE on the training set (N = 435). In most cases, the formula performance tested on the training set outperforms the respective performance on the test set. PE, prediction error; RMSE, root mean squared error.
Strategies for Formula Constant Optimization

Statement of Ethics

This retrospective review of patient data did not require ethical approval in accordance with local guidelines from Ethics Committee of the Bavarian State Chamber of Physicians. Written informed consent from participants was not required in accordance with local guidelines.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Funding Sources

Not applicable.

References


Author Contributions


Data Statement

The anonymized raw data that were analyzed in the study could be provided on request.