



Data Analysis in IOL Power Calculations

6

Giacomo Savini and Kenneth J. Hoffer

When the results of intraocular lens (IOL) power calculation by one or more formulas are reported, different parameters should be provided and a proper statistical analysis should be performed. In recent years, specific guidelines on this subject have been published and updated, and it is likely that further recommendations will be given in the future, as the interest of researchers in this field is increasing [1–4].

Designing the Sample to be Analyzed

The first step of any study on IOL power calculation is the enrollment of an appropriate sample. The following guidelines should be followed:

- For patients who underwent bilateral cataract surgery, only one eye for each patient should be analyzed [1, 4]. Ocular measurements are more alike between fellow eyes than between eyes of different subjects, and measurements from fellow eyes cannot be treated as if they were independent [5]. If the correlation between the right and left eyes of each subject is not accounted for in statistical analysis, there may be errors in the results obtained [6]. It would be preferable to consider only one eye of each individual. In this case, several approaches can be followed such as random selection of one eye (right or left), arbitrary selection of all right eyes, or a clinically based selection (e.g., the eye with the best visual acuity). Alternatively, if both eyes of the same patient are included, appropriate statistical methods (generalized estimating equations), which estimate the correlation and adjust for it in the analysis, may be used [7]. However, in general, the fewer statistical adjustments performed, the better.
- Patients with preoperative and/or postoperative pathologies should be excluded, as well as those with a postoperative corrected distance visual acuity worse than 20/40, because poor acuity lessens the accuracy of the crucial postoperative refractive error [1, 4].
- A uniform sample is preferable. This means that we suggest enrolling eyes that underwent preoperative measurements with the same optical biometer, were operated with the same technique (standard phacoemulsification vs. femtosecond laser-assisted cataract surgery), received the same IOL model, and were refracted using the same method. Exceptions

G. Savini
G.B. Bietti Eye Foundation, Rome, Italy

Studio Oculistico d'Azeglio, Bologna, Italy
e-mail: giacomo.savini@startmail.com

K. J. Hoffer (✉)
St. Mary's Eye Center, Santa Monica, CA, USA
Stein Eye Institute, UCLA, Los Angeles, CA, USA
e-mail: KHofferMD@startmail.com

to this recommendation may be acceptable for studies where it is more difficult to enroll a large sample, such as eyes with keratoconus or eyes with previous corneal refractive surgery.

- The sample size should be sufficient to allow constant optimization. According to Langenbucher et al., at least 100 eyes should be enrolled in order to achieve a stable mean refractive error, at least with formulas with a single constant [8]. The same number was suggested in 2010 by Haigis (personal communication). We agree that such a sample size can be considered sufficient for most studies. Of course, larger samples may be more powerful to disclose statistically significant differences and for this reason we recently suggested a minimum sample size of 200 eyes [4], whereas Holladay et al. suggested a minimum sample size between 300 and 700 eyes [3]. The uncertainty on this issue depends on the fact that a universally accepted parameter to be selected and investigated does not yet exist: the calculated minimum sample size changes if we look at the standard deviation (SD) of the prediction error (PE), the median absolute error (MedAE), or the percentage of eyes with a PE within ± 0.50 diopters (D). Although the SD has been recently advocated as the best parameter [3], there is not yet a global consensus on it. Moreover, the sample size calculation depends on the clinically significant difference that is looked for. Therefore, the help of a statistician is important when designing these studies.
- Also, depending on what is being studied, it is important to be sure that the AL and K ranges of the sample are not skewed toward longer eyes, shorter eyes, or only those with “normal values.”
- Postoperative refraction should be measured when stable. With small-incision surgery and one-piece IOLs, the refraction can be considered to be stable at 1 week from surgery [9–11], but we suggest waiting at least 1 month. Three months may be even better, but no evidence exists for this. Waiting 6–12 months, as recently suggested [3], is quite impractical and leads to an unremarkable advantage. The highest accuracy should be used when assess-

ing the postoperative spherical equivalent: if the patient can read 20/20 for distance without any correction, the examiner should not simply report 0 (plano) as the postoperative refraction but should assess whether adding or subtracting 0.25 D can improve visual acuity further. The testing distance for visual acuity should be standardized. A 6-m (approximately 20-foot) distance, rather than 4 m (approximately 13 feet) or infinity, may be the preferred choice [12]. Refractions at 4 m can be converted to 6 m by adding a value of -0.08 D to the spherical equivalent (e.g., a refraction of 0.00 D at 4 m corresponds to a refraction of -0.08 D at 6 m).

Selecting the Data to be Reported

In addition to the demographics of the study population (age, gender, and ethnicity), the following values should be reported:

- Prediction error (PE): This is defined as the difference between the postoperative spherical equivalent refraction and the predicted refraction (not the target refraction!). It is calculated as the postoperative refraction minus the predicted refraction so that the PE is negative for myopic errors and positive for hyperopic errors. The mean PE with any formula should be zeroed out by means of constant optimization. The latter is a relatively easy task with published formulas [13–16], since it can be carried out on Excel (Microsoft, Redmond, VA), as previously explained [1, 4], or using the internal software of different optical biometers. For the Haigis formula, it is mandatory to optimize all three constants. Constant optimization is more complicated with the latest generation formulas, which are all unpublished and for which it is better to ask for the help of the formula’s authors. Alternatively, it is possible to use specific computer programming languages able to extract data automatically from any database (e.g., Python Software Foundation, Wilmington, DE), enter them into the formula website, and generate a new database containing the predicted refraction for

each eye. Regarding the Holladay 2 formula, which is also unpublished, it is possible to perform optimization using the Holladay IOL Consultant Software & Surgical Outcomes Assessment (www.hicsoap.com).

- There are also some situations where constant optimization should not be carried out or should be carried out with caution. The first scenario is the analysis of specific samples, such as long or short eyes. When evaluating only short eyes, it would be more appropriate to rely on the optimized constants of the whole population (which have to be separately calculated) rather than on the optimized constants specifically calculated for the short eye sample. In the clinical setting, in fact, no one uses separate constants for short and medium eyes. The same approach can be followed for unusual eyes (e.g., those with keratoconus), where it might be more appropriate to use optimized constants obtained from larger samples of healthy eyes rather than from keratoconic eyes. The second scenario is the analysis of eyes with previous corneal refractive surgery: here, constant optimization would be preferable, but the lack of large samples with the same IOL model often precludes it. When more IOL models have to be analyzed simultaneously, it can be acceptable to use (for each IOL) optimized constants from large databases such as those available on the User Group for Laser Interference Biometry (ULIB, <http://ocusoft.de/ulib/c1.htm>, accessed on February 27, 2021) or on the IOLcon website (<https://iolcon.org>, accessed on February 27, 2021).
- Standard deviation (SD) and variance of the PE: SD is the square root of the variance, which is the average of the squared differences from the mean. These values are extremely important as they provides us with the information about how spread out the individual PEs are. Accurate formulas have lower SDs (and variances), whereas higher SDs (and variances) are the consequence of many outliers. SD deviation has been recently indicated as the best parameter to compare the refractive outcomes of different formulas [3].
- Distribution of the PE: The PE has always been considered to be normally distributed [2], but recently this assumption has been negated by Holladay et al. [3] Actually, in previous studies with relatively small sample size our group found a normal distribution of the PE [17, 18], whereas the observation by Holladay and coauthors derives from the largest study ever published [19]. We recommend reporting whether the PE distribution is normal or not because the choice between parametric and nonparametric statistical methods (to be used when comparing the PEs of different formulas) depends on this issue. Additional values that should be provided with the distribution are skewness and kurtosis [3]. The former is related to the symmetry of the PE distribution: the tail may be longer to the left or the right. If skewness ranges between -0.5 and $+0.5$, the distribution is approximately symmetric. The latter describes the tailedness of the sample (and not its peak).
- Median absolute error (MedAE): The absolute prediction error has been considered the most important outcome for many years. Earlier studies reported the mean absolute error (MAE); we then switched our recommendation to the MedAE since Haigis and Norrby showed us that the distribution of the absolute prediction error cannot be normal. The absolute prediction error is still a mandatory outcome measure, especially once constant optimization leads to a mean arithmetic PE of zero.
- Interquantile range: This is the best way to show the spread of the absolute prediction error.
- Percentage of eyes with a PE within a given interval (e.g., ± 0.50 D): This is probably the easiest way to report and remember the accuracy of any IOL power formula. The percentage of eyes with a PE within ± 0.25 D is quite useful to predict the refractive outcomes and expectations for patients receiving multifocal IOLs, where the tolerance to refractive errors is minimal. The percentage of eyes with a PE within ± 0.50 D is the most commonly reported value and can be used as a method to rank formulas.

Analyzing the Data

Once data are collected, they have to be analyzed. As previously stated, the first statistical analysis should investigate whether the PE is normally distributed. For this purpose, Shapiro-Wilk test and Kolmogorov-Smirnov test are probably the two most commonly used tests and are available with the majority of statistical software. Unfortunately, formal normality tests are notoriously affected by large samples in which small deviations from normality yield significant results. In other words, they display a higher probability of rejecting the null hypothesis of normality as sample size increases: for large samples ($n > 300$), these formal normality tests may be unreliable [20]. In this context, it is wise to refer to the central limit theorem [21], according to which in large samples ($n > 30$) the sampling distribution tends to be normal anyway, and to probability-probability plots (P-P plot): these graphs plot the cumulative probability of a variable (the PE) against the cumulative probability of a normal distribution. If values fall on the diagonal of the plot, then the variable is normally distributed [22].

The following questions should then be answered:

- Is the mean PE statistically significant from zero? If data are normally distributed, then one sample *t*-test is recommended; if the distribution is not normal, Wilcoxon rank sum test should be used.
- Is the SD of the PE statistically significant among formulas? Under the assumption that the values for each formula are matched (paired), if data are normally distributed, repeated-measures ANOVA with post-test is recommended; otherwise, the Friedman test with post-test should be used. Recently, the heteroscedastic method has been recommended [3]. This test can be used to compare the SD of different formulas when the PE distribution is not normal and is able to detect statistically significant differences that are missed by the Friedman test. Its main limitation is that it is difficult to use.
- Does the absolute error generated by the formulas under investigation show any statistically significant difference? Since the absolute error never has a normal distribution, nonparametric tests such as the Friedman test should be used.
- Does the percentage of eyes with a PE within ± 0.50 D (or ± 0.25) D show any statistically significant difference among formulas? Cochran's Q test is recommended for this purpose.

References

1. Hoffer KJ, Aramberri J, Haigis W, Olsen T, Savini G, Shammas HJ, Bentow S. Protocols for studies of intraocular lens formula accuracy. *Am J Ophthalmol.* 2015;160:403–5.
2. Wang L, Koch DD, Hill W, Abulafia A. Pursuing perfection in intraocular lens calculations: III. Criteria for analyzing outcomes. *J Cataract Refract Surg.* 2017;43:999–1002.
3. Holladay JT, Wilcox RR, Koch DD, Wang L. Review and recommendations for univariate statistical analysis of spherical equivalent prediction error for IOL power calculations. *J Cataract Refract Surg.* 2021;47:65–77.
4. Hoffer KJ, Savini G. Update on intraocular lens power calculation study protocols: the better way to design and report clinical trials. *Ophthalmology.* 2021;128(11):e115–20.
5. Katz J, Zeger S, Liang KY. Appropriate statistical methods to account for similarities in binary outcomes between fellow eyes. *Invest Ophthalmol Vis Sci.* 1994;35:2461–5.
6. Murdoch IE, Morris SS, Cousens SN. People and eyes: statistical approaches in ophthalmology. *Br J Ophthalmol.* 1998;82:971–3.
7. Hardin JW, Hilbe M. Generalized estimating equations. 2nd ed. New York: Chapman & Hall; 2012.
8. Langenbucher A, Schwemm M, Eppig T, Schröder S, Cayless A, Szentmáry N. Optimal dataset sizes for constant optimization in published theoretical optical formulae. *Curr Eye Res.* 2021;46(10):1589–96.
9. Lyle WA, Jin GJC. Prospective evaluation of early visual and refractive effects with small clear corneal incision for cataract surgery. *J Cataract Refract Surg.* 1996;22:1456–60.
10. Masket S, Tennen DG. Astigmatic stabilization of 3.0 mm temporal clear corneal cataract incisions. *J Cataract Refract Surg.* 1996;22:1451–5.
11. Nejima R, Miyai T, Kataoka Y, Miyata K, Honbou M, Tokunaga T, et al. Prospective inpatient comparison of 6.0-millimeter optic single-piece and 3-piece

- hydrophobic acrylic foldable intraocular lenses. *Ophthalmology*. 2006;113:585–90.
12. Simpson MJ, Chairman WN. The effect of testing distance on intraocular lens power calculation. *J Refract Surg*. 2014;30:726.
 13. Haigis W, Lege B, Miller N, Schneider B. 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.
 14. Hoffer KJ. The Hoffer Q formula: a comparison of theoretic and regression formulas. *J Cataract Refract Surg*. 1993;19(11):700–12. Errata: 1994;20(6):677 and 2007;33(1):2–3
 15. Holladay JT, Prager TC, Chandler TY, et al. A three-part system for refining intraocular lens power calculations. *J Cataract Refract Surg*. 1988;14:17–24.
 16. 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. Errata: 1990;16:528 and 1993;19(5):444–446
 17. Savini G, Hoffer KJ, Balducci N, Barboni P, Schiano LD. Comparison of formula accuracy for intraocular lens power calculation based on measurements by a swept-source optical coherence tomography optical biometer. *J Cataract Refract Surg*. 2020;46:27–33.
 18. Taroni L, Hoffer KJ, Barboni P, Schiano-Lomoriello D, Savini G. Outcomes of IOL power calculation using measurements by a rotating Scheimpflug camera combined with partial coherence interferometry. *J Cataract Refract Surg*. 2020;46:1618–23.
 19. Melles RB, Holladay JT, Chang WJ. Accuracy of intraocular lens calculation formulas. *Ophthalmology*. 2018;125:169–78.
 20. Kim HY. Statistical notes for clinical researchers: assessing normal distribution. *Restor Dent Endod*. 2012;37:245–8.
 21. Kwak SG, Kim JH. Central limit theorem: the cornerstone of modern statistics. *Korean J Anesthesiol*. 2017;70:144–56.
 22. Hoffer KJ, Savini G. Reply. *Ophthalmol*. 2020;128(4):e21–22.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

