Systematic evaluation of wavefront-guided outcomes

Guy M. Kezirian, MD, Bruce Drum, PhD, Malvina Eydelman, MD

Purpose: To present a format for reporting outcomes of aberrometer-guided refractive procedures.

Setting: SurgiVision Consultants, Inc., Scottsdale, Arizona, and Food and Drug Administration, Center for Devices and Radiological Health, Rockland, Maryland, USA.

Methods: Reports of standard refractive and visual outcomes (uncorrected visual acuity, manifest refractive spherical equivalent, best spectacle-corrected visual acuity) should be provided for any refractive surgery report. Comparison of postoperative uncorrected visual acuity to preoperative best spectacle-corrected visual acuity should be included. Aberration reports should convert 2nd-order terms to refractions (measured in diopters) and use standard refractive reporting methods. Changes in coma, spherical aberration, and root-mean-square changes should be described using statistical methods for aggregate data. Underlying statistics should be reported.

Results: Aberration changes are well described by the mean error of the attempted versus achieved outcomes, comparison of the mean changes, and stability over time. Ancillary plots include histogram representation of the postoperative scores. Additional reports of visual function should be included, as appropriate.

Conclusion: Use of standardized tables and graphs permits qualitative and quantitative comparison of outcomes of refractive treatment with wavefront-guided lasers. Modifications of the recommended formats can be expected over time.

J Cataract Refract Surg 2005; 31:1306–1313 © 2005 ASCRS and ESCRS

Aberrometry-guided refractive surgery is currently available with several lasers in the United States. Other manufacturers are investigating or are interested in this option for their lasers. To be well-informed users, clinicians need to have the ability to evaluate and compare outcomes obtained using various treatment modalities.

Adoption of standard graphs and tables for outcome reporting can make this possible. The target audience, clinicians and surgeons, guides the format and content of the proposed standard. For example, the format should follow those already in use in the presentation of refractive surgery outcomes whenever possible. The use of mathematical and optical terms that may be unfamiliar to clinicians should be minimized. Above all, the proposed standard should permit the clinician to determine whether the treatments made the vision better or worse and should permit comparison of different data sets.

This paper suggests a format for the presentation of aberrometry-based outcomes that could provide the basis for the development of such a standard. To promote acceptance, the recommendations presented here incorporate the recommendations of other groups working to develop overall industry standards for the analysis, description, and reporting of wavefront outcomes. Accordingly, the goal is not to replace the work that has been done by the United States Food and Drug Administration (Checklist of Information Usually Submitted in an Investigational Device Exemptions Application for Refractive Surgery Lasers, Diagnostic and Surgical Devices Branch, Division of Ophthalmic Devices, Office of Device Evaluation, Food and Drug Administration, October 10, 1966), the American

Published by Elsevier Inc.
0886-3350/05/$-see front matter
doi:10.1016/j.jcrs.2005.01.028

National Standards Institute (ANSI) (ANSI Z80.11 Draft Standard for Laser Systems for Corneal Reshaping; ANSI Z80-28-200X Draft Standard for Methods for Reporting Optical Aberrations of Eyes), or the Optical Society of America (OSA) subcommittees[1] but to apply that work to create formats for data presentation that can be easily interpreted by clinicians.

Because data are reported for many purposes, it is not possible to anticipate every analysis that might be needed for reporting outcomes. This is not the goal of this project. Rather, the goal is to propose a common, minimum set of information and a format for its presentation to make it easier for clinicians to evaluate outcomes of aberrometer-based treatments of routine refractive errors.

Methods

Inclusion of Standard Outcomes Reports

Any meaningful presentation of refractive outcomes must include reports of the basic visual and refractive outcomes. Reports of clinical outcomes should follow the well-accepted formats proposed by Waring[2] and Koch.[3] They should include reports of uncorrected visual acuity (UCVA), manifest refractive spherical equivalent outcomes, changes in best spectacle-corrected visual acuity (BSCVA), refractive stability, and adverse events as described by ANSI (ANSI Z80.11 Draft Standard for Laser Systems for Corneal Reshaping).

Accepted for publication January 26, 2005.

From SurgiVision Consultants, Inc. (Kezirian), Scottsdale, Arizona, and the U.S. Food and Drug Administration, Center for Devices and Radiologic Health (Drum), and Department of Health and Human Services, Food and Drug Administration, Rockville, Maryland, USA.

Presented at the 5th International Congress of Wavefront Sensing and Optimized Refractive Corrections, Whistler, British Columbia, Canada, February 2004.

No author has a financial or proprietary interest in any material or method mentioned.

The following individuals contributed in the preparation of this manuscript: David A. Atchison, PhD, FAAO (School of Optometry, Queensland University of Technology, Queensland, Australia); Michael Mrochen, PhD (Institute for Biomedical Engineering, University of Zurich and Swiss Federal Institute of Technology, Zurich, Switzerland).

The comments presented here represent the professional opinions of the authors and do not constitute an official document, guidance or policy of the U.S. Government, the Department of Health and Human Services, or the Food and Drug Administration, nor should any official endorsement be inferred.

Reprint requests to Guy M. Kezirian, MD, 28071 North 90th Way, Scottsdale, Arizona 85262, USA. E-mail: Guy1000@SurgiVision.net.

Inclusion of information about basic refractive and visual outcomes provides the clinician with the ability to evaluate new data against known technologies and to consider the results of higher-order aberration treatments in the context of visual and refractive outcomes. If basic refractive outcomes are unacceptable, then consideration of higher-order outcomes is of secondary concern.

Reports of Aberration Outcomes

Reports of outcomes related to the treatment of higher-order aberrations should emphasize the aberrations that are most likely to affect visual function. These include defocus, cylinder, spherical aberration, and coma, using Zernike terms through the 4th order. Inclusion of higher-order terms may be appropriate in some series, but in keeping with the minimum common standard goal of this project, it is not required.

A standard set of reports that describe outcomes for defocus, cylinder, spherical aberration, and coma comprise the major clinical outcome measures for refractive correction. Correlations of these aberrations with visual function are becoming better understood and are clinically relevant.[4,5] Conversely, the functional impact of aberrations such as trefoil, quatrefoil, secondary coma, and spherical aberrations is less well understood and is therefore of secondary concern to the practicing clinician.

Terminology and Reporting Conventions

The working convention for the interpretation and reporting of visual aberrations is to use Zernike polynomials, which by now are familiar to most clinicians. Zernike polynomials describe a series of optical aberrations that are commonly represented in a pyramid with rows representing progressively higher orders (level of the exponent) of terms. The OSA[6] and ANSI (ANSI Z.28-200X Draft Standard for Methods for Reporting Optical Aberrations of Eyes) have recommended the use of a double-notation system for the polynomials and their associated coefficients. This notation should be used when describing aberrations, in preference to the prior single-index notation. Table 1 provides a list of terms and their double-notation formats.

The shape of each aberration is described by its polynomial term; the amount that the aberration contributes to the wavefront is described by the coefficient that precedes it. The coefficients have positive and negative signs to indicate the direction of the aberration, as described in Table 1. These signs are important when performing outcomes analysis because they allow for the assessment of overcorrection or undercorrection. This ability is lost with root-mean-square (RMS) calculations because RMS is the square root of the mean of the squares of each coefficient, and when numbers are squared, they always become positive. The RMS provides

Table 1. Terms and formats for the double-notation system.

Zernike Polynomial	Prior Name	Normalization Factor	Name
Z_2^{-2}	Z3	$\sqrt{6}$	Oblique astigmatism
Z_2^0	Z4	$\sqrt{3}$	Defocus
Z_2^2	Z5	$\sqrt{6}$	With-against-the-rule astigmatism
Z_3^{-3}	Z6	$\sqrt{8}$	Oblique trefoil
Z_3^{-1}	Z7	$\sqrt{8}$	Vertical coma
Z_3^1	Z8	$\sqrt{8}$	Horizontal coma
Z_3^3	Z9	$\sqrt{8}$	Horizontal trefoil
Z_4^{-4}	Z10	$\sqrt{10}$	Oblique quatrefoil
Z_4^{-2}	Z11	$\sqrt{10}$	Oblique secondary astigmatism
Z_4^0	Z12	$\sqrt{5}$	Spherical aberration
Z_4^2	Z13	$\sqrt{10}$	With-against-the-rule secondary astigmatism
Z_4^4	Z14	$\sqrt{10}$	Quatrefoil

The nonshaded rows list the Zernike terms needed to create the data tables and graphs in the proposed standardized format. The first column lists the Zernike polynomials, with the subscript representing the order and the superscript representing the frequency. In the familiar pyramid representation of the Zernike shapes, the order corresponds to the row and the frequency corresponds to position in the row. A frequency of 0 indicates the shape is in the center; negative numbers, shapes on the left side and positive numbers, shapes on the right. Z_4^0, or spherical aberration, is in the center of the fourth row of the Zernike pyramid. The Prior Name column shows the shorthand name previously used to describe the Zernike polynomials; Z_4^0 can also be referred to as Z12. As described by Atchison,[7] multiplying the Zernike coefficient by ''the normalization factor makes the variance of each Zernike polynomial function as 1. This means that the square of the coefficient of the Zernike polynomial function represents the contribution of that polynomial function to the total variance of the wave aberration. Summing the squares of the coefficients for Zernike polynomial functions of a particular order gives the contribution of that order to the variance. Here variance has the same meaning as it does in statistics. Many instruments use the RMS of the wave aberration (the square root of the variance) as a measure of image quality.'' The fourth column provides the description of each Zernike term.

a good assessment of the overall deviation of the wavefront from a planar or ideal form and can be valuable. However, when analyzing results of treatments of specific terms (sphere, cylinder, coma, or spherical aberration), it is useful to retain the signs in the analysis.

Manufacturing Standards

The ANSI is planning to undertake development of a standard for wavefront-guided lasers shortly. At present, however, little standardization exists among the manufacturers of the various lasers and aberrometers. Therefore, comparing results from different technologies is still challenging. The OSA and ANSI have proposed standards for reporting optical aberrations in eyes that are gradually being adopted by most companies. When performing outcomes analysis, it is important to know the methods used by the companies in presenting the aberrometry.

One key item is the use of the OSA normalization terms, which were developed to standardize the variance of each Zernike polynomial to 1 so the square of the coefficient of the Zernike polynomial function represents the contribution of that polynomial function to the total variance of the wave aberration.[7] Not all companies have adopted this standard in their aberrometer software. If they have not, the coefficients should be multiplied by the normalization terms before the analysis is performed. This should be described in the outcomes report.

Refraction Calculations

The coefficients of the 2nd-order terms (defocus and cylinder) can be used to calculate a spherocylinder refraction. As described by Atchison,[7] higher-order terms can also be incorporated into this calculation; this can affect the result if there is a significant amount of spherical aberration present. The pupil size used in the calculations also affects the result because larger pupils allow more peripheral information.

Therefore, when reporting refractive (or higher-order) outcomes calculated from aberrometry, it is important to report the terms and pupil size used in performing the calculations.

Format for Reporting Aberrations

Aberrometry defocus describes the spheroequivalent error. It can be used as such to report changes in spheroequivalent. If reporting sphere and cylinder separately, the sphere can be derived by subtracting half the cylinder amount from the defocus.

There are 2 2nd-order cylinder terms: Z(2, –2) and Z(2, 2), previously referred to as Z3 and Z5. The former describes oblique astigmatism and the latter describes with- and against-the-rule astigmatism. The 2 cylinder terms can be combined into 1 by converting the terms to refraction notation and representing them as a single cylinder amount and axis (cylinder vector). Vector analysis[8,9] can then be used to determine the difference between the preoperative and postoperative cylinder amounts to assess the relationship between the attempted and achieved outcomes.

Coma can also be combined into a single term. However, coma is a localized aberration; unlike cylinder, its orientation must be described by a 360 reference. Combining coma to a single term therefore requires a 4-quadrant representation and terminology that may be new and unfamiliar to clinicians. Therefore, rather than combining coma into 1 term, the recommendation here is to report and analyze the oblique and with and against the rule terms separately.

Data Presentation

Data may be presented in tables or graphs. Both formats include the same information, and sometimes it may be desirable to present both formats. In most presentations, 1 or the other will be preferred.

The recommendation here is to present information graphically whenever possible and to use the formats described below. Accompanying tables may be provided to present underlying or descriptive statistics such as confidence intervals (CIs), ranges, standard deviations SDs, skew, and kurtosis when appropriate.

Histogram representation of data requires the definition of data bins that determine the categories into which data are classified. The present recommendation is for refractive data bins to span 0.50 diopter (D) intervals and for aberrometry data to span 0.2 μm intervals. This is not based on an ideal but reflects the common behavior of current data sets. Finer delineations are permissible, especially in the smaller ranges (such as ±0.25 D outcomes), but larger groupings may prevent differentiation of outcomes and are to be discouraged.

Analysis of "lines of change" of visual acuity should be calculated using logMAR equivalents[10] and grouped using whole-line differences. Partial lines can be considered in the conversion to logMAR, but the results should be grouped into whole-line data groups. Changes of less than 1 line should be ignored.

Analysis of Treatment Efficacy

When analyzing aggregate data that span plus and minus ranges, it can be challenging to differentiate between overcorrections and undercorrections. A common response to this is to use the RMS values, but this introduces the possibility of ambiguity for example, in a population that averaged −0.12 μm of spherical aberration preoperatively and +0.12 μm postoperatively. The RMS values will appear unchanged, even though the result was an overcorrection of 100%.

Scatterplots of attempted versus achieved corrections can graphically depict changes but may be difficult to interpret and compare, especially when the data span 4 quadrants. For example, a coefficient that began as −0.2 μm and became worse will show up in the upper left-hand quadrant of the plot. This can be confusing.

An alternate statistic, which has been called the correction efficacy, is obtained by dividing the achieved correction by the attempted correction. It can be converted to a percentage that is consistent whether the initial values were positive or negative. The statistic can be calculated for each measure (eg, spherical aberration) for each eye and then grouped into categories (undercorrected, on target, overcorrected) for presentation in a histogram.

The correction efficacy statistic can be averaged to summarize aggregate outcomes into 1 value called the mean correction efficacy. This is a single number (with a CI and SD) that represents the mean of the ratios of the achieved to attempted amounts for each eye.

A correction efficacy of 100% indicates a 1:1 ratio of attempted to achieved for the data set. On an attempted versus achieved scatterplot, eyes with a 100% correction efficacy would fall on the 1:1 line. A further refinement of this statistic is to subtract 100% from the result to represent the error that occurred rather than the efficacy ratio. This is called the mean error of attempted versus achieved.

Unlike RMS, the mean error statistics has a sign that indicates whether the result was in the direction desired or the opposite directing. A mean error of 0 indicates the results were, on average, on target. Positive values for the mean error indicate overcorrection, and negative numbers undercorrection. A mean error of −100% indicates that the preoperative and postoperative values are the same; in other words, there was no effect. Numbers less than −100% indicate the procedure made the aberrations worse.

Additional Statistics

When comparing outcomes, especially those involving mean values, it is helpful to know the characteristics of the underlying data. Several statistics are relevant, but only a few are needed to allow adequate interpretation of most data sets. They are the SD and the CI and probability (P) values of observed changes.

The SD is familiar to most clinicians and describes how widely data are scattered about the mean value. Two thirds of the data fall within ±1 SD of the mean and 95% fall within ±2 SDs. The SD of a population does not usually change as more data are added; it is a description of the diversity contained within the data. Reports of mean values should list the SD.

The CI represents the likelihood that the mean value is accurate to a certain level of confidence. It is a function of the SD, the size of the data set, and the desired confidence level. It is common to use the 95% CI when reporting clinical outcomes. Because the CI is a function of the number or data points, it improves with the size of the data set. The CI should also be listed in all reports of mean outcomes.

Probability values should be reported whenever the difference between 2 means is calculated.

Data Selection and Accuracy Measures

The methodology of capturing aberrometry outcomes should be described in all reports. Because imaging reproducibility may vary in some eyes, we recommend that aberrometry examinations should include at least readings to confirm reliable images. Each of the images used for analysis should meet the manufacturer's validation standards. The values from the readings should be averaged and the result used for statistical analysis. An acceptable alternate approach is to use the median value. Either way, the CI of the readings (CI_{EXAM}) should be calculated for each coefficient that is used in the analysis. Evaluation of the mean CI_{EXAM} (the mean sampling CI) for the data points in an aggregate data set indicates the reliability of the aberrometry readings for the coefficient.

Most aberrometers do not currently report the CI_{EXAM} for each coefficient. Manufacturers are encouraged to incorporate this statistic into their software. Ideally, outcomes reports would describe the mean of the CI_{EXAM} values as an indication of the reliability of the information being presented. This valuable information could be calculated by the aberrometer software at the time the readings are taken and averaged. As a further aid, aberrometer software should be made to permit electronic export of all data.

This mean CI_{EXAM} statistic is very important because the measurement of many higher-order aberrations can vary widely from reading to reading. If the CI of the readings is too large, the data cannot be interpreted and should be rejected.

Combining Data for Right and Left Eyes

When combining data for right and left eyes in 1 analysis, both the OSA and ANSI recommend reversing the sign of the coefficients for horizontal terms in left eye. If the analysis is restricted to the 4th order for sphere, cylinder, coma, and spherical aberration, only the horizontal coma term Z(3, 1) is affected.

Results

The data presented in the examples are provided for illustration only. The graphs presented below assume the inclusion of the standard reports of refractive outcomes, as described above.

Figure 1 includes 4 graphs that are recommended for inclusion in all reports of aggregate clinical outcomes for wavefront-guided refractive surgery treatments. As will be shown, it is necessary to interpret all 4 graphs together to understand the outcomes. Together, these plots provide a summary description of the outcomes.

At the top left is a histogram presentation of the postoperative UCVA versus preoperative BSCVA. The *x*-axis labels indicate the lines difference. A table is provided to indicate the mean, SD, and CI statistics. The trend toward slight improvement is demonstrated.

The graph at the bottom left shows the mean error of the attempted versus achieved with an SD bar for each aberration and for the overall RMS values. As noted above, 0% error indicates the mean was on target, and values less than −100% indicate the statistic was worse after the procedure than before. The plot shows that sphere and cylinder were improved on average. Coma terms show small mean errors of attempted versus achieved but with huge SDs. Global RMS values were reduced to 25% of the original value and higher-order RMS (3rd order and above) was worse than preoperatively by 39%.

The plot at the top right of Figure 1 shows the mean values (with their 95% CIs) for each aberration, with a *P* value indicating the significance of the difference in means. Although the mean error of the attempted versus achieved (Figure 1, bottom left) for vertical coma was close to 0, the change in means was not significant, demonstrating the importance of considering the statistical significance of differences when considering the outcomes. Note that it is also possible to superimpose the SD values on this graph, if desired.

The final plot shows the stability of the higher-order RMS values over time along with the CIs for each value. This format can be applied to the mean values for the other aberrations and should be provided if the report includes measurements for more than 1 postoperative interval.

Figure 2 shows a series of histogram plots that depict the distribution of outcomes for each aberration. These plots provide further detail about the information that is summarized in the plots shown in Figure 2. They

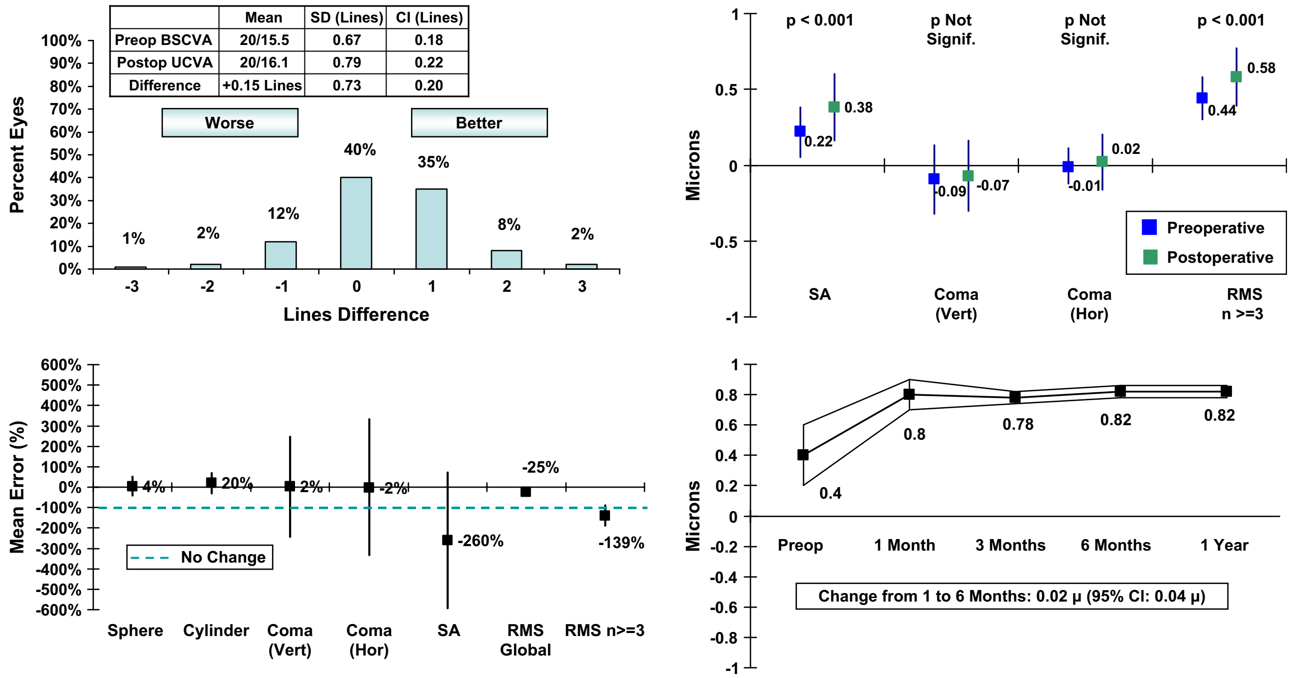

Figure 1. Recommended set of graphs for inclusion in reports of clinical outcomes for aberrometry-guided refractive surgery. *Top left*: Comparison of postoperative UCVA with preoperative BSCVA. Table can be embedded in graph or printed separately, but statistics should be provided. *Bottom left*: Mean correction efficacy. Graph shows mean with SD bars for each value. The large scale in the *y*-axis is needed to represent the data. An alternate format would be to plot the mean with bars representing the CI and provide a separate table to list the supporting statistics. *Top right*: Mean preoperative and postoperative values for each aberration described, with bars representing the 95% CIs. *P* values should be provided whenever means are compared. *Bottom right*: Stability plot showing the mean and 95% CIs for each value. The RMS is plotted here; similar plots may be needed to describe each aberration. The comment should state the change in means for the intervals used to confirm stability.

should be provided whenever the summary plots do not adequately describe the data.

Discussion

Comprehensive analysis of wavefront outcomes is potentially complex. Evaluation of wavefront outcomes at the clinical level does not have to be complex. The evaluation of defocus and cylinder terms is made easier by converting these terms to their familiar refractive equivalents. Outcomes for coma and spherical aberration can be evaluated by considering attempted versus achieved outcomes and comparing mean preoperative and postoperative values.

Any report of clinical outcomes obtained using aberrometer-guided treatments must also describe visual and refractive outcomes. Reports describing perfect higher-order outcomes without information about basic measures of visual function have little clinical relevance for routine refractive corrections.

At present, the higher-order aberrations of primary concern in primary treatment outcomes are spherical aberration and coma. Reports of higher-order aberration outcomes should include summary outcomes that permit qualitative and quantitative evaluation of the surgically induced change. The graphs in Figures 1 and 2, and their underlying statistics, provide the information needed to make this evaluation and to answer the questions: Is it better or worse and by how much?

The recommendations provided here do not describe the complete list of the analyses that could be applied to wavefront outcomes. For example, changes in contrast sensitivity, low contrast acuity, subjective patient questionnaires, and other outcomes add to the understanding of the effects of treatments on visual function and should be reported as appropriate.

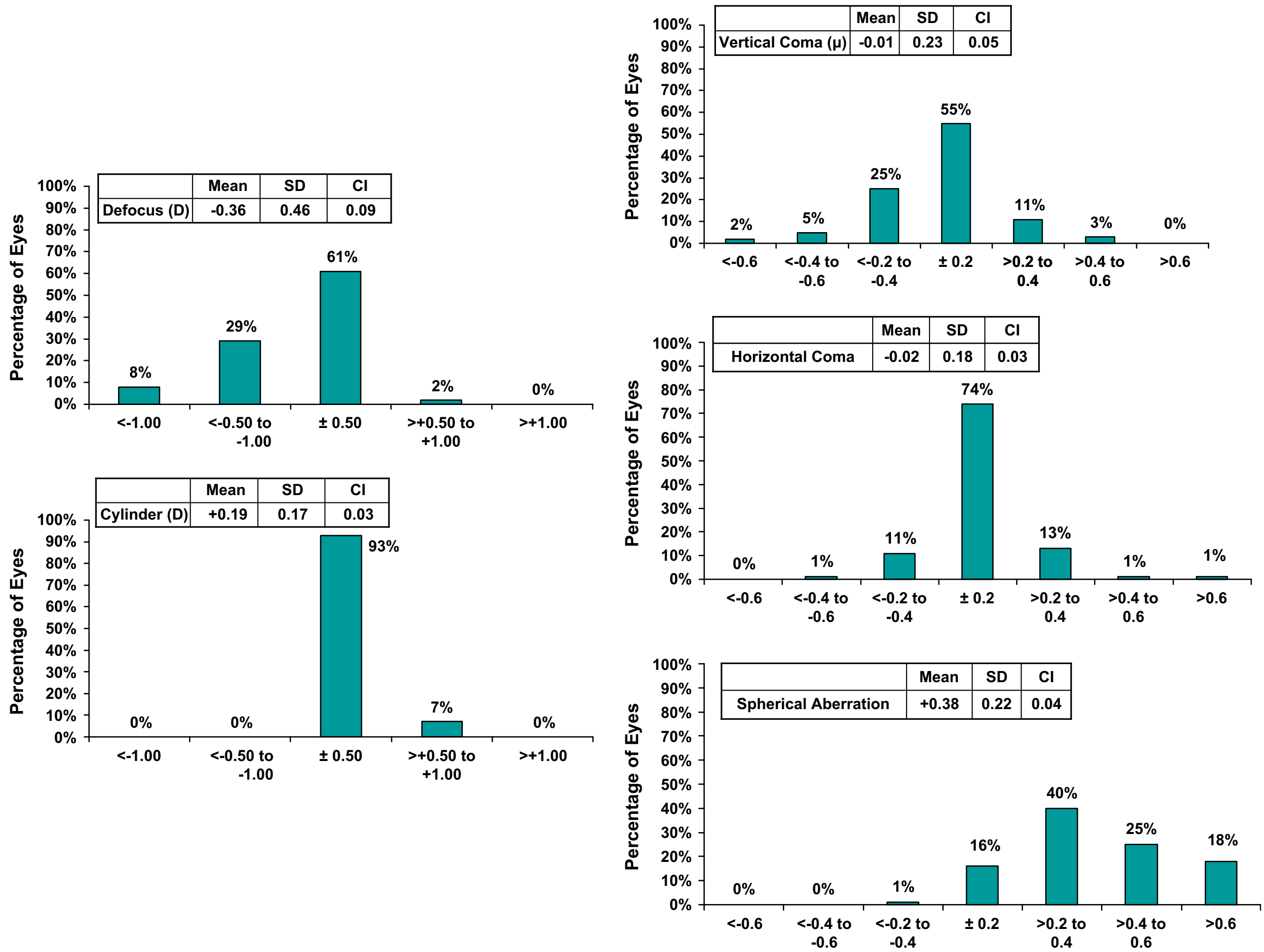

Figure 2. Ancillary plots of the distribution of postoperative aberrations. The mean, SD, and CI values should be listed on the graph.

These recommendations also do not permit a comprehensive description of aberrometer-guided treatments. This is not the goal because comprehensive reports are not usually necessary to assess clinical outcomes or compare results with different technologies. The reports recommended here should be adequate for the purpose of most clinicians. They may not be adequate to judge the safety and effectiveness of devices for regulatory approval or for research and development.

Industry has provided ophthalmologists with many new tools for the treatment of refractive disorders. Clinicians must be able to evaluate the results obtained using new technologies to responsibly decide whether and when to use them. This paper suggests a set of tables for reporting outcomes obtained with aberrometry-guided treatments. It is up to the scientific and clinical communities to decide whether they should become a standard format.

References

1. Thibos LN. Wavefront data reporting and terminology. J Refract Surg 2001; 17:S578–S583
2. Waring GO. Standard graphs for reporting refractive surgery. J Refract Surg 2000; 16:459–466
3. Koch D. Measuring patient outcomes after refractive surgery. J Cataract Refract Surg 2001; 27:645–646
4. Chalita MR, Xu M, Krueger RR. Correlation of aberrations with visual symptoms using wavefront analysis in eyes after laser in situ keratomileusis. J Refract Surg 2003; 19:S682–S686
5. Holladay JT, Dudeja DR, Chang J. Functional vision and corneal changes after laser in situ keratomileusis determined by contrast sensitivity, glare testing, and

corneal topography. J Cataract Refract Surg 1999; 25:663–669

6. Thibos LN, Applegate RA, Schwiegerling JT, Webb R, VSIA Standards Taskforce Members. Standards for reporting the optical aberrations of eyes. J Refract Surg 2002; 18:S652–S660
7. Atchison DA. Recent advances in representation of monochromatic aberrations of human eyes. Clin Exp Optom 2004; 87:138–148
8. Holladay JT, Moran JM, Kezirian GM. Analysis of aggregate surgically induced refractive change, prediction error, and intraocular astigmatism. J Cataract Refract Surg 2001; 27:61–79
9. Alpins N. Analysis of aggregate surgically induced refractive change, prediction error, and intraocular astigmatism. J Refract Surg 2001; 17:705–707
10. Holladay JT. Proper method for calculating average visual acuity. J Refract Surg 1997; 13:388–391