

The Role of Biostatistics in the Quality Improvement of Refractive Surgery

Keith A. Hettinger, MS, MBA

ABSTRACT

PURPOSE: To demonstrate how the large quantity of uniformly collected data available to a corporate refractive surgery provider, Optical Express, is applied to drive improvements in patient outcomes.

METHODS: Optical Express employs a skilled team of biostatisticians to analyze the information in its electronic medical records database of over 5,500,000 patient records. The techniques used to ensure high data quality and the selection of statistical methods used in making data-driven clinical decisions are described. The importance of appropriate statistical methods is demonstrated in an example in which the effect of age on refractive outcomes in low myopes is studied. The use of a corporate database in prospective and retrospective analyses is detailed.

RESULTS: By providing the resources necessary to interpret the information in Optical Express' medical records database, the biostatistics department has helped Optical Express refine its procedures and improve surgical protocols and patient outcomes.

CONCLUSIONS: Biostatistical analyses help transform the large quantities of uniformly collected clinical data available to a corporate surgery provider into information that can be applied to improve clinical practice. Such data-driven process improvements play a key role in improving patient outcomes. [*J Refract Surg.* 2009;25:S651-S654.] doi:10.3928/1081597X-20090611-05

W

ith data on over 5,500,000 patients, the Optical Express clinical outcomes database has the potential to shed light on a wide range of medical and surgical questions. In addition to evaluating the performance of individual surgeons, this database is used to answer broad questions about the safety and efficacy of various refractive surgery procedures and the factors influencing patient outcomes.

Optical Express employs a biostatistics department that ensures the appropriateness of statistical methodology and the quality of data. By ensuring that the conclusions drawn from this data are statistically valid, the biostatistics department is essential to the company's effort to answer a range of questions about both the science and business of refractive surgery.

ENHANCING OUTCOMES WITH DATA-DRIVEN DECISIONS

The primary function of data analysis is to promote evidence-based decisions that improve patient care. By analyzing data from tens of thousands of patients, statistical models can be developed that identify the key factors that impact surgical outcomes. Surgeons can then use these models to modify treatment variables and improve patient outcomes.

In addition to helping with the large-scale analyses reported in this supplement, the biostatistics department performs a variety of other assessments. For example, surgeon reviews are performed biannually to provide quality control information about each surgeon's performance. As part of these reviews, surgeons' performance is compared in a way that takes into account the patient population that each surgeon treats and the types of procedures he or she performs. Thus, the scoring system developed by the biostatistics department gives a valid "apples to apples" comparison.

Finally, the biostatistics department responds to queries regarding the merits of anecdotal information. For example,

From the Biostatistics Department, Optical Express, San Diego, Calif.

The author is an employee of the Optical Express Group.

Assistance in preparing this manuscript was provided by medical writer Kay Downer.

Correspondence: Keith A. Hettinger, MS, MBA, 4520 Executive Dr, Ste 230, San Diego, CA 92121. Tel: 858.597.0897; Fax: 858.597.0896; E-mail: keith.hettinger@opticalexpress.com

if a clinician notices increased induction of cylinder in a small series of patients, the biostatistics department can perform a retrospective analysis to determine whether the findings represent a significant trend or are just a statistical anomaly. If the former, the biostatistics department can perform additional analyses, including prospective studies, to determine the cause of this trend.

IMPORTANCE OF DATA SAFETY AND QUALITY

The Optical Express database is a highly systematized collection of electronic medical records, which uses a scalable private Multi Protocol Label Switching (MPLS) network. The data center where the information is hosted has multiple features such as redundant air supply, redundant power capabilities, and security measures. This system (Tier 4) is considered the highest level by the Telecommunications Industry Association. Security provisions include redundant communication links, fire suppression, intruder detection, private power substation, diesel generators, and 24-hour security. All databases are held in Microsoft SQL 2005 (Microsoft Corp, Redmond, Wash) using Windows authentication and role-based security. There are separate servers in three locations across two continents.

As with any such collection, data accuracy must be maintained to ensure the validity of conclusions drawn from analyses of the data. Optical Express' electronic medical records system includes features that help enhance data quality, including range validations, range restrictions, controls on applicable data types, and comment boxes to verify unexpected values. In addition to these features, everyone who enters data into the Optical Express system is continually reminded to be vigilant about data accuracy.

The Optical Express system is independently audited by Registrat Incorporated (Lexington, Ky), a third-party clinical research organization that specializes in data management. The intention of these independent audits is to provide an unbiased assessment of the data accuracy.

In an effort to further promote accuracy, the biostatistics department also routinely performs a thorough internal audit on all data to identify and query extreme values and other aberrations. For example, by comparing the distribution of preoperative visual acuity measurements, statistical analyses can identify optometrists who may require more training to obtain a precise and accurate refraction for each patient.

IMPORTANCE OF STATISTICAL METHODS

Although data quality is essential for accurate results, using appropriate methods for all analyses is

also crucial. The complexity of statistical questions can vary, but all analyses require proper methods to achieve valid results.^{1,2} To select the correct type of statistical test for a given application, statisticians consider a multitude of issues, including the inherent distribution of the data, the data type available, and whether an assumption of normality is met.

As an example of how different methods can affect the outcome of a statistical analysis, consider the question of patients versus eyes. If a procedure is performed binocularly on 50 patients, should the data be analyzed as 50 patients or 100 eyes? Because the treatments were binocular, one theorized approach is to average both eyes for a single patient and perform the analysis for 50 patients. Another approach is to disregard the fact that data on two particular eyes come from the same patient and perform the analysis on 100 eyes as though each eye was from a different patient.

Intuitively, clinicians know that neither of these extremes is accurate. Because two eyes from a single person are more similar than two eyes from two different people, the statistical analysis should not disregard the fact that eyes come in pairs. But because some patients may have eyes that are different, averaging results from a patient's two eyes will result in a loss of information. Appropriate statistical methodology will therefore account for the similarity between a patient's two eyes while still maintaining each eye as an independent data point. Rather than being an analysis of 50 patients or 100 eyes, a better approach may be an analysis performed on 50 pairs of eyes.

REAL-WORLD APPLICATIONS

To demonstrate this approach, consider a nomogram adjustment designed to account for the possible influence of patient age in the treatment of low myopia (≤ -3.00 diopters [D] with ≤ -1.50 D of astigmatism). For this analysis, the statistician will first set the null hypothesis: age does not affect outcomes for low myopia treatment. Whether this hypothesis is confirmed or rejected depends on the method used to address the "patients versus eyes" question.

As the first step in this analysis, a generalized linear model is created to determine whether patient age significantly affects the postoperative mean spherical equivalent refraction. Using a Simple Random Sample Without Replacement, 1000 randomly selected eyes (507 patients) with low myopia were included in this analysis; patient age ranged from 18 to 65 years. The follow-up time used for the analysis was 1 month for all patients.

Next, the analysis must address the "patients versus eyes" question. The simplest approach is to make no

adjustment to the model and assume that a patient's two eyes behave completely independent of one another. In this approach, 100 eyes from 50 patients would be treated the same as 100 eyes from 100 patients. Using this method, an analysis performed using actual data from the Optical Express database yields $P=.0013$, and the null hypothesis is rejected.

Because of common factors such as identical genes and measurement conditions, however, some correlation between a patient's two eyes is expected. Therefore, a second analytical method takes a "patient centered" approach, assuming a correlation exists between a patient's eyes and accounting for it by using an average of both eyes for the analysis. This approach yields $P=.2178$, which confirms the null hypothesis.

Finally, a third method—which I believe yields the most accurate result—accounts for the possible correlation between a patient's eyes by adjusting the model for intraclass correlation.³ With the use of repeated measures, this method treats each patient as a cluster of data points, and an adjustment is made within that cluster. The result of this method is $P=.0895$, which confirms the hypothesis that age does not significantly affect the treatment of low myopia.

As this example shows, different methods of statistical analysis can yield different conclusions. Compared to the most accurate statistical method (presented last), the first option overestimates the P value and incorrectly rejects the null hypothesis. Likewise, whereas the second method correctly confirms the hypothesis, it nonetheless underestimates the relationship that age may play in the treatment of myopia. In-house biostatisticians ensure the application of appropriate methods for each type of analysis.

DEMONSTRATING IMPROVEMENTS IN CLINICAL OUTCOMES

The clinical benefit that can be achieved through statistical analyses is demonstrated by the following example. Following a comprehensive analysis of the sphere adjustment applied to wavefront-guided treatments, Optical Express surgeons implemented several changes to this adjustment. Specifically, the protocol was changed to specify that the treatment sphere (using a 4-mm wavefront sphere calculation on the aberrometer) should be adjusted to be equivalent to the manifest sphere if both the preoperative wavefront and manifest refraction were of high quality. Previously, the protocol was to either select a wavefront capture that was within the VISX PMA guidelines or to adjust the treatment sphere to be within 0.50 D of the manifest sphere.⁴ The analysis supporting the implementation of this change predicted that this modification would increase refrac-

TABLE
Preoperative and Demographic Characteristics of Patients Treated Before and After a Change in the Sphere Adjustment Applied to Wavefront-guided Treatments

	Before	After
No. of eyes	5799	5885
Gender (M/F, %)	33/67	33/67
Mean age (range) (y)	38.1±10.8 (18 to 69)	37.5±10.6 (18 to 70)
Mean SEQ (range) (D)	-3.77±2.13 (-0.25 to -12.125)	-3.75±2.02 (-0.38 to -12.125)

SEQ = spherical equivalent refraction

tive accuracy and allow a higher percentage of patients to achieve their target refraction.

Following the implementation of this change, the biostatistics team analyzed patient outcomes to evaluate its effect. The sample size for this analysis was extracted from the central Optical Express database using the following selection criteria: 1) all primary wavefront-guided LASIK procedures performed in the 3 months immediately before and 3 months after the change, 2) emmetropia was the refractive goal, 3) treatment of myopia, 4) preoperative cylinder ≤ 2.00 D, and 5) 1-month follow-up. This yielded a total of 11,684 patients; patients treated before the surgical change comprised one cohort ($n=5799$) and those treated following the implementation of the change comprised a second cohort ($n=5885$). This number included all patients in the database at the time period that met the requirements for analysis and had 1-month follow-up. The 1-month follow-up rate for this group was 89%. Preoperative and demographic characteristics were statistically similar for both groups (Table).

The results of this analysis revealed a statistically significant increase (3%) in the percentage of patients achieving 20/20 postoperative uncorrected visual acuity (UCVA) after the surgical protocol was updated ($P<.0001$, Fig). A statistically significant increase (4%) was also noted in the percentage of patients achieving 20/25 postoperative UCVA ($P<.0001$). This result shows that statistical analyses can be used both to drive improvements in treatment technique and to analyze the impact of these improvements.

CHALLENGES

Given the large size of the datasets being analyzed and the nature of the data collection in a real clinical

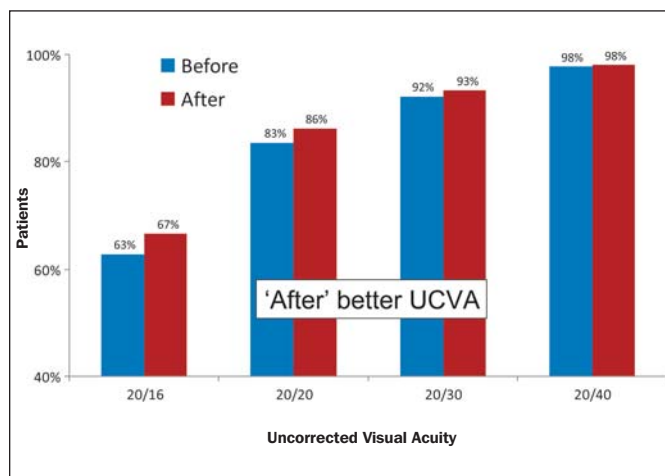


Figure. Patients treated after implementation of a change in treatment technique had better postoperative uncorrected visual acuity (UCVA) compared to patients treated prior to the implementation of this change.

setting, there are challenges that must be addressed. For instance, follow-up times can become a particularly important factor to consider. In a typical clinical setting not all patients will return for a 1-, 3-, 6-, or 12-month postoperative follow-up. Thus, when selecting a follow-up time point for an analysis it is important to examine the characteristics for the group of patients who did not make that clinic visit. Any clinical or statistically significant difference in preoperative or treatment characteristics of the “lost to follow-up” patients must be addressed for selection bias. In many cases, the initial analysis is completed with 1-month data, given the higher rate of follow-up. Additional analysis is then completed at later time points (3, 6, and 12 months) to ensure consistency of conclusions.

Another challenge is that with such a great amount of data comes a great amount of inherent variability. This may be due to a host of issues, such as variability in patient healing responses, differences in examination room chart illumination, or variations in surgeon technique. When analyzing multiple effects for hypothesis testing, building an appropriate model can become complicated. It is imperative to address all identifiable potential independent variables for effect as well as covariance and interaction.^{1,3} A lengthy and stepwise approach is warranted to ensure no lurking variables exist or incomplete conclusions are made.

A unique challenge arises in hypothesis testing when analyzing such large datasets. Historically, hypothesis testing was developed as a tool to infer significant relationships in smaller sample sizes. When these traditional tests are used in the presence of large amounts of data, statistical significance can become prevalent. To address this tendency, it is essential to always include an evaluation of clinical significance. If an explanatory variable is found to be statistically significant in one cohort, it is important to assess whether the mean difference between the two cohorts is clinically meaningful. Another practical way of addressing this tendency is to more regularly employ the use of random sampling. By sampling from these large datasets, more reliable *P* values can be created for inferential purposes. This random sampling testing can be done in a repeated manner.³ This will yield multiple *P* values from multiple random samples that can be analyzed for variability, thus providing a reliability estimate for our hypothesis testing conclusions.

The quantity of data that a large corporate provider such as Optical Express collects is a valuable resource. The company’s biostatistics department plays a key role in interpreting the data. In addition to selecting appropriate statistical methods for each type of analysis, the biostatistics department works to ensure the accuracy and consistency of all data entered into the Optical Express system.

As a result of these efforts, Optical Express can apply its data resources in areas ranging from internal quality control checks to large clinical studies, such as those presented in this supplement. The result is that the quantity of uniformly collected data available in the corporate environment can be used to improve the quality of patient outcomes.

REFERENCES

1. Fleiss JL, Levin B, Cho Paik M. *Statistical Methods for Rates and Proportions*. 3rd ed. Hoboken, NJ: Wiley John & Sons Inc; 2003.
2. Fleiss JL. *The Design and Analysis of Clinical Experiments*. Hoboken, NJ: Wiley-Interscience Publications; 1999.
3. Hansen PL. Large sample properties of generalized method of moments estimators. *Econometrica*. 1982;50:1029-1054.
4. Professional use information. VISX STAR S4 IR™ Excimer Laser System and WaveScan WaveFront® System. CustomVue™ Treatment. Available at: www.accessdata.fda.gov/cdrh_docs/pdf/P930016S020c.pdf. Accessed May 2009.