Introduction to Benchmarking Health Assessment and Outcomes through Applied Statistics

Save to myBoK

by Daniel P. Lorence, PhD

Through the application of some basic statistical techniques, HIM professionals can assist in healthcare benchmarking and management at the organizational level. The author offers an example of how one hospital used data analysis to make decisions that promoted better patient care.

Benchmarking is commonly defined as the process of identifying, understanding, and adapting outstanding practices and processes from organizations anywhere in the world to help an organization improve its performance. Essentially, it's a way to do things better in the workplace. Note that the definition uses the word "outstanding" and not "best." What is best for one organization depends on its unique situation and may not be the best solution adopted by another organization.

In today's healthcare systems, the analysis required to establish a benchmark usually begins with the compilation of standardized healthcare data. Increasingly, HIM professionals are recognized as key members of the outcomes assessment and benchmarking process. Through the application of some basic statistical techniques, HIM professionals can contribute to the burgeoning field of healthcare benchmarking and management at the organizational level, adding to their existing roles and building further career opportunities in one of healthcare's fastest-growing areas.

Measuring clinical processes and outcomes, such as surgical success rates, can be tricky business and normally requires assistance from clinical experts. Historically, physicians have been called upon to form a generalized judgment based on simple review of the patient's chart (or a small number of charts) for a given type of treatment. As patient management became more sophisticated, individual hospitals and clinics accumulated mountains of often-unused clinical data.

Most researchers agree that even the most experienced provider could benefit from this data by comparing his or her patients with thousands of other patients elsewhere, examining data on the types of treatment provided, and comparing resulting outcomes. Most of this data is now available through electronic media. Hospitals, along with state and federal government agencies, collect essentially comparable data sets. HIM professionals are among the few in a typical hospital who have a global outlook on clinical data, being responsible for the collection and management of large volumes of data rather than viewing it as individual patient charts or records. Making sense out of large volumes of data can be a challenge. But with experience, HIM managers can develop an intuitive sense about where there might be information available that leads to answers—or at least focuses the questions—related to clinical outcomes, procedures, and overall healthcare efficacy.

Applying Statistical Analysis to Healthcare Data

One of the biggest challenges HIM managers face when benchmarking with large volumes of data is the application of statistical methods to data analysis. Traditionally, quantitative measurement course work focused on statistical theory, often requiring the use of pen and paper or hand-held calculators to work through statistical formulas. Today, the development of low-cost sophisticated computer software makes high-power quantitative measurement available, so often only a fundamental knowledge of the theories behind the applications is required.

Rather than acting as a statistician, today's health data analyst serves more as a tour guide through the intricacies of hospital information. A favorite analogy used by teachers of health statistics likens statistical application and predictive statistical modeling to driving a tour bus through traffic.

Riding along, the tourist (or data analyst) may notice that every time the bus driver takes a certain action, the bus reacts in a certain way. He can turn the steering wheel left or right, and the bus will move to the left or right. If he steps on the accelerator or brake, the bus will reactively speed up or slow down. In other words, as the driver's actions vary, so does the behavior of the bus (i.e., they co-vary).

It may be useful to predict the behavior of the bus. If, for example, it is about to come to a sudden stop, we need to hold on to something beforehand to avoid being tossed forward. How can we predict the behavior of the bus? By watching the actions of the bus driver. As the driver's actions vary, so does the behavior of the bus (i.e., they co-vary). Measurement of this co-variance can be quantified through mathematical methods as a shorthand way of explaining and predicting relationships.

We should remember that the behavior of the bus may be different than that of a fine sports car. Turning the tiny steering wheel of a Ferrari just a few inches may cause it to turn sharply. Turning the huge wheel of a bus a few inches may only cause a slight change in direction, especially if it has a "loose" feel to it. So a minor action may cause (or predict) a large, noticeable change (or variance) in a vehicle's behavior or a relatively small change (or variance).

Co-variance is the essence of applied statistics, with predictive modeling consisting of simply watching and measuring how one or more causes (e.g., driver actions) influence a given effect (e.g., bus behavior). The causes are traditionally called independent variables and the effects dependent variables. Just how noticeably (or powerfully) each individual cause influences a given effect is measured by a coefficient (how sharply the vehicle is made to turn). So as the driver's actions vary, they may cause a degree of change (variance) in the behavior of the bus (i.e., a weak or a powerful change), as measured by the size of the coefficient.

The coefficient can have either a positive or negative value. So an increase in movement of the accelerator causes a corresponding increase in bus speed, and thus its coefficient will have a positive value. An increase in movement of the brake pedal causes a corresponding decrease in bus speed, and thus its coefficient will have a negative value.

If you can explain most of the behavior of the bus by certain driver actions, you have a good (robust) model. If there are certain key driver actions that you are ignoring, you lack robustness. Just how good your model is at predicting a certain dependent variable is measured by an indicator called the coefficient of determination (or R-squared) value. Unlike an independent variable coefficient, the R-squared value applies to the model as a whole, not just one (predictor) variable.

While this quaint metaphor gives only a basic introduction to statistical theory, it highlights the fact that certain healthcare outcomes can be predicted and optimized by examining key (independent) variables, using data readily available to the HIM professional. The key question: how does each independent, predictor variable (number of medical procedures performed, time in surgery, etc.) co-vary with the dependent variable (the healthcare outcome)?

An Example

When the decision was made to conduct an in-depth study of coronary bypass surgeries at Cook County Hospital in Chicago in late 1997, an organizational task force met to analyze data on 120 patients who had undergone coronary bypass surgery in the past year. Both the medical staff and management were concerned about the falling survival rates of patients who had undergone this operation. Although mortality rates for this procedure had fallen nationwide, Cook County Hospital's death rate for coronary artery bypass surgery remained surprisingly high.

Some task force members suggested that, compared to other local hospitals, the patients at Cook County were simply much worse off than average patients when they received the operation. This hospital had a fairly high proportion of Medicare and Medicaid patients, who as a group were more likely to delay seeking treatment. Some suggested the rates might have been the result of one or two underperforming surgeons who had since moved on to other opportunities. Some suggested that the hospital staff simply did not do enough of these operations to develop the necessary expertise in such a complex area, since everyone knows that "practice makes perfect." In narrowing down possible explanations to such problems, HIM professionals routinely participate in compiling and reviewing patient data and can often provide insight explaining why such a problem was occurring. Since mortality rates for this procedure had decreased in public hospitals across the country, Cook County's three-fold increase in mortality suggested that the problem was growing at an alarming rate.[1]

This type of analysis historically has been limited to the world of healthcare researchers or hospital quality assurance committees. More recently, the availability of data has enabled consumers to apply their own informed judgment in determining the quality of care, and as data becomes more accessible, hospital-specific quality becomes a publicly debated issue.

A group of local newspaper reporters, for example, conducted their own analysis and found that in a 33-month period, 207 patients at County had undergone the surgery. Fifteen of the patients died—a mortality rate of more than 7 percent, compared with a 2 to 3 percent rate at other hospitals treating similar patients.[2] By comparing these data with statistics from public hospitals in New York, Los Angeles, and Boston, they discovered the significantly higher mortality rates. At the Cleveland Clinic, a facility with considerably greater resources than most public hospitals, mortality rates for primary isolated bypass surgery were reported in 1993 as 1.4 percent for 1405 patients who experienced the operation. A more comprehensive nationwide study, using a population of 13,625 patients, found an average mortality rate of 2.0 percent.[3]

Applying a statistical model to this scenario, simple comparisons of quantifiable data were first made—in this case, mortality rates. Next, comparisons with nonpublic hospitals were deemed to provide an unfair comparison and were excluded from the study. Then, a number of factors were hypothesized to be predictive, independent variables that influenced surgical success. This involved, essentially, an in-depth analysis of various independent, predictive factors and how each affected the surgical outcome, addressing questions such as: How did the independent and dependent variables co-vary? Was the variance in operating times similar to the success of the operation? Was the variance among the experience of the surgeons similar to success of the operation? Many additional factors need to be examined in order to conduct a comprehensive study, most of which are measurable from data readily available to the HIM professional. While the process outlined here goes well beyond the traditional role of the medical records manager, it highlights an opportunity available to the manager who desires career growth and increased responsibility, along with the higher salary levels available for this increasingly important work.

The measurement and reporting of such benchmarks is no longer a discretionary practice, since accreditation requirements now mandate that hospitals and health providers establish their own benchmarks and quantify performance of their respective organizations. So not only does internally evaluating performance make good management sense, it may be a deciding factor as to whether or not a provider organization acquires or maintains accredited status.

Could it have been that there were factors which could be associated with better or worse outcomes of the bypass patients, such as the number of operations per hospital, or the number of operations per surgeon, or length of time per operation? HIM professionals have ready access to this information. Was it true that only the "sicker" patients were not surviving the operation? HIM managers have this information at their disposal. Did other hospitals have similar types of patients and similar outcomes? HIM managers can readily obtain this information from other provider organizations or regulatory agencies.

Conclusion

What tangible benefits result from statistical exercises? Generally, the simple act of focusing the attention of the organization on a specific project has a measurable effect. Often, a problem grows not because of deliberate actions taken, but as a result of forces or factors that exist beyond the normal scope of attention for the organization. Beyond this, analyzing a problem using hard data allows an objective, detached view of issues that are often laden with emotional bias and personal interest. Quantitative analysis, in simplest form, has the power to minimize the subjectivity of bias.

At Cook County Hospital, the results were practical and useful. Though hospital management concluded that most of the variation observed in mortality rates was due mainly to an increased severity of illness for the patient population, they nonetheless took steps to increase the amount of procedures performed at the hospital, recognizing the "practice makes perfect" philosophy. They also expanded the interaction of interdisciplinary providers and technicians involved in the procedure. Lastly, they clarified a number of administrative discrepancies that were believed to have an effect on the bypass mortality rates. For the most part, the specific problems identified and the resulting responses were due to objective analysis and scrutiny of data. Though the judgment and experience of the organizational leadership drove the ultimate decision, data analysis was a key tool leading to that decision and provided a defensible rationale for its implementation.

Notes

1. Illinois Health Care Cost Containment Council. "Report on Select Inpatient Procedures." Springfield, IL: 1997.

2. "Bypass Death Rate at County Grows," *Chicago Tribune*, 16 November, 1997.

3. Cosgrove, Delos M., Bruce W. Lytle, Floyd D. Loop, and Eric J. Topol. "Outmigration for Coronary Bypass Surgery in an Era of Public Dissemination of Clinical Outcomes." *Circulation* 93, no. 1 (1996): 30.

Daniel Lorence *is AHIMA's senior research analyst.*

Commonly Used Statistical Tests

A few simple statistical tests often present clues to exploring the types of questions discussed here. While most HIM professionals harbor no illusions of becoming experts on statistical theory, the application of some basic statistical methods could prove useful in pointing to avenues for further discussion or exploration and can highlight areas of interest. Ultimately, it would help providers deliver better care.

t-test

The t-statistic tells you how far away an estimate is from its expected value, relative to a standard error value. This would demonstrate, for example, whether mean values for two different groups are significantly different or different enough to support a claim that two variables being measured are truly different.

Example: You want to know whether the average time spent in surgery between two physician teams is significant. A t-test will show whether or not the variation is likely due to chance. It will not tell you whether or not time in surgery has any effect on the success of the surgery.

ANOVA

Like the t-test, analysis of variance (ANOVA) performs a test to determine whether means taken from two or more samples are equal or significantly different. This goes beyond the t-test by allowing measurement between multiple sample means and different variables at the same time.

Example: You want to know whether the average time spent in surgery between three physician teams is significant, examining separately, surgery with and without complications. ANOVA will show whether or not the variation is likely due to chance, breaking out the surgeries by complications/no complications. Again, it will not tell you whether or not time in surgery has any effect on the success of the surgery.

Chi-squared

Chi-squared (X2) is a statistic that measures the distance between the actual data and expected values computed from a hypothetical, statistical model. If X2 is too large to explain by chance, the data contradicts your model. The definition of large depends on the size of the sample.

Example: You want to know whether a certain continuing education or training program improved surgery success rates. A chi-squared test will allow you to compare observed and expected results, examining a number of likely factors to determine whether or not they are significant.

Correlation

The correlation coefficient is a number between -1 and 1 that indicates the extent of the linear association between two variables. Often, the correlation coefficient is abbreviated as r, or referred to as Pearson's r.

Example: You want to know whether the number of operations per hospital is associated with success (i.e., lower mortality) following surgery. A higher correlation coefficient will indicate success. It will not prove that the number of operations caused the surgery outcome, only that they behave (or co-vary) consistently. The cause of the surgery success may be the result of some other factor that has an effect on both the number of operations and the surgical outcome.

Bivariate Regression

This analytical tool uses a mathematical model (least squares method) to estimate or fit a straight line through a set of observations. Graphically, the slope of this line represents the regression coefficient, and the point where it intercepts the y-axis of a graph represents a constant. The coefficient and the intercept point (plus an error value) represents your predictive model as an algebraic equation. By itself, the coefficient value indicates the strength of the effect that the independent variable has on the dependent variable. The sign of the coefficient (+ or -) shows whether the effect is direct or inverse, i.e., whether the dependent variable changes in the same direction as the independent variable.

Example: You want to know whether the number of operations per hospital causes greater success (i.e., lower mortality) following surgery. A higher regression coefficient will indicate success. Unlike the correlation coefficient, it will suggest that the number of operations is associated with the surgery outcome and that it causes the outcome. It only measures one independent and one dependent variable at a time, so a comparison of the relative effect of each causal factor is not possible without further analysis.

Multiple Linear Regression

This is similar to the bivariate regression test, only now a given independent variable coefficient takes into account, or is modified by, the other independent variables. The practical result is that when you run the test and sequentially add (or step in) one variable at a time, you can see how each independent variable overshadows, or is affected by, other independent variables. This commonly used approach is known as stepwise multiple regression analysis.

Example: You want to know which factors (e.g., number of operations per hospital, operations per physician, or time spent in surgery) cause greater success (i.e., lower mortality) following surgery. You also want to know which has a greater effect, relative to the others. A higher multivariate regression coefficient will indicate success. Unlike the bivariate coefficient, it will show the comparative strength of each of the causal factors and their effect with respect to the dependent variable (surgical outcome). It assumes, however, that an increase in each independent variable will consistently cause a like increase in the dependent variable. This may not always be true.

Non-linear Regression

This test provides a regression coefficient that is calculated from an exponential curve rather than a straight line. This is useful when a non-linear relationship is known or suspected to exist between variables, such as the relationship between age and physical performance.

Example: You want to know if one or more factors (i.e., number of operations per physician) causes greater success following surgery, and if this cause consistently results in a greater effect. It may be that a surgeon reaches an optimal number of operations performed within a given time, after which his or her success rate does not increase as much—or possibly decreases. Or success rates may increase with age to a certain point, after which performance declines. So the cause-and-effect relationship (or covariance) does not behave in a linear fashion, but is different depending on where (in time) you look at it. A variety of non-linear regression tests are available to show this relationship, such as log-linear regression analysis.

Article Citation:

Lorence, Daniel P. "An Introduction to Benchmarking Health Assessment and Outcomes through Applied Statistics." *Journal of AHIMA* 69, no. 6 (1998): 42-46.

Driving the Power of Knowledge

Copyright 2022 by The American Health Information Management Association. All Rights Reserved.