Hot Topics in Orthopaedic Clinical Research Methodology

Matthew Sloan, MD, MS
Keith D. Baldwin, MD, MSPT, MPH

Department of Orthopaedic Surgery
School of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania, USA

Background

Epidemiologic research has increased in prevalence and importance across all medical specialties. Epidemiology shares historic roots with the development of public health practices. The field is focused on determining the cause of diseases among populations and the effectiveness of strategies to control those diseases. In this way prevention and treatment strategies may be tested in order to minimize the impact of disease.

In the modern era of expanding and unsustainable health care costs, government payers and insurance companies in conjunction with quality improvement advocates have promoted "evidence-based medicine." These are clinical practices that have been proven to effectively and efficiently provide safe treatment for patients. This "evidence" is the result of rigorous epidemiologic study. Baldwin et al. previously described in detail how epidemiologic questions are developed, and outlined the fundamentals of study design and data analysis with a focus on orthopaedic surgery research.

In this review, discussion will focus on the importance of clinical research in orthopaedics, and an evaluation of outcome measurement tools, quality improvement initiatives, and clinical trials. Finally, the principles behind construction of a multivariate regression model will be described, in order to illuminate this centrally important, yet seemingly esoteric epidemiologic tool.

Hot Topics

Passage of the Patient Protection and Affordable Care Act in 2010 mandated a program for public reporting of medical outcomes from the individual physician to the hospital level. The aim of this program is to improve quality in two areas: 1) on the provider side through open comparison and value-based repayment, and 2) on the patient side through enhanced consumer decision-making.

The Orthopaedic Core Measurements are most relevant to orthopaedic surgeons. The first version of these measures focuses on total joint arthroplasty, because these procedures cost Medicare $7 billion annually in hospital expenditures, which makes it the top Medicare expenditure.

The specific measures evaluate complications and all-cause 30-day readmission rates following total knee and total hip arthroplasty, and patient satisfaction with orthopaedic surgical care. Looking forward, it is likely that similar Orthopaedic Core Measurements will be developed across all subspecialties with the intent of improved quality and informed consumer choice.

These initiatives have bolstered interest in orthopaedic clinical research and patient reported outcomes. The publication of comparative outcomes among orthopaedic surgeons necessitates both a means for accurate reporting as well as strategies for improving these outcomes. Both of these topics are epidemiologic in nature. The former requires an understanding of the types of outcomes that are measured and the limitations in these measurements. The latter requires an understanding of Quality Improvement (QI) initiatives and Randomized Controlled Trials (RCT).

Outcome Measures

In the development of an epidemiologic study the fundamental question revolves around the relationship between exposures and outcomes. To draw from the CMS Orthopaedic Core Measurements as an example, a study would evaluate the exposure of total joint replacement and the outcome of 30-day all-cause readmission or patient satisfaction with the procedure. Increasing emphasis has been placed on selecting outcomes that are patient-centered, rather than relying on objective data alone. In this way treatments can be tested for their impact on quality of life, rather than improvement in intermediate endpoints (such as knee range of motion), which may be of little interest to the patient. The science of patient-reported outcomes, though initiated by Ernest Codman in the early 20th century with "The end results of health care" is still very fragmented and difficult for clinicians to agree upon.

Physician collected outcomes are measures that are taken by the physician and include subjective symptoms, but also objective measures such as range of motion and physical findings. Harris Hip score is an example of such a measure. In the last several years CMS and other payers have been less interested in these types of outcomes, though many outcomes reported in earlier as well as recent papers use these measures.

Patient-reported outcomes (PRO) are those that come directly from the patient, and describe symptoms. PRO instruments are questionnaires that have been validated for assessment of the symptoms of interest. They must be validated for particular pathologies, and for different cultures and languages. These instruments should demonstrate specificity for the outcome of interest, incorporate questions that are clearly understood and equivalent among diverse patient populations, include an optimal number of questions, and be reproducible. Furthermore, PROs can be either joint specific (such as the DASH for the upper extremity, or FADI for the foot and ankle) or they can be general health measures such as the SF-36 or SF-12.

These instruments are examples of static questionnaires, forms with a fixed set of questions. Issues with this type of outcome measure are that they are vulnerable to ceiling or floor effects, and they need to be specifically validated for each pathology and population in which they are used. Ceiling effects are when the outcome measure does not pick up the high end of performance well because of its questions, and floor effects are the opposite.

Computerized adaptive testing (CAT) has shown promise as a shorter and more effective means for collecting PRO scores in patients with arthritis and other pathologies. This instrument draws from a larger pool of validated survey questions, sequentially choosing items based on the previous response. In this response-based algorithm, fewer questions can be used across a wider spectrum of function to gather more information than longer questionnaires.[9] Theoretically, CAT can overcome the challenges of ceiling and floor effects in static instruments and may be able to detect fine detail at the extremes of function. CAT instruments outperform static instruments at these extremes of the function spectrum by selecting survey items most appropriate to the specific patient, with maximal information ascertained per question.[10]

While great importance is placed on the assessment of patient-reported outcomes, other outcome measures can and should be used in conjunction with patient reports for a complete clinical picture. For example, if you would like to compare patients with specific clinical diagnoses, such as osteoarthritis grade, cardiac ejection fraction, or other clinical indicators, it may be necessary to have physician-reported outcomes. Functional status comparison between physician and patient-reported outcome can also provide important feedback for providers whose opinions may differ dramatically from their patients. Further, predictive models, those that attempt to calculate risk for certain outcomes require a model based on all predictive variables, including physician reported outcomes such as BMI or severity of disease and PRO such as preoperative and postoperative functional status.

Clinical Trials

The different types of epidemiologic study designs, from retrospective to prospective to meta-analysis, has been previously described[1]. The randomized controlled clinical trial is the pinnacle of research methodology, and the only true "experimental" clinical research design. This study design compares at least two randomly assigned groups, one assigned to a control arm and one assigned to an intervention arm. This randomization, if done properly, should sort known and unknown variables equally amongst the intervention and control groups. In this way, the comparison between these groups minimizes bias in a way that non-randomized studies cannot.

Randomized controlled trials (RCT's) in orthopaedic surgery have become more prevalent only in the last decade. In many ways orthopaedics is beginning to follow other medical specialties in terms of research design and methodology; however, orthopaedic RCT design and execution still lags behind other disciplines.

There are shortcomings possible even with the gold standard design. A recent bibliometric analysis showed that among orthopaedic surgery clinical trials that were published, over one-third were underpowered or did not report a power analysis. Of published orthopaedic trials that described significant difference between treatment groups, one-seventh were underpowered. Among the published orthopaedic trials that found no significant difference between treatment groups, over three-quarters were underpowered. The importance of this for the consumer of orthopaedic research is to view outcomes critically, even when the study design is one classically described as a gold standard.

Quality Improvement Initiatives

Quality Improvement (QI) initiatives are clinical projects that attempt to directly improve an identified health care shortcoming. Compared with traditional research, the goal of QI is to elicit rapid change in a complex system, improve outcomes, or improve the patient experience. This has been an area of intense interest over the last several years. Variance in healthcare has been thought to represent inferior quality. Therefore, clinical pathways have been developed to increase value, which is defined as quality relative to cost. Several examples of QI have been initiated in the orthopaedic community.

In adolescent idiopathic spine surgery, researchers at the Children's Hospital of Philadelphia initiated a rapid recovery pathway of adolescent idiopathic scoliosis. This pathway involved a standardized set of medications, OR procedures, and postoperative rehabilitation protocols that resulted in lower patient pain and an average of almost two fewer days of hospitalization.

At Penn Presbyterian Hospital, a multivariate logistic regression analysis model was used to ascertain factors that result in unplanned admission to the intensive care unit (ICU). Following development of this model, patients were risk-stratified into elective admission to the ICU following joint replacement based on those criteria. This pathway led to fewer unplanned admissions to the ICU, and safer postoperative care for these patients.

On the University of Pennsylvania orthopaedic trauma service, a pathway for hip fractures has been generated. It is known that delays to the operating room and longer hospital stays result in increased patient morbidity and mortality, lower quality of care and decreased patient and caregiver satisfaction. The trauma service in conjunction with anesthesiology, geriatric medicine and rehabilitation, generated a protocol to decrease variation and increase expediency of surgery and quality of care. The results are under current investigation.

Multivariate Regression Models

Complex prospective study designs with randomization and blinding require only simple statistics, because confounders have been equally distributed between groups by virtue of randomization. However, non-randomized studies must stratify by significant confounders (in which case the confounders cannot be directly studied) or perform a regression analysis, which eliminates the effect of one variable on another and estimates accurate effect size. If univariate statistics are used alone, there is no evidence that the effect observed is because of the variable of interest or a result of some confounder, either identified or not identified.

Multivariate regression modeling is the solution to this dilemma. Logistic regression is available when the outcome is dichotomous, and linear regression is used when the outcome is linear. Some modeling such as Poisson regression for count variables and multinomial regression for categorical variables is beyond the scope of this discussion.The math behind these models is complex, but the process for developing a model requires only a dataset, a literature review, and a statistical software program. The strength of the multivariate model comes from its ability to provide an association between an outcome and multiple exposures that eliminates the redundant variability that results from each individual factor. Multivariate models are often reported as an association between exposure, x, and outcome, y, that is "controlled for" all other factors studied. For example, a factor, such as age, is entered into the model, and through the regression calculation, the variability between x and y that is due to differences in age among subjects is removed from the association between x and y. This result could be reported as "controlled for age" or "age-adjusted".

The first place to start building a multivariate regression model is with the known or expected variables of interest. These are identified from the literature, from common variables of known effect (such as age or a socioeconomic indicator), and investigatory variables from the study hypothesis. Once a list of intended variables has been compiled it can be helpful to visually assess the relationship between each variable with the outcome of interest by building scatterplots for each variable with the outcome of interest.

With variables of interest identified, model development begins (Figure 1). The ideal model will include the fewest number of variables that explain most of the variability in the outcome. Each variable included in the final model should be highly predictive of the outcome without being highly correlated with another variable in the model. This avoids the issue of multicollinearity, which can make a model cumbersome and difficult to interpret.

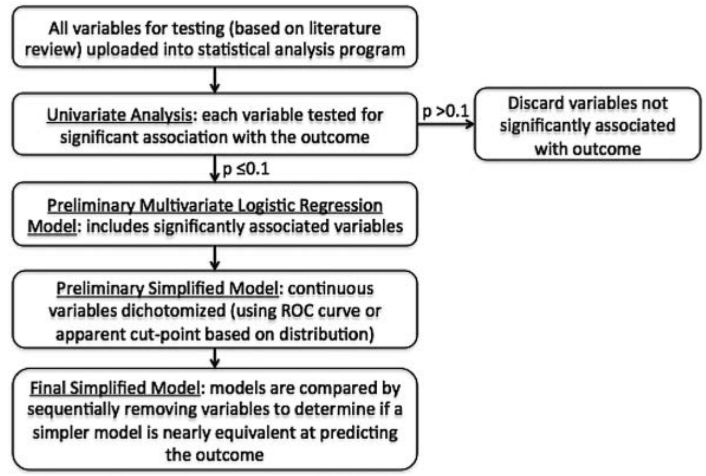

Figure 1. Backwards Model Building Strategy for Multivariate Logistic Regression Model.

Next, a univariate analysis, or a simple linear regression is performed, which by analyzes the association between each variable and the outcome of interest. This will provide a t-statistic and a p-value for significance for each variable with the outcome. Once completed, the model can be developed using a forward or a backwards model building strategy. The forward strategy chooses variables sequentially to add to the model, while the backwards model begins with all variables included in the model and sequentially removes insignificant variables.

Next, continuous variables are dichotomized. This involves choosing cut- points in continuous variables, so that groups such as "high" or "low" within a value range may be compared to one another. This strategy may decrease the precision of the model, but makes the findings easier to understand. For example, imagine the risk of postoperative infection increases 1% with each point in BMI compared to an index value. This may intuitively make more sense if modeled in groups of regular weight, overweight, and obese, where the result could be described as obese patients having a 10% greater risk of postoperative infection compared with regular weight patients. Different cut points may be obtained by simply looking at the frequency of outcome in each group or by more complex methods such as receiver operator characteristics analysis.

Models can be compared using the log-likelihood estimation, which sequentially simplifies the model and determines whether the simpler model differs significantly in predictive value for the outcome compared with the original model. A model that includes many variables is likely not the most parsimonious, or simplest model to predict the outcome.

Summary

The practice of medicine continues to move toward evidence-based practice and payment structures force physicians to improve outcomes and efficiency, Orthopaedic surgeons must adapt to these changes through improved epidemiologic research and enactment of the results of these studies. It is imperative that the modern orthopaedic surgeon understands the results of QI initiatives and clinical trials in order to appropriately incorporate best practices into their patient care.

References

1. Baldwin et al. Orthoepidemiology 101: The Basics of the Art of Outcomes Research in *Orthopaedic Surgery. UPOJ*. 2009;Vol 19.

2. Centers for Medicare and Medicaid Services. Physician Compare Initiative. <www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments>. Accessed 29 Feb 2016.

3. Centers for Medicare and Medicaid Services. "Comprehensive Care for Joint Replacement Model." < https://innovation.cms.gov/initiatives/cjr>. Accessed 15 Mar 2016.

4. Joynt, Karen E., et al. "Contribution of preventable acute care spending to total spending for high-cost Medicare patients." *JAMA* 309.24 (2013): 2572-2578.

5. Centers for Medicare and Medicaid Services. Core Measures. <www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments>. Accessed 29 Feb 2016.

6. Donabedian A. "The end results of health care: Ernest Codman's contribution to quality assessment and beyond." *Milbank Quarterly*, 67(2): 233-256, 1989.

7. Deshpande *et al.* Patient-reported outcomes: a new era in clinical research. *Perspectives in Clinical Research*. Oct-Dec 2011, Vol. 2,Iss. 4:137-44.

8.Acquadro *et al.* Incorporating the Patient's Perspective into Drug Development and Communication: An Ad Hoc Task Force Report of the Patient-Reported Outcomes (PRO) Harmonization Group Meeting at the Food and Drug Administration, February 16, 2001. *Value In Health.* 5 Nov 2003. Vol. 6:522-31.

9. Stewart, Anita L., *et al.* "Functional status and well-being of patients with chronic conditions: results from the Medical Outcomes Study." *JAMA* 262.7 (1989): 907-13.

10. Voleti, Pramod B., *et al.* "Level of evidence of presentations at American Academy of Orthopaedic Surgeons annual meetings." *The Journal of Bone & Joint Surgery* 94.8 (2012): e50

11. Abdullah, Leath, *et al.* "Is There Truly "No Significant Difference"?." *J Bone Joint Surg Am* 97.24 (2015): 2068-2073.

12. Casarett, David, Jason HT Karlawish, and Jeremy Sugarman. "Determining when quality improvement initiatives should be considered research: proposed criteria and potential implications." *JAMA* 283.17 (2000): 2275-2280.

13. Gornitzky, Alex L. "More Than Just Early Discharge: A Rapid Recovery Pathway (RRP) That Improves Pain Control and Reduces Time to Inpatient Recovery Following Posterior Spinal Fusion (PSF) for Adolescent Idiopathic Scoliosis (AIS)." 2015 AAP National Conference and Exhibition. *American Academy of Pediatrics*, 2015.

14. Kamath, A. F., et al. "Intensive care monitoring after total joint replacement."*Bone & Joint Journal* 95.11 Supple A (2013): 74-76.