FOR many decades, the field of anesthesiology has improved patient safety by reducing the incidence of rare, catastrophic events. The proximate and direct linkage of errors such as unrecognized esophageal intubation, vaporizer malfunction, and medical gas misconnections allowed patient-safety experts to improve outcomes by using tools such as root-cause analysis, simulation, human factors redesign, technology, and pervasive checklists. Because these historical anesthesia errors have been transformed into “never” events, the field has expanded its patient-safety efforts to much more common adverse events such as cognitive dysfunction, surgical-site infections, cardiopulmonary complications, and all-cause mortality. As other editorials in this series have noted, anesthesiologists must begin taking responsibility for these comprehensive outcomes to remain relevant in the dynamic healthcare environment.

As a result of this evolution, the tools used by anesthesiology patient-safety experts have also matured. Because the underlying etiology of postoperative adverse events is often multifactorial, the causal relations between intraoperative decisions and long-term outcomes are more difficult to extricate. There are limited prospective interventional data to guide our decisions, resulting in the rise of population level, epidemiological research to fill the knowledge void. This trend is a reaction to the needs of the specialty and our patients, rather than an opportunistic use of existing data sources. The field has invested in the infrastructure necessary to create large observational data sources needed to answer fundamental question: the Anesthesia Quality Institute, Multicenter Perioperative Outcomes Group, Society for Ambulatory Anesthesia database, Pediatric Regional Anesthesia Network, and Society for Cardiovascular Anesthesiologists Adult Cardiac Anesthesia Module all represent investments in current and future patient safety and research.

One specific example of these patient-safety efforts involves surgical-site infections. The financial impact of surgical-site infections is significant, estimated at more than $10 billion annually in the United States alone. More importantly, the impact on patient quality of life, pain, suffering, and function is staggering. The science evaluating the relation between intraoperative anesthetic management and postoperative surgical-site infections has matured from small, prospective, randomized, controlled trials assessing efficacy to large observational studies evaluating effectiveness. During this evolution, many national policies have established specific anesthetic processes of care: timing of prophylactic antibiotics, choice of antibiotic agents, and active warming. Next, a variety of research and quality improvement (QI) efforts focused on measuring processes of care associated with surgical-site infections.

Despite early efficacy literature establishing the value of specific antibiotic timing and active warming, repeated large database analyses have not observed robust effectiveness across hundreds of hospitals. These processes of care remain a vibrant area of large database research and the optimal analytical techniques for large database research are an area of controversy.

Currently, the majority of large database research published in perioperative journals uses classic observational investigation techniques—*a priori* definitions of a primary outcome, creation of control and experimental groups, tools to address selection bias, risk adjustment, and multivariate regression to identify independent associations. It leverages analytical and reporting techniques used for prospective interventional trials and then advances them using epidemiological statistical tools. This body

---

Image: J. P. Rathmell (marble bust by Jean Baptiste Stouf, c. 1791).

Accepted for publication November 13, 2013. From the Department of Anesthesiology, University of Michigan, Ann Arbor, Michigan.

Copyright © 2014, the American Society of Anesthesiologists, Inc. Lippincott Williams & Wilkins. Anesthesiology 2014; 120:523-5
of literature can be described as “hypothesis-driven” research. These hypothesis-driven data are routinely used to guide daily practice and are read and understood by the readers of clinical journals. Given the large sample sizes of these databases, many associations have statistical significance; more stringent registration and reporting requirements may decrease the identification and publication of associations that lack clinical impact. Even well-conducted hypothesis-driven observational research struggles to gain acceptance within the anesthesiology community.5,6 One can only hope that improved registration and reporting may address this skepticism. In addition, advances in statistical techniques to address confounding and new data sources may improve the reproducibility of observational study results.

However, this drive toward hypothesis-driven observational analyses is in conflict with an emerging trend in scientific discovery: “big data analytics.”7 The large amounts of data being created during healthcare delivery represent a unique and unprecedented opportunity to further our understanding of clinical exposures and outcomes. The antithesis of hypothesis-driven research, “data mining” has historically been poorly received by the medical scientific community. However, the value of data mining is already clear for specific healthcare investigations: detection of infectious disease outbreaks via syndromic surveillance and postrelease drug-safety evaluation.8,9 The value of big data extends beyond automated surveillance. The reality of scientific advances is that hypothesis-driven research is based on biologic assumptions. These assumptions reveal not only our current understanding but also our biases. The history of scientific progress is rife with invalid assumptions that, once challenged, gave way to great progress. For example, the discovery of the bacterium Helicobacter pylori was hypothesis driven, yet refuted existing scientific assumptions and biases.10 In a similar manner, the modern use of data mining may expose associations that are transformative, yet hidden due to hypothesis-driven research based on existing biases.

A model for this type of progress already exists: the ubiquitous genome-wide association studies (GWAS). Unthinkable just 10 yr ago, GWAS is a foundation of scientific progress now. GWAS are a prototypical use of big data—by shedding our biases regarding which specific polymorphisms may be associated with a particular phenotype, we allow the data to demonstrate the significant relations. However, GWAS require unique reporting, statistical analysis, and scrutiny. For example, due to the hundreds of thousands of polymorphisms, well-conducted GWAS correct for multiple testing, require minimum effect-size thresholds, and a prospective power analysis.11 Similarly, data-mining investigations of large observational databases will require new statistical and data-validation techniques to address residual confounding, selection bias, and outlier data.

The challenge lies in recognizing this dichotomous path to scientific discovery. Each technique—hypothesis-driven observational analyses and data-mining analytics—has its own strengths and role in scientific progress. However, merging the two techniques is inherently dangerous because the assumptions and biases of one technique do not translate to the other. As a result, it is dangerous to shed the hypothesis-driven approach, yet use hypothesis-driven thresholds for statistical significance or meaningful effect sizes. More importantly, the reporting is fundamentally different. Both techniques result in the observation of associations, but hypothesis-driven research can report that an expected association was tested and found to be present or not. Data-mining analyses can only report that certain associations were observed and now serve as the substrate for future hypothesis-driven research. Mixing these two reporting paradigms merges the worst of each technique and truly exposes the scientific community to “pseudoscience.”

Similarly, we must realize that whether observational data are being used for internal, single-center QI efforts or peer-reviewed academic investigations, the perils and opportunities of observational research remain unchanged. Academic investigations using large observational data are subjected to the peer-review process and receive scrutiny from a variety of academic venues. Some departments have an internal peer-review process for all clinical research, including observational large database research. The process requires a priori primary outcome definitions, patient inclusions and exclusions, and a proposed statistical analysis and power analysis.9 When combined with the rigors of the external peer and statistical review process at major anesthesiology journals, these internal processes result in well-conducted, hypothesis-driven observational analyses with transparency regarding data limitations. However, these internal and external check-steps are often absent during the conduct of QI analyses. QI initiatives include surveillance, efficacy review, cost comparison, and sentinel event analysis. This broad range of functions demands disparate data sources with variant data quality and limitations. Most importantly, integration and analysis of these data sources require the same skill sets as observational research—deep statistical expertise, database querying and aggregation, data-quality assessment, and understanding of the limitations of secondary use of data. However, few anesthesiology QI departments possess the personnel resources to encompass these skills. This challenge is increasing, as recent Maintenance of Certification in Anesthesiology and American Council for Graduate Medical Education guidelines indicate the need to have all diplomats and residents engage in data-driven QI projects. Each of these projects must maintain the same self-restraint and transparent reporting as academic peer-reviewed research. In addition, they must also recognize the perils of changing clinical practice based on a hybrid of hypothesis-driven and data-mining research.

Changes in patient-safety needs, regulatory pressures, and the political environment demand the use of large database research to complement basic science experiments and prospective interventional trials. Differential analytical and reporting requirements of hypothesis-driven versus data-mining investigations are essential to report clinically impactful and scientifically robust results. Only then can we ensure that the one-eyed man is truly a king, and not just a false prophet.
Acknowledgments

All work and funding attributed to the Department of Anesthesiology at the University of Michigan Medical School, Ann Arbor, Michigan.

Competing Interests

The author is not supported by, nor maintains any financial interest in, any commercial activity that may be associated with the topic of this article.

Correspondence

Address correspondence to Dr. Kheterpal: sachinkh@med.umich.edu

References

4. Thomas L, Peterson ED: The value of statistical analysis plans in observational research: Defining high-quality research from the start. JAMA 2012; 308:773–4
7. Larson EB: Building trust in the power of “big data” research to serve the public good. JAMA 2013; 309:2443–4

The Bookplate of Paul M. Wood, M.D.

As a young boy, the future founder of the Wood Library-Museum of Anesthesiology (WLM), Paul Meyer Wood, was tutored in the art of designing bookplates by his mother, Louise Meyer Wood. Mrs. Wood would eventually be appointed as a professor of religious art and architecture at New York City’s Biblical Seminary. In designing his personal bookplate (above), Dr. Paul Wood included images of Morton’s ether inhaler, three books, a portable anesthesia machine, and a syringe, four images symbolizing simultaneously his anesthetic innovations and his collections of anesthesia-related items. Could Dr. Wood have purposefully juxtaposed the “E” and “L” (of “EX LIBRIS”) with the H-like top of the keyed tank of the tallest machine cylinder—to artistically approximate (90 degrees clockwise) the “WLM” acronym of the Wood Library-Museum? (Copyright © the American Society of Anesthesiologists, Inc.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA’s Wood Library-Museum of Anesthesiology, Park Ridge, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.