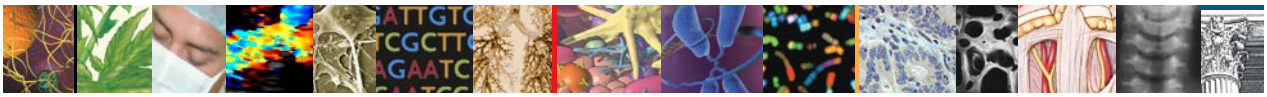

The NEW ENGLAND JOURNAL of MEDICINE

Perspective

JUNE 5, 2014

Learning from Big Health Care Data

Sebastian Schneeweiss, M.D., Sc.D.

The routine operation of modern health care systems produces an abundance of electronically stored data on an ongoing basis. It's widely acknowledged that there is great potential for utilizing these data, within the system that generates them, to inform treatment choices in ways that improve patient care and health outcomes.[1] Imagine entering your office in the morning and finding an e-mail message reading, "Thanks to your new vaccination screening program, as of yesterday your practice had given 120 more vaccinations than similar practices had." Or "As compared with the period before your network's implementation of the new policy of referring patients with atrial fibrillation to the anticoagulation center, seven strokes have been averted, but two additional upper GI bleeds have occurred." Or even "Judging from her track record and the characteristics noted in her medical record, there is an 80% likelihood that Patient C, whom you are about to see, will not fill her prescription for an antihypertensive." In theory, such ongoing structured learning based on routinely collected data could seamlessly augment the knowledge physicians have gleaned from their experience, which involves the same patients and more detailed observations but is less formal in its evaluation processes and more likely to be subject to unintended bias.[2]

Two key "learning" applications of big health care data that hold the promise of improving patient care are the generation of new knowledge about the effectiveness of treatments and the prediction of outcomes. Both these functions exceed the bounds of most computer applications currently used in health care, which tend to offer physicians such tools as context-sensitive warning messages, reminders, suggestions for economical prescribing, and results of mandated quality-improvement activities.

Physicians currently struggle to apply new medical knowledge to their own patients, since most evidence regarding the effectiveness of medical innovations has been generated by studies involving patients who differ from their own and who were treated in highly controlled research environments. But many data that are routinely collected in a health care system can be used to evaluate medical products and interventions and directly influence patient care in the very systems that generated the data.

To facilitate such learning, analytic tools with several key characteristics will be required. First, we need methods that ensure that the patient groups being compared are similar to one another, so that analysts can be sure they are actually studying the effects of care interventions rather than variation in the underlying severity of disease; propensity-score methods, which simultaneously account for many patient characteristics, have proved to robustly reduce confounding biases in studies using health care databases.

Second, most aspects of the analyses need to be automated without loss of validity, so that many research questions can be answered simultaneously and the number of matters investigated can grow as demand increases for quantifying the effectiveness of care. Extensions of propensity-score methods have been developed for automatically adapting to new data sources and reducing confounding.

Third, once analyses have been automated, they should be able to be repeated in rapid cycles tied to data refreshes, which may occur as often as every 24 hours.

Fourth, such software should be easy enough to use that users with little training can set up a learning system fairly quickly and avoid typical pitfalls of database studies that hamper causal interpretations of results — such as failures to designate the timing of the start of treatment and the onset of outcomes, to ensure comparison of similar patients, and to adjust robustly for confounding without adjusting for factors that lie on the causal pathway between exposure and outcome. Most important pitfalls can be avoided with fairly obvious approaches — for instance, by studying patients who have been newly exposed to a given intervention and comparing them with patients newly treated with the next best alternative, assessing patients' characteristics before the intervention was started, and refraining from adjusting for patient factors that arose after the exposure in question began.[3]

Finally, results from such analyses need to be presented in an easily digestible form for a busy clinical audience and further interpreted for patients.

All these components of analytics have been developed, yet our health care system has not been able to systematically integrate them into its work to establish an ongoing learning-and-improvement process. The collection of more data has so far not translated into the generation of more actionable insights into the best ways of treating the patients who are the sources of those data. Given widespread agreement that an effective learning health care system is desirable, why aren't we closer to that goal?

One major impediment is the underuse of existing uniform data standards for electronic medical records. We therefore need analytic approaches that embrace the data turmoil by relying less on standardized data items and having the capacity to process data in any format. Of course, the exposures and clinical outcomes of interest must be clearly identifiable, but for the detailed characterization of patients' health states, which is the foundation for improved control of confounding and for making valid inferences, standardized measurements may not be necessary. Well-measured proxies of a patient's health state — for instance, the use of supplementary oxygen as a proxy for very poor health — can often do as well as complex clinical measures in the prediction of health outcomes. Algorithms can be created to identify such proxies empirically in the data at hand through their observable associations with disease outcomes and then to use those proxies for adjustment. This approach does not require a specific medical interpretation of the proxy factors and can therefore work without the need for data standards and so be implemented rapidly. Such methods have been shown to perform well in studies using health care databases.

Many available data currently reside in separated silos. For example, detailed genetic information is often stored not in the medical record but rather in separate research databases with restricted access — a lack of linkage that's attributable not to technical difficulties but to privacy concerns.[4] Absent a consensus on a resolution for the privacy impasse, we need to accept that portions of patient data will be physically distributed over several databases. In order to conduct multivariate-adjusted analyses, we require better methods for extracting patient information from these distributed databases without making patients identifiable in the process. Such distributed analyses are cumbersome to implement and should be made part of an evidence-generation platform for easy reuse.

Even if such improvements can be made, interpretations of findings from observational studies using secondary health care data will continue to encounter distrust.[5] Although analytic tools such as propensity scores can help to reduce confounding bias, con-

cerns about causal interpretations remain. Randomized studies embedded in routine care that assess patient outcomes by means of electronic medical record databases are cost-effective and reduce residual imbalances in patient characteristics at the start of a study. The Patient Centered Outcomes Research Institute recently launched a major initiative to build a nationwide network of health care systems that will use their infrastructure for such pragmatic randomized trials.

An audio interview with Dr. Schneeweiss is available at NEJM.org

Ultimately, a key to success in learning from big health care data will be to remain focused on our ultimate goal: gaining actionable insights into the best ways to treat the patients in the care system that generated the data. If we work backward from this goal, agreeing on the right analytic methods and the necessary data will be manageable steps, and together we'll be able to negotiate the critical issues of data privacy and standardization.

Dr. Schneeweiss reports serving as a consultant to WHISCON and to Aetion, a software manufacturer in which he also owns shares. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

From the Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston.

1. The learning health care system in America. Washington, DC: Institute of Medicine, 2012 (http://www.iom.edu/Activities/Quality/LearningHealthCare.aspx).
2. Choudhry NK, Anderson GM, Laupacis A, Ross-Degnan D, Normand SL, Soumerai SB. Impact of adverse events on prescribing warfarin in patients with atrial fibrillation: matched pair analysis. BMJ 2006;332:141-5.
3. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323-37.
4. Faden RR, Kass NE, Goodman SN, Pronovost P, Tunis S, Beauchamp TL. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. Hastings Cent Rep 2013;43:Spec Rep:S16-S27.
5. Gabriel SE, Normand SL. Getting the methods right — the foundation of patient-centered outcomes research. N Engl J Med 2012;367:787-90.

DOI: 10.1056/NEJMp1401111

Fostering Responsible Data Sharing through Standards

Rebecca Kush, Ph.D., and Michel Goldman, M.D., Ph.D.

Children with muscular dystrophy and their families make sacrifices to engage in clinical research studies, providing valuable data they expect will contribute to the discovery of a cure, although they know it may not be found in time to help them. This message was emphasized at a recent meeting organized by the Institute of Medicine, where clinical investigators and study sponsors were implored to share research data to fulfill their moral obligation to maximize the chances that patients' contributions translate into therapeutic advances.

The urgent need to build collaborative networks dedicated to data-sharing principles was also underlined at a recent summit on dementia held by the Group of Eight industrialized countries. In fact, the failures encountered in targeting beta-amyloid for the treatment of Alzheimer's disease have led several companies to begin collaboratively developing innovative study designs requiring extensive data sharing.

Unfortunately, the diverse ways in which data are collected and reported in clinical studies make it difficult or impossible to query across data sets, pool and share data, or integrate data for analyses of multiple trials to gain new scientific insights. Yet these problems can be resolved through the use of standard data formats, and the best outcomes would be achieved if data standards were adhered to from the start — within the electronic health record (EHR).

In 1999, the Mars space orbiter exploded when incoming data were misinterpreted: they were assumed to be in SI units when they were actually in U.S. customary units. Without the relevant metadata, such confusion is inevitable. Units and other metadata are critical in medical research as well. Standard data and metadata formats are required for efficient aggregation of patient-level data, trustworthy statistical analyses, and accurately informed clinical decisions. When such standards are not implemented by all parties at the outset of research studies, precious information is lost or, at the very least, time-consuming manual mapping or computer programming is needed to render data comparable.

Data related to cognitive defects in Alzheimer's disease are a classic example. Although there is an Alzheimer's Disease Assessment Scale for testing cognition, various study sponsors use this questionnaire in various ways, preventing accurate comparisons among studies or among patients