Big Data and Predictive Analytics

Applications in the Care of Children

Srinivasan Suresh, MD, MBA

KEYWORDS

• Big data • Predictive analytics • Clinical Informatics • Electronic Health Records

KEY POINTS

- Emerging changes in the United States' healthcare delivery model have led to renewed interest in data-driven methods for managing quality of care.
- Analytics (Data plus Information) plays a key role in predictive risk assessment, clinical decision support, and various patient throughput measures.
- The formation of clinical informatics as a formal board certified medical subspecialty is a key step towards the development of physician informaticists who can leverage information technology to improve the delivery and safety of healthcare.
- The combination of big data and predictive analytics in healthcare has the great potential to positively affect clinical decision support, patient morbidity, and hospital operations such as cost management systems and resource allocation.

"Big data" is a broad term for collection of data sets so large and complex that they are difficult to process using on-hand database management tools or traditional data processing applications.[1] Predictive analytics encompasses a variety of statistical techniques from modeling, machine learning, and data mining that analyze current and historical facts to make predictions about the future, or otherwise unknown events.[2]

Changes in the health care delivery model over the past decade in the United States have led to renewed interest and newer mechanisms in data-driven methods for managing quality of care. Health systems and provider groups are now evaluated using quality metrics for mortality, length of stay, and for hospital readmissions within 30 days. Rankings are increasingly made public, and are being used by insurers to institute pay-for-performance programs.[3] Hospitals need to continually monitor and understand their performance, and determine whether interventions to improve have been successful.

Disclosures: None.
Children's Hospital of Pittsburgh, 4401 Penn Avenue, Room 6415 AOB, Pittsburgh, PA 15224, USA
E-mail address: suresh@chp.edu

Pediatr Clin N Am 63 (2016) 357–366
http://dx.doi.org/10.1016/j.pcl.2015.12.007 **pediatric.theclinics.com**
0031-3955/16/$ – see front matter

DEFINITION OF ANALYTICS AND ITS IMPACT ON HEALTH CARE

Analytics is the systematic use of data combined with quantitative as well as qualitative analysis to make decisions. Health care analytics efficiently applies clinical and administrative data in electronic health records (EHRs), and knowledge of clinical practice standards and guidelines, by way of order sets and protocols, to manage metric-driven quality improvement. The goals of such an initiative are to identify inefficiencies in care delivery and variations from standard practice, explore opportunities for expanding services that may enhance quality of care, efficient targeting of limited resources, and peer benchmarking. The analytical techniques needed to achieve these goals require mature EHR implementations.[4]

EHRs with the required breadth of data increasingly are available at children's hospitals. Institutions may make these data available through a clinical data warehouse, a relational database fed by an EHR's transactional system that supports efficient population queries. National repositories of administrative and sometimes clinical data provide institutions with access to comparative data; for example, the Children's Hospital Association (CHA) Pediatric Health Information System (PHIS). The CHA, through a wide variety of data programs, helps pediatric hospitals drive improved clinical, operational, and financial performance. With shared data sets, potentially better practices across organizations can be identified, ultimately improving children's health care. The PHIS, a comparative pediatric database, includes clinical and resource use data for inpatient, ambulatory surgery, emergency department, and observation unit patient encounters for 45 children's hospitals.[5] PHIS supports a wide range of improvement activities, including clinical effectiveness, resource use, care guideline development, readmission analysis, antimicrobial stewardship, and physician profiling (ongoing professional practice evaluation). Such data sets represent a rich potential source of data for analytics, but they also present substantial challenges in their use. Diagnoses, comorbidities, and procedures typically are represented indirectly as billing codes.[6] Clinical data such as laboratory test results and medication histories may be recorded as local codes or in text, and they may require substantial clinical context to interpret. Efforts are underway to create publicly available libraries of phenotypes defined in terms of patterns in codes and discrete data.[7] Tools for translating these phenotypes into queries of local data warehouses and national repositories could make data-driven quality improvement more broadly practical.

ANALYTICS IN THE FORM OF DECISION SUPPORT AND MONITORING SYSTEMS

Analytics play a key role in predictive risk assessment, clinical decision support, home health monitoring, finance, and resource allocation.[8] They may also be used to enhance less sophisticated, rules-based systems that are already in use. One of the benefits of EHRs has been the integration of clinical decision support (CDS) systems. CDS systems have been shown to reduce errors and improve clinical outcomes in certain settings, such as pediatric intensive care units,[9] and can result in performance improvement on perioperative quality and process measures. Some CDS systems that are designed to prevent medication errors are based largely on commercially available software packages that rely on simple rules. Often, these products do not provide ideal rule sensitivity, and institutions must perform manual reclassification of drug-drug interactions to improve the efficacy of their CDS systems. Analytics may offer a solution to this challenge, because clinicians can use analytics techniques to query and mine the EHR for meaningful connections, and then synergistically combine the knowledge-based rules with analytics applied to EHR data.[10]