Big Data Meets the Electronic Medical Record

A Commentary on "Identifying Patients at Increased Risk for Unplanned Readmission"

Greg de Lissovoy, PhD, MPH

When should a patient be discharged from an inpatient hospital stay? The quick answer might be, only after the course of treatment is complete and further diagnosis, treatment, or monitoring is not required. However, that approach suggests an extended length of stay, whereas other considerations might argue for a more accelerated approach to discharge. For example, the patient may be eager to return home to family or friends willing and able to provide continuing care. The hospital can be a dangerous place; prolonged admission exposes the patient to risks of hospital-acquired infection or other misadventure.[1] When continuing treatment is required, care in a less costly setting such as an outpatient clinic or skilled nursing facility may be feasible. Third-party insurance coverage such as Medicare that pays a set amount per admission gives the hospital a financial incentive to shorten length of stay. In short, the decision to discharge may represent a delicate balance of competing objectives.

Florence Rothman, age 87, was admitted to Sarasota Memorial Hospital in early 2003 with a diagnosis of aortic stenosis and underwent a routine valve replacement procedure. After an unremarkable course of treatment she was discharged in apparent good health. Four days later she collapsed and was taken to the emergency department where she died as a result of complications of surgery.[2,3]

Stunned by this unexpected turn of events, her sons Michael Rothman and Steven Rothman met with medical staff and hospital officials in an effort to understand how their mother's precarious condition could have escaped notice at the time of discharge. They were surprised to learn that the hospital's sophisticated electronic medical records (EMRs) system, while amassing substantial quantities of data over the course of an admission, generated no summary indicators of a patient's condition or "red flags" that could inform decisions about readiness for discharge. The Rothmans vowed to find a way by which future patients might avoid their mother's fate (M. Rothman, personal written communication, June 18, 2013).[3]

Although neither of the brothers was trained in medicine or health services research, they were no strangers to massive electronic databases. Michael Rothman, who holds a PhD in chemistry, spent much of his career as a computer scientist at the IBM Watson Research Laboratory and then became chief information officer for a major bank. Later he founded a company specializing in large data system design and analytics. Steven Rothman is an electrical engineer who began his career at the MITRE Corporation, a contract research organization spun-off from Massachusetts Institute of Technology. While working on aerospace projects in Colorado and New Mexico he became interested in oil and gas exploration and created a company that applied sophisticated mathematics and data visualization to the analysis of seismic data.[2,3]

With active support from the hospital, the Rothmans put their expertise to work exploring the EMR data repository. Their vision was to create algorithms that would

From the Department of Health policy and management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.
The author declares no conflict of interest.
Reprints: Greg de Lissovoy, PhD, MPH, Department of Health policy and management, Johns Hopkins Bloomberg School of Public Health, Hampton House Room 325, 624 North Broadway, Baltimore, MD 21205. E-mail: gdelisso@jhsph.edu.

ISSN: 0025-7079/13/5109-0759

utilize existing EMR data to track patient health status over time, and to present this information as simple and meaningful metrics, available to the clinician in near real time.[3] Several years of effort culminated in the Rothman Index (RI), as described in this issue of *Medical Care*.[4]

Building a model that can predict risk for hospital readmission is not a novel concept. Most of the numerous existing models rely on administrative data generally available only after the patient has been discharged and are most useful for retrospective assessment of quality of care. Only a handful of published models utilize data accessible while the patient is hospitalized, when outputs are available to guide the course of treatment.[5] That situation will likely change rapidly with the marriage of 2 emerging technologies, EMRs and data mining. The EMR makes extensive clinical and demographic data immediately available without the delay required to obtain administrative data such as diagnostic codes that are typically assigned after discharge. Data mining is a general term applied to a wide range of computation-intensive methods for detecting meaningful patterns in very large databases containing disparate data.[6] Development of the RI exemplifies this approach.

Developers of clinical risk prediction models face many challenges. EMRs contain an extensive array of data types ranging from laboratory values to free-form text from which model variables must be constructed. The quantity of data that can accumulate during even a brief hospital stay is substantial; meaningful information must be extracted from what amounts to background noise. Software development is hampered by a lack of standardization across commercial EMR systems.[7] The predictive power of the RI and of other published risk of readmission models is modest at best and generally characterized by high sensitivity and low specificity.[4,8–11]

In a continuing effort to reduce rates of avoidable readmission, the Centers for Medicare and Medicaid Services recently issued more stringent requirements for discharge planning as a condition for hospital participation in the Medicare program.[8] Health status and readmission risk measures generated through real-time analysis of EMR data could potentially improve planning for a patient's transition from inpatient to the next stage of care. Developers and clinical end users must collaborate to learn how to integrate these new health status metrics with existing information on patient clinical and social characteristics to improve quality of care both during the hospital stay and after discharge.

Precision and utility of risk-assessment algorithms will likely improve as the EMR evolves. The process of model building itself can inform the design of more efficient and useful EMR systems. Data mining will help identify factors most closely associated with good and bad outcomes for specific conditions while also revealing essentially irrelevant information now recorded as a legacy of traditional paper-based medical records. Florence Rothman will be watching!

REFERENCES

1. Hauck K, Zhao X. How dangerous is a day in the hospital? A model of adverse events and length of stay for medical inpatients. *Med Care*. 2011;49:1068–1075.
2. PeraHealth Corporation. Available at: http://www.perahealth.com/. Accessed June 23, 2013.
3. HIStalk Interview with Michael Rothman, October 25, 2010. Available at: http://histalk2.com/2010/10/25/histalk-interviews-michael-rothman-president-rothman-healthcare-corporation/. Accessed June 23, 2013.
4. Bradley EH, Yakusheva O, Horvitz LI, et al. Identifying patients at increased risk for unplanned readmission. *Med Care*. 2013;51: 761–766.
5. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. *JAMA*. 2011;306: 1688–1698.
6. Manyika J, Chui M, Brown B, et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. Washington: McKinsey Global Institute. Available at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation. Accessed July 19, 2013; 2011.
7. Cholleti S, Post A, Gao J, et al. Leveraging derived data elements in data analytic models for understanding and predicting hospital readmission. *AMIA Annu Symp Proc*. 2012. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3540449/. Accessed July 13, 2013.
8. Department of Health and Human Services, Centers for Medicare and Medicaid Services. Revision to State Operations Manual (SOM)—Hospital Appendix A. Center for Clinical Standards and Quality/Survey and Certification Group. Available at: http://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/SurveyCertificationGenInfo/Downloads/Survey-and-Cert-Letter-13-32.pdf. Accessed July 17, 2013.
9. Hasan O, Meltzer D, Shaykevich SA, et al. Hospital readmission in general medicine patients. *J Gen Int Med*. 2009;25:211–219.
10. Amarasingham R, Moore BJ, Tabak YP, et al. An automated model to identify heart failure patients at risk for 30-day readmission or death using Electronic Medical Record data. *Med Care*. 2010;48: 981–988.
11. Donzé J, Aujesky D, Williams D, et al. Potentially avoidable 30-day hospital readmissions in medical patients: derivation and validation of a prediction model. *JAMA Intern Med*. 2013;173:632–638.