

Personalized care, comparative effectiveness research and the electronic health record

Richard H. Scheuermann^a and Henry Milgrom^b

^aDivision of Biomedical Informatics and Department of Pathology, U.T. Southwestern Medical Center, Dallas, Texas and ^bDepartment of Pediatrics, National Jewish Health, University of Colorado Denver School of Medicine, Boulder, Colorado, USA

Correspondence to Henry Milgrom, Department of Pediatrics, National Jewish Health, University of Colorado Denver School of Medicine, CO 80206, USA
E-mail: milgromh@njhealth.org

Current Opinion in Allergy and Clinical Immunology 2010, 10:168–170

Personalized medicine, comparative effectiveness research (CER) and the electronic health record (EHR) are three emerging advances that promise to have an enormous impact on the future of healthcare. CER aims to establish best interventions for a given problem, whereas personalized medicine tailors medical treatment to the individual characteristics of each patient. It strives to match the right treatment to the right patient at the right time [1]. Both CER and personalized medicine involve quality evidence-based decisions, but CER evaluates treatment interventions in study cohorts, whereas personalized medicine directs attention to one patient's individual qualities in order to define optimal clinical decisions. A most compelling example of the power of personalized medicine is the evolution of the treatment of acute lymphoblastic leukemia, the most common cancer among children and adolescents [2]. In 1962, only 4% of children with the disease survived 10 years after diagnosis. Today, with the use of genetic screening technologies and improved treatment options, the cure rate has risen to over 80%. It is reasonable to expect that many more successful examples of personalized medicine will be seen in the near future. The challenge, as well as the opportunity, of personalized medicine lies in its capacity to generate and effectively utilize quantitative data that match the complexity of the disease [3].

It might be argued that physicians have always practiced personalized medicine. We have certainly recognized for a long time that patients can respond very differently to the same medication, and we have modified treatments and altered drug dosages and routes of administration to improve the efficacy and tolerability of clinical regimens. However, it is only recently that a patient's molecular information has come to play an important role in the selection of therapy, bringing attention to the opportunities of translating advances in molecular medicine into clinical practice. It appears ever more probable that such information will provide unmatched understanding of a

patient's condition, including susceptibility to disease and its progression, as well as likely drug responsiveness. The ability to classify diseases into distinct molecular subcategories will give rise to new paradigms for treatment. Such information will also form the basis for preventive interventions. The economic rationale for healthcare decisions will be based increasingly on the cost savings realized through preventive approaches to patient management [4].

EHRs promise to transform the way we manage health and disease. Conversion of the traditional paper-based medical chart into an electronic format offers several significant improvements. EHRs are more easily accessible by the numerous members of the healthcare team often involved in the management of complex patients. EHRs are more easily portable for our increasingly mobile patient populations. Decision support systems sustained by EHRs can help to reduce medical errors and to recognize potential adverse reactions to complex drug combinations.

For many years, clinical epidemiology and related fields have sought to improve our understanding of treatment effects at the individual level by analyzing subgroup effects and developing clinical prediction rules. Perhaps the most important goal of CER is to broaden and deepen such information, providing tools for matching medical care much more precisely to individual patients [5^{••}]. In support of CER, EHRs can capture patients' genetic and other individual health information accumulated in the course of routine healthcare, clinical trials and other studies. Secure, ongoing collection of gene-based and other molecular test data from EHRs and clinical laboratories and linking these with population-focused patient registries can support research on relationships among personal traits, interventions and outcomes. In translating CER to practice, health information technology can ensure that evidence pertaining to personalized medicine is present and actionable at the point of decision-making, enabling patients and their physicians to consider patients' personal traits when weighing the risks and benefits of alternative treatment options [6].

The U.S. National Institutes of Health have been stimulating the development of electronic systems to support CER and other types of clinical research through the Clinical and Translational Science Award

(www.ctsaweb.org), which is designed to transform the conduct of clinical research at academic health centers (AHCs) around the country. Systems designed to capture patient health information have focused on individual patient encounters. However, many research questions can only be addressed by aggregating patients into populations with similar characteristics for comparative analysis using statistical methods to identify differences between populations (e.g. between a cohort of patients with a particular disease and healthy controls). For this reason, many health centers have deployed data warehouses, like Informatics for Integrating Biology and the Bedside (i2b2) data warehouse [7], that contain data extracted from EHRs and restructured into dimension and fact tables to support statistical analysis. In considering the use of EHRs for CER and other research purposes, several issues need to be addressed.

Data completeness and accuracy

Different kinds of data are required to address different research questions. For example, disparities research is focused on the comparison of healthcare practices between different racial-segregated, ethnic-segregated, age-segregated or sex-segregated populations. Therefore, it is critical that complete demographic information is captured in the EHR for these purposes. Unfortunately, these data fields may not be required by the EHR data capture system, and data entry professionals focused on patient management may not appreciate the value of capturing these data into the EHR. Indeed, in an evaluation of health records for approximately 10 000 asthma patients at one of our institutions, race was listed as not recorded for 91%. Thus, the availability of EHR can also encourage a dialog between the data providers and the data users about issues related to accuracy and completeness in order to improve the quality of the data in the EHR.

Standardized data versus free text

The availability of common data elements (e.g. sex) with standardized values (e.g. man, woman, unknown) greatly facilitates research analysis. However, much of the valuable data found in EHR are contained within physicians' notes represented in free text. Manual conversion of information in free text into structured data is labor intensive. In order to capture these valuable data embedded in free text physicians' notes, several groups have been developing and evaluating automated natural language processing methods to determine their ability to accurately extract and represent the information contained. Methods focused on the use of rules-based systems armed with dictionaries of relevant medical terms and negation extraction modules are surfacing as the preferred methods to convert the free text found in EHR into structured, standardized representations [8**].

But even structured information can be problematic if the user is unfamiliar with the primary application of the data in the EHR system in which it is captured. For example, ICD-9/10 codes, which correspond to structured representations of disease, can be found in most EHR systems. However, the presence of an International Classification of Diseases (ICD) code in an EHR does not necessarily mean that the patient suffers from the disease represented by the ICD code. In many cases, ICD codes are included in the medical record as part of the differential diagnosis to justify the ordering of a particular test to rule out a disease. Thus, in cases in which the ordered test gives results in the normal range, the patient has actually been determined not to carry the disease. Thus, even with well structured standardized data, it is critical to appreciate the context in which the information is captured and used.

Patient rights

Two basic principles have emerged as the basis of policies about the ethical involvement of human patients in biomedical research—privacy and autonomy [9*]. Although the use of a comprehensive EHR database in a de-identified fashion where all Health Information Portability and Accountability Act of 1996 (HIPAA)-defined protected health information has been censored can protect patients' privacy, data use without informed consent violates the principle of autonomy—the free choice of individuals to participate or not in a particular research project. However, it may be impossible to prospectively obtain informed consent for research that might be conducted in the future. And so, in order to take advantage of the valuable data contained in comprehensive EHR databases, while still respecting patients' autonomy, many AHCs have implemented procedures that make available de-identified EHR data by default but allow patients to opt out of research involving their specimens or data if they so choose. In this way AHCs can still pursue their mission to improve patient health through research while respecting the privacy and autonomy of their patient population.

Conclusion

EHRs will come to play an essential role in the support of both CER and personalized medicine to identify optimal strategies for the management of patients' health and disease. Personalized medicine promises to transform healthcare by using molecular analysis to achieve optimum medical outcomes. As CER informs treatment decisions for cohorts and for the individual patient, it will also advance and direct innovation. CER, personalized medicine and EHR offer complementary advantages of great potential for health. At a time when healthcare systems are coming under increasing financial pressure, this potential must be realized.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

- 1 Haldorsen EMH. The right treatment to the right patient at the right time. *Occup Environ Med* 2003; 60:235–236.
 - 2 Pui CH, Evans WE. Treatment of acute lymphoblastic leukemia. *N Engl J Med* 2006; 354:166–178.
 - 3 West M, Ginsburg GS, Huang AT, *et al.* Embracing the complexity of genomic data for personalized medicine. *Genome Res* 2006; 16:559–566.
 - 4 Conti RD, Veenstra L, Armstrong K, *et al.* Personalized medicine and genomics: challenges and opportunities in assessing effectiveness, cost-effectiveness, and future research priorities. *Med Decis Making*. 2010; 0:0272989X09347014v1.
 - 5 Garber AM, Tunis SR. Does comparative-effectiveness research threaten personalized medicine? *N Engl J Med* 2009; 360:1925–1927.
- This thoughtful perspective addresses the perceived contradiction between the concepts of CER and PM. It finds that CER is a key to individualized care and innovation, not a threat.

- 6 Osheroff JA, Teich JM, Middleton B, *et al.* A roadmap for national action on clinical decision support. *J Am Med Inform Assoc* 2007; 14:141–145.
- 7 Weber GM, Murphy SN, McMurry AJ, *et al.* The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. *J Am Med Inform Assoc* 2009; 16:624–630.
- 8 Uzuner O. Recognizing obesity and comorbidities in sparse data. *J Am Med Inform Assoc* 2009; 16:561–570.
- This paper reports the results of the i2b2 Obesity Challenge evaluating the use of several different natural language processing approaches for the interpretation of clinical discharge notes. The findings suggest that systems that filter the noisy portions of the narratives, project dictionaries of disease names onto the remaining text, apply negation extraction, and process the text through rules provide the best results.
- 9 Malin B, Karp D, Scheuermann RH. Technical and policy approaches to balancing patient privacy and data sharing in clinical and translational research. *J Investig Med* 2010; 58:11–18.

This paper summarizes the key ethical issues related to participation as research subjects in clinical studies. In addition to providing an overview of the concepts of privacy and autonomy, the authors discuss the risks of re-identification using complex data sets, including genetic data, even if the records lack protected health information.