

Electronic Health Records, Medical Research, and the Tower of Babel

Rebecca D. Kush, Ph.D., Edward Helton, Ph.D., Frank W. Rockhold, Ph.D., and C. David Hardison, Ph.D.

We can travel almost anywhere in the world and find a machine that will dispense local currency, taking the money from our home account with the use of a bank card. Yet, when we go from a primary care physician to a specialist in our home town, we must begin at square one, providing the new doctor's office with all our demographic and medical information, often by completing paper forms. If we were traveling abroad and needed access to our health information, we would face formidable difficulties. Financial institutions have for years developed and used standards for the electronic exchange of financial information, but medicine has lagged far behind in the electronic exchange of medical information. Anyone who looks beyond paper medical charts quickly becomes aware of the bewildering array of information standards being used — resulting in a veritable electronic Tower of Babel.

As patients interact with increasing numbers of clinicians and caregivers, it seems reasonable to expect their health care data to follow them. In disasters such as hurricanes or fires, paper records may be lost forever. Fortunately, electronic health records are gradually becoming a reality. When an electronic health record system is implemented, however, it is frequently customized for the users. This means that even record systems developed by the same vendor could collect the same information in different ways for different institutions. For example, the relationship between an adverse event and a particular therapy could be classified as “no,” “unlikely,” “possible,” “definite,” or “probably” by one institution; as “not related,” “doubtful,” “possible,” “very likely,” or “probable” by a second; and as “no,” “yes,” or “unknown” by a third. Without the use of common vocabularies, it is impossible not only for a given hospital's computer system to understand a patient record from another hospital, but also for researchers to compare data across organizations or to collect sufficient data to make informed decisions. Similarly, the Food and Drug Administration (FDA) houses the largest repository of clinical research information in the world, yet it cannot efficiently compare pharmaceutical prod-

ucts of the same class: many of the data exist only on paper or in numerous different electronic formats that cannot be readily integrated into a single database.

Data standards for health care and medical research can permit data to be integrated and analyzed in data sets that are sufficiently robust to inform valid decision making so that health care benefits from adequate research information. Appropriate assessments of a therapy's safety often rely on statistics-based graphic computer tools for signal detection that work best when the data are in a standard format. More generally, data collected in a standard format can be more easily obtained, analyzed, and used for research, thereby supporting the advancement of medical knowledge.

Progress has been made in finding such a common global format or data language (Table 1). Health Level Seven (HL7), a standards-development organization accredited by the American National Standards Institute and headquartered in Ann Arbor, Michigan, has been devising standards for the exchange of health care information for the past two decades. Ten years ago, the Clinical Data Interchange Standards Consortium (CDISC) began to develop standards for medical research data. In 2001, recognizing the large intersection between these two arenas, CDISC and HL7 began to harmonize their standards. Together, they have developed a structured way of describing and documenting the information requirements of medical research through the Biomedical Research Integrated Domain Group (BRIDG) model. This BRIDG model can be used to represent the domain of medical research in work being done on health care standards so that the systems supporting health care and those supporting research will essentially speak the same language.

CDISC has also been working with HL7, the FDA, and an initiative called Integrating the Healthcare Enterprise. This initiative would make it easier for physicians to conduct medical research while practicing medicine and would make it possible for data to be entered only once and then be used for multiple downstream purposes

Table 1. Initiatives for the Exchange of Health Care Information.

Organization or Model	Definition
Health Level Seven (HL7)	One of several standards-development organizations operating in the health care arena that are accredited by the American National Standards Institute; the domain is clinical and administrative data, and the mission is to provide standards for interoperability that improve care delivery, optimize work flow, reduce ambiguity, and enhance knowledge transfer among all of its stakeholders
Reference Information Model	An object model created as a pictorial representation of the clinical data; identifies the life cycle of events that an HL7 message will carry, explicitly representing the connections between information carried in the fields of HL7 messages
Clinical Data Interchange Standards Consortium (CDISC)	A global, neutral standards-development organization with a mission to develop and support global, platform-independent data standards that enable information-system interoperability to improve medical research and related areas of health care
Study Data Tabulation Model	A content standard that is cited as a specification in the final guidance documentation issued by the Food and Drug Administration (FDA); provides a standard format for data in submissions to regulatory organizations for review and to fill cross-study databases containing research results for the FDA and research sponsors
Operational Data Model	A transport standard that allows for the use of any standards-compliant data-collection application or tool to readily exchange information with another such application or tool; supports audit trails needed for FDA compliance
Clinical Data Acquisition Harmonization	Collaborative initiative based on the FDA's Critical Path Initiative Opportunity List (2006); will result in standards for the collection of clinical research data that will facilitate investigators' participation in research studies and also make it easier for downstream integration of clinical research data
Biomedical Research Integrated Domain Group Model	Open, collaborative model of CDISC that is helping to bridge the gap between biomedical research and health care through the harmonization of standards for research with the relevant health care standards of HL7, the FDA, and the National Institutes of Health
eSource Data Interchange	An initiative designed to provide scenarios and recommendations for the use of electronic source data (e.g., from electronic health records) for clinical research in the context of today's regulatory environment; led to the integration profile that permits the collection of data once for multiple downstream uses

such as reporting adverse events and conducting clinical research.¹ This work is of interest to others, including the European Medicines Agency, Japan's Ministry of Health, Labor, and Welfare, and organizations such as the World Health Organization that are promoting the global registration of clinical trials. The FDA is now requesting that regulatory documents be submitted electronically, with the data in a standard CDISC format²; the agency aims to use these data to fill multistudy databases that will facilitate the review of new products and assessments of their safety. In addition, through its Critical Path Initiative,³ the FDA is encouraging the extension of these standards to the collection of research data. A collaboration called Clinical Data Acquisition Standards Harmonization is helping with this effort.

Collecting research data in a standard format and maintaining that format throughout safety reporting, submission to regulatory organizations, and warehousing can benefit all organizations involved in these processes by decreasing their

resource requirements and costs, facilitating communications and data exchange among partners, and improving data quality.⁴ So, if standards are so beneficial, why doesn't everyone use them now? One stumbling block is that many organizations developed proprietary formats before industry standards were available, and changing over from an existing system is never easy. Some institutions may claim to be waiting until the standards are finalized, though of course there will always be enhancements. But the value of global, industry-wide standards is becoming clearer, and increasing numbers of organizations are adopting these standards.

Although substantial strides have been made in the development and harmonization of these standards, support in the form of adoption, participation in the augmentation of standards, funding, and other resources is needed from everyone — hospitals and health care systems, academia, government, pharmaceutical and medical device companies, contract research organizations, ven-

dors of health information technology, and patients — if the standards that are now available are to be fully leveraged. Such broad-based buy-in would ensure that research and health care information could converge and that the data required for making the best health care decisions would become readily available. The ultimate goal, of course, is to facilitate the translation of innovations from research into the daily practice of medicine in ways that enhance the quality of care and the safety of patients.

No potential conflict of interest relevant to this article was reported.

From the Clinical Data Interchange Standards Consortium, Austin, TX (R.D.K.); SAS Institute, Cary, NC (E.H.); Glaxo-

SmithKline, King of Prussia, PA (F.W.R.); and Science Applications International Corporation, San Diego, CA (C.D.H.).

1. Clinical Data Interchange Standards Consortium. Electronic Source Data Interchange (sSDI) Group. Leveraging the CDISC standards to facilitate the use of electronic source data within clinical trials. (Accessed March 28, 2008, at <http://www.cdisc.org/eSDI/eSDL.pdf>.)
2. Study data specifications referenced in final guidance for industry: providing regulatory submissions in electronic format — human pharmaceutical applications and related submissions: using the eCTD specifications. (Accessed March 28, 2008, at <http://www.fda.gov/cder/regulatory/ersr/Studydata.pdf>.)
3. Food and Drug Administration. Critical path initiative. (Accessed March 28, 2008, at <http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html>.)
4. Rozwell C, Kush R, Helton E. Saving time and money. *Appl Clin Trials* 2007;16(6):70-4.

Copyright © 2008 Massachusetts Medical Society.

JOURNAL EDITORIAL FELLOW

The *Journal's* editorial office invites applications for a one-year research fellowship beginning in July 2009 from individuals at any stage of training. The editorial fellow will work on *Journal* projects and will participate in the day-to-day editorial activities of the *Journal* but is expected in addition to have his or her own independent projects. Please send curriculum vitae and research interests to the Editor-in-Chief, 10 Shattuck St., Boston, MA 02115 (fax, 617-739-9864), by September 30, 2008.