Building Sustainable Multi-functional Prospective Electronic Clinical Data Systems

Gurvaneet S. Randhawa, MD, MPH and Jean R. Slutsky, PA, MSPH

Abstract: A better alignment in the goals of the biomedical research enterprise and the health care delivery system can help fill the large gaps in our knowledge of the impact of clinical interventions on patient outcomes in the real world. There are several initiatives underway to align the research priorities of patients, providers, researchers, and policy makers. These include Agency for Healthcare Research and Quality (AHRQ)-supported projects to build flexible prospective clinical electronic data infrastructure that meet the needs of these diverse users. AHRQ has previously supported the creation of 2 distributed research networks as a new approach to conduct comparative effectiveness research (CER) while protecting a patient's confidential information and the proprietary needs of a clinical organization. It has applied its experience in building these networks in directing the American Recovery and Reinvestment Act funds for CER to support new clinical electronic infrastructure projects that can be used for several purposes including CER, quality improvement, clinical decision support, and disease surveillance. In addition, AHRQ has funded a new Electronic Data Methods forum to advance the methods in clinical informatics, research analytics, and governance by actively engaging investigators from the American Recovery and Reinvestment Act-funded projects and external stakeholders.

Key Words: prospective data, comparative effectiveness research, electronic health records, distributed research, governance, patient outcomes, quality improvement, clinical decision support

(*Med Care* 2012;50: S3–S6)

ALIGNING GOALS OF RESEARCH AND HEALTH CARE DELIVERY

Numerous systematic literature reviews conducted by the Agency for Healthcare Research and Quality (AHRQ) and others have identified large gaps in our knowledge on the impact of diagnostic and therapeutic clinical interventions on patient outcomes in the real world.[1–9] There are several reasons for these gaps. The biomedical research enterprise and health care delivery infrastructure have evolved independent of one other, which has led to a lack of alignment between the goals of researchers (to publish original research) and of health care providers (to provide good quality care). In the continuum of research translation—molecular pathway discovery to development of new clinical tools to outcomes and implementation research that ultimately results in improved patient outcomes—there is a heavy emphasis on molecular discovery. For example, a review of National Cancer Institute's 2007 extramural grant portfolio showed >80% of cancer genetic research (827 grants) focused on biomedical discovery and only 1 grant had a population-based focus to reduce burden of disease.[10] The incentive to perform "publishable" research rather than to improve patient outcomes, the one-off, short-term nature of many research projects, and fewer research dollars available for patient outcome-focused translational research are barriers to meet the long-term goal of health care delivery to improve patient outcomes.

Within outcomes research, there is often a lack of focus on the questions most relevant to the patients, providers, and policy makers, who are the end users of this research. There is little incentive within academia and in industry to systematically ascertain unmet users' needs and then to prioritize the research questions that can be feasibly answered. The focus of academic research may change in response to new priorities of research funding organizations and is shaped by the perspectives of peer reviewers. Industry-sponsored research is primarily designed to meet the needs of the marketplace and the regulators and has typically not focused on the most important priorities of patients and their providers. Therefore, the relevant questions are often not asked. In response, AHRQ's Effective Health Care program has made it a priority to ask the relevant questions of patients and providers.[11] The establishment of Patient-Centered Outcomes Research Institute (PCORI) is another step in this direction, reflecting growing awareness of this concern.[12] The need to align the research priorities of the patients, providers, and researchers is also recognized internationally and has led to the development of new initiatives.[13–14]

However, even if researchers are motivated and adequately funded to answer relevant questions, they often lack readily accessible data resources and information infrastructure needed to answer these questions. The challenge of addressing complex questions, such as what affects patient outcomes in a real-world clinical setting, demands a scalable electronic infrastructure that can provide high-quality, clinically rich, prospective, multi-site data for generating internally valid and generalizable conclusions in a timely and efficient manner. It is inefficient to build a data infrastructure from scratch for each

From the Center for Outcomes and Evidence, Agency for Healthcare Research and Quality, Rockville, MD.

The authors declare no conflict of interest.

Reprints: Gurvaneet S. Randhawa, MD, MPH, Center for Outcomes and Evidence, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850. E-mail: gurvaneet.randhawa@ahrq.hhs.gov.

ISSN: 0025-7079/12/5007-00S3

new research project. Further, it is a missed opportunity to create a research infrastructure that does not also meet the time-sensitive needs of clinical care. The development of a scalable electronic infrastructure that is sufficiently flexible to meet the needs of patients and providers, as well as researchers and policy makers, is an essential step to efficiently generate valid answers to questions that are fundamental to improving patient outcomes. This paper highlights AHRQ's activities in building a sustainable, scalable electronic infrastructure designed to meet the needs of diverse users.

BENEFITS OF AN ELECTRONIC HEALTH RECORD (EHR)-BASED INFRASTRUCTURE

The implementation of Health Information Technology for Economic and Clinical Health Act gave a tremendous boost to activities to meaningfully use Health Information Technology.[15] The migration from a paper-based to an EHR-based health care delivery system not only can help in performing current tasks more efficiently, but also it is an opportunity to perform activities that cannot easily be done at present. The latter includes appropriate clinical decision support (CDS) in real time at the point of care, using diverse clinical data for quality improvement (QI), performing comparative effectiveness research (CER), conducting comprehensive disease surveillance, rapid detection of adverse clinical outcomes, and a better understanding of the relationship between surrogate and health outcomes.

Studies using EHR data can mitigate several limitations of current data-gathering approaches based on administrative claims-based data and randomized clinical trials (RCTs). The limitations of administrative claims-based databases severely hamper their potential to be used in large-scale, efficient, and valid observational CER. Because these databases were created for billing and not research, information useful for research is either scant or absent. For example, these databases lack details on type of diagnostic test and actual test results, presence of comorbidity, severity of disease, type of surgical procedure or device used, use of over-the-counter drugs and dietary supplements, details on diet, alcohol use and smoking, and information on patient-reported outcomes (PROs). The lack of each one or a combination of these factors may lead to erroneous conclusions due to different types of bias (eg, selection, classification, and ascertainment) and confounding. In addition, data from these databases are available only after considerable time lag, often several years, which makes them useless for timely decisions on new and emerging technologies. An EHR contains rich clinical information and if data are insufficient, it is feasible to add new data fields and collect new information at point of care. Therefore, linkage to EHRs can overcome several fundamental limitations of claims-based databases.

EHR-based trial can improve the generalizability of RCTs. Well-designed RCTs can yield internally valid conclusions; however, they are typically conducted in highly selected patients and clinical facilities, which hamper generalizability of results to a typical patient in a real-world clinical setting. For example, patients may have more comorbidities or be much older or younger than trial participants. RCTs often focus on surrogate or intermediate end points and not health outcomes. Unless the natural history of a disease is well characterized, one cannot readily extrapolate results from surrogate to health outcomes. RCTs are not suitable for detecting very rare events, long-term outcomes, or interrelationships of multiple coexisting conditions. This study design has not been adapted to answer important clinical and policy-relevant questions such as adoption and impact of a particular technology in different health care delivery sites or by different types of providers, preferences of a patient or provider for a certain technology or outcome, or factors affecting adherence to an intervention. In addition, RCTs tend to be fairly expensive, one-off studies that are often underpowered and of short duration, thereby limiting the usefulness of their conclusions to the health care decision-makers. An EHR-based RCT in a real-world setting can efficiently identify eligible patients using real-time CDS, has the flexibility to link additional data systems as needed, and can yield generalizable results on health outcomes.

AHRQ EXPERIENCE WITH DISTRIBUTED RESEARCH NETWORKS (DRNs)

AHRQ work with DRNs began with 2 projects in 2007. These projects were part of AHRQ's network for Developing Evidence to Inform Decisions about Effectiveness (DEcIDE) (a part of Effective Health Care program), which conducts CER. The DEcIDE DRN projects were initiated to explore the policy and scientific issues relevant to this new approach.

The first DEcIDE project—Distributed Ambulatory Research in Therapeutics Network (DARTNet)—created a new network by linking different EHRs of a practice-based research network based at University of Colorado.[16] This network was designed to serve the needs of the clinicians to improve quality of care and to perform CER. The first project focused on CER of diabetes drugs. DARTNet demonstrated that it is feasible to link different EHRs to each other and to hospital, laboratory, billing, and pharmacy databases. This enabled collection of information on: relevant laboratory test values (ie, glycemic and liver toxicity markers), patient's height and weight, prescriptions and refills, over-the-counter drug and dietary supplement use, and alcohol use. This information can be used for dual purposes: by a clinician to assess which patients are on target clinically and who need a modification of the treatment plan and by a researcher to perform CER on different clinical interventions. It therefore intimately links the delivery of care, QI, and research from data obtained at point of care. Information from DARTNet can be used by a clinical organization to benchmark its performance vis-à-vis its peers, which enables it to learn from better-performing partnering organizations in order to improve quality of care and patient outcomes. This information can also be used by researchers to answer questions ranging from CER to pharmacoepidemiology. Therefore, it is an example of a flexible, EHR-based infrastructure that can meet the needs of patients, providers, researchers, and policy makers.

The other DEcIDE DRN project enhanced an existing HMO research network and outlined the technical design,

key infrastructure components, and organizational structure of this network for conducting CER and demonstrated its capability in performing large scale, population-based CER.[17] In addition, this network has the capability of performing disease and safety surveillance, detecting rare clinical events, and ascertaining utilization of technologies. The ability to have local control of data and share limited, relevant data necessary to answer a research question on a as-needed basis (compared with the traditional model of creating a single centralized data repository from all partner organizations) is critical to ensure that we can perform CER on a wide range of topics in diverse health care delivery sites while protecting both patient privacy and confidentiality as well as the proprietary information of clinical organizations.

It is important but quite challenging to collect PROs in routine clinical care. The US Preventive Services Task Force recommends screening for depression (that necessitates collection of PROs) only in those primary care practices that have an adequate staff support system for screening and follow-up, which are a minority in this nation.[18] For its second project, DARTNet was tasked to work on CER of depression. This project has demonstrated the feasibility of sustainably collecting Patient Health Questionnaire-9 data from primary care clinics (79% of practices continued to collect data after more than a year of end of DEcIDE program funding), if the system of data collection is customized to meet the unique workflow needs of clinicians and the health care delivery team at each site.[19]

Although several technical challenges were encountered when linking information across different electronic systems and adapting current analytical techniques to EHR-based data to perform CER, our experience showed that it is easier to overcome the technical challenges compared with the organizational, sociological, and regulatory challenges encountered when increasing the scale and diversity of a DRN, developing a governance structure with adequate and meaningful participation of relevant stakeholders, and making it a sustainable project. Hence, these were requirements for all new AHRQ-supported research projects that aimed to build prospective electronic clinical data infrastructure.

GOALS OF CURRENT AMERICAN RECOVERY AND REINVESTMENT ACT (ARRA)-CER INFRASTRUCTURE PROJECTS

The ARRA targeted $1.1 billion for CER.[20] More than $100 million of these were invested to build electronic clinical systems for collecting prospective patient-centered outcomes data that can be used for CER, QI, CDS, and surveillance. One portion of this infrastructure investment was used to create a new Electronic Data Methods (EDM) forum to provide diverse stakeholders a venue to share common challenges, to propose solutions, and to advance the informatics and analytical methods and knowledge on governance in order to shape the national dialog with empirical data.

AHRQ released 3 Request For Awards (RFAs) for building electronic infrastructure and performing CER: Prospective Outcome Systems using Patient-specific Electronic data to Compare Test and therapies, enhanced registries for QI and CER, and Scalable DRN,[21–23] and 1 RFA for the EDM forum.[24] The EDM forum was created to advance the interdisciplinary methods in clinical informatics, analytics, governance, and learning health care network by actively engaging investigators from the ARRA-funded projects and a diverse group of external stakeholders.

The common requirements across all infrastructure development RFAs were to:

- Link multiple health care delivery sites (inpatient, ambulatory care, long-term care).
- Connect multiple databases and work across multiple information technology architectures.
- Focus on AHRQ priority populations and CER conditions (including Institute of Medicine's CER topics).
- Collect prospective, patient-centered outcomes.
- Conduct CER.
- Generate valid and generalizable conclusions.
- Focus on governance.
- Plan to be sustainable.

The additional requirements of scalable DRN RFA were to: build multiple cohorts (at least 4 pairs of increasing complexity), conduct CER in at least 2 unrelated priority conditions, collect and analyze data in near-real time (to overcome the challenge of long time lag in obtaining new data in claims-based databases), and demonstrate the ability to scale up the network. The use of open-source software was strongly encouraged.

The additional requirements of the enhanced registry RFA were to: build on existing registry (because the 3-year timeline of the projects did not allow creating a new registry from scratch) and demonstrate the capability of the electronic infrastructure to perform a project on QI and on CER.

Data governance is an important issue common to all projects, because building data systems across different states, diverse health care delivery sites, and different owners of data requires tackling the variability in rules and regulations designed to protect privacy and confidentiality of a patient's information as well as the proprietary interests of partner organizations. However, data governance is only 1 element of governance of a network. The network governance structure ought to allow diverse partner and stakeholder participation in order to understand their needs and requirements, an explicit process to identify and prioritize CER and QI topics, the ability to manage different funding sources and the attendant conflicts of interest, and to be able to balance the needs of researchers with the needs of clinicians, patients, and policy makers. All these considerations, along with the creation of new information useful for decision-making by diverse users, are essential to ensure the long-term sustainability of these networks.

The 11 grants awarded by these 3 RFAs cover the geographic continuum from community-based to state-wide to national projects.[25] The study designs include RCTs, cluster randomized trial, and observational studies to perform CER in a variety of priority populations and conditions. The informatics approaches to create the infrastructure use diverse electronic tools and methods. The goal of this supplement is to share the experience of these projects in development of sustainable prospective electronic data infrastructure useful for

CER and for delivering high-quality clinical care. The authors of the succeeding papers address challenges in 3 domains relevant to conducting CER in electronic databases: analytics, informatics, and governance. Advancing our methods and knowledge in these domains is important to create networks that can meet the needs of patients, providers, researchers, and policy makers. If the information from these systems is valuable to these end users, it increases the chance of their long-term participation, which is essential for sustainability of this new electronic infrastructure.

ACKNOWLEDGMENTS

The authors acknowledge the helpful comments of William Lawrence, Joanna Siegel, and Yen-Pin Chiang (all at AHRQ) and the editorial assistance of Beth Johnson (at AcademyHealth).

REFERENCES

1. Wilt TJ, Shamliyan T, Taylor B, et al. *Comparative Effectiveness of Therapies for Clinically Localized Prostate Cancer. Comparative Effectiveness Review No. 13. (Prepared by Minnesota Evidence-based Practice Center under Contract No. 290-02-0009.)*. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
2. Seidenfeld J, Samson DJ, Rothenberg BM, et al. *HER2 Testing to Manage Patients With Breast Cancer or Other Solid Tumors. Evidence Report/Technology Assessment No. 172. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center, under Contract No. 290-02-0026.) AHRQ Publication No. 09-E001*. 2008; Rockville, MD: Agency for Healthcare Research and Quality.
3. Myers ER, Havrilesky LJ, Kulasingam SL, et al. *Genomic Tests for Ovarian Cancer Detection and Management. Evidence Report/Technology Assessment No. 145. (Prepared by the Duke University Evidence-based Practice Center under Contract No. 290-02-0025.) AHRQ Publication No. 07-E001*. 2006; Rockville, MD: Agency for Healthcare Research and Quality.
4. Bennett WL, Wilson LM, Bolen S, et al. *Oral Diabetes Medications for Adults With Type 2 Diabetes. An Update. Comparative Effectiveness Review No. 27. (Prepared by Johns Hopkins Evidence-based Practice Center under Contract No. 290-02-0018.) AHRQ Publication No. 11-EHC038*. 2011; Rockville, MD: Agency for Healthcare Research and Quality.
5. Matchar DB, Thakur ME, Grossman I, et al. *Testing for Cytochrome P450 Polymorphisms in Adults With Non-Psychotic Depression Treated With Selective Serotonin Reuptake Inhibitors (SSRIs). Evidence Report/Technology Assessment No. 146. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-02-0025.) AHRQ Publication No. 07-E002*. Rockville, MD: Agency for Healthcare Research and Quality; 2007.
6. Bradley L, Palomaki GE, Dotson WD, et al. Can UGT1A1 genotyping reduce morbidity and mortality in patients with metastatic colorectal cancer treated with irinotecan? Prepared for: Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group, Project Number 0208235.036. January 2009.
7. Terasawa T, Issa D, Castaldi P, et al. *Systematic Reviews on Selected Pharmacogenetic Tests for Cancer Treatment: CYP2D6 for Tamoxifen in Breast Cancer, KRAS for anti-EGFR antibodies in Colorectal Cancer, and BCR-ABL1 for Tyrosine Kinase Inhibitors in Chronic Myeloid Leukemia (Prepared by Tufts EPC)*. Rockville, MD: Agency for Healthcare Research and Quality; 2010.
8. Grant M. Angioplasty and stenting of the cervical carotid artery with embolic protection of the cerebral circulation. *Technol Eval Cent*. 2010;12:1–52.
9. Technology Evaluation Center. Special Report: Early Intensive Behavioral Intervention Based on Applied Behavior Analysis among Children with Autism Spectrum Disorders. TEC Assessment. Blue Cross and Blue Shield Association. Available at: http://www.bcbs.com/blueresources/tec/vols/23/special-report-autism.html.
10. Schully SD, Benedicto CB, Gillanders EM, et al. Translational research in cancer genetics: the road less traveled. *Public Health Genomics*. 2011;14:1–8.
11. Agency for Healthcare Research and Quality. Effective Health Care Program [Home-AHRQ Effective Healthcare Program web site]. Available at: http://effectivehealthcare.ahrq.gov/. Accessed November 1, 2011.
12. PCORI [Patient-Centered Outcomes Research Institute web site]. Available at: http://www.pcori.org/. Accessed November 1, 2011.
13. The James Lind Alliance [James Lind Alliance web site]. 2011. Available at: http://www.lindalliance.org. Accessed November 1, 2011.
14. National Institute for Health Research. INVOLVE [INVOLVE web site]. July 25, 2011. Available at: http://www.invo.org.uk. Accessed November 1, 2011.
15. The Office of the National Coordinator for Health Information Technology. HITECH Programs [Health IT web site]. January 27, 2011. Available at: http://healthit.hhs.gov/portal/server.pt?open=512&objID=1487&parentname=CommunityPage&parentid=58&mode=2&in_hi_userid=11113&cached=true. Accessed November 1, 2011.
16. Pace W, West D, Valuck R, et al. *Distributed Ambulatory Research in Therapeutics Network (DARTNet): Summary Report. (Prepared by University of Colorado DEcIDE Center Under Contract No. HHSA290020050037I TO2.)*. Rockville, MD: Agency for Healthcare Research and Quality Available at: http://effectivehealthcare.ahrq.gov/reports/final.cfm; 2009.
17. Brown J, Holmes J, Maro J, et al. *Design Specifications for Network Prototype and Cooperative to Conduct Population-based Studies and Safety Surveillance. Effective Health Care Research Report No. 13. (Prepared by the DEcIDE Centers at the HMO Research Network Center for Education and Research on Therapeutics and the University of Pennsylvania Under Contract No. HHSA290020050033I T05.)*. Rockville, MD: Agency for Healthcare Research and Quality Available at: http://effectivehealthcare.ahrq.gov/reports/final.cfm; 2009.
18. U.S. Preventive Services Task Force. Screening for Depression in Adults [Topic Page web site]. December 2009. Available at: http://www.uspreventiveservicestaskforce.org/uspstf/uspsaddepr.htm. Accessed November 1, 2011.
19. Pace W, Graham D. An Introduction to DARTNet [PowerPoint presentation]. http://www.ctsaweb.org/uploadedfiles/DARTNet%20NIH%20Intro%202011(2).pdf. Accessed November 1, 2011.
20. US Department of Health and Human Services. Comparative Effectiveness Research Funding [HHS Recovery Act web site]. June 2010. Available at: http://www.hhs.gov/recovery/programs/cer/index.html. Accessed November 1, 2011.
21. Agency for Healthcare Research and Quality. ARRA-AHRQ Recovery Act 2009. Limited Competition: PROSPECT Studies: Building New Clinical Infrastructure for Comparative Effectiveness Research (R01) [PROSPECT Funding Opportunity Announcement web site]. December 8, 2009. Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-HS-10-005.html. Accessed November 1, 2011.
22. Agency for Healthcare Research and Quality. ARRA OS Recovery Act 2009. Limited Competition: Enhanced Registries for Quality Improvement and Comparative Effectiveness Research (R01) [Enhanced Registries Funding Opportunity Announcement web site]. January 21, 2010. Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-HS-10-020.html. Accessed November 1, 2011.
23. Agency for Healthcare Research and Quality. ARRA OS: Recovery Act 2009. Limited Competition: Scalable Distributed Research Networks for Comparative Effectiveness Research (R01) [Scalable Distributed Research Networks Funding Opportunity Announcement web site]. January 21, 2010. Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-HS-10-015.html. Accessed November 1, 2011.
24. Agency for Healthcare Research and Quality. ARRA-AHRQ Recovery Act 2009. Limited Competition: Electronic Data Methods (EDM) Forum for Comparative Effectiveness Research (U13) [EDM Forum Funding Opportunity Announcement web site]. December 8, 2009. Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-HS-10-006.html. Accessed November 1, 2011.
25. AcademyHealth. Projects & Investigators [EDM Forum web site]. October 19, 2011. Available at: http://www.edm-forum.org/publicgrant/About/ProjectsInvestigators/. Accessed November 1, 2011.

