

Data Standards in Diabetes Patient Registries

Rachel L. Richesson, Ph.D., M.P.H.

Abstract

Widespread adoption of electronic health records (EHRs) and expansion of patient registries present opportunities to improve patient care and population health and advance translational research. However, optimal integration of patient registries with EHR functions and aggregation of regional registries to support national or global analyses will require the use of standards. Currently, there are no standards for patient registries and no content standards for health care data collection or clinical research, including diabetes research. Data standards can facilitate new registry development by supporting reuse of well-defined data elements and data collection systems, and they can enable data aggregation for future research and discovery. This article introduces standardization topics relevant to diabetes patient registries, addresses issues related to the quality and use of registries and their integration with primary EHR data collection systems, and proposes strategies for implementation of data standards in diabetes research and management.

J Diabetes Sci Technol 2011;5(3):476-485

Introduction

Lack of data standards has been the prime obstacle to routine use of health care data for secondary purposes, such as research or quality monitoring, on a national scale. Identification of health data standards is a priority for the Office of the National Coordinator of Health Information Technology (ONC), created through an executive order in 2004 and legislatively mandated in the Health Information Technology for Economic and Clinical Health Act of 2009.¹ The ONC's strategy for national exchange of electronic health information

includes incentive programs such as the high-profile "meaningful-use" requirements. The anticipated outcome of these incentive programs is the widespread adoption of electronic health records (EHRs) with functionality needed to drive improvements in health care process and patient outcomes.

Concurrently, a proliferation of patient registries is underway, fueled by advances in information technology and by emphasis on patient registries as a key component

Author Affiliation: Department of Pediatrics, University of South Florida College of Medicine, Tampa, Florida

Abbreviations: (AHRQ) Agency for Health Research and Quality, (CDE) common data element, (CDISC) Clinical Data Standards Interchange Consortium, (Diabe-DS) Diabetes Data Strategy, (EHR) electronic health record, (FDA) Food and Drug Administration, (HbA1c) hemoglobin A1c, (HIPAA) Health Insurance Portability and Accountability Act, (HITSP) Healthcare Information Technology Standards Panel, (HL7) Health Level 7, (LOINC) logical observation identifiers names and codes, (ONC) Office of the National Coordinator of Health Information Technology, (PHR) personal health record, (SNOMED CT) Systematized Nomenclature of Medicine—Clinical Terms

Keywords: chronic disease, controlled vocabularies, data standards, diabetes, disease management (systems), medical informatics, patient care management, registries

Corresponding Author: Rachel L. Richesson, Ph.D., M.P.H., Department of Pediatrics, University of South Florida College of Medicine, 3650 Spectrum Blvd., Suite 100, Tampa, FL 33612; email address Rachel.Richesson@epi.usf.edu

of disease-focused research agendas. Reports, sponsored independently by the Agency for Health Research and Quality (AHRQ)² and the California Health Foundation,³ provide robust guidance for development and use of patient registries for comparative effectiveness research, evidence-based medicine, and chronic disease management. The strong interest of these agencies underscores the potential impact of patient registries on health care quality and patient outcomes.⁴

Electronic health record systems can support physician adherence to standards of care as well as promote patient compliance with medications and behavioral modification (and subsequent improvement of patient outcomes) in diabetes. In a 6-month prospective randomized trial of 29 physician teams, better indicators [specifically hemoglobin A1c (HbA1c), low-density lipoprotein cholesterol, and controlled blood pressure] were seen in patients who received customized health promotion letters that were generated using data from a patient registry.⁵ A chronic disease care management program model using a diabetes registry, clinical information systems, guideline implementation, and patient support for self-management, was associated with improved performance and metabolic outcomes in three primary care practice sites in Wisconsin and Minnesota.⁶ Others have demonstrated the coupling of health care systems and registries. Most notably, the U.S. Veterans Health Administration registry is populated continually from their electronic patient record database system.⁷ Some EHRs contain simple chronic disease registry functionality that can enable provider alerts, efficient outreach to patients, and efficient documentation of such outreach within a patient's medical record. Use of patient registries (integrated with EHR functions) as health care tools has the potential to impact tens or hundreds of thousands of individuals with diabetes.

However, any vision of integrating patient registries with EHR functions, or of combining regional diabetes registries to support national or global analyses, will require the use of standards. Data standards can facilitate aggregation of data or analytic results for a variety of purposes—many likely unforeseen at present. Standards can also support the development of new registries by allowing reuse of data elements, definitions, and systems (which often take months or years to develop). The following discussion will provide an introduction to standardization issues as they relate to diabetes patient registries and propose a direction to achieve standards-based patient registries.

Patient Registries: Definitions, Limitations, and Features

A patient registry is an organized program for the collection of a clearly defined set of data on identifiable individuals for a specific purpose.⁸ Patient registries generally use observational research designs to capture data from sampled disease *populations* to better understand disease etiology or to explore patient variation and experience among different treatment options. Registries can also be used for clinical trials recruitment,⁹ patient safety (e.g., post-marketing surveillance), and monitoring of provider performance relative to practice guidelines or target measures. Additionally, registries can support quality improvement, specifically with respect to outreaching to patients through letters, phone, or email. The term patient registry implies follow-up over time. While patient registries can support research, not all research data sets emerge from registry programs, and not all registry programs result in research-quality data sets. In type 1 diabetes, there are several independent registry efforts that have supported new understandings of the epidemiology and mechanisms of the disease. These include the Diamond study¹⁰ worldwide, the EURODIAB¹¹ in Europe, and the SEARCH study/registry in the United States.¹² In addition, there are many independent statewide surveillance programs and registries in the United States organized by state public health agencies.

There are inherent limitations of certain registry designs for certain functions, particularly in the exploration of research questions involving treatment evaluation.^{13,14} To be useful for research, a registry must have high-quality data. Two fundamental concerns related to gauging the quality of registry data include completeness of case ascertainment and validity of values for each data point.¹⁵ The completeness of case ascertainment can be assessed by comparing multiple sources. Authors in the United Kingdom found discrepancies between the prevalence of diabetes derived from epidemiological studies and those reported through the national quality improvement scheme.¹⁶ Another study found that linkage of multiple electronic data sources was significantly more sensitive than general practice registers in identifying both diabetic and high-risk subjects.¹⁷ Using a random sample of 125 charts of Medicare patients, Tang and associates¹⁸ also showed a major discrepancy between diabetes patients identified using clinical data captured in an EHR system and those identified from claims data, resulting in statistically significant differences in the

quality measures for frequency of HbA1c testing, control of blood pressure, frequency of testing for urine protein, and frequency of eye exams for diabetes patients. Others have described the problem of different case definitions and EHR query strategies and are developing an electronic medical record standard definition of diabetes that can be used in different settings.¹⁹ In addition to standardized approaches to utilizing clinical data, it is critical to ensure valid and reliable data for all registry purposes, standardization of case definitions, data definitions, and clinical diagnostic criteria. Detailed examination of representative subsamples can also be conducted to validate large survey results.²⁰

While EHRs can support both registry recruitment and data generation in a registry, the distinction between registries and EHRs must be kept clearly in mind. Registries, sometimes referred to as chronic disease management systems if they are used for that purpose, can serve a number of functions as previously described, but all are designed to collect data on *populations* for predefined purposes, and the collected data are prespecified. A discussion of available registry products for diabetes management, including many variations of the Centers for Disease Control and Prevention-funded diabetes electronic management system, and guidance for different levels of integration with EHRs has been developed by the California Healthcare Foundation.²¹ In contrast to registries, the primary purpose for EHRs is to capture data on *individuals* to support their care. There is some evidence that diabetes care can be improved with EHRs, and subsequent speculation that better EHR implementation models could improve the consistency and impact of these results.²² However, others have shown that widespread implementation of an EHR is not sufficient to improve the quality of diabetes care (as measured by provider compliance with process and treatment guidelines and intermediate diabetes patient outcomes).²³ Lester and colleagues²² synthesized the limitations of information technology in diabetes care to date and provided eight helpful “rules” for designing informatics systems that catalyze change in diabetes care.

There are legal and regulatory ramifications for patient registries, the extent of which depends largely on the purpose of the registry, the type of data collected, who will have access to the data, and what kind of information is shared with the registrant. All patient registries must, at the very least, have a posted privacy policy accessible to potential registrants prior to registration. The privacy policy should detail the purpose of the registry, who will have access to data entered by the registrants, how the

data will be used, how long the data will be maintained, and how the potential registrant can withdraw from the registry [or “unsubscribe” or “withdraw consent;” 45 CFR § 164.530(i)].

Institutions may have both an ethics board (institutional review board) and a privacy board; however, many institutions combine the responsibilities of the ethics board with the privacy board. Registries that send general information related to enrolling research studies or that allow the registrant a conduit through which they may opt to share their information with research sponsors are required to have privacy board [Health Insurance Portability and Accountability Act (HIPAA)] review and approval, as the authorization requirement may need to be altered or, more likely, a waiver is sought [45 CFR § 164.512(i)]. If the registry provides enrollment information on *specific* studies or research programs, then privacy board (HIPAA) review is generally required. (Ethics review for the registry itself is not required, although individual studies might describe the use of the registry in the recruitment plans of ethics-reviewed studies.)

Not all registries require ethics or privacy board review. Contact registries that do not promote or advertise research studies and that collect minimal patient contact information (e.g., name and email address) and minimal health information (e.g., diagnosis and age) can be thought of as information services. These information-sharing registries do not typically require privacy board or institutional review board approval, as these registries do not meet the federal definition of research (45 CFR § 46.102) nor do they typically require a waiver or alteration of the authorization requirement in section 164.512(i) of the privacy rule. However, *any* patient registry that is designed to support any type of research data collection (prospective systematic investigation) will require ethics review and approval and, therefore, privacy board review [45 CFR §§ 46.101 and 46.111(a)(7)].

Data Standards: Definition and Challenges

The interoperability of registries or registry data—including reuse of EHR or personal health record (PHR) data to populate a registry, the use of a registry to populate a patient’s EHR or PHR, or the use of registry data for clinical studies or regulatory submissions for new agents—is dependent upon the use of data standards. The quest for registry standards is complicated by the number of different registries, the variety of purposes that they serve, and the lack of a single governor of registries. Patient registries can

be tied to care functions or to research functions and therefore have two different relevant standards and regulatory communities to consider. The diversity of registry sponsors and objectives also invites confusion regarding the legal and operational definitions, their subsequent evaluation, and the best practices for the use and interpretation of registry data.²⁴ Data standards are consensual specifications for the representation of data from different sources or settings.²⁵ Standards can take many forms: system specifications, messaging syntax, data models, mapping specifications, question-and-answer (value) sets, controlled terminologies, or standardized assessment instruments. Part of the challenge for standards observance is the brutal reality that—often—any given *individual* organization or registry project perceives little immediate benefit or incentive to implement data standards. Standards become vitally important, however, when data is being exchanged or shared,^{26–29} often benefiting a *secondary* user. However, a premise of this article is that standards do indeed benefit the primary user and are worthy of using and promoting to others. Using standardized specifications for registry elements will yield tremendous savings of time and money for individual organizations, both large and small. All too often, clinical subject matter experts have spent precious time establishing homegrown definitions, which, in turn, force data analysts to spend countless hours identifying data sources and models to support those definitions. Moreover, these locally developed registries often do not contain the most valid, up-to-date versions of definitions and value sets, because individual organizations do not possess the necessary expertise. National and statewide registries can be very beneficial to individual organizations if they provide guidance and tools to support standardized data collection, are feasible (i.e., the methods to capture data are clear and sensible), and provide evidence to address meaningful issues.

Because there is no universal standards controller, standards are not of equal status in standing or in requirement. A specification can become a standard by various means (e.g., *ad hoc*, *de facto*, mandate, consensus), as described by Hammond and Cimino.³⁰ In addition, the status of a specification as a standard is entirely dependent upon the setting. For example, the standard representation for a laboratory test result is different in various ONC health care messaging scenarios (use cases) than it is for reporting drug results to the United States Food and Drug Administration (FDA). Further, the FDA and the International Conference on Harmonisation (the regulatory equivalent to the FDA in Europe and Japan)

endorse different terminology standards (Systematized Nomenclature of Medicine—Clinical Terms [SNOMED CT] versus Medical Dictionary for Regulatory Activities) for the submission of clinical and safety data related to investigational agents.

Standards (as mandated by various federal regulators, payors, and certification bodies) for health care providers are complex, subject to interpretation, and not consistently integrated with major EHR systems. Clinical research and patient registries represent less mature standards areas. In those areas, standards selection is even more confusing, as multiple generic and disease-specific standards exist, all of which are continually changing. Even when a professional community agrees on clinical content standards, the actual technical implementations can vary. For example, despite the accepted World Health Organization classification of diabetes disease subtypes, there is still tremendous variation in the representation of diabetes diagnosis types in registries and EHR systems.³¹ Hence, the pursuit of ideal standards for diabetes registries should begin with the knowledge that (1) standards do not exist in a clear sense for registries, in general; (2) data standards do not exist in a clear sense for diabetes, in general; (3) many relevant standards do exist; and (4) a clear understanding of the registry, its intended uses, and relationship to other health data sources (e.g., EHR, FDA, trial registries like ClinicalTrials.gov, and other patient or disease registries) is prudent.

Overlap between Standards and Trends in Clinical Data Standards

Standardized data include specifications for data fields (~variables) and value sets (~codes) that encode the data within these fields. The value sets might be standard lists, such as the United States Office of Management and Budget standard list of racial categories³² or standard lists such as for route of medication administration, required by the FDA.³³ Often, the code sets are whole controlled terminologies, such as SNOMED CT, logical observation identifiers names and codes (LOINC), or RxNorm (some of the U.S.-recommended standards for certain areas and referenced extensively in the various use cases and standards specifications developed by the Healthcare Information Technology Standards Panel [HITSPP] for the ONC).³⁴ An information model is a broad term that includes database models, domain models, and formalized concept models that stipulate the slots or data fields. The “standard” information models referenced by research and health care domains are wildly different. For regulated research submissions and

drug safety, the operational data model and study data tabulation model, and corresponding vocabulary slots and code lists, developed by the Clinical Data Standards Interchange Consortium (CDISC) are the standard.³⁵ For clinical data exchange between health care providers, the Health Level Seven (HL7) messaging and structured document models are the standard in the United States and many developed countries.

The use of code sets or controlled terminologies, and consequently their designation as standards, is intrinsically entangled with the nature of the information model (i.e., a “container”) in which they are used. Terminologies often represent multiple constructs that can overlap with the same constructs in an information model and cause confusion. For example, each uniquely identifiable test in the LOINC standard for laboratory test names includes the following information components: laboratory test name, laboratory test analyte, laboratory test measurement scale, and laboratory test units.³⁶ When those constructs are also included in an information model (e.g., the CDISC study data tabulation model for FDA regulatory data submissions has a specified field for “lab test units”), there is a need for guidance on how the terminology should fit into an information model to eliminate duplication or contradiction. Similarly, the concept “left arm” exists in SNOMED CT as a single concept code but also can be constructed from component concepts “arm” + “left.” The appropriate SNOMED CT concept would vary upon whether the information model context had a single specified slot for body site, or distinct data slots for body site and laterality. Similar examples abound for family history concepts (e.g., “maternal history of heart disease”) and drug concepts (e.g., “oral insulin”). These terminology–information model interactions are a major challenge to any sort of health care data exchange and have long been a dominant clinical informatics research problem.³⁷ Consequently, effective diabetes data standards should specify information model or data elements, as well as the associated coding systems.

Because possible synergies between EHRs and diabetes registries are so strong, and because there is overlap in the types of data collected, standards initiatives for diabetes registries should consider adopting EHR data standards to maximize any likelihood of sharing data between applications. These same synergies are the topic of national (disease-agnostic) discussions related to planning for national data exchange competencies. Specifically, the HITSP (an American National Standards Institute-administered cooperation of public and private sector stakeholders and a primary advisory committee

to the ONC) has organized standards recommendations by 18 topic areas (e.g., EHR laboratory results reporting, newborn screening, patient–provider secure messaging). Each topic area includes detailed narratives of envisioned interoperability scenarios (i.e., “use cases”) and the subsequent functional and data requirements.³⁴ In particular, the HITSP clinical research use case defines specific functionality required for information interchanges between registries and EHRs, though not the specific data fields.³⁸ In a separate effort, the Public Health Data Standards Consortium, a nonprofit association of federal, state, and local public health organizations, developed an informative report on approaches and benefits for interoperability between EHRs, clinical registries, and public health registries within regional Health Information Exchanges.³⁹ The report includes a case study on diabetes, including suggested diabetes-specific patient-reported data on diabetes management and care.

Content Standards: Important “Common” Diabetes Data Elements

One way around the pervasive and problematic terminology–information model interactions and other complexities of standards mechanics is to agree on the *content* that should be collected in the form of common data elements (CDEs). This has been done by other scientific and professional communities, including the American Association of Cardiology and the American Heart Association, who have a long history of successful registries—and sets of well-defined consensus-based “key data elements”—to support both research and quality monitoring.^{40–44} Common data elements are the important “units of data collection”—they are meaningful to users and relevant and usable for various purposes. There are few resources for methodology of CDE development.^{45–47} One CDE project that is specifically focused on diabetes at the point of capture (i.e., EHR) is the Diabetes Data Strategy (Diabe-DS) demonstration project.^{48,49} The project was formed in early 2009 in the HL7 EHR Working Group (with representatives from academia, professional societies, government, EHR developers, and pharmaceutical industry) to develop a repeatable process that identifies important data elements for clinical care and secondary use. This project developed narrative user scenarios (called “use cases”) to describe the capture and use of data elements in primary (patient care) and secondary (research and reporting) settings. Data elements were collected from a variety of sources and are being mapped to the U.S.-based HITSP specification and HL7 EHR system functions for patient care, clinical research, and quality measurement. The Diabe-DS

has defined a set of over 100 important elements. The relationships of these elements to other standard specifications are being formalized, though the Diabe-DS elements have not been formally vetted or endorsed by diabetes stakeholders as of this writing.

The Future of Diabetes Data Standards and Patient Registries

The U.S. AHRQ commissioned a comprehensive report on the role of patient registries for scientific, clinical, and policy purposes.¹⁵ This report, updated in late 2010, provides the most comprehensive and relevant set of best practices for registry design and framework for assessing quality of registry data for evaluating patient outcomes. These guidelines address the importance of integrated registries and EHR systems and should inform any new registry endeavor. The AHRQ guidelines also provide a broad standards strategy that focuses on finding and leveraging existing standards and on the essential role of explicit registry objectives and stakeholder consensus. Strategies for achieving content standards within a given disease area, however, have not been well defined. Some suggested themes are presented here.

Focus on Content and Consensus, not Technical Specifications.

The most difficult and important issue regarding data standards, especially in the context of this special issue on diabetes registries and technologies, is achieving consensus on the content that should be collected. Once important data are identified, the data collection formats and relationships to existing standards can be defined easily and made standard by informed technical experts. Therefore, diabetes stakeholders should focus on identifying the content. A necessary first step is to identify envisioned scenarios and *data sharing* requirements important to the field. Issues related to formal (i.e., computer-readable) representations of important diabetes concepts and data elements are complex but tractable. Technical experts can comply with technical specifications (controlled terminologies, data models, messaging syntax) as needed. The critical elements for data collection cannot be defined, however, without diabetes experts and stakeholders, and these elements will never become standard (i.e., widely or uniformly used) without consensus within the diabetes domain.

Look Around

The assessment of existing standards and activities is truly worth the effort. The intent of this article is to

inform diabetes researchers and providers that other standards efforts that partially overlap are underway and that other disease groups or practice areas (e.g., pediatrics, cardiovascular, emergency medicine) are struggling with similar issues and might provide some relevant content standards. This review has identified some, but certainly not all, resources for content that the diabetes community can leverage. It will be most fruitful to first look (hard) and leverage what has been done and focus diabetes-specific standards development efforts only in areas where there truly is no standard—recognizing that achieving consensus on standards in a treatment community this large will take much time. The types of data collected in diabetes patient registries are not unlike those for other chronic conditions, and therefore, the same types of data standards apply. For example, many patient registries collect data on patients (demographics and identifiers), various risk factors, medical history, family history, clinical observations, and laboratory values. Dietary data are particularly important in studying the etiology of diabetes and have been explored in cancer research and in the international TEDDY (The Environmental Determinants of Diabetes in the Young) project,⁵⁰ where food composition databases of participating countries have been harmonized to ensure comparability of nutrient intake estimates.^{51,52} The issues related to measurement and conditions are not trivial and will require research and experimentation to define best practices and/or methodological standards (e.g., HbA1c measurement⁵³). These issues can be identified only by expert diabetes consortia. However, they can be informed and prioritized by the labors and successes of other scientific and professional communities who are addressing the very same problems.

Focus on Purpose on Long-Term Objectives

Registries can serve various purposes, which ultimately guide the data that should be collected. As registries continue to evolve and serve multiple functions (patient care, research, community surveillance, and population health), it can become increasingly difficult to define the necessary data elements and how they should interact with various information systems (such as EHRs) to support the various registry functions. In 1998, Elwyn and colleagues⁵⁴ questioned whether registries should be considered as a clinician tool or a public health tool. Since then, diabetes registries have proliferated and are being used for both functions, underscoring the need for data capture standards in primary health care settings.

Clearly, secondary uses of health care data are facilitated by having appropriate information collected at the source.

The primary data capture (in clinical settings) should be granular enough to support one or many secondary uses, including patient registry functions, and therefore, future EHR data collection requirements should be driven by secondary use data considerations. Obviously, if registries are going to relate to EHRs, then standards in EHRs are in order, and these standards should be informed by intended secondary users. The diabetes stakeholder community can and should drive these efforts.

If there are multiple secondary uses (e.g., public health, clinical research, quality monitoring), then their requirements must be harmonized if we truly want one “collect once, reuse many” health information strategy. The diabetes stakeholder community, not standards bodies, should be defining these requirements. It is a challenge because the task (diabetes data standards) requires the cooperation of multiple highly specialized diabetes communities (e.g., endocrinologists, pediatricians, internal medicine, statistics, psychology, and immunology) in addition to various specialized technical disciplines (e.g., information technology, informatics, quality measurement, population health, research, and policy). The Diabe-DS project, a pilot project supported by the American Health Information Management Association and HL7, is focused on harmonizing these secondary uses in order to provide a minimal set of type 1 diabetes data elements that can be considered in primary care EHR data collection modules.^{49,55} The American Association of Cardiology is a great model, having developed accessible and vetted standards^{42–44,56,57} and registries to support observational research and drive evidence-based care.⁴⁰ The challenge of harmonizing clinical care data requirements to support other functions (such as registry, quality, and billing) and simplify clinical documentation is neither easy nor intuitive. If the task were easy, it would already be done. If the strategy were intuitive, the path would be clearly defined by national health experts. The diabetes community should watch ONC efforts and begin to articulate how various use cases (e.g., emergency responder, quality, and consumer preferences) can be customized to address the scenarios—and important data elements—to support diabetes care and management on individual and population levels.

Develop a “Home” for Diabetes Standards Discussion and Dissemination

Because clinical practice and scientific knowledge are continuously evolving, data standards related to diabetes will also be dynamic. Policy leaders should recognize that ongoing maintenance will be a necessity, and models for continued collaboration, inclusion, consensus,

and transparency are required to keep diabetes stakeholders engaged and committed to using these standards. Standards for diabetes data need to be current, accessible, open, easy to use, and useful. Standards will only be used if they are available and accessible and easy to implement. A single point of information would benefit the developers of diabetes registries. Easy identification and access to standards that are usable and understandable will promote the standardization of data collected by various registries.

Because registry and EHR data collection are related, the standards for both can be considered within a diabetes data standards home. The standards resources and recommendations should include definition of registry functions and EHRs and the unique standards requirements and best practices for each. Part of these standards will be best practices and also an eye for data and operational standards that will evolve with emerging models of care, regional health data collection, and the evolving national health information infrastructure. An articulated vision of primary care and registry products and the relationship to other health care processes, including drug development, patient education, continuity of care, and quality measurement, will be important to guide EHR developers and registry providers, as well as the purchasers of those products.

Develop Consensus Standards for Registries

This standards home can develop, support, and promote CDEs for diabetes-specific patient registries and enable future opportunities for sharing or comparing data across registries. Building upon established guidance, such as the AHRQ guidance for development of registries to evaluate patient outcomes,² a centralized leadership should identify key features and considerations that are critical to support diabetes registries of various types. The group can also identify consensus steps for any registry development effort, which should include the following: planning (including documenting explicit goals and success measures and termination criteria for the registry); scope (international, national, regional, or local); design specifications (including completeness of case ascertainment, type of data collected, verification of data validity, and patient follow-up); data standards for various types of registries; policies for registry governance and oversight; sampling and recruitment strategies; quality assurance methods; as well as analysis, reporting, and dissemination of findings. Particular features of new registries (e.g., interoperability), approaches to various data sources, and subsequent technical or standards requirements could be defined

centrally and support an infinite number of local applications.

Conclusion

A premise of this article is that the use of data standards is a requisite for a quality registry program and for achieving integration functions with health information systems that will support changes in provider behavior and improved patient outcomes in diabetes. The use of data standards will enable regional and provider-based diabetes registries to support continuity of care, quality of care, population monitoring, and global research. It is hoped that the overview of standards issues presented here clearly brings to light the need for standards specialists and will inspire a commitment to standards in the design and implementation of any diabetes registry or data collection program. It is also hoped that the vital need for consensus communities of diabetes stakeholders is made equally clear. New forums for collaboration between diabetes registry stakeholders and technical standards experts will drive new capabilities for registry and EHR interoperability in diabetes and, by extension, national information infrastructure that can effectively support important national activities in patient-oriented care, public health, quality monitoring, and research advances for diabetes.

Funding:

This work was supported in part by contract HHSN267200800019C to the Type 1 Diabetes TrialNet Study Group, a clinical trials network funded by the National Institutes of Health through the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Allergy and Infectious Diseases, the National Institute of Child Health and Human Development, and the General Clinical Research Centers Program with support of the Juvenile Diabetes Research Foundation International and the American Diabetes Association.

Disclaimer:

The contents of this work are solely the responsibility of the author and do not necessarily represent the official views of the National Institutes of Health, the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Allergy and Infectious Diseases, the National Institute of Child Health and Human Development, TrialNet, the Juvenile Diabetes Research Foundation International, or the American Diabetes Association. The views expressed in written materials or publications do not necessarily reflect official policies of the Department of Health and Human Services, nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. government.

Acknowledgments:

The author is grateful to Ms. Kate Paulus for her expertise regarding the regulatory and legal aspects of registries. The author also thanks Drs. Kendra Vehik and Craig Beam for their constructive reviews of this work, as well as two anonymous external reviewers for their valuable comments.

References:

1. Health Information Technology for Economic and Clinical Health (HITECH) Act, title XIII of division A and title IV of division B of the American Recovery and Reinvestment Act of 2009 (ARRA), Pub. L. No. 111-5 (Feb. 17, 2009).
2. Agency for Healthcare Research and Quality. Registries for evaluating patient outcomes: a user's guide. Rockville: Agency for Healthcare Research and Quality; 2010.
3. California HealthCare Foundation. Using computerized registries in chronic disease care. <http://www.chcf.org/publications/2004/02/using-computerized-registries-in-chronic-disease-care>. Accessed April 22, 2011.
4. Dreyer NA, Garner S. Registries for robust evidence. *JAMA*. 2009;302(7):790-1.
5. Stroebel RJ, Scheitel SM, Fitz JS, Herman RA, Naessens JM, Scott CG, Zill DA, Muller L. A randomized trial of three diabetes registry implementation strategies in a community internal medicine practice. *Jt Comm J Qual Improv*. 2002;28(8):441-50.
6. Montori VM, Dinneen SF, Gorman CA, Zimmerman BR, Rizza RA, Bjornsen SS, Green EM, Bryant SC, Smith SA; Translation Project Investigator Group. The impact of planned care and a diabetes electronic management system on community-based diabetes care: the Mayo Health System Diabetes Translation Project. *Diabetes Care*. 2002;25(11):1952-7.
7. F O Kern E, Beischel S, Stalnaker R, Aron DC, Kirsh SR, Watts SA. Building a diabetes registry from the Veterans Health Administration's computerized patient record system. *J Diabetes Sci Technol*. 2008;2(1):7-14.
8. Brooke EM. The current and future use of registers in health information systems. Geneva: World Health Organization; 1974. http://whqlibdoc.who.int/offset/WHO_OFFSET_8.pdf. Accessed April 22, 2011.
9. Richesson RL, Lee HS, Cuthbertson D, Lloyd J, Young K, Krischer JP. An automated communication system in a contact registry for persons with rare diseases: scalable tools for identifying and recruiting clinical research participants. *Contemp Clin Trials*. 2009;30(1):55-62.
10. DIAMOND Project Group. Incidence and trends of childhood type 1 diabetes worldwide 1990-1999. *Diabet Med*. 2006;23(8):857-66.
11. EURODIAB ACE Study Group. Variation and trends in incidence of childhood diabetes in Europe. *Lancet*. 2000;355(9207):873-6.
12. Mayer-Davis EJ, Bell RA, Dabelea D, D'Agostino R Jr, Imperatore G, Lawrence JM, Liu L, Marcovina S; SEARCH for Diabetes in Youth Study Group. The many faces of diabetes in American youth: type 1 and type 2 diabetes in five race and ethnic populations: the SEARCH for Diabetes in Youth Study. *Diabetes Care*. 2009;32 Suppl 2:S99-101.
13. Green SB, Byar DP. Using observational data from registries to compare treatments: the fallacy of omnimetrics. *Stat Med*. 1984;3(4):361-73.
14. Miettinen OS. The need for randomization in the study of intended effects. *Stat Med*. 1983;2(2):267-71.
15. Goldberg J, Gelfand HM, Levy PS. Registry evaluation methods: a review and case study. *Epidemiol Rev*. 1980;2:210-20.
16. Bagheri A, Sadek A, Chan T, Khunti K, de Lusignan S. Using surrogate markers in primary electronic patient record systems to confirm or refute the diagnosis of diabetes. *Inform Prim Care*. 2009;17(2):121-9.

17. Morris AD, Boyle DI, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, MacDonald TM. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. DARTS/MEMO Collaboration. *BMJ*. 1997;315(7107):524–8.
18. Tang PC, Ralston M, Arrigotti MF, Qureshi L, Graham J. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *J Am Med Inform Assoc*. 2007;14(1):10–5.
19. Harris SB, Glazier RH, Tompkins JW, Wilton AS, Chevendra V, Stewart MA, Thind A. Investigating concordance in diabetes diagnosis between primary care charts (electronic medical records) and health administrative data: a retrospective cohort study. *BMC Health Serv Res*. 2010;10:347.
20. Cronenwett JL, Likosky DS, Russell MT, Eldrup-Jorgensen J, Stanley AC, Nolan BW, VSGNNE. A regional registry for quality assurance and improvement: the Vascular Study Group of Northern New England (VSGNNE). *J Vasc Surg*. 2007;46(6):1093–2.
21. California HealthCare Foundation. Chronic disease registries: a product review. <http://www.chcf.org/publications/2004/05/chronic-disease-registries-a-product-review>. Accessed April 22, 2011.
22. Lester WT, Zai AH, Chueh HC, Grant RW. Diabetes information technology: designing informatics systems to catalyze change in clinical care. *J Diabetes Sci Technol*. 2008;2(2):275–83.
23. Crosson JC, Ohman-Strickland PA, Hahn KA, DiCicco-Bloom B, Shaw E, Orzano AJ, Crabtree BF. Electronic medical records and diabetes quality of care: results from a sample of family medicine practices. *Ann Fam Med*. 2007;5(3):209–15.
24. Richesson RL, Vehik K. Patient registries: utility, validity and inference. In: de la Paz MP, Groft SC, eds. *Rare diseases epidemiology*. Dordrecht: Springer; 2010, 87–140.
25. Richesson RL, Krischer J. Data standards in clinical research: gaps, overlaps, challenges and future directions. *J Am Med Inform Assoc*. 2007;14(6):687–96.
26. Chalmers RJ. Health care terminology for the electronic era. *Mayo Clin Proc*. 2006;81(6):729–31.
27. Dudeck J. Aspects of implementing and harmonizing healthcare communication standards. *Int J Med Inform*. 1998;48(1-3):163–71.
28. American Medical Informatics Association; American Health Information Management Association Terminology and Classification Policy Task Force. Healthcare terminologies and classifications: an action agenda for the United States. <http://www.ncvhs.hhs.gov/080221p4.pdf>.
29. Institute of Medicine of the National Academies. To err is human: building a safer health system. Washington DC: Institute of Medicine, National Academy of Sciences; 1999. <http://www.iom.edu/Reports/1999/To-Err-is-Human-Building-A-Safer-Health-System.aspx>. Accessed April 22, 2011.
30. Hammond WE, Cimino JJ. Standards in biomedical informatics. In: Shortliffe EH, Cimino JJ, eds. *Biomedical informatics: computer applications in health care and biomedicine*. New York: Springer; 2006, 265–311.
31. Rollason W, Khunti K, de Lusignan S. Variation in the recording of diabetes diagnostic data in primary care computer systems: implications for the quality of care. *Inform Prim Care*. 2009;17(2):113–9.
32. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. Federal Register 1997. http://www.whitehouse.gov/omb/fedreg_1997standards. Accessed April 22, 2011.
33. U.S. Food and Drug Administration. FDA resources for data standards. <http://www.fda.gov/forindustry/datastandards/default.htm>. Accessed December 6, 2010.
34. Healthcare Information Technology Standards Panel (HITSP); American National Standards Institute. <http://www.hitsp.org/>. Accessed November 23, 2010.
35. Clinical Data Standards Interchange Consortium. <http://www.cdisc.org/>. Accessed December 13, 2005.
36. McDonald CJ, Huff SM, Suico JG, Hill G, Leavelle D, Aller R, Forrey A, Mercer K, DeMoor G, Hook J, Williams W, Case J, Maloney P. LOINC, a universal standard for identifying laboratory observations: a 5-year update. *Clin Chem*. 2003;49(4):624–33.
37. Chute CG. Medical concept representation. In: Chen HC, Fuller SS, Friedman C, Hersh W, eds. *Medical informatics: knowledge management and data mining in biomedicine*. New York: Springer; 2005, 163–82.
38. Healthcare Information Technology Standards Panel. Clinical research interoperability specification. HITSP/IS158. http://hitsp.org/ConstructSet_Details.aspx?&PrefixAlpha=1&PrefixNumeric=158. Accessed April 22, 2011.
39. Public Health Data Standards Consortium. Standards for public health data exchange: functional requirements standard for diabetes care management and surveillance. http://www.phdsc.org/health_info/pdfs/Standards-for-Public-Health-PHDC-FINAL-Report.pdf. Accessed April 22, 2011.
40. Bhatt DL. Advancing the care of cardiac patients using registry data: going where randomized clinical trials dare not. *JAMA*. 2010;303(21):2188–9.
41. Hendel RC, Budoff MJ, Cardella JF, Chambers CE, Dent JM, Fitzgerald DM, Hodgson JM, Klodas E, Kramer CM, Stillman AE, Tilkemeier PL, Ward RP, Weigold WG, White RD, Woodard PK; American College of Cardiology (ACC); American Heart Association (AHA). ACC/AHA/ACR/ASE/ASNC/HRS/NASCI/RNA/SAIP/SCAI/SCCT/SCMR/SIR 2008 Key data elements and definitions for cardiac imaging: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Clinical Data Standards for Cardiac Imaging). *J Am Coll Cardiol*. 2009;53(1):91–124.
42. Buxton AE, Calkins H, Callans DJ, DiMarco JP, Fisher JD, Greene HL, Haines DE, Hayes DL, Heidenreich PA, Miller JM, Poppas A, Prystowsky EN, Schoenfeld MH, Zimetbaum PJ, Heidenreich PA, Goff DC, Grover FL, Malenka DJ, Peterson ED, Radford MJ, Redberg RF; American College of Cardiology; American Heart Association Task Force on Clinical Data Standards; (ACC/AHA/HRS Writing Committee to Develop Data Standards on Electrophysiology). ACC/AHA/HRS 2006 key data elements and definitions for electrophysiological studies and procedures: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (ACC/AHA/HRS Writing Committee to Develop Data Standards on Electrophysiology). *J Am Coll Cardiol*. 2006;48(11):2360–96.
43. Radford MJ, Arnold JM, Bennett SJ, Cinquegrani MP, Cleland JG, Havranek EP, Heidenreich PA, Rutherford JD, Spertus JA, Stevenson LW, Goff DC, Grover FL, Malenka DJ, Peterson ED, Redberg RF; American College of Cardiology; American Heart Association Task Force on Clinical Data Standards; American College of Chest Physicians; International Society for Heart and Lung Transplantation; Heart Failure Society of America. ACC/AHA key data elements and definitions for measuring the clinical management and outcomes of patients with chronic heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Heart Failure Clinical Data Standards): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Failure Society of America. *Circulation*. 2005;112(12):1888–916.

44. Cannon CP, Battler A, Brindis RG, Cox JL, Ellis SG, Every NR, Flaherty JT, Harrington RA, Krumholz HM, Simoons ML, Van De Werf FJ, Weintraub WS, Mitchell KR, Morrisson SL, Brindis RG, Anderson HV, Cannom DS, Chitwood WR, Cigarroa JE, Collins-Nakai RL, Ellis SG, Gibbons RJ, Grover FL, Heidenreich PA, Khandheria BK, Knoebel SB, Krumholz HL, Malenka DJ, Mark DB, McKay CR, Passamani ER, Radford MJ, Riner RN, Schwartz JB, Shaw RE, Shemin RJ, Van Fossen DB, Verrier ED, Watkins MW, Phoubandith DR, Furnelli T. American College of Cardiology key data elements and definitions for measuring the clinical management and outcomes of patients with acute coronary syndromes. A report of the American College of Cardiology Task Force on Clinical Data Standards (Acute Coronary Syndromes Writing Committee). *J Am Coll Cardiol*. 2001;38(7):2114–30.
45. Ahmadian L, Cornet R, Kalkman C, de Keizer NF; NVA Working Group “Minimal Dataset for Preoperative Assessment”. Development of a national core dataset for preoperative assessment. *Methods Inf Med*. 2009;48(2):155–61.
46. Winget MD, Baron JA, Spitz MR, Brenner DE, Warzel D, Kincaid H, Thornquist M, Feng Z. Development of common data elements: the experience of and recommendations from the Early Detection Research Network. *Int J Med Inform*. 2003;70(1):41–8.
47. Mohanty SK, Mistry AT, Amin W, Parwani AV, Pople AK, Schmandt L, Winters SB, Milliken E, Kim P, Whelan NB, Farhat G, Melamed J, Taioli E, Dhir R, Pass HI, Becich MJ. The development and deployment of common data elements for tissue banks for translational research in cancer - an emerging standard based approach for the Mesothelioma Virtual Tissue Bank. *BMC Cancer*. 2008;8:91.
48. Richesson RL, Mon D, Kallem C, Gunter P. A strategy for defining common data elements to support clinical care and secondary use in clinical research. Presented at: 2010 AMIA Clinical Research Informatics Summit. http://crisummit2010.amia.org/files/symposium2008/100_richesson.pdf. Accessed April 22, 2011.
49. Health Level 7. EHR Diabetes Data Strategy. http://wiki.hl7.org/index.php?title=EHR_Diabetes_Data_Strategy. Accessed November 23, 2010.
50. Hagopian WA, Lernmark A, Rewers MJ, Simell OG, She JX, Ziegler AG, Krischer JP, Akolkar B. TEDDY--The Environmental Determinants of Diabetes in the Young: an observational clinical trial. *Ann N Y Acad Sci*. 2006;1079:320–6.
51. Uusitalo U, Kronberg-Kippilab C, Aronsson CA, Schakeld S, Schoene S, Mattisson I, Reinivuog H, Silvish K, Sichert-Hellert W, Stevensj M, Norrisk JM, Virtanen SM. Food composition database harmonization for between-country comparisons of nutrient data in the TEDDY Study. *J Food Comp Anal*. To be published.
52. Moyers S, Richesson R, Krischer J. Trans-Atlantic data harmonization in the classification of medicines and dietary supplements: a challenge for epidemiologic study and clinical research. *Int J Med Inform*. 2008;77(1):58–67.
53. Weykamp C, John WG, Mosca A. A review of the challenge in measuring hemoglobin A1c. *J Diabetes Sci Technol*. 2009;3(3):439–45.
54. Elwyn GJ, Vaughan NJ, Stott NC. District diabetes registers: more trouble than they're worth? *Diabet Med*. 1998;15 Suppl 3:S44–8.
55. Kallem C, Richesson R, DuLong D, Sison L, Van Dyke P, Mon DT. Advancing secondary data uses through data standards: HL7 project advances the “collect once, use many times” paradigm. *J AHIMA*. 2011;82(4):38–9.
56. McNamara RL, Brass LM, Drozda JP Jr, Go AS, Halperin JL, Kerr CR, Lévy S, Malenka DJ, Mittal S, Pelosi F Jr, Rosenberg Y, Stryer D, Wyse DG, Radford MJ, Goff DC Jr, Grover FL, Heidenreich PA, Malenka DJ, Peterson ED, Redberg RF; American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Data Standards on Atrial Fibrillation); American Collage of Cardiology; American Heart Association. ACC/AHA key data elements and definitions for measuring the clinical management and outcomes of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Data Standards on Atrial Fibrillation). *Circulation*. 2004;109(25):3223–43.
57. Hendel RC, Budoff MJ, Cardella JF, Chambers CE, Dent JM, Fitzgerald DM, Hodgson JM, Klodas E, Kramer CM, Stillman AE, Tilkemeier PL, Ward RP, Weigold WG, White RD, Woodard PK. ACC/AHA/ACR/ASE/ASNC/HRS/NASCI/RSNA/SAIP/SCAI/SCCT/SCMR/SIR 2008 key data elements and definitions for cardiac imaging: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (Writing Committee to Develop Clinical Data Standards for Cardiac Imaging). *Circulation*. 2009;119(1):154–86.