Perspective

Learning health systems, embedded research, and data standards—recommendations for healthcare system leaders

Rachel L. Richesson

Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, Michigan, USA

Corresponding Author: Rachel L. Richesson, PhD, MPH, Department of Learning Health Sciences, University of Michigan Medical School, 1111 E Catherine Street, Ann Arbor, MI 48109, USA; richessr@med.umich.edu

Received 4 June 2020; Revised 24 July 2020; Accepted 8 September 2020; Editorial Decision 1 September 2020

ABSTRACT

Learning health systems that conduct embedded research require infrastructure for the seamless adoption of clinical interventions; this infrastructure should integrate with electronic health record (EHR) systems and enable the use of existing data. As purchasers of EHR systems, and as critical partners, sponsors, and consumers of embedded research, healthcare organizations should advocate for EHR system functionality and data standards that will increase the capacity for embedded research in clinical settings. As stakeholders and proponents for EHR data standards, healthcare leaders should support standards development and promote local adoption to support quality healthcare, continuous improvement, innovative data-driven interventions, and the generation of new knowledge. “Standards-enabled” health systems will be positioned to address emergent and critical research questions, including those related to coronavirus disease 2019 (COVID-19) and future public health threats. The role of a data standards officer or champion could enable health systems to realize this goal.

Key words: data standards, data governance, electronic health record systems, pragmatic clinical trials, embedded research, healthcare systems

LAY SUMMARY

The use of data standards in electronic health record (EHR) systems can enable the use of clinical data to improve patient care and to increase our understanding of which treatments work best in the real world. However, data standards can be complicated and costly to implement. Data standards officers could be employed by healthcare systems to promote the value of data standards and coordinate their use in EHRs, ultimately enabling “research-ready” EHRs and learning health systems.

ELECTRONIC HEALTH RECORD SYSTEMS ENABLE EMBEDDED RESEARCH

High healthcare costs and inadequate knowledge about the real-world effectiveness for most healthcare services are spurring the need for embedded research to learn which treatments and services are most effective. Researchers at NYU Langone Health System recently reported their instantiation of a learning health system by conducting randomized embedded trials to measure the impact of existing, often routine, activities and taking appropriate action. They have identified ineffective activities (e.g., post-discharge telephone calls, appointment reminder letters), saving millions of dollars that were redirected to activities with greater impact. In the current state, few health systems can recognize activities with greater impact than those that address the urgent and unanswered clinical questions related to coronavirus disease 2019 (COVID-19) treatment approaches, screening, diagnosis, prediction, prevention, and prioritization for personal protective equipment and ventilators. The research investigations to answer these questions must include a...
broad range of patients and providers in healthcare delivery settings and will not have the luxury of narrow eligibility criteria or extensive data collection. In other words, these investigations will require research designs that are truly embedded in healthcare systems.

Embedded research is pragmatic by design, testing healthcare interventions that can be smoothly integrated into clinical workflows (often using electronic health record [EHR] systems), and collecting outcomes through usual care with little or no additional data collection. The authors of the NYU exemplar attribute their success to the careful selection of questions that could be explored within existing workflows and with data already captured in the EHR. The experience of other organizations conducting embedded clinical trials also confirms that the availability of data from EHR systems is an essential factor.

Despite the potential of EHRs to enable embedded research, their complexity and variability can hinder the rapid configuration and conduct of studies required to address clinical and health delivery questions as they emerge. Variation in data and functionality from one EHR system to another means that organizations cannot easily replicate research or quality improvement studies without costly and time-consuming modification of code to configure interventions (eg, using alerts, pre-configured orders, or custom displays) or query data for reporting and analysis. In the NIH Health Care Systems Research Collaboratory, for example, many multi-site pragmatic trials reported that programming for data queries, data quality metrics, automated alerts and dashboards to support the trial, and analytic code had to be created anew for every health system participating in the trial. Unfortunately, this inability to easily conduct embedded research in different settings, sites, and patient populations also prevents the accumulation of collective knowledge of real-world effectiveness of healthcare interventions. COVID-19 is shining a bright spotlight on this national incapacity for embedded and multi-site research.

The COVID-19 pandemic has ignited national interest in standardized data and an appreciation for biomedical research. The audacious public endorsement of hydroxychloroquine by our President without supporting evidence is a sobering illustration of the critical need for national capacity to conduct efficient randomized clinical trials with affected populations to generate results that are swift, trustworthy, relevant, and actionable. The customization of EHR systems and supporting data standards for embedded clinical trials is the next great challenge for the field of informatics.

Reeves et al recently reported the quick optimization of the EHR to support screening, decision support, and reporting for COVID-19 in a statewide health system, and multiple national and international collaborations have formed impressively fast in response to the pandemic. However, the current lack of EHR interoperability has hindered the mobilization of large multi-site studies. Academic medical centers with data already mapped to a common data model are more able to contribute to national reporting and research efforts, but smaller healthcare organizations without this data readiness require time and technical support to connect into national efforts. Multiple registries, observational studies, and clinical trials for COVID-19 have emerged, but single site or even regional studies might not be sufficiently powered to address the questions that we need to answer—highlighting the need for interoperable systems and data standards that can support multi-site investigations on large or national scale.

STANDARDS SUPPORT USE OF EHR DATA FOR RESEARCH AND INNOVATION

If health systems used standard specifications for representing data collected in EHRs, then programs to query, report, and analyze clinical data could be reused within and across organizations. To date, only a handful of data types are coded using nationally recognized standards (eg, diagnoses using International Classification of Diseases [ICD] codes, procedures using Current Procedural Terminology [CPT] codes). The widespread and continued use of these code systems for clinical, epidemiologic, and health services research—despite their known limitations—demonstrates the value of standardized data. Standardization of other data types, such as symptoms, patient features, goals, services, and outcomes, could facilitate research within and across health systems. Data standards can empower any clinician with a good idea (eg, bilirubin chart, cardiac risk calculator, custom growth chart display, or patient-directed medication management tool) to create an “app” and integrate it into any EHR system. Such standards-enabled innovation will offer providers and patients choices for how to view and act upon health data. The US Office of the National Coordinator is promoting standards to accelerate the use of health data, but the relevance and impact of these standards will ultimately depend upon the engagement of health systems whose providers and researchers will use them.

RECOMMENDATIONS FOR HEALTH SYSTEMS TO EMBRACE STANDARDS AND ADVANCE THEIR MISSIONS

Using finite resources, health systems have an imperative to deliver the best care possible and need to continuously innovate, evaluate, and prioritize the services and apps they provide. As purchasers and custodians of EHR systems, healthcare system leaders can be engaged consumers who recognize the central role of data standards in a dynamic health and research ecosystem. To catalyze this, organizations should:

1. Identify a data standards officer or champion to investigate, promote, and coordinate the use of standards and reduce unnecessary variation in data collected across the health system. For example, lab tests could be natively coded in Logical Observation Identifiers Names and Codes (LOINC) for rapid comparison and querying of results (eg, microbiology, glucose). Pain assessments could be standardized. (A recent study of 10 organizations revealed dozens of structured data elements for pain, including many variations on the classic 0–10 pain intensity scale.) There may be multiple operational definitions (ie, queries) for specific conditions across research studies, quality measurement reports, and registries, making it difficult to retrieve and reuse queries to identify patients with diabetes or other conditions (eg, chronic obstructive pulmonary disease, substance abuse, chronic pain, terminal illness, pregnancy). Designated individuals with a broad organizational view can identify opportunities where data standards will create efficiencies and enable the reuse of analytic programs and executable applications such as clinical decision support. In parallel, they can ensure that the organization is able to participate in networked research and leverage external data sources (eg, community-level health metrics, environmental data, or the national death index).
Incorporate data standards into data governance and develop policies around new data collection based upon the principle of using and promoting data standards. Regular examination of the utility and burden of data collection can reduce clinician frustration with EHRs and documentation requirements. Governance structures can address the thoughtful selection of new data elements, and retirement of data collection that serves no clear purpose. Health-care organizations should strive for a set of core data elements that support many purposes—including embedded research, population health management, clinical decision support, and quality measurement. Local researchers should be able to propose additional data elements to support research in specific disease or clinical domains, but a standards officer could act as a data sentry to ensure that those new elements are represented in a standardized format, promoting consistent question modeling and leveraging the formal semantics of controlled terminologies. Similarly, common approaches to collecting patient-reported outcomes can support organizational goals for quality and value-driven care across different disease populations and health services.

Participate in national standards efforts. The standards officer can champion the role of data standards in disseminating research and successful innovation for adoption by other organizations. A local standards officer can coordinate organizational activities with standards development organizations (SDOs), and ensure that their EHR systems, data, and applications address local needs and comply with standards implementation approaches that will complement national research strategy and registries or research networks. Standards officers can participate in SDOs to request and organize the extension of existing standards as needed. There are many SDOs, and health systems and professional medical societies can coordinate their participation for stronger impact. For example, the standards officer can encourage or assign others to actively participate in standards working groups that align with their interests (eg, clinical decision support, laboratory standardization, patient reported outcomes (PROs), and coordinate the sharing of information and cooperation with standards delegates within or across organizations to better understand standards with which they are less familiar.

Develop infrastructure to implement emerging standards, including education, coordination, and communication with other stakeholders. Health systems leaders need to capitalize on the national paradigm shift of the US Food and Drug Administration toward active surveillance, real-world evidence, and embedded research—all of which require data standards. The emerging HL7 Fast Healthcare Interoperability Resources (FHIR) standard is a modern and extensible data exchange standard that has garnered interest from big industry like Apple, Google, Microsoft, and Amazon, and endorsement by CMS. The NIH has formally encouraged researchers to use FHIR to access EHR data and to share research data sets. The NLM is exploring and promoting FHIR to extract genomic and phenotype data from NIH databases and to integrate data from genetic test results into EHRs.

Adoption of the FHIR specification, however, will not eliminate the need to coordinate and standardize data collection. For example, if one organization collects and codes smoking status one way (eg, Current, Former, Never) and another codes it differently (eg, Yes, No), then combining or comparing data between organizations will require transformation and be subject to variation or error. Similarly, using FHIR to evaluate an opioid treatment program will likewise be problematic if the intervention codes or pain assessment values vary across units or sites. Because FHIR is a new standard, there are ambiguities on how to integrate complex coding systems such as the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). Furthermore, a proliferation of FHIR profiles with different specifications for coding systems or required elements can lead to confusion and duplication of effort if left unchecked. Organizations can use their standards officer to identify optimal and consistent approaches for representing data using FHIR, including composite measures computed from existing data and external data sources. Furthermore, organizations can help develop and inform methods to query and transmit population data (ie, “batch FHIR”) that is needed for research and population health. Consistent approaches for integrating FHIR with other standards (eg, CDS Hooks) can facilitate the development use of EHR-integrated decision support to increase the adoption of treatments shown to be effective. (CDS Hooks identifies common actions [eg, open record, create care plan, order medication] that can be used to “hook” CDS alerts or reminders into the EHR and hence the clinical workflow.)

Undoubtedly, many healthcare organizations have implemented various data standards and are addressing the above recommendations through existing data governance structures. However, these organizations will benefit from a designated standards advocate role, occupied by a person who can bring the technical knowledge about various standards structures and personal connections with SDOs to advance a vision for a standards-enabled and research-ready organization. This person can ensure that when new data collection is proposed—eg, for mobile apps, questions about emerging risk factors for new syndromes or infectious diseases such as COVID-19—appropriate existing standards are applied or a path to standardization that others can adopt is quickly identified.

The activities of a standards officer could take many forms and the role does not necessarily require a full-time effort or a single person. One or more individuals with some allocated time and dedicated perspective of standards implementation could be coordinated to maximize the efficiency and strategic use of standards to facilitate organizational goals and national collaborations. This proposed role of standards officer complements existing Chief Medical Information Officer or Chief Research Informatics Officer roles but is markedly different in its focus on the science and practice of data standards development, implementation, and evaluation. A data standards officer can bring deep technical knowledge of an array of standards for different data types and clinical disciplines, an understanding of dependencies and unintended consequences, and access to professional networks with varied experience implementing and using standards in different settings and programs. Leaders from the C-suite can certainly champion for standards, but their commitment to organizational operations and priorities leaves limited time for the detailed understanding of data standards needed for successful implementation. A standards officer’s focus on broader goals of system interoperability, sustainability, and scalability can facilitate not only the development of internal research infrastructure and learning health system capabilities, but also can foster external collaborations and interactions (eg, with other research or learning networks or other data sources), offering extended opportunities to deliver and assess high-value healthcare and ultimately achieve the Quadruple Aim.
of any size to participate in both observational and interventional research studies, ultimately enabling sufficient sample size and diversity to meaningfully impact population health.

The recommendations around standards presented here are not novel, but they have not been widely adopted due to socio-technical challenges long familiar to informaticians, including the tension between system designers and end-users around the burden of clinical documentation relative to the perceived utility of the data collected. The value propositions for data standards themselves have been a challenge for informatics, as standards-enabled data aggregation and exchange appear to impact the “public good” rather than those that purchase EHR systems or collect the data. For some data standards, such as ICD-10-CM, financial incentives for organizations are quite clear. The value propositions for more complex standards such as SNOMED CT or FHIR, or those relevant to only specific disease areas (eg, common data elements, PRO questions), are even more challenging to develop. The recommendations around data standards presented here are not tenable without the realignment of incentives that drive the behavior of health systems. However, the trajectory toward value-based care payment is established and gaining momentum, and as we progress toward that goal, health system leaders, stakeholders, and policymakers will increasingly expect and demand adequate data to measure and compare the value of different interventions and approaches to care. In the meanwhile, informatics professionals focused on standards and working with health systems can and should create the infrastructure for data and embedded clinical trials that is needed and expected, and will inevitably be demanded.

Embedded research will require near real-time outcomes and access to outcome data outside of visits. Standards support this and are a key component for “research readiness” of an organization. Healthcare systems are positioned to demand and adopt data representation and exchange standards that will facilitate healthcare innovation and subsequent evaluation using embedded research. Consensus and coordination will be required to ensure that the standards evolve to support the needs of learning health systems. Standards are not “one and done.” They are dynamic and will require user communities to define—and refine—how they should be applied in different contexts. New standards will always be needed, especially in areas of innovation (eg, new devices, patient-generated, and patient-reported data). These innovations and emergent health issues will create a continuous need for evaluation driven by data. Data standards will support these evaluations and are critical and dynamic component of learning health systems. Using one or more designated data standards officers, healthcare systems can and should actively participate in the ongoing development and promotion of data standards that will support quality healthcare delivery, continuous quality improvement, and the generation of new knowledge.

FUNDING

This work was supported in part by the National Institutes of Health (NIH) Common Fund, through a cooperative agreement (U34 AT007748) from the Office of Strategic Coordination within the Office of the NIH Director. The opinions expressed in this paper do not represent the official views of the NIH.

AUTHOR CONTRIBUTIONS

The sole author RLR developed the ideas in this manuscript, wrote and edited the text, and revised the manuscript.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. IOM. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Washington, DC: Institute of Medicine, The National Academies Press; 2013.
2. Horwitz LT, Kuznetsova M, Jones SA. Creating a learning health system through rapid-cycle, randomized testing. N Engl J Med 2019; 381 (12): 1175–9.
3. Louden K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2013; 350 (1): h2147.
4. Richesson RL, Hammond WE, Nahm M, et al. Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. J Am Med Inform Assoc 2013; 20 (e2): e226.
5. Richesson RL, Green BB, Laws R, et al. Pragmatic (trial) informatics: a perspective from the NIH Health Care Systems Research Collaboratory. J Am Med Inform Assoc 2017; 24 (5): 996–1001.
6. Reeves JJ, Hollandsworth HM, Torriani FJ, et al. Rapid response to COVID-19: health informatics support for outbreak management in an academic health system. J Am Med Inform Assoc 2020; 27 (6): 853–9.
7. HSPC. HSPC’s Gallery Secondary HSPC’s Gallery. 2019. https://www.dev.hpsconsortium.org/gallery Accessed June 3, 2020.
8. SMART HealthIT. Something New and Powerful: SMART on FHIR®. Secondary Something New and Powerful: SMART on FHIR®. 2014. http://smarthealthit.org/smart-on-fhir/ Accessed June 3, 2020.
9. Mandl KD, Kohane IS. Data Standards May Be Wonky, but They Will Transform Health Care. STAT. 2019. https://www.statnews.com/2019/03/03/data-standards-wonky-transform-health-care/ Accessed June 3, 2020.
10. ONC. U.S. Core Data for Interoperability (USCDI); from 2019 ISA Reference Edition. Secondary U.S. Core Data for Interoperability (USCDI); from 2019 ISA Reference Edition 2018. https://www.healthit.gov/isa/us-core-data-interoperability-uscdi Accessed June 3, 2020.
11. ONC. Interoperability Standards Advisory. Secondary Interoperability Standards Advisory. 2019. https://www.healthit.gov/isa/ Accessed June 3, 2020.
12. ONC. National Health IT Priorities for Research: A Policy and Development Agenda. Washington, DC: Office of the National Coordinator for Health IT; 2019.
13. Regenstrief. Logical Observation Identifiers Names and Codes (LOINC®). Secondary Logical Observation Identifiers Names and Codes (LOINC®). 2005. http://www.regenstrief.org/loinc/ Accessed June 3, 2020.
14. Westra BL, Johnson SG, Ali S, et al. Validation and refinement of a pain information model from EHR Browsedata. Appl Clin Inform 2018; 9 (1): 185–98.
15. Richesson RL, Rusinovitch SA, Wisted D, et al. A comparison of phenotype definitions for diabetes mellitus. J Am Med Inform Assoc 2013; 20 (e2): e319–26.
16. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014; 12 (6): 573–6.
17. Richesson RL, Lynch CO, Hammond WE. Developing and promoting data standards for clinical research. In: Richesson RL, Andrews JE, eds. Clinical Research Informatics. Cham: Springer International Publishing; 2019: 403–31.
18. CMS. Fact Sheet. CMS Advances Interoperability & Patient Access to Health Data through New Proposals. Secondary Fact Sheet. CMS Advances Interoperability & Patient Access to Health Data through New Proposals 2019. https://www.cms.gov/newsroom/fact-sheets/cms-advances-interoperability-patient-access-health-data-through-new-proposals Accessed June 3, 2020.
19. NIH. Fast Healthcare Interoperability Resources (FHIR®) Standard. Notice Number: NOT-OD-19-122. Bethesda, MD: National Institutes of Health, Office of the Director; 2019.
20. Brennan PF. Accelerating Innovation in Science. Secondary Accelerating Innovation in Science. 2019. https://nldirector.nlm.nih.gov/2019/07/30/accelerating-innovation-in-science/ Accessed June 3, 2020.
21. HL7. CDS Hooks. Secondary CDS Hooks. 2019. https://cds-hooks.hl7.org/ballots/2018May/hooks/ Accessed June 3, 2020.