About SDC

SDC Initiative Charter

Background

With electronic health record (EHR) adoption rising across the U.S., the volume and detail of information captured by healthcare organizations and providers will grow exponentially. Although health care providers and others use various sources and methods to capture and synthesize patient-level data, EHRs have been recognized as the data source with the highest potential to provide timely and relevant data in a form that is quickly usable for quality and safety improvement, population health, and research (sometimes labeled “secondary” use or "reuse"). EHR data obtained during episodes of care will become increasingly valuable to healthcare organizations striving to leverage electronic information to drive efficiency and quality. Of particular interest are efforts to leverage clinical data captured during episodes of care and link the clinical data to supplemental data collected for other purposes including: 1) research, 2) patient-safety event reporting, 3) public health reporting, and 4) Determination of Coverage. Once captured, aggregated and analyzed, these combined data can be used to identify trends, predict outcomes and influence patient care, drug development and therapy choices.

For example, a clinician treating a patient who is participating in a clinical trial or comparative effectiveness research (CER) would select, through the EHR, the applicable electronic case report form (eCRF) and would be able to complete the form with a combination of information extracted automatically from his or her most recent entry into the patient’s EHR and new information added by the clinician in direct response to the eCRF. Data captured would be stored and ultimately aggregated and transferred to the end users.

Challenge Statement

The utility of EHR data for supplemental purposes has been limited due to a lack of uniformity in the terminology and definitions of data elements across EHRs. This limitation is compounded by the fact that clinician workflow often records patient information in unstructured free-text data well after the episodes of care. Linking EHR data with other data in a uniform and structured way could accelerate quality and safety improvement, population health and research.

A multi-pronged approach is warranted. Various clinical and health services research groups and specialty societies are already engaged in independent initiatives to standardize data collection across projects in their domains in order to maximize the utility of the resulting datasets for subsequent research. Many important efforts that focus on this level of standardization, such as the Patient Reported Outcomes Measurement System (PROMIS), PhenX(consensus measures for Phenotypes and eXposures), and the Federal Interagency Traumatic Brain Injury Research (FITBIR) Informatics System, are funded by the National Institutes of Health (NIH) and other Federal sources. The National Library of Medicine (NLM) is working with the NIH research community and others to identify and coordinate research initiatives that use standardized patient assessment instruments and structured data definitions, also known as Common Data Elements (CDEs). Work is also beginning, under the auspices of NLM and the Department of Health and Human Services (HHS), to consider how to incorporate these CDEs more directly into the data infrastructure for patient-centered outcomes research (PCOR) using EHRs. The Agency for Healthcare Research and Quality (AHRQ) has developed a comparable library of terms and Common Formats to standardize data collected and reported for patient safety events. With such CDEs and standardized assessment instruments, data captured within an EHR could be consistently defined and collected, thereby improving its validity and usability not just in retrospective analysis but also in prospective observational or interventional research, comparative effectiveness research and patient safety monitoring.

The Office of the National Coordinator for Health IT (ONC) has demonstrated its value as a forum for addressing complex data architecture challenges, particularly as a means to publicly develop and test alternatives in advance of their inclusion in the regulations implementing the 2009 Health Information Technology and Economic and Clinical Health (HITECH) Act. The Structured Data Capture Initiative will build on both the results of and lessons learned in prior efforts to bring consensus to this next critical aspect of our collective health data infrastructure.

Scope Statement

To define the necessary requirements (including metadata) that will drive the identification and harmonization of standards to facilitate the collection of supplemental EHR-derived data.

This initiative will develop and validate a standards-based data architecture so that a structured set of data can be accessed from EHRs and be stored for merger with comparable data for other relevant purposes to include:

- The **electronic Case Report Form** (eCRF) used for clinical research including Patient Centered Outcomes Research (PCOR)
- The **Incident Report** used for patient safety reporting leveraging AHRQ ‘Common Formats’ and FDA form 3500/3500a
- The **Surveillance Case Report** Form used for public health reporting of infectious diseases
- The collection of patient information used for Determination of Coverage, as resources permit.

The infrastructure will consist of four new standards that will enable EHRs to capture and store structured data. These will consist of:

- a standard for the CDEs that will be used to fill the specified forms or templates;
- a standard for the structure or design of the form or template (container);
- a standard for how EHRs interact with the form or template; and
- a standard to enable these forms or templates to auto-populate with data extracted from the existing EHR

The standards will facilitate the collection of data in such a way that any researcher, clinical trial sponsor and/or reporting entity can access and interpret the data in electronic format. They will also support development of concise, architectural guidance using easy-to-understand documentation, user-friendly tooling and formal models to assist vendors in applying technical requirements for the customized use of specified forms or templates. For the purposes of this initiative, the data collected will not be stored within the EHR system. In sharing this data, ONC recognizes that certain forms of data may be subject to particular state or federal laws regulating use and disclosure. Standard specifications will incorporate the tools necessary for driving interoperability such as XML and the CDISC/ IHE integration profile Retrieve Form for Data Capture (RFD) ; this does not, however, imply any constraints on data formats that can be used during data capture and processing, as long as they do not prevent interoperability. The RFD integration profile is currently used within the research community to embed structured electronic forms with common data elements within the EHR to facilitate collection of research data. The SDC Initiative will align with and leverage other initiatives of ONC . It will also build upon external initiatives that are focused on improving the comparability and utility of data derived from independent collection efforts through standardizing definitions of data elements and tools, such as PROMIS, PhenX, caDSR, and other initiatives identified in the NLM-NIH Common Data Element repository, and the AHRQ 'Common Formats' and Electronic Data Methods (EDM) Forum for Comparative Effectiveness Research (CER).

Value Statement

Given the significant Federal investments made in EHR adoption in the last 4 years, structured data capture within EHRs is poised to be a critical component of a variety of health services, quality measurement and clinical and health services research. Stage 3 Meaningful Use (MU) will focus on creating a learning health system to support quality, research, and improve public and population health. This initiative will lead the national vision to design the trusted mechanisms to enable patient information to flow securely from the system it was collected—the EHR—to other systems, such as research consortia, registries, bio repositories and public health systems, with an authorized use for it. Information will be shared in compliance with policy, regulation, and Patient Consent Directives (e.g., 42 C.F.R Part 2 Confidentiality of alcohol and drug abuse patient records; and 38 USC § 7332-Confidentiality of certain medical records). The identification and harmonization of standards for structured data capture within EHRs will not only help achieve this vision, but they will also help reduce the:

- Data collection burden on health care providers by enabling secure, single-point data entry that populates to multiple systems
- Need to make site-specific modifications to EHR system capabilities in order to enable participation in important reporting and research activities
- Barriers to volunteer adverse event reporting on medical products to public health agencies leading to improvements in population health

These efforts will identify a standard for structured data, whether it is used for a clinical trial, Determination of Coverage, or to report on a patient safety event, which can be collected in a timely manner, then readily compared and aggregated improving the overall quality, value and utility of these data. Furthermore, the development of a national infrastructure will improve access to standardized electronic versions of data collection instruments relevant for use in research and patient care such as validated instruments for collecting data on pain, fatigue, physical function, depression, anxiety and social function. It will be easier to integrate these instruments into EHRs in ways that will ultimately reduce duplicate data entry. Likewise, data collected will be more comparable and therefore more useful in ascertaining what works best for different patient populations.

Target Outcomes & Expected Deliverables

The SDC Initiative will provide an infrastructure to standardize the capture and expanded use of patient-level data collected within an EHR. In the short term, specification of standards for data reuse will support and spur development and implementation of software and pilots that will inform refinement of these standards, prior to their consideration for inclusion in Meaningful Use and EHR certification requirements. In the longer term, the additional functionality will support enhancements and efficiencies in such diverse domains as patient-centered outcomes research and clinical trials, adverse event reporting and public health monitoring and surveillance, Determination of Coverage and patient care.

The value of this initiative will be measured through the attainment of the following immediate and long-term outcomes:

1. Identification of functional requirements from a **Use Case** describing key conditions and business rules to enable the capture and storage of specified forms or templates, while protecting privacy and confidentiality
2. Development of concise consensus-driven architectural **guidance** using easy-to-understand documentation, user-friendly tooling and formal models to assist researchers, patient safety personnel, software vendors and others in applying technical requirements for the customized use of specified forms or templates. Guidance will be updated and versioned appropriately to allow ubiquitous access to all parties.
3. Execution of one or more **pilots** to evaluate the use of the specified form standard in specific contexts, such as patient-centered outcomes research and patient safety event reporting. The pilots will examine the application of specified-form standard for the conduct of PCOR and patient safety event reporting, drawing upon the NLM/NIH common data elements repository, as well as AHRQ 'Common Formats' for patient safety events.
4. Proliferation and use of NIH-identified and curated **CDEs** for PCOR and AHRQ 'Common Formats' for patient safety event reporting
5. Development or identification of four national **standards** specific for: CDEs used to fill electronic forms or templates, the structure or design of the form or template, the standardized functions for how EHRs interact with those standards, and the specifications that enable these forms or templates to auto-populate with data extracted from the existing EHR.
6. Alignment and integration to other health IT infrastructure (through a Learning Health System) to support effective maintenance, distribution, and use of specified forms or templates
7. Enhancement of patient care through improvements in quality and safety interventions, population health, and research
8. Improvement in provider experience and workflow when using EHRs for patient care and other purposes

Standards

Standardization efforts established by other projects will be leveraged. These include, but are not limited to:

- NLM/NIH repository of PCOR-related projects with common data elements and standardized assessment tools, such as:
    - Common data elements developed/supported by NIH Institutes, e.g., National Cancer Institute (NCI), National Institute of Neurological Disorders and Stroke (NINDS).
    - Standards and measures developed by the NIH funded project, Patient Reported Outcomes Measurement Information System (PRO MIS)

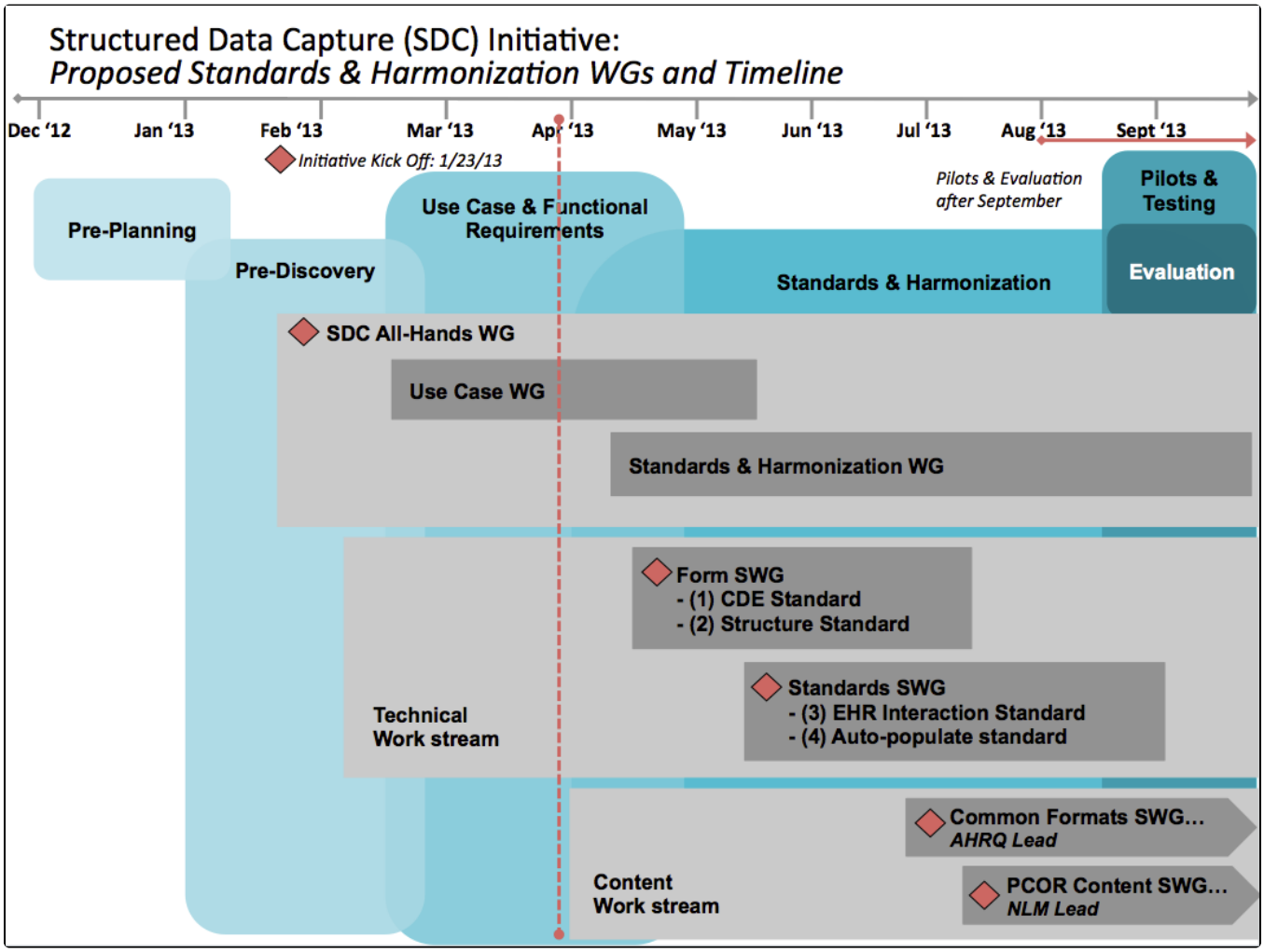
    - PhenX Toolkit measures for use in genome-wide association studies and other large-scale genomic research
    - CDISC CDASH
- Other projects using common data elements and standardized assessment tools, such as:
    - Agency for Healthcare Research and Quality (AHRQ) defined 'Common Formats' for Patient Safety Event Reporting, which reside in the United States Healthcare Information Knowledgebase (USHIK)
    - CDISC Clinical Data Acquisition Standards Harmonization (CDASH)
    - Biomedical Research Integrated Domain Group (BRIDG)
    - Food & Drug Administration (FDA) Adverse Spontaneous Triggered Events Reporting for Devices (ASTER-D)
- Integrating the Healthcare Enterprise (IHE)/ CDISC Integration ProfileRetrieve Form for Data Capture (RFD)
- IHE Clinical Research Data Capture (CRD) and Retrieve Clinical Knowledge (RCK) Profile
- Value sets derived from accepted Meaningful Use standards, as represented in the Value Set Authority Center
- Standards identified Standards Catalog
- Standards and approaches identified through the ONC Challenge Grant Program, Theme 5: Fostering Distributed Population-Level Analytics, and SHARPn Project 'Secondary Use of EHR Data'
- ONC Direct Project (as a transport standard, in relation to the send form functionality)

Stakeholders and Interested Parties

Stakeholders and interested Parties include but are not limited to the following:

- Healthcare Providers and Clinical Informaticians
- Clinical/PCOR Research Community and CER/PCOR Thought Leaders and organizations, such as: Patient-Centered Outcomes Research Institute (PCORI), EDM Forum, CDISC, Electronic Medical Records and Genomics Network (eMERGE), Distributed Ambulatory Research in Therapeutics Network (DARTNet), Electronic Patient-Reported Outcome (ePRO) Consortium, ASTER-D
- Patient Safety Organizations (PSOs)
- Privacy and Security Experts
- Patient Advocates
- Pharmaceutical Firms
- Device manufacturers
- Government Agencies:
    - Food & Drug Administration (FDA), Assistant Secretary for Planning and Evaluation (ASPE), NIH (NLM & other Institutes/Centers), AHRQ, Centers for Disease Control (CDC), Centers for Medicare & Medicaid Services (CMS), Indian Health Services, Human Resources and Services Administration (HRSA), Institute of Medicine (IOM), Veterans Administration (VA), Department of Defense (DoD), Social Security Administration (SSA), Department of Transportation (DoT), State Medicaid Programs
- Vendors:
    - EHR/EMR systems, Health Information Exchange (HIE), Data Warehouse/Data Mart, Electronic Data Capture (EDC) and Patient Safety Event Reporting System
- Standards-Related Organizations: Standards Development Organizations (SDOs), vocabulary/terminology organizations, standards setting organizations
- Healthcare payers, particularly those with robust research, quality improvement (QI), patient safety and public health activities, including professionals involved in registries and surveillance at a reginonal and national level
- Professional liability carriers
- Healthcare Professional associations

Proposed Timeline



This timeline represents notional estimates of planned activities and will be adjusted as work moves forward through the initiative phases.

Risks

- Identification of content models for standard that do not compete or overlap with existing or developing models
- Insufficient engagement and participation by vendor communities, and eventual adoption in Certified EHR products
- Insufficient commitment by various vendors and organizations to participate in SDC pilots
- Accessibility of a CDE library or repository in time for pilots and completion of pilots in time for future stages of Meaningful Use
- Competing definitions of Common Data Element and CDE libraries
- Identification of a useful but parsimonious set of data elements for forms/templates that can auto-populate with data extracted from the existing EHR in time for pilots
- Overwriting and/or duplicate data entry
- Solutions for research, patient safety, and public health reporting may vary; one standard solution may not apply
- Standards and solutions may not scale to small vendors and small practices
- Proposed project timeline does not reflect actual deadlines in relevant standards or regulatory bodies
- Not all EHR systems offer or can support pre-population functionality; if the auto-populate standard is prescriptive, the SDC timeline may be delayed as vendors implement the functionality in their systems
- Providers may not 'fill' all CDEs displayed on an electronic form
- Accuracy of EHR data and reliability of diagnosis codes for reuse in clinical research
- Patient privacy: de-identification of EHR data for clinical research

Success Metrics and/or Success Criteria

- Upfront engagement in development of use case requirements and solution from potential pilot sites
- Balloted standard(s) are included as EHR certification criteria
- Proliferation and use of NLM-identified and curated **CDEs** for PCOR and AHRQ 'Common Formats' for patient safety event reporting
- Harmonization of CDEs among federal agencies and research communities in other countries

SDC Community Members

Charter Consensus Voting

Voting on the SDC Project Charter is now closed.

1. Kahn, M., Raebel, M., Glanz, J. M., Riedlinger, K., & Steiner, J. (July 2012). A Pragmatic Framework for Single-site and Multisite Data Quality Assessment in Electronic Health Record-based Clinical Research. Medical Care, S21-S29.
2. American Medical Informatics Association. (2008) *Advancing the Framework: Use of Health Data—A Report of a Working Conference of the American Medical Informatics Association.*
3. PricewaterhouseCoopers. (2009). *Transforming healthcare through secondary use of health data.*