CDISC SHARE, a Global, Cloud-based Resource of Machine-Readable
CDISC Standards for Clinical and Translational Research

Samuel Hume, DSc 1, Anthony Chow 1, Julie Evans 2, Frederik Malfait 3, Julie Chason 1, J. Darcy Wold 1, Wayne Kubick, MBA 4, Lauren B. Becnel, PhD 1, 5
1 Clinical Data Interchange Standards Consortium (CDISC), Austin, TX, USA;
2 Samvit Solutions, LLC, Reston, VA, USA;
3 TransCelerate Biopharma, Conshohocken, PA, USA;
4 Health Level7 (HL7) International, Ann Arbor, MI, USA;
5 Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX, USA

Abstract

The Clinical Data Interchange Standards Consortium (CDISC) is a global non-profit standards development organization that creates consensus-based standards for clinical and translational research. Several of these standards are now required by regulators for electronic submissions of regulated clinical trials’ data and by government funding agencies. These standards are free and open, available for download on the CDISC Website as PDFs. While these documents are human readable, they are not amenable to ready use by electronic systems. CDISC launched the CDISC Shared Health And Research Electronic library (SHARE) to provide the standards metadata in machine-readable formats to facilitate the automated management and implementation of the standards. This paper describes how CDISC SHARE’s standards can facilitate collecting, aggregating and analyzing standardized data from early design to end analysis; and its role as a central resource providing information systems with metadata that drives process automation including study setup and data pipelining.

Introduction

Billions of dollars are spent on clinical and translational research annually. Due in part to the deficient application of data standards, data from these studies exist in non-interoperable silos. Siloed, non-standard data cannot be easily reused for meta-analyses, hypothesis generation, safety signal detection, or other secondary uses. The Clinical Data Interchange Standards Consortium (CDISC) is a non-profit standards development organization (SDO) that develops and supports global, vendor-neutral, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare. For 20 years, CDISC and its community of thousands of volunteers from over 435 member organizations and the research community at large has developed standards for collecting, tabulating, exchanging, submitting and archiving clinical and translational research, including clinical trials as well as preclinical, nutrition, comparative effectiveness, epidemiology, safety surveillance and other public health research. Each standard and its implementation guide contains normative content, which are the prescriptive elements that describe the scope and the metadata that comprise the standards, and informative content that helps implementers to understand how to apply the standard in their organization. These standards are developed and maintained using robust processes that foster stakeholder engagement, community comment, governance and maximal transparency. CDISC standards have been downloaded in approximately 100 countries and are indexed within both BioPortal 1 and FAIRsharing.org (formerly BioSharing.org) 2.

Beginning December 2016, both the US Food and Drug Administration (FDA) 3, 4 and Japan’s Pharmaceuticals and Medical Device Agency (PMDA) 5, 6 began requiring new trials’ submissions to comply with certain CDISC standards. These standards are recommended by the European Medicines Agency (EMA) Clinical Trial Advisory Group 7 and the Chinese FDA 8. CDISC also has alliances across the National Institutes of Health (NIH) including a long-standing collaboration with National Cancer Institute (NCI) Enterprise Vocabulary Services (EVS) team 9, National Institute for Allergies and Infectious Diseases (NIAID), which has used CDISC standards for pharmacovigilance studies and meta-analyses and requires them for new HIV research protocols and in Brazil, Russia, India, China and South Africa (BRICS) studies; the National Institute for Diabetes and Digestive and Kidney Disorders (NIDDK), which was part of the consortium that developed CDISC’s polycystic kidney disease TA; and others.

CDISC’s standards are open and freely available via the CDISC Website 10 as PDFs. These files are human readable, but not easily computable by software, so implementers had to extract normative metadata from each standard to incorporate it into their CDISC-compliant information systems. To obviate this issue, CDISC created the CDISC Shared Health and Research Electronic Library (SHARE), a metadata management platform that includes a metadata repository (MDR) and ecosystem of tools and services that, together, support standards development teams in creating
content and standards implementers (e.g., researchers, industry) by making its standards available in a variety of electronic formats including RDF, XML, JSON, CSV, and PDF.

Here we describe CDISC SHARE, including methods to retrieve machine-readable standards metadata and examples of real world applications.

Materials and Methods

**Standards Development Process.** CDISC standards are internally and publicly reviewed through a standardized operational process where releases are made available on the CDISC public Atlassian Confluence Wiki website for comment during a specified interval of time. Any member of the public may create an account in the CDISC Wiki and mark up candidate release documents with comments that are automatically integrated into CDISC’s instance of JIRA. Comments are then compiled and reviewed, and changes are made to the model where persuasive. The final model, its documentation and all comments (along with their disposition/resolution) are made available to the public. For controlled terminology, approved packages are managed within NCI’s EVS teams’ systems, where their curators perform quality control and annotate the terminologies with links to synonyms and related terms indexed within EVS. The CDISC SHARE team extracts these annotated, curated terminologies through an automated pipeline where the content is imported into the CDISC SHARE MDR, and the CDISC SHARE curators assess post-import logs for any failures or other issues that require resolution prior to final publication in the MDR.

**Metadata Repository (MDR).** CDISC SHARE provides a cloud-based CDISC standards specific MDR for the curation, management, and publication of the standards metadata in machine-readable formats. CDISC SHARE is currently comprised of a commercial Java Enterprise Edition application backed by an Oracle database, running in JBoss middleware. The MDR implements the Object Management Group’s Reusable Asset Specification as a base model upon which the SHARE metamodel is layered. The CDISC SHARE metamodel is based on the International Standards Organization’s (ISO) standard for metadata registries, ISO/IEC 11179, and forms the foundational basis to structure information about metadata assets, their relationships, and versioning. The CDISC SHARE RESTful Application Programming Interface (API) provides an interface for software applications to query and retrieve standards metadata from the MDR. For implementers not using CDISC SHARE API-enabled software, CDISC SHARE provides an export feature based on the API libraries that enables the CDISC SHARE biocurators to publish full standards for download by implementers. Implementers download the content from the CDISC Website in several formats including XML, RDF, and CSV; and manually import the files into local metadata management systems for use by downstream research information systems.

**CDISC SHARE Software Ecosystem.** The CDISC SHARE Ecosystem is a software tools platform that complements the CDISC SHARE MDR by extending additional services to standards developers and implementers. The CDISC SHARE Ecosystem includes CDISC’s instance of Atlassian Confluence Wiki for active standards development and publication support; JIRA for issue tracking and community public review; Bitbucket for collaborative development and version control; and numerous bespoke software applications for loading and extracting MDR content.

**Quality Control (QC) Pipeline.** After internal and public reviews are completed, final standards versions enter the CDISC SHARE QC pipeline. Normative content is extracted from the standards and assessed by expert human curators. A variety of QC scripts are then run against the metadata to ensure that datatypes, character limitations, and other quality items are appropriate. Where issues are uncovered they are corrected with support from standards development teams, as required. QC’ed metadata are then imported into the CDISC SHARE metadata repository, where a second set of scripts are run against the data. These scripts enforce CDISC SHARE model rules of referential integrity identifying cross-standard issues as well as violations of CDISC SHARE versioning rules. Once loaded, content is exported and compared to the original import content to ensure correctness.

**Communications.** As new standards are imported into the MDR and curated, CDISC Communications notifies the public by updating the CDISC SHARE Exports webpage, releasing notifications in the CDISC Newsletter, sending email blasts, posting updates to CDISC’s LinkedIn groups, and tweeting from the CDISC SHARE Twitter account. CDISC SHARE Workshops and CDISC SHARE Showcases that highlight innovative software solutions that implement CDISC SHARE content are communicated via these same channels. The public can register for the Newsletter, public webinars, and International CDISC Conferences on the CDISC Website.

**Results**

**CDISC Standards Organization.** CDISC standards (Figure 1) are categorized as Foundational Standards, Semantics, Therapeutic Area (TA) Standards, and Transport Standards. Foundational Standards include Standard for Exchange...
of Nonclinical Data (SEND)\textsuperscript{14} for the collection and tabulation of animal model and other pre-clinical data, Protocol Representation Model (PRM)\textsuperscript{15}, Analysis Data Model (ADaM)\textsuperscript{16} for defining analysis datasets, Clinical Data Acquisition Standards Harmonization (CDASH)\textsuperscript{17} that provide a minimal set of data elements common to essentially all studies, the Study Data Tabulation Model (SDTM)\textsuperscript{18} for data tabulation, and others.

![End to End Standards & Process Automation](image)

**Figure 1.** CDISC End-to-end Standards including Foundational Standards, Semantics, and TA Standards

Except for Protocol, the foundational standards and their implementation guides are organized into classes of related categories of information. Within SDTM for example, each class is subdivided into domains, and each domain is comprised of variables. Semantics include CDISC Controlled Terminologies, developed in collaboration with NCI EVS, and the Biomedical Research Integrated Domains Group (BRIDG) domain information model. TA User Guides provide indication-specific “slices” of foundational standards and their terminologies, with examples of how to use these standards for a given TA. Transport Standards support submissions, data interchange, study archival, and automated study setup\textsuperscript{19}.

CDISC implementation and user guides contain normative and informative content. Normative content includes the classes, domains, variables, definitions, and other core metadata of a standard. Informative content includes examples, assumptions and other information that help adopters to determine how to best implement the standard within their organization. While tools within the CDISC SHARE Ecosystem such as the Confluence Wiki and JIRA are utilized to create and maintain both normative and informative content, the CDISC SHARE MDR contains the normative standard metadata to avoid confusion of an actual standard versus examples of implementation.

While each CDISC standard represents a different aspect of the clinical research data lifecycle and is developed as an independent standard, the CDISC standards are also intended to work together across the lifecycle. When the standards are organized to represent the flow of information across the clinical research data lifecycle this is called the end-to-end standards representation.

**CDISC SHARE Model Version 1.0.** Normative CDISC standards metadata in the MDR conform to a version 1.0 CDISC SHARE model. SHARE content is organized into layers of increasingly concrete models, starting with the Open Management Group (OMG) Reusable Asset Specification (RAS) as the foundational model, next implementing International Standards Organization (ISO)/International Electrotechnical Commission (IEC) 11179 as the metamodel, and then adding a CDISC standards model as top-layer. ISO/IEC 11179 forms the foundational basis of CDISC SHARE’s metamodel to structure information about metadata assets, versioning, ownership and other core metadata information. RAS and ISO/IEC 11179 are standards commonly implemented in metadata management systems, but the CDISC model is novel as it is the first time CDISC has structured its contents standards into a one common model.

The CDISC SHARE metamodel maintains the relationships needed to represent the end-to-end CDISC model that captures how one standard flows into the next in the clinical research data lifecycle. These are relationships that exist between each of the individual foundational standards. For example, a CDASH variable may include a maps-to relationship to an SDTM variable. This is a two-way relationship that also provides a maps-from relationship linking the SDTM variable back to the CDASH variable. The CDISC SHARE end-to-end standards model represents the first time the standards have been explicitly represented as one full-lifecycle standard as opposed to distinct standards intended to work together. To represent the model adequately, relationships implicit in the CDISC foundational standards have been made explicit in the model. The number of relationships currently maintained in CDISC SHARE exceeds the number of metadata elements by a ratio of approximately 5:1. The end-to-end standards representation provides support for standards metadata-driven automated data pipelining.

The CDISC model is organized within SHARE’s metamodel as Metadata Elements (MDEs), which are discrete pieces of metadata such as an individual variable, and Metadata Element Sets (MDESs), which are categories of MDEs such
as a domain or groupings of MDESs such as a class that categorizes multiple domains. The metamodel also represents semantics which are organized into value domains that are in turn comprised of one or more domain values. These metamodel elements are used to represent the code lists and individual terms found in the CDISC Controlled Terminology.

Within each foundational standard the metadata are represented in the SHARE metamodel as a hierarchy of related components rooted in the parent top-level MDES (Figure 2). This MDES contains both class and domain MDESs. Class MDESs can include both domain MDESs and their own variable MDEs, though most variable MDEs are children of the domain MDESs. Variables often have an associated CDISC Controlled Terminology code list captured by the represented-by relationship.

![Figure 2. CDISC SHARE metamodel hierarchy](image)

**CDISC SHARE Application Programming Interface (API).** The CDISC SHARE metadata are exposed for retrieval through a RESTful API. The API’s schema and other documentation are available on the CDISC Website, and this schema’s endpoints represent CDISC SHARE metamodel MDEs and MDESs (Figure 2). Each CDISC SHARE metadata object has a unique object identifier (OID). API end users can request an entire standard, including all of its domains, variables and URLs to retrieve the latest terms for each code list. Alternatively, the API service recursively traverses this standards hierarchy to fulfill requests for content at various levels of granularity. To request the metadata for a specific domain, for example, an end user “drills down” through the standards hierarchy to that domain, first identifying the OID for a specific standard, then the OID for the domain’s parent class, and finally the OID for the domain itself. The API supports discovery by providing the resources to retrieve either the next level of detail within the standards hierarchy or the full top-level standard.

Several optional query parameters exist for the API’s endpoints. For full, top-level standards such as the SDTM Model and Implementation Guide, end users can specify the returned media type as either XML or RDF-XML, Turtle or N-Triples. For components of these standards, either XML or JSON may be requested. API users can specify whether they would like to retrieve standards with a lifecycle status of Approved Final or Approved Provisional. Finally, end users can request that returned metadata remain uncompressed or be compressed using gzip or zip. Details on attaining access to the CDISC SHARE API are available on the CDISC Website.

The availability of a standards MDR containing metadata representing the full clinical research data lifecycle with an API available to support process automation is uncommon among biomedical research and healthcare data standards. **Table 1** lists well-known data standards applied to healthcare and methods of accessing the standards metadata, including whether the standards are consumable via an API.
Table 1. Comparison of Methods for Accessing Standards Metadata (see Table 2 for acronym definitions)

| Standard            | Methods of Accessing Computable Standards Metadata                                               | Online Repository? | API? |
|---------------------|-------------------------------------------------------------------------------------------------|--------------------|------|
| **Content Standards**                                                                                           |                     |        |      |
| CDISC               | • RESTful API                                                                                   | Metadata Repository | Yes  |
|                     | • Content standard downloads in XML, JSON, and RDF                                               |                     |      |
|                     | • XML schemas and conformance rules                                                             |                     |      |
| DDI                 | • XML schema                                                                                     | No                 | No   |
|                     | • XML and RDF vocabularies                                                                       |                     |      |
| **Data Exchange Standards**                                                                                       |                     |        |      |
| HL7 FHIR®           | • XML, JSON, and RDF schemas                                                                    | N/A†               | N/A† |
| HL7 CDA® R2         | • XML schemas, schematron, and XSLT                                                               | No                 | No   |
| **Terminology / Semantics Standards**                                                                            |                     |        |      |
| LOINC®              | • RELMA® application to search the LOINC database and map to LOINC codes                         | Yes                | Yes* |
|                     | • The LOINC table in Microsoft Access or CSV                                                     |                     |      |
|                     | • FHIR RESTful API pilot                                                                         |                     |      |
| SNOMED CT®          | • Open source SNOMED CT online Browser                                                           | Terminology Browser| Yes**|
|                     | • SNOMED CT machine readable concept model                                                        |                     |      |
|                     | • Mapping tool                                                                                   |                     |      |
|                     | • RESTful API - not for healthcare system use                                                     |                     |      |
| WHO ICD             | • Download ICD-10 in Classification Markup Language                                             | Terminology Browser| No   |
|                     | ICD-10 online browser                                                                            |                     |      |
| AMA CPT®            | • Microsoft Access MDB and CSV download                                                           | Terminology Browser| No   |
|                     | • CPT® Developer's Tool Kit for download only                                                    |                     |      |
|                     | • CPT® Assistant Online                                                                         |                     |      |
|                     | • SNOMED CT® to CPT® Rules-Based Cross Maps                                                       |                     |      |

† Schemas and API are fully documented and multiple reference implementations exist online.
* LOINC is piloting a FHIR-based RESTful API
** SNOMED CT provides a RESTful API for lookups, but it is not available for use in production healthcare systems

**CDISC SHARE Exports.** CDISC SHARE Exports make CDISC standards accessible via download on the CDISC Website in a variety of electronic formats including CSV, XML such as ODM-XML and Define-XML, PDF and RDF. CDISC SHARE Exports help users find, understand and use rich metadata and controlled terminologies relevant to clinical studies more efficiently and consistently, and improve integration and traceability of clinical data from protocol through analysis. CDISC SHARE Export files are available to CDISC member organizations and to individual academic researchers upon request. Details on using CDISC SHARE Exports are available on the CDISC Website.

Three categories of CDISC SHARE Export files exist. Normative content in one or more formats are made available as metadata. PDF guides, though not exclusively normative, are made available as documents provided to complement the machine-readable normative metadata and to make the standards artifacts available for download from one location on the CDISC Website. Diff files, or difference files, display differences between two versioned standards by listing the additions, deletions and updates from one version of a standard to the next. Diff files are intended to support impact analysis for adopters to assess what changes are present and how that would impact their systems. Diff files can be used to drive incremental loading of new standards versions into a standards MDR.

**Discussion**

Despite substantial annual investment in research, the time to bring new medical therapies to the bedside remains too long. New approaches need to be taken to better leverage existing sources of health and research data. Data from tens of thousands of interventional and non-interventional studies performed each year cannot be aggregated and analyzed without considerable cost, largely because the data are not collected using a standard format and semantics.
Without data standards, commercial software tools that analyze the data cannot feasibly be produced requiring the development of bespoke software applications that contribute to higher costs, longer time-lines, and increased data quality concerns. Siloed, non-standard data may lead to redundant research, where previously-published discoveries not highlighted by Medline or search engines are likely to remain “hidden” in non-interoperable datasets that require informatics expertise to fully interpret. This redundancy and inefficiency ultimately prolongs the process of translation from bench to bedside to provide new treatments that may improve human health. Many scientists have also cited limited reproducibility as a major challenge to research utility that is exacerbated by limited access to study data that can be effectively used by others to verify study results. CDISC standards and other community-reviewed standards, ontologies, and code-sets provide tools that can help implementers solve these issues. Application of any standard requires an up-front investment on the part of an implementer, but once adopted these standards save resource time and money.

Due in part to regulatory requirements, the biopharmaceutical industry including major pharmaceutical companies, medical device companies, clinical research organizations, IT vendors, regulators, and others have implemented CDISC standards within their organization. Even before the December 2016 FDA mandate, over 70% of submissions to FDA were SDTM-compliant. CDISC standards implementation is not limited to regulated research sectors, which comprise only a segment of the total volume of clinical and translational research. Foundational and transport standards have been implemented within essentially all popular clinical research data systems, including the implementation of CDISC’s Operational Data Model (ODM-XML) in commercial applications and open source systems such as the OpenClinica and REDCap electronic data capture platforms; within an importer to i2b2; and several others. CDISC SHARE will support more efficient incorporation of CDISC standards into tools such as these, helping to drive automation and innovation. CDISC’s SDTM and CDISC SHARE are being evaluated for their potential role in the to-be-developed NCI Clinical and Imaging Data Commons.

**Examples of CDISC SHARE’s Use.** In September 2016, CDISC collected survey data from over 200 CDISC SHARE users. This survey showed that over 50% of respondents used CDISC SHARE (1) to populate CDISC standards libraries in their own organizational MDRs, (2) to assess the normative content within standards, (3) to support impact analyses when considering adoption of a new version of a standard, and (4) to directly utilize the standards in software and reporting implementations.

**CDISC SHARE Pilots.** In 2015 and 2016, six organizations volunteered to participate in a CDISC SHARE API Pilot team that developed the first version of the API. The organizations included Business and Decision Life Sciences, EDETEK, Entimo, Accenture, Fujitsu, and the European Translational Information & Knowledge Management Services (eTRIKS), a European Union Innovative Medicines Initiative collaborative. All the pilot participants tested the CDISC SHARE API and many have since implemented interfaces into their commercial products. Many of the vendors participated in CDISC SHARE Showcases demonstrating software that loaded CDISC SHARE content via the API or using content from CDISC SHARE Exports. XML4Pharma, a CDISC SHARE Showcase participant, noted that updating their SDTM-ETL (Extract, Transform and Load) software with a new version of SDTM was reduced from 2 weeks to 2 hours using CDISC SHARE content.

**eTRIKS.** The Innovative Medicines Initiative (IMI)-funded European Translational Information and Knowledge Management Services (eTRIKS) project delivered an open, sustainable, standards-based translational research informatics and knowledge management platform and suite of tools and services for IMI-funded research and the greater community. eTRIKS utilized CDISC standards as part of its Standards Starter Pack and CDISC SHARE content within the metamodel of its Harmonization Service (eHS) Platform. This platform was created to aggregate and integrate clinical and biomedical data collected over the course of multiple, disparate protocols and projects. SDTM metadata were imported into the eTRIKS eHS to support data aggregation from multiple different studies.

**Trace-XML.** In order to support high quality data review and reproducibility in clinical research, it is critical to have access to an unbroken chain of data, ideally beginning at the electronic health record (EHR), that includes audit information. The FDA has identified a lack of traceability as one of the top 7 data standards issues, yet essentially no solutions historically existed. Trace-XML has been developed to visualize traceability across the clinical research data lifecycle for an entire study. Using SHARE metadata, Trace-XML enables standardized clinical study metadata to be represented as an interactive graph displaying the full, interconnected history of each data element. Trace-XML makes full study traceability computable, and uses the CDISC SHARE API to add the additional metadata needed to instantiate the relationships between standards, such as linking a CDASH variable to the associated SDTM variable. Using the CDISC SHARE API enables Trace-XML to dynamically find and add the relationships between new CDISC content standards versions without requiring the software to maintain a local data store of that standards metadata.
**CDISC SHARE Ongoing Development.** The SHARE 1.0 model (Figure 2) provides a simplified view of CDISC standards and their components that is heavily abstracted. CDISC standards are interconnected to each other and to conformance rules, biomedical concepts, controlled terminology subsets, study activities, and other related metadata. To improve support for this additional metadata and to better document the relationships among them, the SHARE team has embarked upon a version 2.0 of the model that is anticipated for release in mid-2018. Examples of new features that will be available in CDISC SHARE 2.0 include (1) CDISC biomedical concepts (defined below), (2) machine-readable conformance rules, and (3) version 2.0 of the CDISC SHARE API. Requests from the community, including expanding SHARE to manage metadata from implementers and using the repository as a neutral mechanism by which to share protocol information, will be explored in 2018.

**CDISC Biomedical Concepts.** Normative CDISC standards content is generally represented as two-dimensional tables that render well in PDF and match reporting and aggregation needs, but do not always represent the way that organizations collect and store data. The MDR makes it possible for CDISC to expand those two-dimensional connections between variable and domain, for example, and to layer on multi-dimensional associations that reflect the way researchers and clinicians think about biomedical elements. These representations, created to model the rich complexity of translational research, are called biomedical concepts. A small number of more complex relationships are already available in CDISC SHARE for variables that bridge multiple domains, and a growing number of biomedical concepts exist in concept maps created using CMAP to represent relationships visually and template-based Excel files to represent relationships in a tabular format. Each biomedical concept includes a set of related concepts, defined relationships among concepts, subsets of controlled vocabularies, rules and logic. An example concept map from the CDISC Breast Cancer TA User Guide (Figure 3) highlights the unique observations and results that form the basis of associated biomedical concepts. Biomedical concept inputs may also be drawn from concept models developed for routine healthcare, such as the Clinical Information Modeling Initiative (CIMI).

**Figure 3. CDISC Breast Cancer TA Concept Map for Biomedical Concepts**

The CDISC SHARE 2.0 model has been extended to manage these biomedical concepts to provide sophisticated, machine-readable standards that will provide explicit value sets to populate drop-down lists in electronic case report forms, specify datatypes for a given concept, define which data must be collected together (e.g., a measurement and its unit of measure), and provide logic and validation rules.
**Machine-readable Conformance Rules.** CDISC Foundational and Transport Standards development teams have created conformance rules for their standards. The validation rules perform an important role in data quality within the CDISC standards as these rules often represent constraints specified or implied by the individual standards. The CDISC SHARE team will incorporate into SHARE 2.0 conformance rules required for different organizations to implement the standards consistently. These rules have become a necessary component for data exchange including regulatory submissions to the FDA and PMDA.

**Modern Transport Methods and Formats.** Regulators currently require the use of SAS version 5 transport files for submission, though this will eventually be deprecated in favor of more modern format. SHARE makes transport methodology or file format more immaterial, as CDISC SHARE provides the standards metadata in a range of modern electronic data formats. Via the CDISC SHARE API, software can be created by the community to consume the MDR’s content dynamically and present that content in whatever structure they choose. SHARE and ODMv2 provide API support that will enable more efficient and innovative ways to exchange clinical research data and metadata, including alignment with other advances in healthcare data exchange such as HL7’s FHIR.

**Conclusion**

The CDISC SHARE MDR provides normative CDISC standards metadata to increase the pace and quality of standards development, as well as to increase the efficiency and effectiveness of organizations implementing the standards. CDISC SHARE provides a novel model-based representation of the CDISC standards that has dramatically increased the number of relationships explicitly rendered within the normative standards and includes an end-to-end clinical research data lifecycle representation of the standards. The standards metadata in the CDISC SHARE model combined with the RESTful API to retrieve content enable the development of software tools that automate an increasing portion of the clinical research data lifecycle. New versions of the CDISC SHARE model and API will expand both the standards content as well as the services provided by the API to expand the opportunities available for software developers to create innovative metadata-driven applications.

**Glossary**

**Table 2. Glossary of acronyms**

| Acronym | Definition | Acronym | Definition |
|---------|------------|---------|------------|
| ADaM | Analysis Dataset Model | JSON | Java Script Object Notation |
| AMA | American Medical Association | FDA | US Food and Drug Administration |
| API | Application Programming Interface | FHIR | Fast Healthcare Interoperability Resources |
| BRIDG | Biomedical Research Integrated Domain Group | MDB | Microsoft Database |
| CDA | Clinical Document Architecture | MDR | Metadata Repository |
| CDASH | Clinical Data Acquisition Standards Harmonization | NCI | National Cancer Institute (US) |
| CDISC | Clinical Data Interchange Standards Consortium (SDO) | NIH | National Institutes of Health (US) |
| CIMI | Clinical Information Modeling Initiative | ODM | Operational Data Model |
| CPT | Current Procedural Terminology | PDF | Portable Document Format |
| CSV | Comma Separated Values | PMDA | Japan Pharmaceuticals and Medical Devices Agency |
| DDI | Data Document Initiative | PRM | CDISC Protocol Representation Model |
| Define-XML | CDISC standard for dataset metadata in XML format | RDF | Resource Description Format |
| EHR | Electronic Health Record | RELMA | Regenstrief LOINC Mapping Assistant |
| EMA | European Medicines Agency | HL7 | Health Level 7 (SDO) |
| EVS | Enterprise Vocabulary Services (NCI) | ICD | International Classification of Diseases |
| ISO | International Standards Organization | IMI | Innovative Medicines Initiative |
Table 2 – Continued. Glossary of acronyms

| Acronym | Definition | Acronym | Definition |
|---------|------------|---------|------------|
| LOINC  | Logical Observation Identifiers Names and Codes | SNOMED CT | Systematized Nomenclature of Medicine - Clinical Terms |
| REST   | REpresentational State Transfer | TA | Therapeutic Area |
| SDO    | Standards Development Organization | WHO | World Health Organization |
| SDTM   | Study Data Tabulation Model | W3C | World Wide Web Consortium |
| SEND   | Standard for Exchange of Nonclinical Data | XML | eXtensible Markup Language (W3C) |
| SHARE  | Shared Health and Research Electronic library | XSLT | eXtensible Stylesheet Language Transformations |

Funding

CDISC SHARE is partially funded through FDA U01 FC005854, U24 FD005243, and U24 FD005643; NIAID U01 FD005854 and from CDISC. Initial CDISC SHARE funding and ongoing support was provided by TransCelerate BioPharma. LB also received funding from the NCI Cancer Center Support Grant P30CA125123.

References

1. Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, et al. BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic acids research. 2011;39(suppl_2):W541-W5.
2. McQuilton P, Gonzalez-Beltran A, Rocca-Serra P, Thurston M, Lister A, Maguire E, et al. BioSharing: curated and crowd-sourced metadata standards, databases and data policies in the life sciences. Database. 2016;2016.
3. FDA. Providing Regulatory Submissions in Electronic Format - Standardized Study Data: Guidance for Industry. Secondary Providing Regulatory Submissions in Electronic Format - Standardized Study Data: Guidance for Industry. In: FDA, editor. 2014.
4. FDA. Study Data Technical Conformance Guide V3.3. In: FDA U, editor. fda.gov 2017.
5. PMDA. Notification of Practical Operations of Electronic Study Data Submissions. Secondary Notification of Practical Operations of Electronic Study Data Submissions. Japan: PMDA; 2015.
6. PMDA. Technical Conformance Guide on Electronic Study Data Submissions. Secondary Technical Conformance Guide on Electronic Study Data Submissions Japan: PMDA; 2015.
7. EMA-CTAG. Final advice to the European Medicines Agency from the clinical trial advisory group on clinical trial data formats. In: Group CTA, editor.: EMA; 2013.
8. Wang L. Global Data Standards Become a Hot Topic in China. CDISC [Internet]. 2009. Available from: https://www.cdisc.org/system/files/all/article/application/pdf/global_data_standards_in_chinafinal.pdf.
9. CDISC Terminology - National Cancer Institute. NCI Enterprise Vocabulary Services. 2017. Available from: https://www.cancer.gov/research/resources/terminology/cdissc.
10. CDISC. Clinical Data Interchange Standards Consortium. 2017. Available from: http://www.cdisc.org/.
11. CDISC Operating Procedure COP-001 Standards Development. Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/sites/default/files/resource/CDISC-COP-001-Standards-Development-2017.pdf.
12. Reusable Asset Specification. Object Management Group. 2005. Available from: http://www.omg.org/spec/RAS/.
13. ISO/IEC. ISO/IEC 11179 Part 3: Registry metamodel and basic attributes. International Organization for Standardization / International Electrotechnical Commission; 2013.
14. Standard for Exchange of Nonclinical Data (SEND). Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/standard/foundational/send.
15. CDISC. CDISC Protocol Representation Model Version 1.0. cdisc.org: CDISC; 2010.
16. Analysis Data Model (ADaM). Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/standards/foundational/adam.
17. Clinical Data Acquisition Standards Harmonization (CDASH). Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/standard/foundational/cdash.
18. Study Data Tabulation Model (SDTM). Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/standards/foundational/sdtm.
19. Hume S, Aerts J, Sarnikar S, Huser V. Current applications and future directions for the CDISC Operational Data Model standard: A methodological review. Journal of Biomedical Informatics. 2016.
20. CDISC SHARE API User Documentation. Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/share/api-user-documentation.
21. Operational Data Model (ODM-XML). Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/standards/transport/odm.
22. Define-XML. Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/standards/transport/define-xml.
23. CDISC SHARE Exports User Documentation. Clinical Data Interchange Standards Consortium. 2017. Available from: https://www.cdisc.org/share/exports-user-documentation.
24. Kondro W. Drug approvals taking too long?: Can Med Assoc; 2002.
25. Cosgrove L, Vannoy S, Mintzes B, Shaughnessy AF. Under the influence: The interplay among industry, publishing, and drug regulation. Accountability in research. 2016;23(5):257-79.
26. Bierer BE, Li R, Barnes M, Sim I. A global, neutral platform for sharing trial data. New England Journal of Medicine. 2016;374(25):2411-3.
27. Becnel L, Darlington Y, McOwiti A, Naumov A, Kankanamge W, Dehart M, et al. Transcriptomine: hierarchical pathway and biosource mapping support visualization of nuclear receptor signaling transcriptional networks. Science Signaling ePub. 2016.
28. Baker M. 1,500 scientists lift the lid on reproducibility. Nature News. 2016;533(7604):452.
29. Loannidis JP. What most clinical research is not useful. PLoS medicine. 2016;13(6):e1002049.
30. Getz K. Study on the Adoption and Attitudes of Electronic Clinical Research Technology Solutions and Standards: Tufts Center for the Study of Drug Development; 2007.
31. rwslib: A Medidata Rave Services Library. Medidata Solutions, Inc. 2017. Available from: http://rwslib.readthedocs.io/en/latest/.
32. InForm Clinical Data API Guide. Oracle. 2013. Available from: http://docs.oracle.com/cd/E67648_01/doc.601/E38596.pdf.
33. Chappell E. What can be achieved with CDISC ODM? Standardized libraries of content are just the beginning,. 2014. Available from: http://www.formedix.com/2014/03/05/what-can-be-achieved-with-cdisc-odm-standardized-libraries-of-content-are-just-the-beginning/.
34. Industry Standards: A Commitment to Enhance Value: CDISC Standards. Omnicomm. 2016. Available from: http://www.omnicomm.com/standards.
35. Camacho RJ, Stäubert S, Löhle M. Automated Import of Clinical Data from HL7 Messages into OpenClinica and tranSMART Using Mirth Connect. Studies in health technology and informatics. 2016;228:317-21.
36. Schrimpf D, Haag M, Pilz L. Possible combinations of electronic data capture and randomization systems. Methods of information in medicine. 2014;53(3):202-7.
37. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics. 2009;42(2):377-81.
38. Klipin M, Mare I, Hazelhurst S, Kramer B. The process of installing REDCap, a web based database supporting biomedical research: the first year. Applied clinical informatics. 2014;5(4):916.
39. Lyon JA, Garcia-Milian R, Norton HF, Tennant MR. The use of Research Electronic Data Capture (REDCap) software to create a database of librarian-mediated literature searches. Medical reference services quarterly. 2014;33(3):241-52.
40. Obeid JS, McGraw CA, Minor BL, Conde JG, Pawluk R, Lin M, et al. Procurement of shared data instruments for research electronic data capture (REDCap). Journal of biomedical informatics. 2013;46(2):259-65.
41. Mendis M. CDISC ODM to i2b2 Importer Project2012. Available from: http://community.i2b2.org/wiki/display/ODM2i2b2/Home.
42. eTRIKS. eTRIKS. 2017. Available from: https://www.etriks.org/.
43. Curcin V, Miles S, Danger R, Chen Y, Bache R, Taweeel A. Implementing interoperable provenance in biomedical research. Future Generation Computer Systems. 2014;34:1-16.
44. FDA. CDER common data standards issues document version 1.1. FDA; 2011.
45. Chhatre D, Malla A. CDER/CBER’s top 7 CDISC standards issues. FDA 2012.
46. Hume S, Sarnikar S, Becnel L, Bennett D. Visualizing and Validating Metadata Traceability within the CDISC Standards. AMIA Summits on Translational Science Proceedings. 2017;2017:158.
47. Cañas AJ, Hill G, Carfi R, Suri N, Lott J, Gómez J, et al., editors. CmapTools: A knowledge modeling and sharing environment. Proceedings of the First International Conference on Concept Mapping; 2004 September 14-17, 2004; Pamplona, Spain: Editorial Universidad Pública de Navarra.
48. Clinical Information Modeling Initiative. 2017 [cited 2017]. Available from: https://www.opencimi.org/.