BRIDGing CDASH to SAS:
How Harmonizing Clinical Trial and Healthcare Standards May Impact SAS Users

Clinton W. Brownley, Cupertino, CA

ABSTRACT

The Clinical Data Interchange Standards Consortium (CDISC), Health Level Seven (HL7), and other stakeholders have come together in the Biomedical Research Integrated Domain Group (BRIDG) project to harmonize existing clinical trial and healthcare standards. This paper will focus on ways in which the ongoing efforts to harmonize clinical trial and healthcare standards change the connections and requirements among the standards and how these changes may impact SAS users in clinical research.

INTRODUCTION

The efficacy of the healthcare and clinical trial industries depends, in part, on the timely exchange of accurate information. Healthcare providers need access to patients' historical medical and prescription records in order to more accurately diagnose conditions and prescribe medications. Clinical trial researchers need access to similar information in order to disqualify potential trial participants from inappropriate studies and monitor for possible adverse effects. Despite the similarity of data collected by both industries and the recognized importance of timely, accurate information exchange, communication between the industries has remained hindered by separate electronic data management systems and different definitions and protocols.

Recently, there has been growing momentum to remove these two hindrances. The idea of connecting separate electronic data management systems so that data can quickly and easily be passed back and forth (i.e. making the systems interoperable) has been popularized and championed by the Obama administration, which has made electronic health records a priority and included funds in the stimulus plan to encourage their adoption. Given the political support for electronic records and interoperability, at least within the healthcare industry, several companies, including Google and IBM, are working to develop online personal health records and connect healthcare data management systems across the country[1].

Designing electronic records for the healthcare and clinical trial industries is only one important step on the path to efficient interoperability and efficacy. A complementary step is developing universal definitions and protocols so that the two industries can ensure the right data are passed back and forth across their systems. "Without the use of common vocabularies, it is impossible not only for a given hospital's computer system to understand a patient record from another hospital, but also for researchers to compare data across organizations or to collect sufficient data to make informed decisions"[2]. This step requires Standards Developing Organizations (SDOs) from each industry to collaborate to harmonize their pre-existing definitions and protocols. An important SDO in the healthcare industry, Health Level Seven (HL7) has developed standards for interoperability within the domain of clinical and administrative data. Similarly, within the clinical trial industry, the Clinical Data Interchange Standards Consortium (CDISC) has developed standards to "support the acquisition, exchange, submission and archive of clinical research data and metadata". Today, HL7, CDISC and other key stakeholders are working together to harmonize their standards to enable efficient interoperability between the healthcare and clinical trial industry's data management systems[3].

The efforts to harmonize these pre-existing standards have led to interesting insights, as well as pivotal changes to data definitions and models that will provide significant benefits to data users. Data users will benefit from the discovery that there are several similar data elements on case report forms (CRFs), which could eliminate the need for double data entry. They will also benefit from the reduced study start-up time and resources and improved protocol quality and integrity. Moreover, by being able to develop software products that take advantage of the harmony between the healthcare and clinical trial data management systems, data users may be able to expedite the identification or confirmation of drug safety issues[4].

CURRENT HARMONIZATION EFFORTS

As one of the most active SDOs involved in the harmonization of clinical trial and healthcare standards, CDISC has initiated several projects that address different aspects of these harmonization efforts[5]. Three of CDISC's most important projects are the Clinical Data Acquisition Standards Harmonization (CDASH), the Study Data Tabulation Model (SDTM), and the CDISC Healthcare Link Initiative (HLI). Together, these three projects make significant headway toward the harmonization of healthcare and clinical trial standards.

The Clinical Data Acquisition Standards Harmonization (CDASH) initiative is a project supported by a collaborative group of organizations, and led by CDISC, to "make it easier for sites to conduct clinical research; to collect data once for multiple purposes; and to improve data quality and patient safety"[6]. To do so, CDASH focuses on developing "global, consensus-based 'content standards,' including element name, definition, and metadata for a basic set of global data collection fields on case report forms (CRFs) based on the CDISC SDTM model[7]. The CDASH standard describes recommended data collection sets for 16 domains, including demographic, adverse events, and other safety domains that are common to all therapeutic areas and types of clinical research". The standard is important to both the clinical trial and healthcare industries because it streamlines data collection in a way that promotes improved data interoperability and enhances the interface with healthcare and electronic health records.

Whereas CDASH defines a standard for data collection on CRFs, the Study Data Tabulation Model (SDTM) defines a standard structure for study data tabulations that are to be submitted as part of a product application to a regulatory authority such as the United States Food and Drug Administration (FDA)[8]. The SDTM standard is based on a model that focuses on observations. In the SDTM, observations collected during a study can be divided among three general classes: Interventions, Events, and Findings. Within these classes, each observation can be described by a series of variables, which are classified into one of five major roles: Identifier, Topic, Timing, Qualifier, and Rule. By providing this general framework for describing the organization of information collected during human and animal studies and submitted to regulatory authorities, the SDTM minimizes the need for industry to submit the same data to regulators in multiple formats, enables regulators to work more efficiently and effectively by using software tools to evaluate standardized datasets, and "supports the FDA's efforts to develop a repository for all

submitted studies and a suite of standard review tools to access, manipulate, and view the study data"[9].

While CDASH and SDTM define standards for data collection and data submission within the clinical trial industry, the Healthcare Link Initiative (HLI) takes the next step and focuses on "interoperability between electronic health records (EHRs) and clinical research systems. HLI has developed integration standards for the two systems so that data that is relevant for clinical research can easily be retrieved from existing EHR systems. By linking the two industry's systems, HLI has enabled researchers to access EHR data and use it for research purposes. In so doing, it eliminates the need for double data entry and improves data quality and the timeliness of data sharing, which is critical in safety reporting.

Like CDISC, HL7 is an important SDO actively involved in the harmonization of clinical trial and healthcare standards. HL7 developed the Reference Information Model (RIM), "a large pictorial representation of the clinical data (domains) that identifies the life cycle of events that a message or groups of related messages will carry and explicitly represents the connections that exist between the information carried in the fields of HL7 messages," to develop "coherent, extendible standards that permit structured, encoded health care information to be exchanged between computer applications while preserving meaning"[10]. HL7 also has several work groups that use the RIM to develop standards across different healthcare domains. As the group that works in the clinical research domain, Regulated Clinical Research Information Management (RCRIM) develops standards to "improve or enhance information management during clinical research and regulatory evaluation of the safety, efficacy, and quality of therapeutic products and procedures" by defining "messages, document structures, and terminology to support the systems and processes used in the collection, storage, distribution, integration, and analysis of clinical research information". Given that their missions are similar, RCRIM, under a memorandum of understanding, provides liaison to CDISC. Most importantly, RCRIM works with CDISC and other stakeholders to push forward the harmonization of clinical trial and healthcare standards and the interoperability of their data management systems. By collaborating to develop standards for interoperability, HL7 hopes to "improve care delivery, optimize workflow, reduce ambiguity, and enhance knowledge transfer among stakeholders".

As mentioned in the previous paragraph, CDISC and HL7 have been collaborating with the National Cancer Institute (NCI), the Food and Drug Administration (FDA), and other stakeholders to develop the Biomedical Research Integrates Domain Group (BRIDG) model. The goal of the BRIDG project is to "develop a shared view of the data, relationships, and processes which collectively define the domain of protocol-driven research and its associated regulatory artifacts" in order to achieve semantic interoperability (i.e. shared meaning) among both people

and systems[11]. In other words, the BRIDG model should provide a standard computable clinical trial protocol so that "protocol information can be repurposed across multiple clinical research documents, databases, and systems from study start-up through reporting and regulatory submissions". Starting with a set of data elements CDISC identified as common to regulated clinical research protocols, the Protocol Representation (PR) group, composed of CDISC, HL7, and the other stakeholders, developed the BRIDG model as a Unified Modeling Language (UML) diagram that is consistent with both CDISC's clinical trial standards and HL7's RIM model[12]. Since developing the BRIDG model, the PR group has initiated three priority use cases to further develop the machine-readable protocol: (1) harmonizing BRIDG with CDISC's SDTM model for data submission to the FDA, (2) creating a model to support protocol and trial tracking and the various clinical trial registries, and (3) working with technology providers to develop specialized applications based on the BRIDG model[13]. The PR group believes the BRIDG model, as a standard computable clinical trial protocol, and the ongoing use cases will improve healthcare delivery and clinical trial research by increasing the quality and reducing the costs of software development, knowledge acquisition, communication, and collaboration[14].

THE IMPACT OF HARMONIZATION ON SAS USERS
The efforts to harmonize healthcare and clinical trial standards, including the CDASH, SDTM, HLI, and BRIDG projects, have led to pivotal, synergistic changes to data definitions and models that will provide significant benefits to SAS users and data analysts.

As one initial change toward harmonization, the CDASH project describes data elements that are common across case report forms and recommends that the standard case report form comprise those data elements[15]. This change affords several direct and indirect benefits to SAS users and data analysts. First, a standard case report form minimizes the need for double data entry. This, in turn, reduces significant inefficiencies stemming from manual data manipulation and lessens the psychological injuries highly-educated scientists feel when required to enter the same data twice[16]. Second, a standard case report form enables SAS users and data analysts to develop standard SAS programs and macros that exploit the standardized data elements to automatically manipulate, transfer, and analyze the data[17]. Automating these processes further reduces the inefficiencies of manual data manipulation and chances for error[18]. Finally, with standard data elements, researchers can use electronic health records to pre-populate case report forms. Pre-populating case report forms with data already available in electronic health records reduces the cost of data entry and the chance of errors without entangling the records in the regulatory aspects of clinical trial research or interfering with patient care.

Like the CDASH project, the SDTM project's standards for data submission to regulatory agencies, like the Food and Drug Administration, afford benefits to SAS users and data analysts from both sides of the regulatory review process. Like the CDASH standards, the SDTM standards enable SAS users and data analysts to develop standard SAS programs and macros that exploit the standardized data elements to automatically prepare, transfer, and evaluate the data. Automating these processes has the potential to reduce inefficiencies and errors and allow researchers to work on different tasks that can improve public health and generate more value for their companies. In addition, SDTM's collaboration with the BRIDG project catalyzes the development of applications that read standard SDTM information, which will enable regulators to review research data more efficiently. Finally, the efficiencies that stem from BRIDG's

standard, machine-readable protocol and SDTM's data submission standards help researchers begin studies with less start-up time and resources and may also improve protocol quality and integrity.

While the CDASH and SDTM projects afford several benefits to SAS users and data analysts, the significant benefits come from incorporating these standardization efforts into CDISC's Healthcare Link Initiative and BRIDG projects. The goal of the Healthcare Link Initiative project is to link electronic health records and clinical trial systems and make them interoperable while the goal of the BRIDG project is to "make clinical trial protocol information backward-referenceable and forward-reusable within and across multiple clinical trial documents, databases, and systems." Together, these projects define standard data elements for data collection and submission, develop a standard, machine-readable protocol, and link electronic health records and clinical trial systems. From reduced inefficiencies and chances for error to increased automation and standard software applications, SAS users and data analysts stand to reap significant benefits from the harmonization of existing clinical trial and healthcare standards.

CONCLUSION
The efficacy of the healthcare and clinical trial industries depends on the timely exchange of accurate information. In an effort to ultimately improve the efficacy of healthcare and clinical trial research, CDISC, HL7, and other stakeholders have come together in the BRIDG project to harmonize existing healthcare and clinical trial industry standards by defining universal data element definitions and developing a standard computable clinical trial protocol. These harmonization efforts, which have already provided significant benefits to data users, promise to move us closer to more efficient and effective healthcare and clinical trial research.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

REFERENCES
1. Vascellaro, J. and W. Bulkeley (2009) "Google, IBM Promote Online Health Records," *The Wall Street Journal*, pp. B4.
2. Kush, R. et al. (2008) "Electronic Health Records, Medical Research, and the Tower of Babel," *The New England Journal of Medicine*, 358(16), pp. 1738-1740.
3. Neuer, A. (2009) "Workgroup Sets Priorities to Harmonize Standards for EHRs and Research," *eCliniqua*, Available at: http://www.bio-itworld.com/news/2009/01/05/workgroup-sets-EHR-standards.html?terms=r+kush.
4. Helton, E. and G. Mann (2007) "Using CDISC Models to Analyze Drug Safety Data," *SAS Free On-Demand Video Webcast*, Available at: http://www.jmp.com/landing/cdisc_webcast.shtml.
5. Clinical Data Interchange Standards Consortium (CDISC): http://www.cdisc.org/

6. Adams, T. (2007) "Standardized CRF Data Elements – An Idea Whose Time Has Come?" *ACRP Monitor*, pp. 11-12.
7. ClinPage (2008) "Carbonic e-Case Report Forms," *ClinPage*, Available at: http://www.clinpage.com/article/carbonic_e_case_report_forms/C12.
8. Kush, R. et al. (2008) "The Future of ODM, SDTM, and CDISC," *ClinPage*, Available at: http://www.clinpage.com/article/the_future_of_odm_sdtm_and_cdisc/C12.
9. Kubick, W. et al. (2008) "Study Data Tabulation Model (Version 1.2)," *Clinical Data Interchange Standards Consortium*, pp. 1-35.
10. Health Level Seven (HL7): http://www.hl7.org/
11. Biomedical Research Integrated Domain Group (BRIDG): http://bridgmodel.org/
12. Kush, R. (2008) "At the Heart of Clinical Research: The Protocol," *Master Control, Inc.*, Available at: http://www.mastercontrol.com/newsletter/pharmaceutical/clinical_research_protocol_0908.html?source_n3w5.
13. Willoughby, C. et al. (2007) "A Standard Computable Clinical Trial Protocol: The Role of the BRIDG Model," *Drug Information Journal*, Vol. 41, pp. 383-392.
14. Willoughby, C. et al. (2007) "A Standard Computable Clinical Trial Protocol: The Role of the BRIDG Model," *Drug Information Journal*, Vol. 41, pp. 383-392.
15. ClinPage (2008) "CDASH Case Report Form Push," *ClinPage*, Available at: http://www.clinpage.com/article/cdash_case_report_form_push/C9.
16. ClinPage (2009) "Trial-Friendly EHR Systems," *ClinPage*, Available at: http://www.clinpage.com/article/trial-friendly_ehr_systems/C9.
17. Neuer, A. (2009) "Workgroup Sets Priorities to Harmonize Standards for EHRs and Research," *eCliniqua*, Available at: http://www.bio-itworld.com/news/2009/01/05/workgroup-sets-EHR-standards.html?terms=r+kush.
18. de Montjoie, A. (2008) "CDISC Data Standards: Reaching Out to a Global Audience," *World Pharma Network*, pp. 10-12.

ACKNOWLEDGMENTS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Clinton W. Brownley, PMP
21682 Lindy Lane
Cupertino, CA 95014
Phone: (412) 973-4769
Email: clintonbrownley@yahoo.com