Journal of Biomedical Informatics 57 (2015) 88–99

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Standardizing data exchange for clinical research protocols and case report forms: An assessment of the suitability of the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM)

Vojtech Huser [a,*], Chandan Sastry [b], Matthew Breymaier [c], Asma Idriss [c], James J. Cimino [a,d]

[a] Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, United States
[b] National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health, Bethesda, MD, United States
[c] National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, United States
[d] Laboratory for Informatics Development, National Institutes of Health Clinical Center, Bethesda, MD, United States

ARTICLE INFO

Article history:
Received 24 October 2014
Revised 27 March 2015
Accepted 26 June 2015
Available online 15 July 2015

Keywords:
Clinical research informatics
Clinical protocol
Standards
Clinical trial
electronic Case Report Forms

ABSTRACT

Efficient communication of a clinical study protocol and case report forms during all stages of a human clinical study is important for many stakeholders. An electronic and structured study representation format that can be used throughout the whole study life-span can improve such communication and potentially lower total study costs. The most relevant standard for representing clinical study data, applicable to unregulated as well as regulated studies, is the Operational Data Model (ODM) in development since 1999 by the Clinical Data Interchange Standards Consortium (CDISC). ODM's initial objective was exchange of case report forms data but it is increasingly utilized in other contexts. An ODM extension called Study Design Model, introduced in 2011, provides additional protocol representation elements.

Using a case study approach, we evaluated ODM's ability to capture all necessary protocol elements during a complete clinical study lifecycle in the Intramural Research Program of the National Institutes of Health. ODM offers the advantage of a single format for institutions that deal with hundreds or thousands of concurrent clinical studies and maintain a data warehouse for these studies. For each study stage, we present a list of gaps in the ODM standard and identify necessary vendor or institutional extensions that can compensate for such gaps. The current version of ODM (1.3.2) has only partial support for study protocol and study registration data mainly because it is outside the original development goal. ODM provides comprehensive support for representation of case report forms (in both the design stage and with patient level data). Inclusion of requirements of observational, non-regulated or investigator-initiated studies (outside Food and Drug Administration (FDA) regulation) can further improve future revisions of the standard.

Published by Elsevier Inc.

1. Introduction

There are increasing pressures to lower the cost of conducting human clinical studies. One way to achieve this is to streamline the communication of clinical study protocol information to study sites and other stakeholders, such as trial registries or institutional review boards (IRBs). Because completed clinical studies represent a significant past investment and the need to re-analyze the data is common, many institutions that are consolidating data from clinical studies into larger repositories will benefit from such streamlining as well. A meta-analysis [1] reported that between 9% to 49% of randomized control trials report on outcomes that were not declared in a trial registry. This indicates that post hoc analyses can be quite frequent. Public pressure for comprehensive sharing of clinical study data will likewise benefit from improved exchange of study data and metadata [2].

Abbreviations: CDISC, Clinical Data Interchange Standards Consortium; ODM, Operational Data Model; SDM, Study Design Model; PRM, Protocol Representation Model; CTDMS, Clinical Trial Data Management System; SDTM, Study Data Tabulation Model; eCRF, electronic Case Report Form.

* Corresponding author at: Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health Clinical Center, 10 Center Dr, Room 6-5561, Bethesda, MD, United States. Tel.: +1 (240) 753 0405.
E-mail address: vojtech.huser@nih.gov (V. Huser).

http://dx.doi.org/10.1016/j.jbi.2015.06.023
1532-0464/Published by Elsevier Inc.

Table 1
Parts of study documentation and study results with relevant policies and stakeholders.

Study stage	Data elements	Relevant policies	Receiving entity
Study design documentation	Study protocol	*Study registration information:* US law, WHO list of required elements, ICMJE *Study protocol:* ICH Good Clinical Practice guideline	*Study registration information:* trial registry, medical journal *Study protocol:* IRB
	Case report forms	US law (submission to FDA; 21CFR11), HHS Interoperability Specification, HHS Structured Data Capture initiative	IRB (some CRFs), study sites, research coordinators
Study results	Basic summary results (public disclosure)	US law, EU law (EMA regulations)	trial registry, medical journal, reviewers, authors of meta-analyses
	Individual patient level data (disclosure limited to regulator or research team)	US law (submission to FDA), NIH data sharing guideline	Regulatory authority (approval of new products), trial results sharing platforms (eg, dbGaP)

Abbreviations: ICMJE: International Committee of Medical Journal Editors; ICH: International Conference on Harmonization; WHO: World Health Organization; HHS: US Department of Health and Human Services; FDA: Food and Drug Administration; EMA: European Medicines Agency.

The US National Institutes of Health (NIH) shares all of the above motivations for standardizing protocol information. We use an example of one NIH protocol to examine the issues of standardization and explore the suitability of the Clinical Data Interchange Standards Consortium (CDISC) Operational Data Model (ODM) standard to facilitate exchange of clinical protocol data. In contrast to prior studies, we evaluate the use of a single format that could cover the complete study life-cycle from study inception to termination and sharing of study results data.

2. Background

2.1. Protocol structure

A clinical study (or *protocol*) goes through several stages, including protocol drafting by a research team, protocol submission and approval by one or more IRBs, study registration within a clinical trial registry, study recruitment pre-screening (*pre-screen recruitment questions*), actual patient recruitment at the study site (*in-person recruitment questions*), and collection of study data, perhaps using electronic Case Report Forms (eCRF) within a Clinical Trial Data Management System (CTDMS) or an electronic health record (EHR).

Ideally, protocols should be represented in a format that supports common protocol data elements, such as study title, locations or enrollment goal, as well as stage-specific data and metadata elements, such as information required for IRB approval, sharing study details and study eCRFs with all study sites (including clinical research organizations or sponsors), and submission of final data to a statistician, regulator or data sharing platform. A single format for all these tasks, from study drafting (prior to enrolling the first patient) through study completion and follow-up (after the last patient's data have been collected) would be preferable.

We define a study *protocol* as a detailed document that is typically 10–80 pages long and includes the study schedule, detailed description of all study events as well as other elements defined by the Good Clinical Practice guideline (E6) [3] from the International Conference on Harmonization. This guideline, which was created with input from US as well as EU authorities, standardizes numerous protocol sections, such as, withdrawal criteria, blinding or adverse event reporting. Within the protocol, we also distinguish a short set of study metadata elements that we refer to as *study registration information* (such as title, principal investigator, research sites, study design, or enrollment goal). These elements are typically required by trial registries or internal study administration systems.

Table 1 provides an overview of different study documentation components, as well as relevant policies for each component. We consider design of case report forms to be an important attachment to the study protocol and an integral part of good study documentation. Study protocols are also needed when study results are obtained and communicated, and hence we also include in Table 1 study results data elements. These include summary data that are required by some trial registries (and by US law) and individual patient level data, that are needed for submission to a regulator or to some data sharing platforms (eg, TrialShare from National Institutes of Allergy and Infectious Diseases (NIAID) Immune Tolerance Network) [4]. The resulting representation will need to be computable, for example, an Extensible Markup Language (XML) file that can be consumed both by information systems and by humans (after transformation into formats such as Hypertext Markup Language (HTML) or Portable Document Format (PDF).

Existing standards cover to some extent the representation of study registration information and CRFs; however, there are no standards capable of transporting (in a structured sense) the full protocol document. The dominant standard development organization (SDO) for creating clinical research informatics standard is the Clinical Data Interchange Standards Consortium (CDISC), which was established in 1999. CDISC standards development efforts are organized into different workgroups, with the Protocol Representation Group (formed in 2002) working to create such a standardized format. CDISC intends to use this or a similar format to standardize submission of protocol data to clinical trial registries. For example, currently it is not possible to use a single format to submit clinical study data to USA's clinical trial registry (ClinicalTrials.gov) and EU's registry run by European Medicines Agency (EMA).

2.2. Prior work in protocol representation

2.2.1. ODM-based work

Protocol representation formats have been the subject of several prior studies. The most relevant standard is the CDISC Operational Data Model. Prior studies using ODM focused mostly on case report forms, rather than strictly on protocol representation. Bickel et al. developed an i2b2-based tool that can import the CDISC ODM formatted data into an i2b2-based data warehouse [5]. Dugas at al. developed a Web-based platform [6] and an R package [7] supporting exchange of empty eCRFs in ODM format. Karam et al. from World Health Organization developed an ODM extension focused on clinical trial registration [8]. ODM was one of the first standards developed by CDISC and it was meant from the start to be a foundation standard with building blocks for capturing a range of clinical study data. ODM was initially created in 1999 with updates to version 1.3 in 2005 and a small update to