

Data standards for clinical research data collection forms: current status and challenges

Rachel L Richesson,¹ Prakash Nadkarni²

¹Department of Pediatrics, University of South Florida College of Medicine, Tampa, Florida, USA

²Yale University, New Haven, Connecticut, USA

Correspondence to

Dr Rachel L Richesson, University of South Florida College of Medicine, Department of Pediatrics, 3650 Spectrum Boulevard, Suite 100, Tampa, FL 33612, USA; richestrl@epi.usf.edu

Received 5 November 2008

Accepted 8 February 2011

ABSTRACT

Case report forms (CRFs) are used for structured-data collection in clinical research studies. Existing CRF-related standards encompass structural features of forms and data items, content standards, and specifications for using terminologies. This paper reviews existing standards and discusses their current limitations. Because clinical research is highly protocol-specific, forms-development processes are more easily standardized than is CRF content. Tools that support retrieval and reuse of existing items will enable standards adoption in clinical research applications. Such tools will depend upon formal relationships between items and terminological standards. Future standards adoption will depend upon standardized approaches for bridging generic structural standards and domain-specific content standards. Clinical research informatics can help define tools requirements in terms of workflow support for research activities, reconcile the perspectives of varied clinical research stakeholders, and coordinate standards efforts toward interoperability across healthcare and research data collection.

INTRODUCTION

Data collection for clinical research involves gathering variables relevant to research hypotheses. These variables ('patient parameters,' 'data items,' 'data elements,' or 'questions') are aggregated into data-collection forms ('Case Report Forms' or CRFs) for study implementation. The International Organization for Standardization/International Electro-technical Commission (ISO/IEC) 11179 technical standard¹ defines a data element as 'a unit of data for which the definition, identification, representation, and permissible values are specified through a set of attributes.' Such attributes include: the element's internal name, data type, caption presented to users, detailed description, and basic validation information such as range checks or set membership.

Data element and CRF reuse can reduce study implementation time, and facilitate sharing and analyzability of data aggregated from multiple sources.²⁻³ In this paper, we summarize relevant CRFs standards and their limitations, and highlight important unaddressed informatics-standardization challenges in optimizing research processes and facilitating interoperability of research and healthcare data.

BACKGROUND AND SIGNIFICANCE

CRFs support either primary (real-time) data collection, or secondarily recorded data originating elsewhere (eg, the electronic health record (EHR) or

paper records). EHR and Research data capture differ in that the latter records a subset of patient parameters—the research protocol's variables—in much greater depth and in maximally structured form; narrative text is de-emphasized except to record unanticipated information.

Historically, CRFs were paper-based. While primary electronic data capture (EDC) has steadily increased,⁴ paper is still used when EDC is infeasible for logistic or financial reasons. The existence of secondary EDC also influences manual workflow processes related to verification of paper-based primary data, for example, checks for completeness, legibility, and valid codes. The present limbo between paper and EDC complicates standardization efforts.

CRF standards: current activities

Currently, no universal CRF-design standards exist, though conventions and some 'best' practices do.⁵⁻⁹ The Clinical Data Interchange Standards Consortium (CDISC, <http://www.cdisc.org>), which focuses primarily on regulated studies, has proposed such standards. However, these proposals, while valuable for general areas such as drug safety, do not address broader issues of clinical research, including observational research, genetic studies, and studies using patient-reported experience as key study endpoints.

Clinical Data Standards Acquisition Standards Harmonization

In response to the Food and Drug Administration (FDA)'s 2004 report, 'Innovation/Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products,' a CDISC project, Clinical Data Standards Acquisition Standards Harmonization (CDASH), addresses data collection standards through standardized CRFs.¹⁰ Initial CDASH standards focused on cross-specialty areas such as clinical-trials safety. Disease- or therapeutic-specific standards are now being considered, along with tools and process development to facilitate data-element reuse across diseases.

OpenEHR archetypes

The OpenEHR foundation has proposed archetypes¹¹⁻¹² as a basis for HL7 Clinical Document Architecture templates.¹³ Archetypes are agreed-upon specifications that support rigorous computable definitions of clinical concepts. For example, the archetype for Blood Pressure measurement includes type of measure (eg, diastolic, systolic, mean arterial), measurement conditions (eg, activity level, position), body site where measured, time of day when measured, and measurement

units. While Logical Observation Identifiers, Names, and Codes (LOINC) covers similar ground, it does so through numerous, unlinked concepts rather than a unified template. Also, measurement aspects idiosyncratic to BP—for example, body position—are handled by incorporation into the concept's Component Name, for example, 'INTRAVASCULAR SYSTOLIC ^ SITTING.' OpenEHR, by contrast, allows the semantic model's structure to vary with the parameter being described.

Clinical researchers have been specifying parameter measurement with precision long before 'archetypes' were conceived. For example, to accurately compare two medications for a chronic illness, one must control all the conditions that can influence a parameter's measurement.

Standards for medical subdomains

Domain-specific common data elements (CDEs) are emerging from groups such as the American College of Cardiology,^{14 15} the National Cancer Institute (NCI)'s Cancer Bioinformatics Grid (caBIG),^{16 17} NIH-Roadmap-Initiative interoperability demonstration projects,¹⁸ the National Institute of Neurological Disorders and Stroke,^{19 20} the Consensus Measures for Phenotypes and EXposures (PhenX) project (for clinical phenotyping standards for Genome-wide Association Studies),²¹ the Diabetes Data Strategy (Diabe-DS) project,^{22 23} and the Health Information Advisory Committee (HITAC) effort.^{24 25}

Data-element repositories

The NCI's Cancer Data Standards Repository (caDSR) uses the ISO/IEC 11179 design to 'bank' CRF questions and answer sets.^{26–28} Criticisms of caDSR include lack of curation, redundancy, and the absence of a representation of CRFs.²⁶ CaDSR's utility will possibly improve as NCI redesigns it with requirements from collaborating organizations. CDISC is using the ISO/IEC 11179 design for CSHARE, a repository for domain-specific research questions and answer sets.

The Agency for Healthcare Research and Quality (AHRQ)-hosted United States Health Information Knowledgebase (USHIK) metadata registry²⁹ includes artifacts emerging from federal healthcare-standardization task forces, such as information models, data elements, data-collection requirements, functional requirements, system specifications, and supporting documentation. Controlled terminologies such as LOINC and the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), also possess certain characteristics of data-element repositories: for example, LOINC encodes questions and answers for many patient-directed surveys and clinical patient assessment instruments.

Harmonization of healthcare and clinical research data standards

Clinical information models such as the Health Level 7 (HL7) Reference Information Model (RIM) use terminologies differently than the research-oriented CDISC Operational Data Model (ODM).³⁰ While HL7 interoperability depends in part on mapping data elements to concepts in standard terminologies, ODM only cares that a terminology may act as a source for a data element's contents—for example, an element 'ICD9CM_Code' is populated with terms from ICD-9-CM, 2010 edition. ODM does not support mapping of data elements themselves (eg, serum total cholesterol, systolic BP) to terminologies. Consequently, although intended to support data interchange, ODM cannot address the mapping problem, where semantically identical data elements may have different names across different systems.

The Biomedical Research Integrated Domain Group (BRIDG) domain-analysis model was developed jointly by FDA, NCI, CDISC, and HL7 to overcome this gap.³¹ BRIDG, however, still does not specify use of standard terminologies, and to date only pilot applications have been developed using BRIDG.

LIMITATIONS OF EXISTING STANDARDS AND METHODOLOGY CDASH best-practice recommendations

While valuable overall, some CDASH recommendations reflect historical paper-based workflows and off-line or non-electronic operations. For example, for parameters that must be computed in real time, the CDASH specifications advocate worksheets that require data-entry staff to use calculators, instead of programming computations directly into electronic CRFs.

CDASH controversially recommends not providing coding dictionaries for adverse events, medications, or medical history to research staff when interviewing patients, supposedly to minimize potential bias. This advice, if followed, risks introducing errors (eg, misspelled drug names owing to faulty patient recall) that can only be resolved by recontacting the patient, whereas online drug-name lists are searchable with algorithms such as Double Metaphone that support spelling-error recovery.³² Further, online access is almost mandated for adverse-event grading using NCI's Common Terminology Criteria for Adverse Events (CTC AE), where adverse-event severity grades are defined unambiguously to minimize interobserver variation but are far too numerous to commit to human memory.

ISO/IEC 11179 data model

ISO/IEC 11179, used by CaDSR and CSHARE, was originally intended for descriptive-metadata registries. Applying it to the significantly different problem of clinical research has unearthed numerous concerns. While able to capture isolated data elements' semantics reasonably well, ISO/IEC 11179 cannot represent interelement constraints—for example, the sum of the differential white-blood-cell-count components must equal 100, and systolic blood pressure must be greater than diastolic. There is no concept of higher-level element groupings (such as CRFs), of element order within groupings, of calculated elements, or of rules where certain elements are only editable conditional to specific values being entered in previous elements (so-called skip logic).

The standard has a limited concept of data-element presentation: the relationship between an element and its associated caption and explanation text is modeled as one to one. However, for many research applications, the relationship is actually one to many. For example, in multinational clinical studies, the same CRF may be deployed in different languages. Here, data elements, while having fixed internal names, will have alternative (language-specific) captions and explanations.

Finally, while ISO/IEC 11179 repositories are effectively thesauri, their data model differs radically from the standard concepts—terms—relationships design used for thesauri such as the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) or the Unified Medical Language System (UMLS). While concepts (with or without associated definitions) are central to thesauri, in ISO/IEC 11170, data elements (equivalent to terms) are central: concepts merely categorize elements, and narrative definitions are (incorrectly) associated with data elements instead of concepts.

The Extended Metadata Registry (XMDR) consortium, <http://xmdr.org>, aims to extend ISO/IEC 11179 to address terminology issues. This group, however, appears to be

inactive—its last publicly posted group meeting was in late 2007—and its impact is uncertain.

CDISC operational data model

The XML-based CDISC operational data model (ODM), a metadata- and data- interchange model, rectifies many limitations of ISO/IEC 11179. It explicitly models CRFs, and addresses the multilanguage issue through a `TranslatedText` element—a Unicode string plus a language identifier. The ODM also partially addresses calculated parameters and cross-element-based validation. However, it is not comprehensive enough to allow receiving systems to use imported metadata directly for CRF generation.

Computations and validation expressions

Computations and validation expressions need to be expressed in the syntax of a specific programming language. Unless both systems use the same language, manual modifications of the metadata by programmers are required. The ODM accepts this unavoidable limitation. The `FormalExpression` element used to specify computations must contain a ‘context’ subelement naming the language used—for example, ‘Oracle PL/SQL.’

Script/instructions

For CRFs used as interview questionnaires, cross-interviewer variation is minimized through standardized scripts—sentences spoken verbatim to the subject to elicit the desired information. Instructions provide guidance in CRF usage and data-gathering. Electronically, scripts/instructions are typically displayed on demand during data capture but are hidden during review or modification. ODM lacks both script and instruction definitions.

Validation of text

ODM lacks support for regular-expression validation of text data elements.

Interelement dependencies

Interchange models are generally simpler than data-storage models.³³ Metadata essential to storage-model robustness, but irrelevant for interchange purposes, may be omitted from the interchange model. An example of ODM omission is interelement dependencies: an element being validated, computed, or skipped depends on other elements. Dependencies may be complex—for example, calculation of renal function status depends on computation of estimated glomerular filtration rate, which in turn depends on serum creatinine, age, sex, and race.

Dependency checking prevents accidental removal or renaming of independent elements from a CRF, which would cause the CRF to operate incorrectly.

Other important ODM omissions include ownership and context of use. Practically all CRFs used in autism research, for example, are copyrighted. Copyright information must be part of the CRF definition. Context of use includes documentation about the clinical conditions where the CRF applies and prerequisites for the CRF’s use.

CRF STANDARDS CHARACTERIZATION AND STATUS

CRF standards can be conceptualized at several levels: form, group, section, and item. We summarize the areas of agreement and dispute at each level, and also consider aspects of CRF-design processes that impact consistent research data collection.

Form level

Little consensus exists on of the choice and content of CRF standardization candidates. Few CRFs can be reused unchanged

across all protocols. Even for seemingly common activities such as physical exam and medical history, structured data capture—explicit recording of findings—varies vastly by disease and protocol. For example, gastrointestinal bleeding or hepatic encephalopathy is recorded explicitly in a cirrhosis study, but not in schizophrenia.

Within a tightly defined disease domain, standard CRFs seem feasible and useful, though their content may change with future variations in study designs. For example, the venerable Hamilton Depression Rating Scale originated in 1960 as a 17-item questionnaire.³⁴ Later, some researchers created different subsets, while others incorporated additional questions.³⁵ Many proposed ‘standard’ CRFs may well meet a similar fate. Long-term content stability may be one measure of CRF-standard success.

The segregation of data items relevant to a research protocol into individual CRFs is often based on considerations other than logical grouping, and may vary with the study design. For example, in a one-time survey, one may well designate a single CRF to capture all items if these are not too numerous. In a longitudinal study, however, items recorded only once at the start of the study are placed in a CRF separate from items that are sampled repeatedly over multiple visits.

One concern about ‘standard’ CRF use is that users should not be pressured to collect parameters defined within the CRF that are not directly related to a given protocol’s research objectives: such collection costs resources and violates Good Clinical Practice guidelines.³⁶ Even instructing research staff to ignore specific parameters constitutes unnecessary information overload: presenting extraneous parameters onscreen is poor interface design. Dynamic CRF-rendering offers one way out of this dilemma: protocol-specific CRF customization allows individual investigators to specify, at design time, the subset of parameters that they consider relevant. Web-application software can read the customization metadata and render only applicable items.

Group level

A group is a set of semantically closely related parameters. For example, a Concomitant Medications group would include the medication name; how recorded (eg, generic or brand name); dosage details—numeric value, units, frequency, and duration; a start date, end date, whether this was a continuation of previous therapy, therapeutic indications, and possibly compliance information.

Other parameter groupings, such as the components of a differential white-blood-cell count or a liver-function panel, occur naturally in medicine. Typically, a group is associated with a single time-stamp that records when the event (eg, a blood draw) related to its parameters occurred, or two time-stamps to record the start and end of events that have a duration (eg, a course of radiotherapy).

Explicit associations between related parameters within the group include skip logic and expressions for calculated elements. Both LOINC and PhenX standards consider groups (‘panels’) as a series of observations. OpenEHR archetypes can also be used as section building-blocks.

Section level

A section encompasses one or more groups. The division of CRFs into sections is often arbitrary. In paper-based data capture, CRFs consisting of a single, giant section are not unknown. For example, the 1989 revision of the Minnesota Multiple Personality Inventory for psychiatric assessment has 567 questions. In real-time EDC, by contrast, subdivision into smaller sections is

generally preferred, allowing (or requiring) the subject to save data changes before moving to another section. This minimizes the risks of inadvertent data loss due to failure to save, timeouts, or service interruption. Section size is often determined by the number of items that can be presented on a single desktop-computer screen.

The requirement for CRF-content flexibility to deal with disease and protocol variations impacts the involved sections/groups. It is doubtful whether section names/captions should be standardized. The designation of section headings and explanation that serve to describe the section's purpose is, we believe, best left to individual investigators.

Item level

Standardization of items is non-controversial, being the linchpin of semantic interoperability. Survey Design and Measurement Theory provides well-accepted best practices for design of good items such as mutually exclusive and exhaustive answer choices,³⁷ non-leading question text,^{7, 8} and consistency of scale escalation in answer sets.⁶ A review of the literature, including the CDASH recommendations, gives useful general guidance on constructing yes/no questions, scale direction, date/time formats, scope of CRF data collection, prepopulated data, and collection of calculated or derived data.⁵⁻⁹

All the standards discussed earlier emphasize use of narrative definitions for items. Such definitions need to be made maximally granular—that is, divided into separate fields—because different parts of the definition such as explanatory text, scripts, instructions, and context of use serve different purposes.

Certain items (especially questionnaire-based ones) have a discrete set of permissible values (also called 'responses' or 'answers'). The set elements may be unordered (eg, 'Yes, No, Don't Know') or ordered (eg, severity grades such as 'Absent, Mild, Moderate, Severe' or Likert scales). One must record whether enumerations are unordered or ordered, because they impact how data based on these items can be queried. Thus, one can ask for patients who had a severity greater than or equal to 'Moderate,' but data based on unordered enumerations can only be compared for equality or inequality to a value.

CRF development process

The notion of process as vital to quality metrics and outcomes is reinforced through standards such as ISO 9000^{38, 39} and the health-outcomes research literature.⁴⁰⁻⁴² While CRF content is necessarily variable, consensus regarding standards for explicit processes for identification or development of quality data is more readily reached.

The CDASH standards document, 'Recommended Methodologies for Creating Data Collection Instruments,' presents important and necessary features of the CRF development process. The techniques described include: adequate and 'cross-functional' team review; version control, and documented procedures for design, training, and form updates. The FDA also requires rigor in the development, validation, and use of data elements related to patient-reported outcomes as study endpoints in investigational new drug studies.⁴³

FUTURE CHALLENGES FOR CLINICAL RESEARCH INFORMATICS

As the field of clinical research informatics matures, it will need to move from a mode of primarily reacting to clinical researchers' needs through service provision, to one of active leadership by suggesting directions for standardization. We now identify several challenges for clinical research informatics related to data element and CRF definition and data capture.

Articulating the data-collection standards needs for all of research

The limited focus of disease-specific consortia makes comprehensive coverage of individual areas more likely. However, it may lead to proliferation of multiple, possibly incompatible, definitions for overlapping subject areas, such as tobacco exposure or dietary history. Similarly, researchers would benefit from a clear understanding of the extensive overlap of various clinical terminologies (eg, SNOMED CT and LOINC, SNOMED CT and RxNorm), as well as advice regarding which standards are appropriate for a particular research context.

CDISC's focus on regulated research leaves many standardization issues unaddressed. An AMIA Clinical Research Informatics group could be well poised to identify the gaps and devise strategies to fill them. They would also be able to address relationships between clinical research data collection standards and EHR specifications, as well as the broad issue of secondary use of clinical data for research. Additional tasks could include the review of standards and their scope, and relating them to needs of clinical research.

Banking of research-data elements

Reuse of standard CRF and higher-level groupings can be facilitated by publicly available repositories. A greatly extended database counterpart of the CDISC ODM may possibly meet the requirements of the repository data model. The comprehensive documentation of individual items and groupings, as well as links between these and concepts in standard biomedical terminologies, will increase usability and utility.⁴⁴ When additionally supported by robust search tools, the repositories can serve as educational tools for researchers. As suggested by Brandt *et al*,⁴⁵ item repositories can reduce the burden on new investigators to create their own items, because existing, validated items or sets of items can be reused.

We now discuss some significant unsolved challenges for such repositories.

Modeling of questions

Repositories must distinguish between apparently identical items that have different presentations, and provide detailed recommendations for choosing from these. Consider a questionnaire regarding past history of several clinical conditions (eg, diabetes, myocardial infarction, etc), where the response to each can be 'Yes,' 'No,' or 'Don't Know.' A second questionnaire presents the same clinical conditions with check boxes, which can be either checked (Yes) or unchecked (No).

Because both healthcare and research generally require recording unknowns explicitly, CDASH correctly recommends representation 1. However, for paper-based CRFs, if the list of clinical conditions is extensive with most responses expected to be 'No,' CDASH recognizes that representation 2 (a series of checkboxes) is significantly more ergonomic, at the risk of introducing some data-capture error for Don't Know's. This risk depends on the patients under study, being less for highly self-knowledgeable patients. (CDASH, however, does not currently document that if primary EDC is an option, one can use representation 1 and still support good ergonomics. An electronic CRF could present all items during initial data entry with the default 'No' preselected, with onscreen instructions to click 'Yes' or 'Don't Know' as applicable.)

While 'best practice' recommendations clearly depend on the clinical setting, repositories that are intended to guide investigators must also include recommendations and guidance.

Value set or terminology 'binding'

The linking of repository elements to concepts in clinical terminologies presents several challenges.

Clarity regarding the motivation and strategy for clinical coding

Best-practice approaches for employing controlled terminologies must be defined and documented. For example, while SNOMED CT has a complex concept model, its use can involve a simple approach if the use case supports it. For example, the Patient Registry Item Specification and Metadata (PRISM) project, which applies SNOMED CT for data elements related to rare disease registries, employs only certain SNOMED CT hierarchies and does not require post coordination for situational context.⁴⁶ Other uses cases, particularly those that involve interoperability between disparate systems, could increase mapping complexity and mandate post coordination.

Different ways of encoding the same item

Because many data items contain question and answer components, there are multiple approaches to use them. In SNOMED CT, for example, one could use the Concept ID 'Abnormal Breath Sounds' (concept ID #301273002) with the qualifiers 'Present' (#52101004) or 'Absent' (#2667000). Alternatively, the combination of question-plus-answer ('Abnormal breath sounds=absent') could be represented by a single SNOMED CT concept 'Normal Breath Sounds' (#48348007).

Any standardization effort will need to specify guidelines for consistent use of SNOMED CT, to help eliminate most of the terminology—information model interactions that plague standards implementation in healthcare.^{47–49} However, a fully modeled SNOMED CT expression to represent the question and assign its semantic aspects, relying on existing SNOMED CT modeling guidelines,⁵⁰ is probably unwarranted.

Clinical research informaticians may need to create a SNOMED CT extension to support fully modeled expressions. While coordination of multiple parallel efforts may become an issue, identifying and comparing modeling, implementation, and coding strategies is a high priority.

Absence of clear standards for psychosocial assessment items

While most data elements related to clinical disease might be expected to match to SNOMED CT concepts exactly, this is not the case for questionnaires that deal with psychiatric/psychosocial areas. For example, the 24-question Center for Epidemiological Studies Depression (CES-D) Scale⁵¹ is widely used for self-rating by patients undergoing cancer therapy. One CES-D question is 'You were bothered by things that usually Don't bother you,' with the four-level ordinal response set, '0, Rarely; <1 day/week'; '1, Some of the time (1–2 days/week)'; '2, Moderately (3–4 days/week)' and '3, Most/all of the time (5–7 days).' Responses to all items are summed to yield an overall depression score.

Trying to fully define either the question or the answers to existing SNOMED CT concepts using post coordination is a formidable challenge, especially given that SNOMED CT's compositional model does not support the NOT operator (SNOMED CT User Guide 2010, appendix B; Negation). Representing CES-D items through pre coordination would require the creation of $24 \times 4 = 96$ new SNOMED CT concepts. While many applications of SNOMED CT aspire to use it as an ontology—a collection of information that supports reasoning—it is doubtful if either approach would enable useful reasoning either with the concepts themselves or with data indexed by them (eg, 'identify patients with a score greater than 2'). SNOMED CT, like most terminologies, currently has no idea either of ordered sets, or of

numeric operations. Similarly, trivial reasoning problems, such as determining that a score of 3 is worse than a score of 2, are impossible with SNOMED CT's current knowledge representation but would be readily addressed with a modestly augmented 11179-based representation.

An alternative standard for representing observations and measures is LOINC, which currently (June 2010) contains 58 967 terms with 15 608 clinical terms, including those from standardized assessment instruments. While CES-D is not currently included, Bakken *et al* have verified LOINC's suitability for ordinal scales.⁵² Hopefully, the LOINC—IHTSDO cooperation will include coordination of content for such items, as well as discourage redundant efforts by independent researchers and consortia.

Data aggregation

Automated or semiautomated facilitation of meta-analysis of multiple data sets by electronic inspection of element definitions is an open problem. Unless element definitions across two or more studies are determined to be semantically identical—same terminology mapping, same data type, units, enumeration—or allow a mathematical transformation into a common grain, sometimes with minor or major information loss, it is not possible to combine elements across studies. To illustrate a worst-case information-loss scenario, if one study measures smoking by number of years smoked currently and in the past, while another measures the same by cigarettes per week, all that the merged data can tell us is whether a given individual is a non-smoker, ex-smoker, or current smoker.

While terminological mappings can facilitate intervariable comparisons in theory, the practical issues with terminological binding discussed above create formidable challenges. Unless non-redundant terminology subsets are created for clinical research, and the capabilities of terminologies significantly enhanced, this problem cannot even begin to be tackled except in straightforward cases.

CONCLUSION

Data-capture standards can facilitate efficacious development and implementation of new studies, element reuse, data quality and consistent data collection, and interoperability. Because of the protocol-centric nature of clinical research, opportunities for shared standards at levels higher than individual items are relatively limited compared with item-level standards. Nevertheless, disease-specific CRF standardization efforts have helped identify standard pools of data items within focused research and professional communities, and consequently helped achieve research efficiencies within their application areas. It will be interesting to see whether disease-specific efforts such as the NCI CRF standardization initiatives can remain in harmony with evolving national research standards specifications.

Of more immediate and widespread (pan-disease) relevance are standardization efforts toward the development of sound processes and workflow for CRF and CRF section development, as well as data collection and validation. Such development should also emphasize the use of terminologies to facilitate semantic interoperability. As good CRF design principles and community collaboration become best practices in clinical research, the structure and content of individual CRFs/sections can be left reasonably flexible to allow adaptation to individual protocol requirements.

Acknowledgments We wish to thank our colleagues at the University of South Florida, M Nahm of Duke University for insightful comments on early drafts, and

T Patrick of University of Wisconsin-Milwaukee for his stimulating comments on terminology aspects.

Funding Funding and/or programmatic support for this project was provided by Grant Numbers RR019259-01 and RR019259-02 from the National Center for Research Resources and National Institute of Neurological Disorders and Stroke, respectively, both National Institutes of Health components, and the National Institutes of Health Office of Rare Diseases Research.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Iso-lec.** *Information Technology—Metadata Registries (MDR), in Part 1: Framework*. Geneva: International Organization for Standardization/ International Electrotechnical Commission, 2004. <http://metadata-stds.org/11179/#A1>.
2. **Souza T**, Kush R, Evans JP. Global clinical data interchange standards are here! *Drug Discov Today* 2007;**12**:174–81.
3. **Levin R.** *Data Standards for Regulated Clinical Trials: FDA Perspective*. 2005. http://www.cdisc.org/pdf/2004_06_14_cdisc.pdf (accessed 21 Jul 2006).
4. **Welker JA.** Implementation of electronic data capture systems: barriers and solutions. *Contemp Clin Trials* 2007;**28**:329–36.
5. **Clinfource I.** *E-Training for Clinical Trials. On-line Training Session*. 2008. <https://www.clinfource.com/documents/brochurell.pdf> (accessed 14 Mar 2011).
6. **Pocock SJ.** *Clinical Trials: A Practical Approach*. New York: Wiley and Sons, 1983:247.
7. **Gore SM.** Assessing clinical trials—record sheets. *BMJ (Clin Res Ed)* 1981;**283**:296–8.
8. **Crewson Applegate KE PE.** Fundamentals of clinical research for radiologists. *Am J Reontgenol* 2001;**177**:755–61.
9. **Lu Z.** Technical challenges in designing post-marketing eCRFs to address clinical safety and pharmacovigilance needs. *Contemp Clin Trials* 2010;**31**:108–18.
10. **CDISC.** *Clinical Data Acquisition Standards Harmonization: Basic Data Collection Fields for Case Report Forms*. Draft version 1.0. <http://www.cdisc.org/cdash> (accessed 1 Sep 2010).
11. **Leslie H.** International developments in openEHR archetypes and templates. *HIM J* 2008;**37**:38–9.
12. **Kalra D**, Beale T, Heard S. The openEHR foundation. *Stud Health Technol Inform* 2005;**115**:153–73.
13. **Browne E.** *Archetypes for HL7 CDA Documents*. 2008. http://www.openehr.org/wiki/download/attachments/3440870/Archetypes_in_CDA_4.pdf (accessed 14 Mar 2011).
14. **McNamara RL**, Brass LM, Drozda JP, et al. ACC/AHA key data elements and definitions for measuring the clinical management and outcomes of patients with atrial fibrillation: A report of the American College of Cardiology/American Heart Association task force on clinical data standards (Writing Committee to Develop Data Standards on Atrial Fibrillation). *J Am Coll Cardiol* 2004;**44**:475–95.
15. **Buxton AE**, Calkins H, Callans DJ, et al. ACC/AHA/HRS 2006 key data elements and definitions for electrophysiological studies and procedures: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (ACC/AHA/HRS Writing Committee to Develop Data Standards on Electrophysiology). *Circulation* 2006;**114**:2534–70.
16. **Ohmann C**, Kuchinke W. Future developments of medical informatics from the viewpoint of networked clinical research. Interoperability and integration. *Methods Inf Med* 2009;**48**:45–54.
17. **NCI.** *caBIG. Cancer Biomedical Informatics Grid. Data Standards. 2006 01-04-2008*. https://cabig.nci.nih.gov/workspaces/NCDE/Data_Standards/index_html/ (accessed 25 May 2006).
18. **Nahm M**, McCourt B, Walden A, et al. *Cardiovascular and Tuberculosis Data Standards, Release 1.0, Package 1*. 2008. <http://www.cdisc.org/standards/cardio/index.html> (accessed 23 Nov 2010).
19. **Stone K.** NINDS common data element project: a long-awaited breakthrough in streamlining trials. *Ann Neurol* 2008;**68**:A11–13.
20. **NINDS.** *NINDS Common Data Elements. Harmonizing Information. Streamlining Research. Project Overview*. 2010 Aug 19 <http://www.commondataelements.ninds.nih.gov/ProjReview.aspx> (accessed 20 Aug 2010).
21. **Stover PJ**, Harlan WR, Hammond JA, et al. PhenX: a toolkit for interdisciplinary genetics research. *Curr Opin Lipidol* 2010;**21**:136–40.
22. **Richesson RL**, Mon D, Kallam C, et al. *A Strategy for Defining Common Data Elements to Support Clinical Care and Secondary Use in Clinical Research, in 2010 AMIA Clinical Research Informatics Summit*. San Francisco, 2010. http://crisummit2010.amia.org/files/symposium2008/100_richesson.pdf. (accessed 14 Mar 2011)
23. **HL7.** *Diabe-DS Project Wiki—EHR Diabetes Data Strategy*. 2010 http://wiki.hl7.org/index.php?title=EHR_Diabetes_Data_Strategy (accessed 23 Nov 2010).
24. **ANSI.** *Healthcare Information Technology Standards Panel (HITSP). Enabling Healthcare Interoperability*. 2010. <http://www.hitsp.org/> (accessed 23 Nov 2010).
25. **Kuperman GJ**, Blair JS, Franck RA, et al. Developing data content specifications for the nationwide health information network trial implementations. *J Am Med Inform Assoc* 2010;**17**:6–12.
26. **Nadkarni PM**, Brandt CA. The common data elements for cancer research: remarks on functions and structure. *Methods Inf Med* 2006;**45**:594–601.
27. **Warzel DB**, Andonaydis C, McCurry B, et al. Common data element (CDE) management and deployment in clinical trials. *AMIA Annu Symp Proc* 2003:1048.
28. **Covitz PA**, Hartel F, Schaefer C, et al. caCORE: a common infrastructure for cancer informatics. *Bioinformatics* 2003;**19**:2404–12.
29. **AHRQ.** *The United States Health Information Knowledgebase (USHIK)*. 2010. <http://ushik.ahrq.gov/index.html?Referer=Index> (accessed 23 Nov 2010).
30. **Richesson RL**, Krischer JP. *Data standards in clinical research: gaps, overlaps, challenges and future directions*. *J Am Med Inform Assoc* 2007;**14**:687–96.
31. **Fridsma DB**, Evans J, Hastak S, et al. The BRIDG project: a technical report. *J Am Med Inform Assoc* 2008;**15**:130–7.
32. **Phillips L.** *The Double Metaphone Search Algorithm, in C/C++ Users Journal*. 2000. <http://portal.acm.org/citation.cfm?id=349132> (accessed 18 Mar 2011).
33. **Blaha M.** Data store models are different from data interchange models. *Electr Notes Theor Comp Sci* 2004;**94**:51–8.
34. **Hamilton M.** Rating depressive patients. *J Clin Psychiatry* 1980;**41** (12 Pt 2):21–4.
35. **Williams JBW**, Link MJ, Rosenthal NE, et al. *Structured Interview Guide for the Hamilton Depression Rating Scale, Seasonal Affective Disorders Version (SIGHSD)*. New York: New York Psychiatric Institute, 1988.
36. **EMA.** *ICH Topic E 6 (R1) Guideline for Good Clinical Practice; CPMP/ICH/135/95*. 59. 2002. <http://www.emea.europa.eu/pdfs/human/ich/013595en.pdf> (accessed 14 Mar 2011).
37. **Aday LA.** *Designing and Conducting Health Surveys*. 2nd edn. San Francisco, CA: Jossey-Bass, 1996:560.
38. **Poksinska B**, Kahlgaard JJ, Antoni M. The state of ISO 9000 certification: A study of Swedish organisations. *TQM Mag* 14, 2002. doi:10.1108/09544780210439734. <http://dx.doi.org/10.1108/09544780210439734>, <http://www.emeraldinsight.com/journals.htm?articleid=842040&show=abstract>.
39. **Tsim YC**, Yeung VWS, Leung ETC. An adaptation to ISO 9001:2000 for certified organisations. *Managerial Auditing Journal* 17, <http://dx.doi.org/10.1108/02686900210429669>. 2002. <http://www.emeraldinsight.com/journals.htm?articleid=868561&show=abstract>.
40. **Donabedian A.** Quality assurance. Structure, process and outcome. *Nurs Stand* 1992;**7**(11 Suppl QA):4–5.
41. **Donabedian A.** *Explorations in Quality Assessment and Monitoring: Vol. 1. The Definition of Quality and Approaches to its Assessment*. Ann Arbor, MI: Health Administration Press, 1980.
42. **Donabedian A.** Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;**44**:166–206.
43. **FDA.** *Guidance for Industry. Qualification Process for Drug Development Tools. DRAFT GUIDANCE.F.a.D.A.C.f.D.E.a.R. (CDER)*. 2010. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf> (accessed 14 Mar 2011).
44. **Stausberg J**, Löbe M, Verplancke P, et al. Foundations of a metadata repository for databases of registers and trials. *Stud Health Technol Inform* 2009;**150**:409–13.
45. **Brandt CA**, Cohen DB, Shifman MA, et al. Approaches and informatics tools to assist in the integration of similar clinical research questionnaires. *Methods Inf Med* 2004;**43**:156–62.
46. **Richesson R**, Shereff D, Andrews J. [RD] PRISM Library: patient registry item specifications and metadata for rare diseases. *J Libr Metadata* 2010;**10**:119–35.
47. **Chute CG.** Medical concept representation. In: Chen H, Fuller SS, Friedman C, et al, eds. *Medical Informatics. Knowledge Management and Data Mining in Biomedicine*. New York: Springer, 2005:163–82.
48. **Mead CN.** Data interchange standards in healthcare IT: computable semantic interoperability: now possible but still difficult, do we really need a better mousetrap? *J Healthcare Inf Manag* 2006;**20**:71–8.
49. **Rector AL.** The Interface Between Information, Terminology, and Inference Models. In: *Tenth World Conference on Medical and Health Informatics: MedInfo-2001*. London, 2001.
50. **IHTSDO.** *SNOMED CT Style Guide: Situations with Explicit Context?* International Healthcare Terminology Standards Development Organization, Copenhagen, 2008.
51. **Raloff LS.** The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Measure* 1977;**1**:385–401.
52. **Bakken S**, Cimino JJ, Haskell R, et al. Evaluation of the clinical LOINC (Logical Observation Identifiers, Names, and Codes) semantic structure as a terminology model for standardized assessment measures. *J Am Med Inform Assoc* 2000;**7**:529–38.