
LOINC®: a universal catalogue of individual clinical observations and uniform representation of enumerated collections

Daniel J. Vreeman*

Indiana University School of Medicine,
410 W. 10th Street, Suite 2000,
Indianapolis, IN 46202, USA

and

Regenstrief Institute, Inc,
410 W. 10th Street, Suite 2000,
Indianapolis, IN 46202, USA
E-mail: dvreeman@regenstrief.org

*Corresponding author

Clement J. McDonald

Lister Hill National Centre for Biomedical Communications,
National Library of Medicine,
National Institutes of Health,
8600 Rockville Pike, Bldg. 38-A,
Rm. 7N704, Bethesda, MD 20894, USA
E-mail: clemmcdonald@mail.nih.gov

Stanley M. Huff

University of Utah,
26 South 2000 East,
Room 5775 HSEB,
Salt Lake City, UT 84112, USA

and

Intermountain Healthcare,
4646 W. Lake Park Blvd.,
Salt Lake City, UT 84120, USA
E-mail: Stan.Huff@imail.org

Abstract: In many areas of practice and research, clinical observations are recorded on data collection forms by asking and answering questions, yet without being represented in accepted terminology standards these results cannot be easily shared among clinical care and research systems. LOINC contains a well-developed model for representing variables, answer lists and the collections that contain them. We have successfully added many assessments and other collections of variables to LOINC in this model. By creating a uniform representation and distributing it worldwide at no cost,

LOINC aims to lower the barriers to interoperability among systems and make this valuable data available across settings when and where it is needed.

Keywords: clinical observations; framework; health information technology; patient data; patient assessments; data sets; public health; research; standards; terminology.

Reference to this paper should be made as follows: Vreeman, D.J., McDonald, C.J. and Huff, S.M. (2010) 'LOINC®: a universal catalogue of individual clinical observations and uniform representation of enumerated collections', *Int. J. Functional Informatics and Personalised Medicine*, Vol. 3, No. 4, pp.273–291.

Biographical notes: Daniel J. Vreeman received a BA in Biological Sciences from Cornell, and a DPT from Duke. He completed a two-year NLM-sponsored Medical Informatics Research Fellowship at Regenstrief Institute and an MS in Clinical Research with an emphasis on medical informatics. His primary research focus is on the role of standardised clinical vocabularies to support electronic health information exchange. As Associate Director for Terminology Services at Regenstrief, he directs the development of LOINC, an international standard for laboratory and clinical observations, and provides leadership and oversight to terminology services that undergird the Indiana Network for Patient Care, a regional health information exchange in central Indiana.

Clement J. McDonald, distinguished physician-scientist, is one of the nation's most accomplished, productive experts in electronic health record systems. Previously, he was Professor of Medicine, IU School of Medicine and Director, Regenstrief Institute, a privately endowed, internationally renowned research institute. He developed the Regenstrief Medical Record System, directed its use in clinical trials and improved patient care. He created the Indiana Network for Patient Care, now considered a national model for regional health information exchange. He pioneered development of health data standards including LOINC, and received numerous research grants and contracts. He directed an NLM-funded Informatics Training Center at IU.

Stanley M. Huff is the Chief Medical Informatics Officer at Intermountain Healthcare, and a Professor (Clinical) in Medical Informatics at the University of Utah School of Medicine. He is responsible for the architecture and functions of all clinical information systems at Intermountain Healthcare. His academic interests centre on medical vocabularies, clinical information models and medical database architectures, fields upon which his career has centred for over 20 years. He is Chair of the Clinical LOINC Committee, and teaches a course in medical vocabulary and data exchange standards at the University of Utah. He is currently working on the design and implementation of a next-generation Electronic Medical Record system, focusing on the issues of linking standard terminologies and ontologies to detailed clinical models.

1 Introduction

The healthcare we deliver continues to be hampered by the incompleteness of patient data available to providers when and where they need it (Smith et al., 2005; van Walraven et al., 2008). Coalescing the many varied sources that produce and store health

information is especially difficult because of the plethora of idiosyncratic local conventions for representing clinical concepts in different electronic systems. We can build bridges across these islands of data much more efficiently by using data exchange standards (McDonald, 1997). LOINC® (Logical Observation Identifiers Names and Codes) is a universal code system for identifying laboratory and clinical observations that facilitates exchange and pooling of results for the clinical care, research and outcomes management (McDonald et al., 2003). When used in conjunction with widely adopted messaging standards such as Health Level 7 (HL7), vocabulary standards like LOINC can be an essential ingredient for efficient electronic processing and storage of clinical data that comes from many independent sources.

In many areas of practice and research, clinical observations are recorded on data collection forms by asking and answering questions. Survey instruments, questionnaires, case reports and other forms are an important and ubiquitous method of measuring a wide range of health attributes and other aspects of care delivery. They are widely used to screen, assess and monitor aspects as diverse as health-related quality of life, functional status, mental capacity, social participation and caregiver support. Yet, without being represented in accepted terminology standards, these results cannot be easily shared among clinical care and research systems.

All of the potential advantages of health information technology are constrained by the scope of the data available within them. LOINC intentionally covers a circumscribed domain, namely, observation identifiers. The LOINC Committee focused on this domain for several reasons (Huff et al., 1998). In particular, because many systems were electronically sending procedure and measurement results using institution-specific names and codes, such a standard would have immediate benefits. LOINC is an openly developed standard that divides its work into two divisions: the Laboratory division focuses on the observations and measurements that can be made on specimen and the Clinical division focuses on the observations and measurements that can be made on patients. Many areas of LOINC such as clinical laboratory testing (Vreeman et al., 2007; McDonald et al., 2003), radiology reports (Vreeman and McDonald, 2005, 2006) and clinical note titles (Dugas et al., 2009; Hyun et al., 2009) have been found to have good content coverage in live clinical systems. Over time, we have continued to expand LOINC's content in many areas. The current LOINC version (Version 2.34, December 2010) contains 61,255 terms, of which 44,511 are laboratory terms and 16,744 are clinical terms.

Since its inception, Regenstrief has developed LOINC as an open standard and distributed it at no cost worldwide. LOINC has been widely adopted in both the public and the private sector, in the USA and internationally. Since 2008, LOINC worldwide adoption has continued to grow at the fast pace of 9 new users per day or more than 280 months. There are presently users in 140 different countries. Several countries (including Brazil, Canada, Germany, the Netherlands and Mexico) have adopted LOINC as a national standard, and there are large-scale health information exchanges using LOINC in Spain, Singapore and Korea as well. There are currently efforts underway in 18 countries to translate LOINC into 13 languages. Within the USA, LOINC has been adopted by many large national reference laboratories, health information exchanges, healthcare organisations, insurance companies, research applications, and also by several national standards. Notably, the Department of Health and Human Services adopted LOINC as the standard across federal agencies for laboratory result names, laboratory test order names, and federally required patient assessment instruments. LOINC has long

been source vocabulary included in the National Library of Medicine's Unified Medical Language System. This past year, the HITECH Act of the ARRA stimulus bill authorised the Centres for Medicare and Medicaid Services (CMS) to give reimbursement incentives for eligible providers and hospitals that become 'meaningful users' of certified Electronic Health Record (EHR) technology, and subsequently LOINC was adopted as the standard for lab orders and results in these meaningful use and standards certification criteria (Health Information Technology Final Rule, 2010).

Structured collections of observations are one important area where we have focused recent development efforts. Within LOINC, we make a distinction between

- *panels* such as the 'complete blood count' or 'Braden scale', which are collections that have enumerated discrete contents
- *documents* such as a 'physical therapy evaluation note' or 'discharge summary', which are general information collections whose contents are not definitively enumerated (McDonald et al., 2010).

Our focus in this paper is on clinical (non-laboratory) panels and their contents. We use the term 'panel' in a general sense that encompasses survey instruments, questionnaires, standardised patient assessments, data collection sets and other kinds of 'forms'. For the purpose of this paper, we use 'variable' to refer to one of the items in a panel (which in some contexts may be thought of as a question or a data element), and will use 'answer' to refer to the result of a variable (which for quantitative variables would be a number, but for categorical variables may be thought of as an answer or choice). As a corollary, we use 'answer list' to mean the set of allowable answers, values, or choices for a particular variable.

Within our work on structured collections of variables, we have put a special emphasis on extending LOINC's representation of standardised patient assessment instruments. LOINC's goal in including assessment content is to provide a 'master question file' and uniform representation of the entire instrument's essential aspects. In this way, we could, for example, enable a depression severity score to be shared with the same exchange, storage and processing infrastructure as health information systems use for communicating the results of a complete blood count or set of vital signs.

The purpose of this paper is to describe LOINC's model for representing panels with the variables and answer lists they contain, highlight the scope of current coverage for clinical panels and to discuss some of the key lessons learned along the way.

2 Methods

2.1 Overview of LOINC

LOINC constructs 'fully specified' names according to an established model that contains six main axes (Table 1). The fully specified LOINC names contain sufficient information to distinguish among similar clinical observations, but do not carry all possible information about the testing procedure and result. Guided by the pragmatics of usual convention, tests and measures that have different columns on a clinical report or significantly different reference ranges are assigned separate LOINC names and codes.

Thus, different LOINC codes are assigned to observations that measure the same attribute but have different clinical meanings.

LOINC is distributed at no cost from the LOINC website (<http://loinc.org>) as a database table (available in several formats) containing the LOINC codes, names, and many additional attributes like synonyms, alternate names, example units and reference ranges, etc. New versions of the standard are published at least twice yearly (typically in June and December). In addition, Regenstrief develops and distributes at no cost a software program called RELMA (the *RE*genstrief *LOINC* Mapping Assistant) that contains functions for searching the LOINC database, reviewing the detailed accessory content, and for mapping local terms to LOINC.

Table 1 LOINC formal name model

<i>Axis</i>	<i>Name</i>	<i>Description/example</i>
1	Component	The analyte or attribute being measured or observed. E.g., potassium, haemoglobin
2	(Kind of) Property	Distinguishes among different kinds of quantities relating to the same substance. e.g., mass concentration, catalytic activity
3	Time (Aspect)	Identifies whether the measurement is made at a point in time or a time interval. e.g., 24 h for a urine sodium concentration
4	System	The sample, specimen, body system, patient, or other object of the observation. e.g., serum, urine, radial artery
5	(Type of) Scale	The scale or precision that distinguishes among observations that are quantitative, ordinal (ranked choices), nominal (unranked choices), or narrative
6	(Type of) Method	An optional axis that identifies the way the observation was produced. It is only used to distinguish observations that have clinically significant differences in interpretation when made by different methods

2.2 Representing panels in LOINC

We have built a robust model for representing panels in LOINC through iterative and collaborative development. The methods used in developing this model to cover the complexities of standardised assessment instruments have been described previously (Vreeman et al., 2010), so here we present summary of the model's key features. Because of the important psychometric properties of standardised assessments, LOINC's model captures not only the overall hierarchical organisation of the instrument (panel), but also many other additional attributes of each variable (often a question), such as the exact question text and answer list. In this way, the LOINC database not only serves as a master question file, but also provides a standardised representation of each instrument as a whole.

2.2.1 Hierarchical panel structure

LOINC represents panels, whether laboratory batteries or assessment instruments, by creating LOINC panel terms that are linked to an enumerated set of child elements. A complete hierarchical structure can be represented because the child elements themselves can be panel terms, which enables full nesting. For each collection of

variables that can be used as an independent package, we create a LOINC panel term and build its complete hierarchical structure. The fully specified LOINC names for panel terms are constructed according to the usual LOINC model, but typically have the name of the data set or assessment (or section header in the case of nested sets) in the Component, and a ‘-’ for the Property and Scale because the child elements vary in these attributes.

2.2.2 *Attributes of individual variables*

The main LOINC table contains the LOINC code, fully specified name, and fields for many other additional attributes about the terms. Table 2 presents a subset of these attributes that are important in representing the essential aspects of content from questionnaires, standardised assessments and other data sets. These LOINC term attributes are optionally populated where appropriate; some of the fields are used almost exclusively by terms from assessment instruments (e.g., question text, question source) whereas others are used by LOINC terms from many domains. Because many assessment instruments are copyrighted and made available under specific terms of use, the ability to identify and store the exact text of the external requirements was an important evolution of the LOINC data structure.

Table 2 LOINC term attributes important for variables in panels

<i>LOINC attribute</i>	<i>Description</i>
Question text	Exact text of survey question
Question source	Assessment name and question number
External copyright	External copyright notice and terms of use
Definition/description	Defining or describing narrative text
Example units	Example units of measure
HL7 field sub ID	HL7 message field where the content should be delivered (if Null, assume OBX-5)
HL7 v2 data type	HL7 version 2 data type
HL7 v3 data type	HL7 version 3 data type

2.2.3 *Answer lists*

The clinical meaning of many questions on assessment instruments is inextricably tied to the allowable answer options, and thus LOINC contains a data structure for linking LOINC observation codes to answer lists. Table 3 lists the key attributes about answer lists and their allowed answer options that are represented in the current LOINC model. For each answer list that has enumerated options stored in LOINC, Regenstrief generates a unique identifier for each answer option and an OID to identify the collection of answers into an answer list. For variables whose values may be drawn from a large external terminology such as the International Classification of Diseases or Current Procedural Terminology, those lists are not enumerated within LOINC. Rather, we indicate the presence of an external answer list with a flag and identify the code system and OID for the list.

Table 3 LOINC answer list and answer item attributes

<i>LOINC answer list attribute</i>	<i>Description</i>
LOINC answer list OID	Object identifier (OID) for the answer list as a collection
LOINC answer list external link	Link (e.g., URL) to external system that officially controls or provides additional information about this answer list
LOINC answer ID	LOINC-generated unique identifier for this answer item
LOINC answer string	The exact text of this answer item
LOINC answer sequence	Number indicating the position of this item in the answer list
LOINC answer local code	Local (original form) code for this answer item
LOINC answer score	Score value for this answer item if it is used in scoring algorithm
LOINC answer global code	Alternate identifier for this answer item from another standard terminology
LOINC answer global code system	Code system for alternate identifier (e.g., SNOMED CT or UMLS)

2.2.4 Panel-specific attributes of variables

The LOINC model for representing assessment content not only supports reuse of variables across panels, but also enables some attributes to be stored at the level of the instance of the variable within a particular panel. This feature allows these non-defining attributes (e.g., local code, help text, branching logic, etc.) to vary for the same LOINC code used in different panels. A sample of these panel-specific item attributes is listed in Table 4. We also use this mechanism to handle the circumstances where the same clinical observation has different labels across instruments, e.g., ‘Body Mass Index’, ‘BMI’ and ‘Body Mass Index (BMI)’, in a Display Name Override field.

Table 4 Panel-specific LOINC term attributes

<i>LOINC attribute</i>	<i>Description</i>
Display name override	Display name for item in this panel
Cardinality	Allowable number of repetitions for item
Observation ID in form	Local code or identifier for the item
Skip logic	Narrative text of branching logic
Data type in form	Panel-specific data type
Answer sequence override	Override of default answer sequence
Consistency checks	Validation rules for item
Relevance equation	Equation for determining the relevance of the item in this panel
Coding instructions	Directions for answering this item

2.3 Special export of panel content in LOINC distribution

All of the panel content in LOINC (both laboratory and clinical) is made available at no cost in the standard LOINC release formats and within RELMA. Additionally, beginning with LOINC version 2.26 (January 2009), the contents of many panels have also been released in a special export format as a separate download. This export is distributed in a

spreadsheet that includes three separate worksheets for the three tables defining the full panel construct: one for the hierarchical structure and panel-specific attributes, one for the LOINC concepts and associated variable-level attributes, and another that defines the answer list associated with each concept (where appropriate).

3 Results

3.1 *A growing universal instrument and item bank*

With the iterative refinements made to the LOINC model for representing panels, we have successfully represented a wide variety of content. Over time, we have continued to add new content to LOINC, including many patient assessments. Table 5 lists the assessment instruments that are available in the structured export format of the current LOINC release (version 2.34, December 2010). This export contains more than 42,000 terms from 58 different panels. The LOINC model has been successfully used to represent collections that are patient-reported (e.g., howRU) and clinician-observed (e.g., Morse Fall Scale), clinically focused (e.g., Confusion Assessment Method) and administratively focused (e.g., Nursing Management Minimum Data Set). We have put special effort into representing the instruments required for payment by the US Federal Government for assessing functioning and disability in post-acute care settings. LOINC now includes the full representation of the Minimum Data Set (MDS) version 2 and version 3 (used in skilled nursing facilities), the Outcome and Assessment Information Set (OASIS) version B1 and version C (used in home health settings), the Mental and Physical Residual Functional Capacity assessments (used by the Social Security Administration to support disability claims), and the Continuity Assessment Record and Evaluation (CARE) instrument that is being developed for use across all post-acute care settings. Figure 1 shows a representation of a sample item from the MDS version 3, and Figure 2 shows the display from RELMA for the corresponding LOINC term that illustrates some of the rich assessment content in LOINC.

Table 5 Assessments available in LOINC 2.34 structured export format

<i>Assessment name</i>
Brief Interview for Mental Status (BIMS)
Continuity Assessment Record and Evaluation (CARE)
Clinical Care Classification (CCC)
Confusion Assessment Method (CAM)
Geriatric Depression Scale (GDS)
Geriatric Depression Scale (GDS) – short version
HIV Signs and Symptoms (SSC) Checklist
howRU
Living with HIV (LIV-HIV)
Mental Residual Functional Capacity (RFC) Assessment Form
Minimum Data Set (MDS) version 2
Minimum Data Set (MDS) version 3

Table 5 Assessments available in LOINC 2.34 structured export format (continued)

Assessment name
Morse Fall Scale
Nursing Management Minimum Data Set (NMMDS)
Omaha System
Outcome and Assessment Information Set (OASIS) – B1
Outcome and Assessment Information Set (OASIS) – C
Patient Health Questionnaire (PHQ) – 9
Patient Health Questionnaire (PHQ) – 2
Patient Reported Outcomes Management Information System (PROMIS)
Phenotypes and eXposures Measures (PhenX)
Physical Residual Functional Capacity (RFC) Assessment Form
Quality Audit Marker (QAM)
Test of Infant Motor Performance (TIMP)
US Surgeon General Family Health Portrait

Figure 1 Original item E1100 from MDS version 3 form

Figure 2 RELMA details view (partial screenshot) of the LOINC term for item E1100 from MDS version 3

54696-0		Change in behavioral or other symptoms in last 7 days MDSv3				
NAME						
Fully-Specified Name:	Component	Property	Time Aspect	System	Scale	Method
	Change in behavioral or other symptoms in last 7D	Fund	Pt	*Patient	Ord	MDSv3
STATUS						
Active.						
Last updated: 2009/01/20.						
OBSERVATION ID IN FORM						
E1100						
FORM CODING INSTRUCTIONS						
Consider all of the symptoms assessed in items E0100 through E1000.						
BASIC ATTRIBUTES						
Class/Type:	SURVEY.MDS/Survey					
NORMATIVE ANSWER LIST:						
Source:	Change in behavioral or other symptoms					
Code System OID:	1.3.6.1.4.1.12009.10.1.34					
SEQ#	Answer	Code	Answer ID			
1	Same	0	LA11009-0			
2	Improved	1	LA65-8			
3	Worse	2	LA11011-6			
4	N/A because no prior MDS assessment	3	LA11012-4			
SURVEY QUESTION:						
Text:	Change in behavioral or other symptoms - How does resident's current behavior status, care rejection, or wandering compare to prior assessment (OBRA or PPS)?					
Source:	MDSv3.E1100					
MEMBER OF THESE PANELS						
54800-6	Minimum Data Set - version 3.0: -: Pt: *Patient: -:					

The LOINC model accommodates panels with categorical variables that have enumerated answer lists as well as other clinical variables that report physical quantities, like height, weight, or systolic blood pressure using the typical LOINC terms. In addition to the content available in the structured export, LOINC also includes several other standardised collections of variables in the same data model. For example, LOINC includes the full set of variables for the standard HIV care and antiretroviral therapy specified by the World

Health Organization for patient monitoring (World Health Organization, 2011), various health tracking data sets for use by consumers in patient health records, the National Centre for Injury Prevention and Control's Data Elements for Emergency Department Systems (National Center for Injury Prevention and Control, 2011), the Medical Event Reporting System – Total Health System (MERS – International, 2011), the Pathology Laboratory Electronic Reporting (Volume V) and Data Standards and Data Dictionary (Volume II) standards published by the North American Association of Central Cancer Registries (North American Association of Central Cancer Registries, 2011), and others.

Regenstrief is also creating LOINC content in collaboration with developers of two innovative clinical research variable sets: the *Phenotypes* and *eXposures* (PhenX) measures and the Patient-Reported Outcomes Measurement Information System (PROMIS). PhenX (PhenX, 2011) is funded by the National Human Genome Research Institute to develop and distribute a set of high-priority measures that will enable cross-study comparisons and analyses in genome-wide association and other clinical studies. PROMIS (PROMIS, 2011) is funded by the National Institutes of Health Roadmap for Medical Research Initiative to develop publicly available computer-adaptive tests for measuring patient-reported symptoms such as pain, fatigue, physical functioning and other aspects of health-related quality of life across a wide variety of chronic conditions. The current LOINC release (version 2.34, December 2010) contains a representation of four PhenX domains (360 terms) and all of the items in the PROMIS version 1.0 item banks (660 terms organised into 21 item banks and 21 short forms). Representing these variables in LOINC will promote data sharing across settings by integrating the wide spectrum of patient observations from laboratory tests to research assessments into a unified standard.

By collecting the details about individual variables and the panels that contain them, LOINC makes it easy for system implementers to access the content in a common format. The Personal Health Record being developed by the National Library of Medicine is an early example of a system that has the capability to read the LOINC panel definition and dynamically create electronic data collection forms (Lister Hill National Center for Biomedical Communications – US National Library of Medicine – National Institutes of Health, 2011). Having a standard for patient observations of all kinds also makes it possible to construct interoperable electronic result messages that blend routine clinical data with results from formal research questionnaires. Furthermore, LOINC's no-cost worldwide distribution keeps the barriers to adoption very low.

3.2 Enabling interoperability together with other health information technology standards

LOINC's standardised representation of assessment content is an important enabling component of interoperable exchange between electronic systems and has been adopted by several large US initiatives. The National Committee on Vital and Health Statistics endorsed LOINC's model for assessments based on the recommendations of the Consolidated Health Informatics workgroup on Functioning and Disability (National Committee on Vital and Health Statistics, 2011). These recommendations adopt LOINC as the standard for federally required

- questions and answers
- assessment forms that include functioning and disability content.

Additionally, the LOINC model was incorporated into the HL7 Draft Standard for Trial Use ‘CDA Framework for Questionnaire Assessments and CDA Representation of the Minimum Data Set Questionnaire Assessment’ (Health Level Seven International, 2011). HL7’s questionnaire assessment draft standard filled an important gap by providing an implementation guide for patient assessments. This standard includes both an internationally applicable component that supports exchange of any assessment represented in LOINC and a detailed guide for implementing the exchange of the MDS version 3 that is required for use in nursing homes (effective October 2010) in the USA by the Centres for Medicare and Medicaid Services (CMS) (Centers for Medicare & Medicaid Services, 2011). The Health Information Technology Standards Panel, a cooperative partnership advancing interoperability in support of clinical care and public health, incorporated the HL7 draft standard with its support for the LOINC assessment model into the C83 CDA Content Modules Component (Health Information Technology Standards Panel, 2011).

In addition to patient assessments, LOINC’s model for representing variables and their answer lists has been adopted in other contexts as well. We previously mentioned adoption of LOINC in the cancer registration standards produced by the North American Association of Central Cancer Registries and the Data Elements for Emergency Department Systems developed by the National Centre for Injury Prevention and Control. The National Immunisation Program of the Centres for Disease Control (CDC) has adopted LOINC as standard identifiers for all the variables related to immunisation scheduling and forecasting. The recently published ‘Implementation Guide for Immunisation Messaging Release 1.0’ containing these LOINC codes has been adopted as part of the Standards and Certification Criteria that support the achievement of meaningful use Stage 1 by eligible professionals and eligible hospitals under the Medicare and Medicaid EHR incentive programme (US Department of Health and Human Services, 2011). Similarly, the developers of the Nursing Management Minimum Data Set (NMMDS) worked with Regenstrief and the LOINC Committee to represent all of the NMMDS variables and associated answer lists in LOINC (Westra et al., 2010). The NMMDS has been recognised by the American Nurses Association, and provides a minimum set of essential standardised management data to support nursing management and administrative decisions for quality improvement.

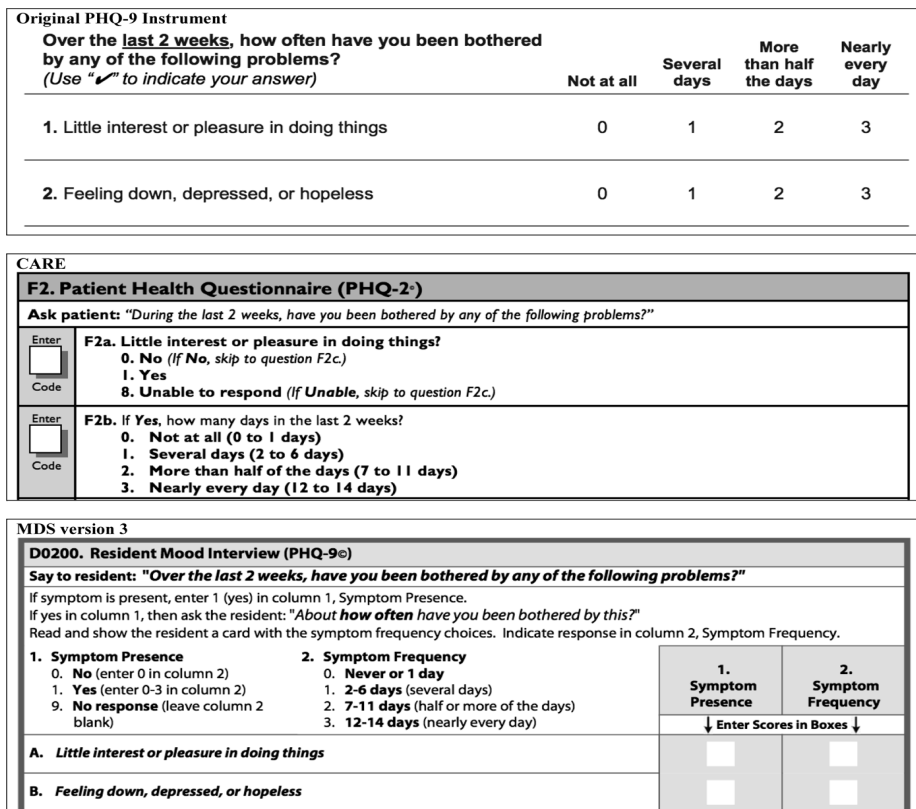
3.3 Variation across panels

When adding the content from these assessments to LOINC, we found substantially more variation across panels than we had initially expected, and some of it could have been avoided. Many of the variables in the OASIS, MDS and CARE instruments are very similar, but not directly comparable. For example, although many of the variables from MDS version 3 were similar to those in CARE, the look-back reference period differs (seven days vs. two days). The lack of direct comparability is also present between different versions of the same instrument. Common changes we observed between instruments included considerably modifying the question wording, adding or removing answers from the answer list, as well as adding or removing whole variables from the set. To illustrate, consider the MDS version 3, which we modelled in LOINC

after representing the OASIS-B1, MDS version 2 and CARE instruments. Of the 710 variables in the MDS version 3, 72 LOINC terms were reused from MDS version 2, 13 LOINC terms were reused from the CARE instrument and four LOINC terms (height, weight, birth date and discharge date) already existed in LOINC from other sources.

Moreover, some of the differences we observed might have been prevented. For example, both CARE and MDS version 3 include two items from the PHQ (Kroenke et al., 2001), which is a standardised, validated and copyrighted instrument. Figure 3 shows the different representation of these items between the three instruments. CARE and MDS version 3 differ from the original PHQ by breaking each question into two responses, and differ from each other in their answer lists. Likewise, the MDS version 2, OASIS-B1, MDS version 3 and CARE instruments all ask clinicians to record the number of pressure ulcers that a patient has at a given stage. Table 6 shows the different coding instructions given on these four instruments. As a final example, consider the commonly assessed attribute of pain frequency. Table 7 shows the variations in the questions and associated answer lists about pain frequency among the MDS version 2, MDS version 3, CARE, OASIS-B1 and OASIS-C.

Figure 3 Variations in questions from the PHQ in the original instrument, CARE, and MDS version 3



Developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke and colleagues, with an educational grant from Pfizer Inc. No permission required to reproduce, translate, display or distribute.

Table 6 Variations in variables about number of pressure ulcers at a given stage from MDS version 2, OASIS B1, CARE, and MDS version 3

<i>Assessment instrument</i>	<i>Coding instructions</i>
MDS version 2	Code 9 = 9 or more
OASIS-B1	Code 4 = 4 or more
MDS version 3	N/A
CARE	Code 8 = 8 or more ulcers, 9 = 'Unknown'

Table 7 Variations in variables and answer lists for pain frequency from MDS version 2, MDS version 3, CARE, OASIS-B1, and OASIS-C

<i>Assessment instrument</i>	<i>Question stem</i>	<i>Answer list</i>
MDS version 2	Frequency with which resident complains or shows evidence of pain (in last 7 days)	No pain, pain less than daily, pain daily
MDS version 3	How much of the time have you experienced pain or hurting over the last 5 days	Almost constantly, frequently, occasionally, rarely, unable to answer
CARE	Have you had pain or hurting at any time during the last 2 days	Yes, no, unable to respond
OASIS-B1	Frequency of Pain Interfering with patient's activity or movement	Patient has no pain or pain does not interfere with activity or movement, less often than daily, daily but not constantly, all of the time
OASIS-C	Frequency of Pain Interfering with patient's activity or movement	Patient has no pain, patient has pain that does not interfere with activity or movement, less often than daily, daily but not constantly, all of the time

4 Discussion

LOINC contains a well-developed model for representing variables, answer lists and the collections that contain them. With continued growth, LOINC is expanding as a large 'master observation file' that provides a uniform representation of the essential attributes for items in data collection forms. The level of standardisation achieved by modelling this content in LOINC provides an important component of enabling interoperable data exchange, storage and processing. By creating a uniform representation and distributing it worldwide at no cost, LOINC aims to lower the barriers to interoperability among systems and make this valuable data available across settings when and where it is needed.

Many promising opportunities exist for continuing to expand the rich content already present in LOINC. The CDC has several ongoing initiatives that are also adopting LOINC as the standard for variables, including the National Health and Nutrition Examination Survey that includes examinations and interviews of about 5000 nationally representative participants (Centers for Disease Control and Prevention, 2011; Bonander and Gates, 2010) and the Case Reporting Standardisation workgroup (Case Reporting Standardization Workgroup, 2011) that is harmonising the variables used in case

reporting of national notifiable conditions. In addition, Regenstrief is also engaging in early conversations with the American Psychological Association and American Physical Therapy Association about including widely used instruments for assessing mental health and movement impairments in LOINC.

LOINC's inclusion of assessments aims to achieve a convergence of codes and vocabulary for observations by providing a uniform and standardised representation. This approach complements the current efforts to build metadata repositories and other clinical information models by providing the *lingua franca* that can populate the models and be used for exchanging data between and among clinical and research systems. One such metadata repository is the National Cancer Institute's Cancer Data Standards Registry and Repository (caDSR), which is a database and a set of Application Programming Interfaces and tools to create, edit, control, deploy and find common variables (Covitz et al., 2003). The names, definitions, answer lists and other variable attributes from LOINC could be used to populate the metadata in caDSR. Similarly, LOINC's assessment content has already been represented in the data model of the CEN/ISO 13606 standard, which makes them usable and editable in this archetype format by software tools with features like GELLO code to automatically calculate anion gap, or automatic generation of an HL7 v2 message (Medical Objects, 2011). The Clinical Data Interchange Standards Consortium (CDISC) has produced a study data tabulation model standard (CDISC, 2008a, 2008b) for reporting data sets to regulatory authorities that supports and recommends LOINC as the universal identifier for observations. CDISC has also developed the Clinical Data Acquisition Standards Harmonisation (CDASH) for enabling more efficient data collection, including a structure for grouping questions into collections, specifying the exact question text, and listing coded response values (CDISC, 2008a, 2008b). Although the CDASH specification does not currently contain a domain neutral structure for representing the full content of assessment instruments, it seems feasible that the panel content in LOINC could also be imported into this structure for the domains that are covered. The ability to insert LOINC into other data models makes available a wider range of tools and services for implementers. Clinical study data management systems such as TrialDB (Brandt et al., 2003) and REDCap (Harris et al., 2009) are one such type of application that we believe could also leverage LOINC's universal identifiers and complete representation of clinical variables and assessment content to more easily exchange data between clinical and research systems.

4.1 Lessons and recommendations

To inform future work in the informatics of metadata, questions and answer lists, we have synthesised the observations made in developing the LOINC representation of this diverse panels into a set of recommendations and lessons learned.

4.1.1 Variation abounds and limits comparability

As we modelled various assessment instruments in LOINC, we were struck by the degree of variation among observations measuring similar clinical characteristics. In some cases, there may be good justification for making entirely new instruments or considerably modifying the questions of an existing instrument. Indeed, many of these variations were intentional choices of the assessment developers, but we also noticed other differences

that seemed arbitrary and might have been avoided. The lack of comparability between the assessment instruments required for payment by CMS in post-acute care settings creates obstacles for caring for often-fragile patients; the information on one assessment cannot be used to directly populate another.

We urge clinical researchers and other potential data set developers to look closely at existing collections and variables. Before inventing yet another variant, the possible benefits should be weighed against the loss of data comparability. The larger the amount and generalisability of the existing data, the more carefully we should consider any potential modifications. Having a large collection of panels and variables in LOINC's uniform format should make it easier to review and reuse the existing content. Brandt et al. (2004) have described a set of approaches and informatics tools that can be used with such a master collection to assist researchers in integrating disparate research questionnaires.

Furthermore, collaboration between data collection developers and standards developers could smooth the process. Our starting point for building the LOINC model of most assessment instruments was typically a paper form, although some had their own unique software programs and data structures. In the journey to transform the content into LOINC's uniform model, we were forced to reconcile many potential discrepancies and ambiguities that might have been clarified if the uniform data model was a component of the initial conceptualisation. Some of these issues included:

- How were the answers of 'unknown', 'undetermined', or 'unable to answer' represented
- For variables with a possible answer of 'other specified', how was that answer and the blank line value stored
- Are units of measure implied for any of the variables and if so, how is it presented to the user
- Which text on the form is really the variable or 'question' and which is just supplementary (and perhaps could be presented differently to the user).

Some of the large differences we observed in question style impact both the user experience and how that data could be stored in an electronic record. For example, some instruments asked users to specify yes, no, or unknown to a very long list of potential diseases whereas others instructed them just to list the active ones. Starting with the LOINC model and an eye towards exchanging the results with widely adopted messaging standards like HL7 may help elucidate some of these latent challenges. The best practice recommendations in the CDASH (CDISC, 2008) are an important step in this regard, and could be complemented with the content and uniform representation of LOINC.

4.1.2 Intellectual property restrictions can be large barriers

Prior to being able to include a copyrighted instrument in LOINC, Regenstrief must negotiate separate (often resource consuming) agreements with each copyright holder. The LOINC structure allows us to provide a copy of the terms-of-use, attribution, descriptions, links to papers and reference material and other notices. Many copyright owners require attribution and specification of the terms of use that protect against changing the variables, which are sensible. However, other owners limit use in difficult

ways like requiring royalties for each use. These restrictions present a large barrier to widespread interoperable exchange of their results, and may even be unknown to most users (Powsner and Powsner, 2005). Consistent with LOINC's overall distribution aims of free worldwide use, we have included content (with permission or under applicable terms of use) that allows the content to be used and distributed at no cost for clinical, administrative and research purposes. We strongly recommend that organisations that fund development of standardised data collections (especially validated ones) require that they be made available with unrestrictive licences.

4.1.3 A master catalogue and uniform representation is a step forward

Building a master catalogue of panels and variables in LOINC is an enabling step towards interoperable data exchange, but much work remains. Many opportunities remain for expanding the content represented in LOINC to other domains. The uniform format that LOINC's model provides should make it more efficient to build data collection interfaces and processing components that in turn make it easier to collect and manipulate these data. The instruments that can be administered directly to patients may be of special importance because they limit the amount of data entry time required of clinicians. Furthermore, representing the variables from various collections in the same standard vocabulary as laboratory and other clinical measurements encourages their use in other potentially beneficial health information technology applications such as clinical decision support and quality reporting systems. We have lamented the amount of variation in the variables we modelled, but it is difficult to know which differences are meaningful without empirical analysis. Such analyses will be easier to conduct if the data can be pooled by a common exchange infrastructure. And finally, future valuable work would be to develop efficiency-gaining processes like automatically populating a standardised form based on the existing data from an EHR.

5 Conclusion

LOINC contains a well-developed model for representing variables, answer lists and the collections that contain them. We will continue adding high-priority new content as part of LOINC's open development process. By creating a uniform representation and distributing it worldwide at no cost, LOINC aims to lower the barriers to interoperability among systems and make this valuable data available across settings when and where it is needed.

Acknowledgements

This work was performed at the Regenstrief Institute, Indianapolis, IN, and was supported in part by contracts HHSN276200800006C from the National Library of Medicine, FORE-ASPE-2007-5 from the AHIMA Foundation, 0-312-0209853 from RTI International, and 3UL1RR025761-02S6 from the National Center for Research Resources.

References

- Bonander, J. and Gates, S. (2010) 'Public health in an era of personal health records: opportunities for innovation and new partnerships', *Journal of Medical Internet Research*, Vol. 12, No. 3, Available at: <http://www.jmir.org/2010/3/e33/> [Accessed 18 March, 2011].
- Brandt, C.A., Deshpande, A.M., Lu, C., Ananth, G., Sun, K., Gadagkar, R., Morse, R., Rodriguez, C., Miller, P.L. and Nadkarni, P.M. (2003) 'Center for medical informatics. TrialDB: a web-based clinical study data management system', *AMIA Annu Symp Proc. 2003*, Washington DC, 794. PubMed PMID: 14728299; PubMed Central PMCID: PMC1480035.
- Brandt, C.A., Cohen, D.B., Shifman, M.A., Miller, P.L., Nadkarni, P.M. and Frawley, S.J. (2004) 'Approaches and informatics tools to assist in the integration of similar clinical research questionnaires', *Methods of Information in Medicine*, Vol. 43, No. 2, pp.156–162.
- Case Reporting Standardization Workgroup (2011) Public Health Case Reporting – phConnect, Available at: <http://www.phconnect.org/group/phcasereporting> [Accessed 18 March, 2011].
- Centers for Disease Control and Prevention (2011) *NHANES – National Health and Nutrition Examination Survey Homepage*, Available at: <http://cdc.gov/nchs/nhanes.htm> [Accessed 18 March, 2011].
- Centers for Medicare & Medicaid Services (2011) *MDS 3.0 for Nursing Homes and Swing Bed Providers*, Available at: http://www.cms.gov/nursinghomequalityinits/25_nhqimds30.asp [Accessed 18 March, 2011].
- Clinical Data Interchange Standards Consortium (CDISC) (2008a) *Study Data Tabulation Model Version 1.2*, Available at: <http://www.cdisc.org/sdtm> [Accessed 18 March, 2011].
- Clinical Data Interchange Standards Consortium (CDISC) (2008b) *Clinical Data Acquisition Standards Harmonization (CDASH) Standard Version 1.0*, Available at: <http://www.cdisc.org/cdash> [Accessed 18 March, 2011].
- Covitz, P.A., Hartel, F., Schaefer, C., De Coronado, S., Fragoso, G., Sahni, H., Gustafson, S. and Buetow, K.H. (2003) 'caCORE: a common infrastructure for cancer informatics', *Bioinformatics*, Vol. 19, No. 18, pp.2404–2412.
- Dugas, M., Thun, S., Frankewitsch, T. and Heitmann, K.U. (2009) 'LOINC codes for hospital information systems documents: a case study', *Journal of the American Medical Informatics Association: JAMIA*, Vol. 16, No. 3, pp.400–403.
- Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N. and Conde, J.G. (2009) 'Research electronic data capture (REDCap) – a metadata driven methodology and workflow process for providing translational research informatics support', *Journal of Biomedical Informatics*, Vol. 42, No. 2, pp.377–381.
- Health Information Technology Standards Panel (2011) *C83 – CDA Content Modules Component v2.0.1*, Available at: http://www.hitsp.org/ConstructSet_Details.aspx?&PrefixAlpha=4&PrefixNumeric=83 [Accessed 18 March, 2011].
- Health Information Technology (2010) 'Initial set of standards, implementation specifications, and certification criteria for electronic health record technology', *Federal Register*, Vol. 75, No. 144, p.44650.
- Health Level Seven International (2011) *HL7 DSTU Comments*, Available at: <http://www.hl7.org/dstucomments/> [Accessed 18 March, 2011].
- Huff, S.M., Rocha, R.A., McDonald, C.J., De Moor, G.J.E., Fiers, T., Bidgood Jr., W.D., Forrey, A.W., Francis, W.G., Tracy, W.R., Leavelle, D., Stalling, F., Griffin, B., Maloney, P., Leland, D., Charles, L., Hutchins, K. and Baenziger, J. (1998) 'Development of the logical observation identifier names and codes (LOINC) vocabulary', *Journal of the American Medical Informatics Association: JAMIA*, Vol. 5, No. 3, pp.276–292.
- Hyun, S., Shapiro, J.S., Melton, G., Schlegel, C., Stetson, P.D., Johnson, S.B. and Bakken, S. (2009) 'Iterative evaluation of the health level 7–Logical observation identifiers names and codes clinical document ontology for representing clinical document names: a case report', *Journal of the American Medical Informatics Association: JAMIA*, Vol. 16, No. 3, pp.395–399.

- Kroenke, K., Spitzer, R.L. and Williams, J.B. (2001) 'The PHQ-9: validity of a brief depression severity measure', *Journal of General Internal Medicine*, Vol. 16, No. 9, pp.606–613.
- Lister Hill National Center for Biomedical Communications – US National Library of Medicine – National Institutes of Health (2011) *The NLM Personal Health Record (PHR)*, Available at: <http://tinyurl.com/NLMPHROverview> [Accessed 18 March, 2011].
- McDonald, C.J. (1997) 'The barriers to electronic medical record systems and how to overcome them', *Journal of the American Medical Informatics Association: JAMIA*, Vol. 4, No. 3, pp.213–221.
- McDonald, C.J., Huff, S., Mercer, K., Hernandez, J.A. and Vreeman, D.J. (2010) *Logical Observation Identifiers Names and Codes (LOINC®) User's Guide*, Available at: <http://loinc.org/downloads/files/LOINCManual.pdf> [Accessed 18 March, 2011].
- McDonald, C.J., Huff, S.M., Suico, J.G., Hill, G., Leavelle, D., Aller, R., Forrey, A., Mercer, K., DeMoor, G., Hook, J., Williams, W., Case, J. and Maloney, P. (2003) 'LOINC, a universal standard for identifying laboratory observations: a 5-year update', *Clinical Chemistry*, Vol. 49, No. 4, pp.624–633.
- Medical Objects (2011) *Templates*, Available at: <http://wiki.medical-objects.com.au/index.php/TEMPLATES> [Accessed 18 March, 2011].
- MERS – International (2011) *Medical Event Reporting System – Total Health System*, Available at: http://www.mers-international.com/mers_usa.html [Accessed 18 March, 2011].
- National Center for Injury Prevention and Control (2011) *DEEDS Publication – National Center for Injury Prevention and Control*, Available at: <http://www.cdc.gov/ncipc/pub-res/deedspage.htm> [Accessed 18 March, 2011].
- National Committee on Vital and Health Statistics (2011) *Consolidated Health Informatics Standards Adoption Recommendation: Functioning and Disability*, Available at: <http://www.ncvhs.hhs.gov/061128lt.pdf> [Accessed 18 March, 2011].
- North American Association of Central Cancer Registries (2011) Welcome to The North American Association of Central Cancer Registries (NAACCR, Inc.), Available at: <http://www.naacr.org/Home.aspx> [Accessed 18 March, 2011].
- PhenX (2011) *PhenX Home Page*, Available at: <https://www.phenx.org> [Accessed 18 March, 2011].
- Powsner, S. and Powsner, D. (2005) 'Cognition, copyright, and the classroom', *The American Journal of Psychiatry*, Vol. 162, No. 3, pp.627–628.
- PROMIS (2011) *PROMIS Home Page*, Available at: <http://www.nihpromis.org/default.aspx> [Accessed 18 March, 2011].
- Smith, P.C., Araya-guerra, R., Bublitz, C., Parnes, B., Dickenson, L.M., Van Vorst, R., Westfall, J.M. and Pace, W.D. (2005) 'Missing clinical information during primary care visits', *JAMA: The Journal of the American Medical Association*, Vol. 293, No. 5, pp.565–571.
- US Department of Health and Human Services (2011) *Standards and Certification Criteria Final Rule*, Available at: <http://tinyurl.com/standardscertfinalrule> [Accessed 18 March, 2011].
- Vreeman, D.J., Finnell, J.T. and Overhage, J.M. (2007) 'A rationale for parsimonious laboratory term mapping by frequency', *AMIA Annual Symposium Proceedings/AMIA Symposium, AMIA Symposium*, Chicago, IL, pp.771–775.
- Vreeman, D.J. and McDonald, C.J. (2006) 'A comparison of intelligent mapper and document similarity scores for mapping local radiology terms to LOINC', *AMIA Annual Symposium, Proceedings/AMIA Symposium, AMIA Symposium*, Washington DC, pp.809–813.
- Vreeman, D.J. and McDonald, C.J. (2005) 'Automated mapping of local radiology terms to LOINC', *AMIA Annual Symposium Proceedings/AMIA Symposium, AMIA Symposium*, Austin, TX, pp.769–773.
- Vreeman, D.J., McDonald, C.J. and Huff, S.M. (2010) 'Representing patient assessments in LOINC®', *AMIA Annual Symposium Proceedings/AMIA Symposium, AMIA Symposium*, Washington DC, pp.832–836.

- van Walraven, C., Taljaard, M., Bell, C.M., Etchells, E., Zarnke, K.B., Stiell, I.G. and Forster, A.J. (2008) 'Information exchange among physicians caring for the same patient in the community', *CMAJ: Canadian Medical Association Journal*, Vol. 179, No. 10, pp.1013–1018.
- Westra, B.L., Subramanian, A., Hart, C., Matney, S.A., Wilson, P.S., Huff, S.M., Huber, D.L. and Delaney, C.W. (2010) 'Achieving 'Meaningful Use' of electronic health records through the integration of the nursing management minimum data set', *The Journal of Nursing Administration*, Vol. 40, Nos. 7–8, pp.336–343.
- World Health Organization (2011) *WHO – Patient Monitoring Guidelines for HIV Care and Antiretroviral Therapy*, Available at: <http://www.who.int/hiv/pub/imai/patientguide/en/> [Accessed 18 March, 2011].