Recommendations for the Submission of LOINC® Codes in Regulatory Applications to the U.S. Food and Drug Administration

LOINC Working Group:

U.S. Food and Drug Administration (FDA), U.S. National Institutes of Health (NIH), Clinical Data Interchange Standards Consortium (CDISC), and Regenstrief Institute

November 2017

Table of Contents

1 LOINC Working Group (WG) Members and Affiliations

Member Name	Professional Affiliation
Swapna Abhyankar, M.D.	Regenstrief Institute/LOINC
Lauren Becnel, Ph.D.	CDISC
Boris Brodsky	FDA
Ron Fitzmartin, Ph.D.	FDA
Margaret Haber	NCI
Virginia Hussong	FDA
Clement McDonald, Ph.D.	NLM
Erin Muhlbradt, Ph.D.	NCI
Eileen Navarro, M.D.	FDA
Vaishali Popat, M.D.	FDA
Colleen Ratliffe	FDA
Lilliam Rosario, Ph.D.	FDA
Jessica Voqui, Pharm.D.	FDA
Daniel Vreeman, D.P.T.	Regenstrief Institute/LOINC
Alan Williams, Ph.D.	FDA
Lawrence Wright, Ph.D.	NCI

2 Background

2.1 Purpose of this Recommendation Document

Regulated clinical research is carried out and supported across multiple sectors including academic medical centers, biopharmaceutical companies, biotechnology companies, clinical research organizations, federal agencies, information technology vendors and others. The FDA will soon begin requiring the inclusion of a medical laboratory and observation terminology coding system, Logical Observation Identifiers Names and Codes (LOINC®). While this requirement will help align semantics between healthcare and regulated clinical research, readiness for this requirement and understanding of LOINC can vary dramatically across these sectors.

In January 2017, the FDA convened a Committee, including FDA, NIH, CDISC, and Regenstrief Institute, to evaluate the use of terminology standards, with a particular focus on LOINC. The Committee formed a LOINC Working Group to: (1) review the relationship between the CDISC laboratory codes and the Regenstrief LOINC laboratory codes, (2) address the concerns raised by the CDISC community in the 2017 survey, and (3) develop recommendations for the submission of LOINC within the CDISC data exchange format in applications to the FDA.

This document, prepared by CDISC with inputs from the Working Group members, serves to provide the community with information about LOINC and to make recommendations to the FDA on the initial required scope.

2.2 Introduction

Standards for biomedical research data exchange and terminology support data retrieval, system interoperability and research reproducibility by providing common structures and shared meaning for the collection, aggregation, analysis and exchange of clinical and non-clinical research data between or among electronic systems. Per the FDA website's Study Data Resource page, "Having standard, uniform study data enables FDA scientists to explore many new research questions by combining data from multiple studies. Data standards also help FDA receive, process, review, and archive submissions more efficiently and effectively."

The FDA Data Standards Catalog explicitly states which standards the FDA's Center for Biologics Evaluation and Research (CBER), and, Center for Drug Evaluation and Research (CDER), accept and/or require for regulatory submissions, alongside effective dates for each data exchange or terminology standard. Standards referenced in the FDA catalog include the CDISC Study Data Tabulation Model (SDTM) data exchange/tabulation standard and LOINC, a terminology standard maintained by the Regenstrief Institute. LOINC is a freely available international standard for identifying health measurements, observations, and documents that is ubiquitous in health data systems worldwide and is an essential ingredient of system interoperability. Presently, LOINC is utilized in more than 172 countries and has been adopted as a national standard in nearly 30 countries, including the United States.

The FDA Data Standards Catalog currently specifies CDISC data exchange standards for the submission of clinical study datasets to CBER and CDER. It also includes the requirement for LOINC codes to populate the "LBLOINC" variable in the laboratory (LB) domain specification when study data are submitted to the FDA using CDISC data exchange standards. This requirement applies to new drug applications (NDAs), abbreviated new drug application for generics (ANDAs), biologics license applications (BLAs), and e-submissions of some investigational new drugs (INDs) in accordance with section II.A of the FDA guidance document "Providing Regulatory Submissions in Electronic Format – Standardized Study Data"[1]. The latest version of the FDA Data Standards Catalog at the time of writing this recommendations document is v4.5.2 (04-13-2017). The requirement for LOINC codes begins for studies that start after March 15, 2020 for NDAs, ANDAs and certain BLAs, and on March 15, 2021 for certain INDs.

2.3 CDISC Member Survey Regarding LOINC Readiness

When the FDA requirement for the use of LOINC was published, some CDISC members expressed concern about inclusion of LOINC in electronic submissions. CDISC launched a member survey in the first quarter of 2017 to better understand these concerns, their prevalence and member readiness. This survey is informational only, as it does not necessarily represent the full breadth of global clinical researchers' knowledge and use of LOINC.

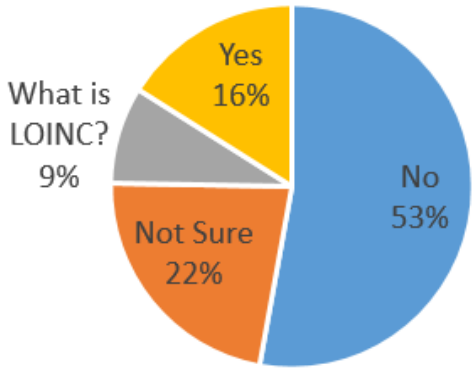

Figure 1. Organizations engaged in regulated research that are currently collecting LOINC codes. N = 180.

Data were collected from 180 respondents (Figure 1), which showed that 16% were currently collecting LOINC codes in at least some studies, whereas 84% were not. Nearly 10% of respondents reported never having heard of LOINC. For those that were collecting LOINC codes, they reported doing so for 15-50% of their total volume of trials and noted that LOINC codes were useful, when they were received, as they provided a more precise description of the collected assay. They also mentioned that local and academic laboratories were unable to provide LOINC codes in many cases.

Among the 85% of respondents who did not collect LOINC codes, some indicated that they had been requesting LOINC codes from central laboratories, but the codes had not been provided. Other concerns included realm specificity for sponsors engaged in global trials, availability of LOINC codes for complex new tests and from smaller laboratories, and inconsistency among different laboratories with regard to the application of LOINC codes. In general, the respondents were concerned about the lack of clear guidance with regard to FDA requirements for use of LOINC and the potential for each organization implementing bespoke solutions that may not truly meet regulatory need.

The FDA/Clinical Data Interchange Standards Consortium (CDISC)/United States National Institutes of Health (NIH)/Regenstrief Institute LOINC Working Group ("Working Group") was created to address some of the respondents' concerns and provide specific recommendations for the use of LOINC in clinical trial data submissions to the FDA.

2.4 Pre-Requirement Submissions of LOINC Codes to FDA

Currently, FDA supports, and in the near future will require, the use of LOINC codes in clinical trials. Until required, the Working Group encourages sponsors to voluntarily follow these recommendations and any others that the FDA may issue when submitting laboratory data to FDA.

3 Working Group Considerations

It is important to recognize that LOINC includes two major divisions, Laboratory and Clinical, and that the FDA's current support of LOINC for submitting clinical trial data to CBER and CDER, as identified in the FDA Data Standards Catalog, is limited to the laboratory division. The LOINC

Laboratory division generally aligns with the CDISC SDTM LB (laboratory) domain.

One of the primary differences between LOINC and the CDISC LB domain is that LOINC terms are pre-coordinated and CDISC LB concepts are post-coordinated. According to Stevens and Sattler, "A pre-coordinated [set of information or terminologies] has all the terms and relationships between them needed by an application; it is static; 'what you see is what you get'. A post-coordinated [set of information or terminologies] has the building blocks for the terms needed in an application such that they can be built as required, and so that their relationship can be determined as required; it is dynamic; the ontology is much more than 'what you see', since you can compose new expressions from the given building blocks.[3]" This means that in LOINC, a single concept represents the combination of several different aspects of a laboratory test, such as the analyte being measured, the timing (e.g., point in time versus 24-hour collection), and the specimen type, while the CDISC LB domain contains three separate concepts for the analyte, timing, and specimen type.

Pre-coordination and post-coordination both offer benefits. Pre-coordinated data support more rapid data entry and reuse. For example, it is far easier in busy clinical environments to enter a single piece of data that contains multiple data elements, rather than spend additional time entering each data element individually. LOINC fully specified names are pre-coordinated using a six part model that distinguishes among tests with clinically different meanings. The six parts, also called Dimensions, are the Component, Property, Timing, System, Scale, and Method. All axes other

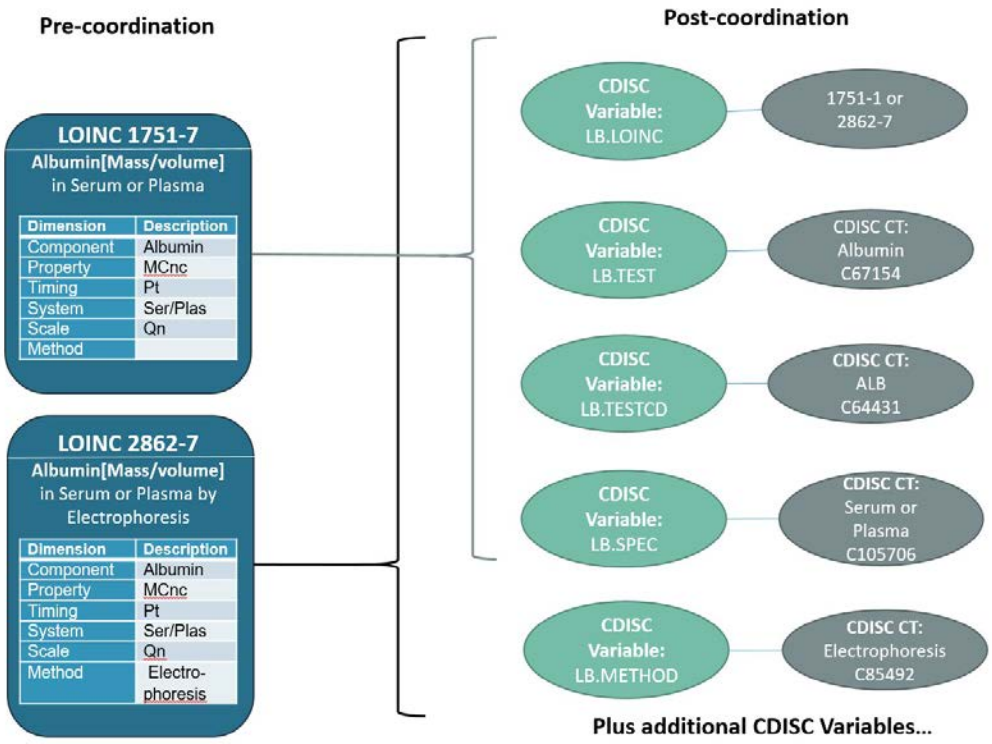

Figure 2. Comparing pre-coordinated LOINC and post-coordinated CDISC LB Variables for a laboratory test measuring Albumin in serum or plasma. Two pre- coordinated LOINC codes (blue boxes, left) are shown. Each LOINC code is a single entity comprised of six primary Parts. For Post-Coordinated CDISC LB variables (green, right) and associated controlled terminologies (grey, right) for the albumin test are shown with "C-code" identifiers from the National Cancer Institute's Enterprise Vocabulary Services.

than the Method have a value for every LOINC term. The Method may or may not have a value. LOINC's pre-coordinated names match the level of specificity of common laboratory test ordering and reporting paradigms where there are separate fields for indicating the test name, the test result value, units of measure, reference range, etc. As of version 2.61 (June 2017), the LOINC release now contains an artifact linking the pre-coordinated LOINC terms to the set of LOINC parts of which it is composed[3]. Having these codes and relationships facilitates secondary aggregation and analysis of primary data coded with LOINC.

On the other hand, CDISC is post-coordinated for ease of granular data analysis. Each aspect of the test (e.g., analyte, specimen, etc.) has a different identifier, which makes it easier to aggregate and analyze data based on the specific combination of concepts that are relevant to the research question. See Figure 2 for a comparison of pre-coordination and post-coordination, where multiple CDISC SDTM LB domain variables map to a single LOINC code. Related LOINC codes may require representation by different variables and CDISC controlled terms. Figure 2 includes two LOINC codes: one with data in all six major parts, 2862-7, and a less specified, "method-less" LOINC code, 1751-7. The related code, 2862-7, specifies albumin measurements were made by electrophoresis, which maps to CDISC's LB.METHOD variable. For the method-less LOINC, LB.METHOD cannot be mapped. Similarly, some LOINCs do not specify a specific System, which maps to LB.SPEC, and instead indicate that the System is 'XXX' to denote that it can be one of several locations or specimen types. End users should utilize the most specific LOINC possible, where known information matches the LOINC part values for that code, without over specifying by selecting a LOINC code that has one or more part values that are more granular than what is known for the result data (Figure 3). End users should not over specify LOINCs. If the method is not known, for example for a laboratory test as shown in Figure 3, mapping to LOINC 2862-7 would be over specifying because it includes a defined method, and LOINC 1751-7 would be a more appropriate match because it is method-less.

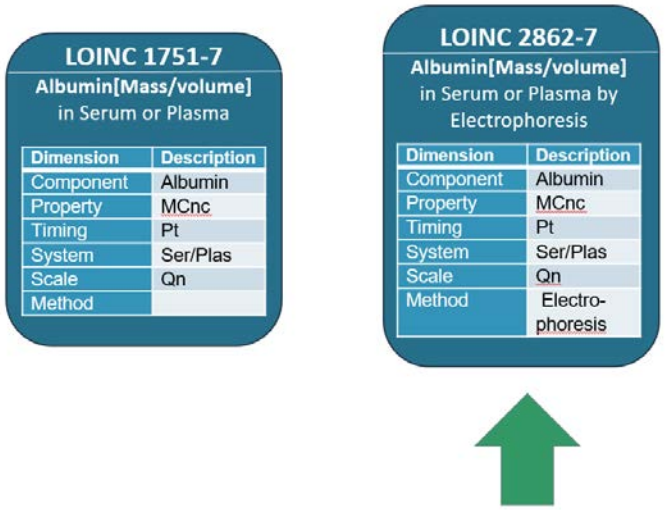

Figure 3. Selecting the most specific LOINC available without over specifying. For an albumin test from either serum or plasma where the analyte was measured by the electrophoresis method and is reported in g/dL, LOINC 2862-7 is appropriate as indicated by a green arrow.

A mapping between the six major LOINC parts and CDISC LB variables is provided in Table 1. Harmonizing the CDISC lab domain variable values with the LOINC Part values represents a substantial volume of work. In the absence of clear use cases for such harmonization, the Working

Group declared it to be out of scope for this initial set of recommendations, but it may be considered in future work.

Table 1. A comparison of LOINC Parts to CDISC Laboratory (LB) Domain Variables

LOINC Part (entire column represents a single LOINC code)	CDISC Standard Equivalent (each row represents a single CDISC code)
Component	LBTEST/CD + other variables
Property	-
Time	MULTIPLE: Various Timing Variables
System	SPEC+LOC
Scale	-
Method	METHOD + other variables

4 Working Group Recommendations

In general, the Working Group recommends that where a valid LOINC code exists for a given lab value, the community should provide the code in LBLOINC. Many laboratories have mapped their internal codes to LOINC codes and will make such mappings available in electronic messages or as standalone mapping files. Increasingly, in vitro Diagnostic (IVD) vendors are making mappings[4] of their internal test codes to LOINC codes available for their products, which will increase the efficiency and consistency of identifying appropriate LOINC codes. If such LOINC codes are not available from the laboratory, the Working Group recommends that sponsors not attempt to derive a mapping to LOINC as they create the SDTM files. Lab test names are notoriously ambiguous, and making assumptions in the mapping process without confirming with the test performer can lead to serious data quality issues.

Based on CDISC 2017 LOINC survey data, the Working Group also feels it is important that the requirement for LOINC code submissions start with a subset of common diagnostic lab tests that are also used within clinical research in order to ease researchers' adoption of the FDA requirement. To create this list, the CDISC Lab Working Group assessed the most commonly utilized lab tests in clinical trials as defined by three major central laboratories, the LOINC 2,000+ and data summarizing LOINC codes that have been received by the FDA. These LOINC codes are predicted to represent ~80-90% of the routine care labs also utilized within clinical trials. A mapping of this subset of common LOINC codes to specific combinations of CDISC LB domain variables and controlled terminologies has been created by the CDISC Lab Working Group. At the time of writing this recommendation document, this mapping document was undergoing

review by the CDISC Lab Working Group and key members from the CDISC Controlled Terminologies Team and this LOINC Working Group. The mapping document and an implementation guide will be made available from the CDISC Lab Working Group on the CDISC and Regenstrief websites.

The Working Group specifically recommends that:

- Regardless of whether a LOINC code is available or not for a given lab test record, submitters must still submit lab data in CDISC format, including the use of CDISC controlled terminology for lab variables.
- Sponsors should include LOINC codes for any test where a valid LOINC code exists for new trials' electronic submissions per the FDA Data Standards Catalog requirement dates.
    - LOINC codes that contain specified methods should be utilized over similar LOINC codes that do not have a method specified in the Method part as appropriate. See Figures 2 and 3 for detail.
    - When the specimen type for a lab test is known and/or where the reference range differs by specimen types, a LOINC code with a specific System (such as Ser/Plas, Urine, CSF, etc.) should be used instead of a similar LOINC code that has "XXX" as the System.
    - Sponsors may submit LOINC codes for lab tests outside of the common subset described above when it becomes available. The Working Group recommends that sponsors be encouraged to do so in order to allow the community and regulatory officials to adapt to the new requirement. When provided, LOINC codes for other lab tests should be accepted by FDA.
    - The most specific LOINC code should always be used without over-specifying details that are not known by the sponsor, such as method or scale.
    - LOINC codes with a status of deprecated or discouraged should not be submitted.
        - Appropriate status for a LOINC code is defined in the LOINC Manual Section 11.2 Classification of LOINC Term Status.
- For any lab test where a LOINC code is not submitted, the reason for its omission should be noted in the clinical Study Data Reviewers Guide.
    - The Working Group proposes that a starter set of reasons be predetermined (perhaps as CDISC terms) for consistency of reporting, including:
        - Performing laboratory unable to determine if appropriate LOINC code exists
        - Performing laboratory indicates that no appropriate LOINC code currently exists
- LOINC code specified in LB.LOINC applies to LB.ORRES (original result) rather than to LB.STRES(C/N) (standardized result). Such an approach is consistent with the recommendation that LB.LOINC should be taken from original laboratory result (i.e., the data transfer itself), and not be derived/converted into another form.
- LOINC codes should be submitted for human clinical studies only.

5 Potential Future Phases

The Working Group may reconvene in the future to make further recommendations. The Working Group considered topics such as the inclusion of additional LOINC codes for tests that are not routinely performed, LOINC codes for labs performed on non-human subjects, and codes for more clinically-focused variables. Though it is highly desirable for regulators to easily and routinely analyze standard of care lab tests that have assigned LOINC codes, challenges exist to this approach for research lab tests, most of which will not have an assigned LOINC code.

Regarding the pre-coordinated LOINC model and post-coordinated CDISC model, a harmonization of CDISC LB domains with LOINC represents a substantial volume of work. In the absence of clear use cases for such harmonization, the Working Group determined that it was out of scope for this initial set of recommendations, but it may be considered in future work.

6 References

1. Providing Regulatory Submissions In Electronic Format – Standardized Study Data. United States Food and Drug Administration https://www.fda.gov/downloads/Drugs/Guidances/UCM292334.pdf. Accessed May 15, 2017.
2. Robert Stevens and Uli Sattler (2013) Post-coordination: Making things up as you go along. Ontogenesis. http://ontogenesis.knowledgeblog.org/1305. Accessed June 23, 2017.
3. Dan Vreeman. (2017) LOINC Version 2.61 and RELMA Version 6.20 Available. https://loinc.org/news/loinc-version-2-61-and-relma-version-6-20-available/. Accessed June 23, 2017.
4. IICC News. (2017) IVD Industry Connectivety announces LIVC specification for digital publication of LOINC to vendor IVD test results in clinical laboratories. http://ivdconnectivity.org/iicc_announces_livd_specification/. Accessed June 26, 2017.

7 Appendices

7.1 Abbreviations

Abbreviation	Description
ANDA	Abbreviated New Drug Application
BLA	Biologics License Applications
CBER	Center for Biologics Evaluation and Research
CDER	Center for Drug Evaluation and Research
CDISC	Clinical Data Interchange Standards Consortium
CRO	Contract Research Organization
FDA	United States Food and Drug Administration
IND	Investigational New Drug
LB	Laboratory
LOINC	Logical Observation Identifiers Names and Codes
MB	Microbiology
NCI	United States National Cancer Institute
NDA	New Drug Application
NIH	United States National Institutes of Health
NLM	United States National Library of Medicine
SDRG	Study Data Reviewer's Guide
SEND	Standard for Exchange of Non- Clinical Data
SDTM	Study Data Tabulation Model

7.2 Helpful Links

CDISC Homepage: http://www.cdisc.org

CDISC SDTM webpage: https://www.cdisc.org/standards/foundational/sdtm

CDISC SDTM Implementation Guide webpage:
https://www.cdisc.org/standards/foundational/sdtmig

LOINC® Homepage: https://loinc.org/

FDA Study Data Standards Resources, including the Data Standards Catalog:
https://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm

7.3 Educational Resources and Events

7.3.1 CDISC

CDISC – SDTM002A – SDTM Basics, CDISC – SDTM003A – Basics of the SDTM Implementation Guide, and CDISC – SDTM010A – Findings Domains: https://secure.trainingcampus.net/uas/modules/trees/store/wcatalog.aspx?cat=1029&ci=10&vtype=2&pi=0

CDISC Public Courses: https://www.cdisc.org/education/public-courses

SDTM Standard in PDF: https://www.cdisc.org/standards/foundational/sdtm

SDTM IG in PDF: https://www.cdisc.org/standards/foundational/sdtmig

SDTM in Other Machine-Readable Formats (CDISC Members Only and Individual Academic Researchers Upon Request): https://www.cdisc.org/members-only/share-exports

7.3.2 LOINC

LOINC Getting Started Guide: https://loinc.org/get-started/ LOINC overview video: https://youtu.be/OJYQ2xHNIUU LOINC and RELMA Workshops: https://loinc.org/meetings

LOINC Essentials eBook, which covers the basics of LOINC and specific mapping strategies: https://danielvreeman.com/loinc-essentials/