Paper DS08

Clearing the Fog Surrounding LOINC

Jordan Bristow, PHASTAR, London, UK

ABSTRACT

Logical Observation Identifiers Names and Codes (LOINC) is a growing database containing a universal standard for identifying health measurements, observations and documents. This paper will give an overview of LOINC and provide a foundation for the reader in a developing area of data management. This paper aims to spur the reader into becoming informed of a method of standardising data for a global audience that they may have been unaware of and providing clarity in an area that may seem to be intractable at first glance.

INTRODUCTION

Beginning 15th March 2020, submissions of studies for New Drug Applications (NDAs), Abbreviated New Drug Applications (ANDAs), and certain Biologic License Applications (BLAs) will be required to include Logical Observation Identifiers Names and Codes (LOINC®) as specified by the Food and Drug Administration (FDA) Data Standards Catalog; certain INDs (Investigational New Drug) will have the same requirement from 15th March 2021. Implementation of these standards is being set in place to allow for succinct interoperability between clinical data systems.

The aim of this paper is to give the reader a foundational overview of LOINC and provide an explanation as to the usefulness and practicality of including these codes within studies for future submissions and has been created to document the findings of the PHASTAR working group on LOINC.

This paper provides an example of a LOINC term and explains its constituent parts. While seemingly intractable from an initial standpoint, LOINC provides a rich database for clinicians and researchers to draw from. This ensures that once the standards are applied to their data, local or proprietary terms are then able to be transferred to other institutions with ease; allowing for a swift exchange of data when required.

Each LOINC term is comprised of up to six subsections, where the function of each subsection provides a unique identity for every test in the LOINC database. The components that make each term unique will be elaborated upon and a few examples provided of differing results that could be contained within each.

The global scope of LOINC and the growing community that is becoming increasingly prominent in data standards governance makes this a fast-evolving and exciting area of clinical development.

This paper will be limited to the foundational level, with suggested reading for the reader on further information regarding LOINC's developments.

LOINC'S PURPOSE

First established in 1994 by the Regenstrief Institute, LOINC codes are developed and maintained in a freely available database encompassing both medical laboratory test names and clinical observations.

The key issue that LOINC aims to remedy is that the aggregation of many different sources of data may inherently contain a myriad of ways to identify the same tests.

While a specialist may be able to observe these and interpret that they all refer to the same measurement, a computer has no such knowledge. Local computer systems can apply test names that can be viewed as ambiguous when transferred across to other systems; the naming convention of one system may not necessarily be the same as another. They can use different methods of identifying the same test or measurement which can create difficulties with data exchange that may then impact the speed at which data can be interpreted by these systems. The key idea behind the inception of LOINC is to provide interoperability between different systems; allowing communication, exchange of data and use of the information that has been exchanged. The final result of this implementation is that clinical data can be available when and where it is needed without the obfuscation caused by differing system's local codes being confused with each other.

LOINC TERM

A LOINC term can be considered to be a representation of a question about clinical phenomena that can be observed or measured. There are two essential pieces to a LOINC term: the LOINC Code and the LOINC Fully Specified Name (FSN).

The LOINC Code is a unique permanent identifier that acts as a computer processable representation of a LOINC term. The LOINC code itself is not intended to be understood by humans; the codes are assigned sequentially, so only infer the relative age in the database.
Once these codes are published, they are never removed from the database and are permanently available.

Linked to each code is the LOINC Name which is the human-readable text version of the LOINC code.
The **FSN**'s fundamental purpose is to include the essential defining characteristics of a term and is most useful for mapping to LOINC. It is designed to include all information necessary to distinguish between clinically important differences and is essential to defining the uniqueness of each LOINC term. The FSN is comprised of five or six components that will be described in more detail in the next section of this paper. It makes use of abbreviations in naming that could be difficult to interpret for a clinician unfamiliar with the terminology.

In addition, there are 2 other main LOINC Name types that can be used in other contexts, but are no substitute for the FSN in providing the defining characteristics of a term:
Long Common Name (LCN)
Short Name

The **LCN** can be thought of as the primary display name of a LOINC term; a more descriptive and easily interpretable name than the FSN that would be loaded into a database. It removes abbreviations and uses common names for a general display of a term that allows a reader to understand a test name without needing to read the FSN for the full details of the test.

The **Short Name** is mainly used for reporting purposes, e.g. in column headers or report labels in systems that have character length limitations. The target character limit for the Short Name is 40, although due to the nature and complexity of some tests requiring multiple components, this limit is not always practicable. Not all LOINC terms feature a Short Name and as a result, the Short Name cannot be used as an identifying key in any database.

In development is a fourth and fifth name type that aim to be more easily understood. These names are currently known as **Display Name** and **Consumer Name** and are in beta and alpha stages respectively at the time of writing. Display Name aims to be a provider-friendly name and Consumer Name is intended for the consumer. The criteria surrounding the consideration of what makes a name 'provider friendly' and 'consumer-friendly' are currently being investigated, hence the early stages of development for these names.
These are available for viewing when searching LOINC codes, but, as with the Short Name, not all LOINC codes currently have these names.

See below for examples of the names displayed in the LOINC database for Sodium in Urine and Sodium in Blood:

LOINC Code	FSN	LCN	Short Name	Display Name	Consumer Name
2955-3	Sodium:SCnc:Pt:Urine:Qn	Sodium [Moles/volume] in Urine	Sodium Ur-sCnc	Sodium (U) [Moles/Vol]	Sodium, Urine
2947-0	Sodium:SCnc:Pt:Bld:Qn	Sodium [Moles/volume] in Blood	Sodium Bld-sCnc	Sodium (Bld) [Moles/Vol]	Sodium, Blood

Anything that is not essential for the identification of a test is not included in the LOINC Name, e.g., testing instrument, where testing was done (geographic location), who carried out the test etc..

While mapping to LOINC the focus should be on the FSN. However, after this has been completed, there are often other considerations that must be taken into account.
If displaying lab test results to a clinician, the best practice may be to use the local term that the clinician is familiar with. As the LOINC terms have already been mapped, the LOINC code will be available in the background and standardisation assured when transferring to other systems.

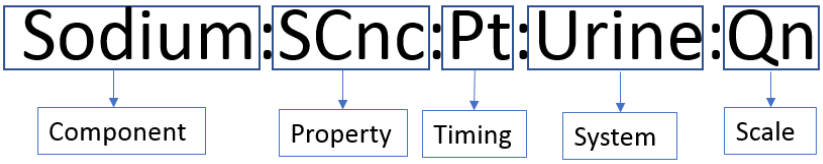
Software applications that make use of LOINC are required to include the six components of a LOINC FSN, as well as the Short Name where applicable, by the LOINC License. It is also advised that the LCN is included.
In this context it is recommended that the LCN is used as the to display the primary name in a system data dictionary.

As mentioned previously, while a LOINC code can never be removed from the system, it can be either rendered obsolete or superseded by new terms.
This could be due to ambiguity in the term. When this occurs, the code's status is changed to 'deprecated', and a description or reason is provided as to why and a direct linkage to superseding terms is provided. There may be multiple superseding terms that are a direct result of the original term's ambiguity and two differing terms may have been mixed together in the original. The linked terms provide a more specific and accurate representation of the result that also allow for the archival of retired terms – thereby preserving the database, and upholding the standard of non-removal of any terms created.

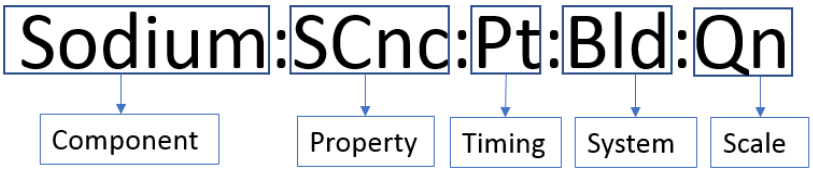
ANATOMY OF A TERM
For a LOINC term, along with the LOINC code, a LOINC FSN is comprised of five or six major axes: Component, Property, Timing, System, Scale and Method. These are delineated from each other by ":" in the FSN.

Using the previous two examples in the LOINC database of Sodium in Urine and Sodium in Blood, the FSN can be split into these components:

For LOINC Code 2955-3, the FSN is as follows:

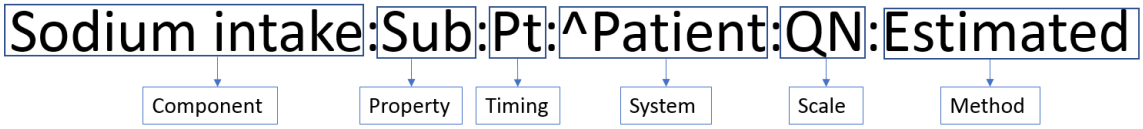


In a similar way, for LOINC code 2947-0, the FSN is:



Neither of these examples have a Method specified.
An example of a LOINC code with a populated Method can be seen below:

For LOINC code 9086-0, Sodium intake Estimated with FSN:
"Sodium intake:Sub:Pt:^Patient:QN:Estimated".



The **Component** contains the analyte that is being measured or observed.
Within the component there can be component subparts, taking the form:
"Analyte Name^Challenge^Adjustments".

The formal analyte name allows for specification of subclasses of an analyte.
The Challenge, if needed, specifies chemical or physiological stimulus before a measurement of an analyte takes place.

There are only a few LOINC terms that make use of Adjustments, but it can be used to add an additional consideration taken for adjustment of a term e.g. adjusted to patient's actual body temperature.
The "^" delineates between each subpart within the Component section of a LOINC term name. If there is no "^" then the analyte has no subparts.

Property describes the particular characteristic or attribute of that analyte.
A few examples of Property are below:

LOINC Abbreviation	Property
MCnt	Mass content
MCnc	Mass concentration
SCnc	Substance concentration

LOINC terms are not distinguished by units of measure. There is not a different LOINC code for all the varying units that a test would have. The general idea behind Property is that it should relate to a distinguishing characteristic of a name.
Because of this, the Property of a LOINC term is considered to be the most difficult area of consideration when mapping.

Timing is only used for distinction between measurements completed at a single point in time versus those that have been collected over a period of time. Examples of variations of the Timing aspect are below:

LOINC Abbreviation	Timing
Pt	Point in time
24H	24 hour collection
7D	7 days (lookback period)

System contains the information needed to determine the specimen or system on which the observation was made.
A few examples of the LOINC abbreviations for System are displayed below:

LOINC Abbreviation	System
Tiss	Tissue
Ser/Plas	Serum or Plasma
XXX	Unknown

The System can also have a subpart: Super System.
This would be displayed as "System^SuperSystem".
The Super system can be used to identify a sample from a source that is not from a patient, while the default value is taken to be Patient and is otherwise not shown in the LOINC term.

Scale is used to distinguish between different data types such as: quantitative, ordinal or nominal. Examples of the LOINC abbreviations are displayed:

LOINC Abbreviation	Scale
Ord	Ordinal
Nom	Nominal
Qn	Quantitative

Method is the only axis that is optional and not all LOINC terms have a method attached; it can be used to describe the technique used to make the measurement at a general level. It is only needed when the technique used for an observation could affect the clinical interpretation of the results. A couple of examples of this are below:

LOINC Abbreviation	Method
CT	Computer tomography
IA	Immunoassay

FDA INCLUSION

A LOINC working group recommendation for the submission of LOINC codes with new studies specifies that, while LOINC codes can be broadly categorised into Laboratory codes and Clinical codes, the FDA will only require Laboratory codes to be submitted with new regulatory applications.

The LOINC laboratory codes are generally considered to be most closely in line with the CDISC SDTM LB domain.
A key difference between LOINC and the LB domain is that LOINC is pre-coordinated and CDISC LB is post-coordinated.
A pre-coordinated set of data is considered to be static, i.e., the information presented contains all the information necessary to be used in an application.
A post-coordinated set of data has the components required and these can be compiled together as needed.

With this in mind, one LOINC term can represent a concept through the components described above. These are unchanging once the LOINC code is applied; the correct LOINC code can only be applied to one concept.
In contrast, the LB domain contains several dimensions that are needed for accurate identification of a test and its results that are compiled as required.

As such, LOINC and the LB domain variables share characteristics in their results that can coincide with each other.
While the LB.LBLOINC variable is currently already supported in the LB domain, its inclusion is currently permissible, but will be required for new studies after 15th March 2020.
The current mapping between the LOINC parts and the CDISC LB domain variables are shown below:

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

Source: https://www.fda.gov/media/109376/download

However, these two systems are not always aligned in this way.
The LOINC components of Scale and Property do not have a comparable analogue in the LB domain.
Additionally, the LB.LBLOINC variable should be taken from the original result (LB.LBORRES) and not the converted form i.e. LB.LBSTRES(C/N). This is consistent with the working group recommendation that LB.LBLOINC is taken from the original laboratory result and that the LOINC code should be applied during the creation of the raw data.

GLOBAL ADOPTION

At the time of writing, LOINC version 2.66 is available and features over 91,000 terms. This is the result of a continual input by the LOINC community that comprises over 81,000 registered users from 175 countries with over 20 translations made available.
There are major releases for LOINC twice per year in June and December that allow for a continually evolving database of codes. These could be requested due to a number of reasons e.g. new lab tests to take into consideration, new data models and even information that had previously been unstructured being transitioned to structured formats that need new LOINC codes to be added to accommodate these new formats.
Development is propelled forward by the LOINC community as new terms are added at the request of the users. As these new terms are continually added to the database, the collection becomes greater and encompasses more of an institution's local test names. This allows for even greater adoption of the standard and enhances the usability of the standard as this occurs.

The only major prohibition stipulated for the use of LOINC is that it cannot be used to develop a new set of order or observation standards to provide the same function that LOINC does. This would defeat the purpose of creating the standard in the first place; the goal being to establish one standard with one code for these observations.
The working group established by the FDA may make additional recommendations as required for topics such as:

requests for inclusion of LOINC codes for tests that are not routine, or inclusion of non-human lab LOINC codes. Additionally, the complete alignment of the pre-coordinated data models from LOINC and the post-coordinated models from the CDISC LB domain may be considered for future developments.

CONCLUSION

To conclude, LOINC's goal to unify Laboratory and Clinical data results across institutions is enabling separate institutions to make fast comparisons of their results that would potentially be delayed and more difficult to decipher without the standard applied. This interoperability goal of LOINC thereby allows a patient's data to be shared quickly and allows studies to progress at a faster and more efficient rate than without.
The rapidly increasing LOINC membership and increasing number of requests for new terms across the globe signifies that this standard is becoming more widely adopted; with the increasing number of codes, the greater the ease at which the codes can be implemented by institutions into their own databases. This ensures that the LOINC database becomes more valuable as the database expands.

Development for LOINC is a continuous process and encompasses such enhancements as updates to the database and website allowing for faster searching and returning of term results, and improvements to allow for a more streamlined approach when requesting new terms.
LOINC's active community engagement allows for suggestions to be provided by its members that can then be implemented to make processing or viewing of LOINC codes more efficient. LOINC's development is further driven by a user survey for suggested improvements that is conducted twice annually.

In summary, LB.LBLOINC will become a required variable within FDA submission datasets, but the codes contained therein are not intended to be understood by humans. Instead the LOINC Names are designed to provide information that is interpretable by humans. The LOINC FSN provides the full definition of a term that is split into five or six component parts. It is recommended that sponsors do not derive mappings to LOINC for the creation of SDTM files. This could lead to serious data issues without confirming with the test performer due to the ambiguity surrounding lab test names. PHASTAR are currently preparing for compliance with this requirement by compiling and disseminating this information to the wider community.
For further information regarding LOINC, please see the references section of this paper, these will be able to provide a deeper insight into the intricacies surrounding LOINC as well as the integration of LOINC in tandem with other data standards.

REFERENCES

1. LOINC Working Group (2017). Recommendations for the Submission of LOINC Codes in Regulatory Applications to the U.S. Food and Drug Administration. (https://www.fda.gov/media/109376/download)
2. LOINC Annual Report (2018): https://loinc.org/files/reports/2018%20LOINC%20Annual%20Report.pdf
3. Scope of LOINC: https://loinc.org/get-started/scope-of-loinc/
4. Presentations about LOINC: https://loinc.org/slideshows/
5. LOINC presentation videos: https://loinc.org/videos/
6. Which LOINC name should I use for what?: https://danielvreeman.com/loinc-name-use/
7. An Introduction to LOINC (MedInfo 2019): https://danielvreeman.com/presentation/an-introduction-to-loinc-medinfo-2019-version/
8. Veena Nataraj and Diane Piper. Delivering LOINC Codes in Future – Bridging Gaps within clinical lifecycle. CDISC Interchange 2018 - Paper. (https://www.cdisc.org/system/files/all/event/restricted/2018_US/DeliveringLOINCCodesinFuture_Nataraj_Piper.pdf)
9. LOINC code for Sodium in urine: https://s.details.loinc.org/LOINC/2955-3.html?sections=Comprehensive
10. LOINC code for Sodium in blood: https://s.details.loinc.org/LOINC/2947-0.html?sections=Comprehensive
11. LOINC code for Sodium intake estimated: https://s.details.loinc.org/LOINC/9086-0.html?sections=Comprehensive

ACKNOWLEDGEMENTS

The author would like to thank Rosalind Hobson, Alastair Sword and Jigme Namgyal from PHASTAR for their advice and support during the creation of this paper.

CONTACT INFORMATION

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

Jordan Bristow
PHASTAR
Unit 2, 2A Bollo Ln, Chiswick
London / W4 5LE
Email: jordan.bristow@phastar.com
Web: https://phastar.com/