What Can LOINC Do for SDTM?

CDISC/NCI provide controlled terms for observations such as lab test names (LBTESTCD, LBTEST, etc.). So why would another classification system be useful? The simple answer lies in the difference between the classification systems. The CDISC/NCI controlled terms for lab tests are not unique. Instead they require additional information to differentiate the lab tests. But LOINC codes are unique, so only the code is required to identify the test.

Glucose Example:

From the CDISC NCI SDTM Controlled terms:

LBTESTCD (C65047) Term "GLUC" (C105585)
LBTEST (C67154) Term "Glucose" (C105585)

However this is simply "A measurement of the glucose in a biological specimen". Therefore the "specimen" is required to identify the measurement. CDISC NCI SDTM also provides a list of terms for specimen (see below):

LBSPEC (C78734 Specimen Type)
"URINE" (C13283)
"BLOOD" (C12434)

Within the specimen types, the measurements are also classified by the testing process or method used. CDISC NCI SDTM also provides controlled terms for the different types of measurement, e.g. test strip (dipstick) vs. quantitative measurement, so LBMETHOD is also needed:

LBMETHOD (C85492 Method)

"DIPSTICK" (C106516)

The table below shows how the CDISC SDTM Labs (LB) domain requires several columns populated to correctly identify a particular clinical measurement.

LBTESTCD	LBTEST	LBSPEC	LBMETHOD
GLUC	Glucose	BLOOD	
GLUC	Glucose	URINE	
GLUC	Glucose	URINE	DIPSTICK

To represent a lab test in SDTM, several columns (LBTESTCD, LBSPEC, LBMETHOD etc.) are required, each having a different set of controlled terms. As the source data may not contain sufficient detail, it can be very difficult to correctly classify a lab test directly into CDISC SDTM – especially partial data such as local lab results.

However, LOINC provides a unique "code" for each test, along with a mapping to specimen type, method, units etc. Below is a breakdown of the LOINC codes for the glucose example:

Glucose (urine) returns 27 tests, classified by time (timepoint 6, 8, 10, 24 hours) and method (quantitative, test strip, test strip automated).

So for [timepoint], [test strip] there are three possible tests, depending on the units e.g. "positive/negative", "mmol/L", "mg/dL".

Most results in SDTM are submitted in the SI units, in this case "mmol/L" and "mg/dL" being the conventional units. If the result is simply "positive/negative", this would be submitted as a "character" result.

From the LOINC table and the CDISC NCI controlled terms, we can construct a code table to translate between the two schemes.

LOINC		LBTESTCD	LBTEST		LBSPEC		LBMETHOD		Unit
5792-7	C105585	GLUC	Glucose	C13283	URINE	C106516	DIPSTICK	C67015	mg/dL
22705-8	C105585	GLUC	Glucose	C13283	URINE	C106516	DIPSTICK	C64387	mmol/L
25428-4	C105585	GLUC	Glucose	C13283	URINE	C106516	DIPSTICK		

Note the code table should map the LOINC code (5792-7) to the CDISC code (C1005585) not the free text (glucose). This is because the free text is subject to revision between CDISC controlled term versions, but the CDISC code stays constant for a given concept.

With the code table in place, a lab result would be processed as follows:

LOINC	TSTCD	TEST	RESULT	UNT
5792-7	URGLUV	Urinary Glucose	180	MG/DL

First get the SDTM LBTESTCD/LBTEST string pair:

LOINC	NCICode
5792-7	C105585

NCICode	LBTESTCD	LBTEST
C105585	GLUC	Glucose

Next get the specimen type:

LOINC	NCICode
5792-7	C13283

NCICode	LBSPEC
C13283	URINE

Method (LBMETHOD):

LOINC	NCICode
5792-7	C106516

NCICode	LBMETHOD
C106516	DIPSTICK

Original Unit (LBORRESU)

It is common for the unit string supplied by the EDC or central lab to not match the CDISC NCI string, e.g. "MG/DL" instead of "mg/dL". This requires manual mapping correction. However with the LOINC code table we know what the correct unit string should be (mg/dL) so can insert the value from the code table. A simple programmed check can compare both strings and trigger an alert for manual review.

LOINC	NCICode
5792-7	C67015

NCICode	LBORRESU
C67015	mg/dL

Original Unit (LBORRESU)

We can extend the LOINC code table to assign the "SI Unit" and "Conventional Unit" and then document the conversion factor between the two.

LOINC	NCICode	CVUNIT	NCICode	SIUNIT
5792-7	C67015	mg/dL	C64387	mmol/L

Example lab (LB) output:

LBTESTCD	LBTEST	LBORRES	LBORRESU	LBSPEC	LBMETHOD
GLUC	Glucose	180	mg/dL	URINE	DIPSTICK

This process can be applied to other controlled terms; it is not restricted to just CDISC NCI SDTM. Additional information such as "lab category" (LBCAT), can also be assigned using the LOINC code.

Lab Category (LBCAT) example:

LOINC	Test	LBCAT
5792-7	Glucose [Urine]	Urinalysis

Lab Toxicity Grade (LBTOXGR)

To calculate LBTOXGR, the measurement or lab test must first be identified and the appropriate grade threshold calculations found from the "Common Terminology Criteria for Adverse Events (CTCAE)" documentation. Having the LOINC code makes this process more reliable, as the grade calculation may depend on several factors (specimen, unit etc). Additional information, such as the subject's gender, may have to be included to select the correct calculation.

Example "Conventional Unit" to "SI Unit" Conversion

Normalising "original result" (--ORRES) to standard result numeric (--STRN) is common for most SDTM findings datasets. This process relies on the correct identification of the measurement being standardised. We often find different conversion factors are required depending on the original unit, specimen etc. So LOINC codes are very helpful in controlling this process. Note this process can cover lab, ECG, vital signs etc. as these all have available LOINC codes.

To drive the normalisation process, we first prepare a code table aligning the LOINC code to the units.

Conversion table:

LOINC		ORRESU		STRNU	
LOINC	Text	NCICode	Text	NCICode	Text
8329-5	Core Body Temperature	C44277	F	C42559	C
8329-5	Core Body Temperature	C42559	C	C42559	C

Conversion table cont.

CVNU	SI Unit	Constant A	Constant B	Constant C	Constant K	Comment
F	C	5	9	-32	0	
C	C	1	1	0	0	

Use the equation below to calculate the normalised value:

$$Standard\ Value = \left(\frac{Constant\ A}{Constant\ B}\right) \times (Collected\ Value + Constant\ C) + Constant\ K$$

If the conversion factors are not fractions, e.g. 0.55 instead of 5/9, we simply set "Constant A" to the multiplier (0.55) and set "Constant B" to one. This will introduce a rounding error, which should be documented in the conversion table "comment" column. Additional information, such as LOINC table version number, measurement name, specimen type or additional notes, should be included to provide context.

In preparation for the normalisation process, we must align the "original units" to the CDISC/NCI terminology. This is done using a code table to convert the collected "unit string" into the CDISC NCI Unit NCI Code. It's recommended to "clean" the collected unit first, e.g. upper-case the string and remove trailing spaces etc. This reduces the number of entries in the alignment code table and also makes the conversion more robust. CDISC NCI also provide a list of synonyms for each submission value. These can be imported into the alignment table from the NCI documentation. For example:

NCI Code	CDISC Submission	CDISC Synonym(s)
C44277	F	Degree Fahrenheit
C44277	F	Fahrenheit
C42559	C	Degree Celsius
C42559	C	Celsius

Example of data cleaning process:

LOINC	VSORRES	VSORRESU	Cleaned Unit	Unit NCICode	Unit Text
8329-5	39.5	Centigrade	CENTIGRADE	C42559	C
8329-5	38.2	c	C	C42559	C
8329-5	38.5	C	C	C42559	C
8329-5	100	F	C	C44277	F
8329-5	101	fahrenheit	FAHRENHEIT	C44277	F

With the aligned VSORRESU, for each input data row we perform a lookup using the LOINC code and "Unit NCICode" to get the conversion factors (Constant A, Constant B, Constant C, Constant K) and apply the normalization formula:

LOINC	Unit NCICode	VSORRES	Conversion	Result
8329-5	C42559	39.5	(1/1) x (39.5 + 0) + 0	39.5
8329-5	C42559	38.2	(1/1) x (38.2 + 0) + 0	38.2
8329-5	C42559	38.5	(1/1) x (38.5 + 0) + 0	38.5
8329-5	C44277	100	(5/9) x (100 - 32) + 0	37.77
8329-5	C44277	101	(5/9) x (101 - 32) + 0	38.33

Example output:

LOINC	VSORRES	VSORRESU (modified)	VSSTRN	VSSTRNU
8329-5	39.5	C	39.5	C
8329-5	38.2	C	38.2	C
8329-5	38.5	C	38.5	C
8329-5	100	F	37.77	C
8329-5	101	F	38.33	C

This approach has several advantages:

- Correct identification of measurement (body temperature).
- Uses the "unit" NCI code – not the collected text – which provides a "submission ready" unit.
- Allows different conversion factors for each measurement. For example, it is quite common for conversion factors to depend on specimen type or method.
- Provides full documentation of the normalisation process (often requested for FDA submissions).
- Standardises the process across all studies – but neutral to CDISC NCI SDTM terminology version, by using the NCI code, not "unit" text.

Advantages of LOINC

Using the above approach gives a consistent assignment of several e.g. SDTM LB columns, and provides a mechanism to standardise the unit normalisation process.

For local lab data or results without LOINC codes, there is a robust coding process available with several tools for assignment of LOINC.

LOINC also provides the option of more granular and reliable grouping of results for analysis and reporting, as each lab test is identified.

Standard processes, such as normalisation to SI units, assignment of "category", calculation of "LBTOXGR" etc. can all leverage the LOINC code to provide a fully documented and standardised process. This can be applied across all study SDTM conversions.

Finally, LOINC codes are available for several SDTM findings datasets e.g. labs (LB), vital signs (VS), ECG (EG) etc.

Ed Chappell

Switching from the electronics industry to the software industry in 2001, Ed Chappell first worked in electronic publication and machine readable documentation for USPTO Patent filings with Thomson Scientific. Seven years later he transitioned into the pharmaceutical industry with Formedix. As a senior consultant at Formedix, Ed provides training on CDISC Standards and XML technology to customers.