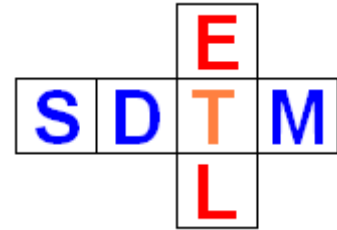


SDTM-ETL 4.0: Using the LOINC-SDTM-LB Mapping and similar functions

Author: Jozef Aerts, XML4Pharma

Last update: 2021-04-20



Introduction

Many within the CDISC world consider [LOINC](#) coding as a burden, due to the new FDA requirements that for laboratory tests at least, the LOINC code must be delivered in LBLOINC.

However, LOINC coding is not a burden, it is a blessing and a great opportunity.

LOINC is a pre-coordinated coding system for all kinds of tests, so not only for laboratory tests as many think. LOINC is used all over the world, and mandated to be used in laboratory reporting in many countries. So, if you get your lab results from the hospital or an external lab, there is a high probability that the LOINC code can be provided. But of course, one has to ask.

Due to the new FDA requirement, CDISC developed a mapping between the 2000+ most used LOINC codes and the SDTM lab domain. This mapping also allows to automatically populate LBTESTCD, LBTEST, LBSPEC, LBLOC, LBMETHOD etc., and that is exactly what we have implemented in SDTM-ETL. This also means that when one has the LOINC code of the lab test (to be delivered by the lab of course), there is no need anymore to spend hours or days to puzzle out the categorization into the SDTM variables. One can simply auto-populate these variables starting from the LOINC code.

Getting the LOINC code and populating LBLOINC

If one has the LOINC code of the lab test, it usually is already stored into the EDC system and thus can be exported into the ODM export. Using the normal drag-and-drop procedure, followed by using the wizards, one can then easily generate the mapping for LBLOINC. In case LBLOINC was not provide it by the template, you can still add it using the menu "Insert – new SDTM Variable". Do also not forget to assign an "external codelist" (named "LOINC" of course) to the properties of the LBLOINC variable.

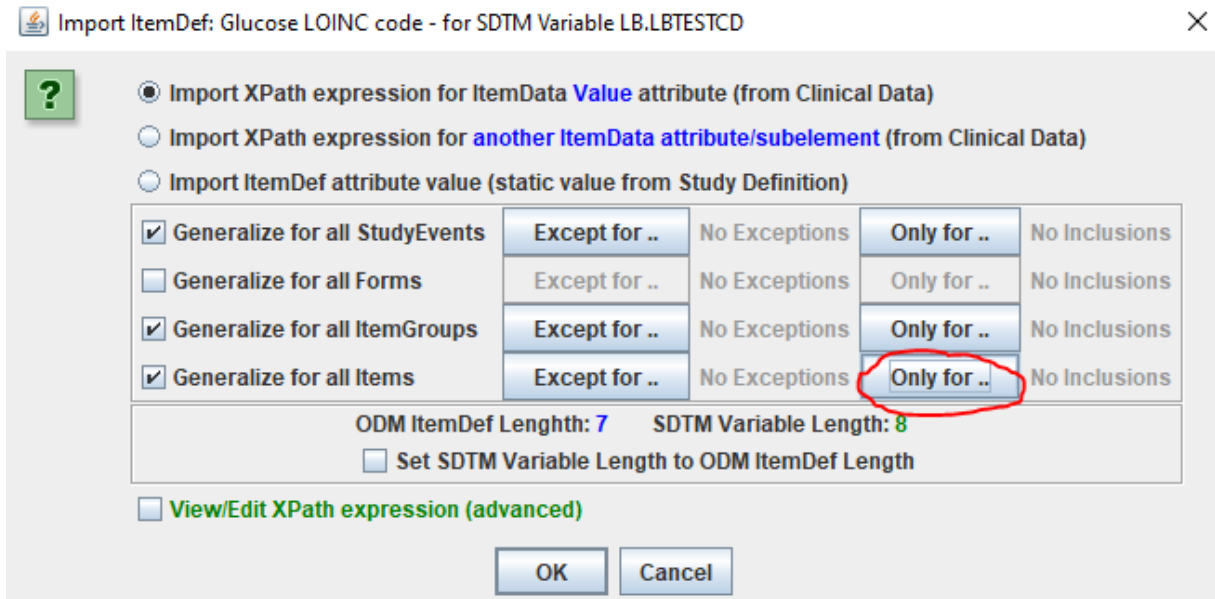
After that, you can choose to set the structure of the dataset either to the classic "one record per LBTESTCD per ... per ... per subject", or to "one record per LBLOINC per ...". The latter may be a better choice when you have a mixture of quantitative and non-quantitative measurements for the same analyte. A typical example is "glucose in urine", some results being quantitative (a concentration), and other ordinal (+1, +2, ...).

Generating the mappings for LBTESTCD, LBTEST, LBSPEC, LBMETHOD...

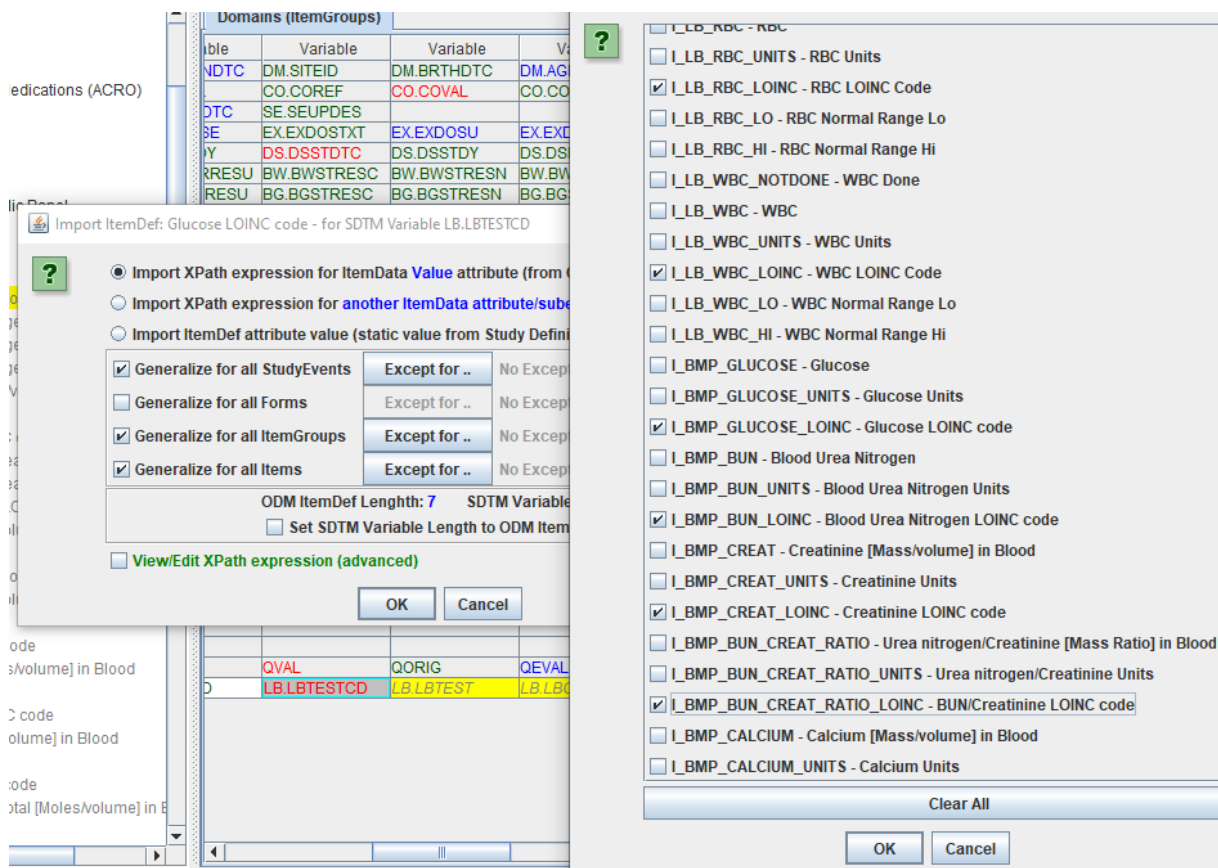
As the variable LBLOINC has been put near the end of the identifiers in LB (a major design errors in our opinion, as the LOINC code is the real identifier of the test, and NOT LBTESTCD), one cannot simple reuse \$LB.LBLOINC in the mappings for LB.LBTESTCD etc. So, one will need to just copy-paste the mapping for LBLOINC into the cells for LBTESTCD, LBTEST etc., or drag-and-drop again from the field in ODM for the LOINC code into the cells for LBTESTCD, LBTEST etc.. For example:

- ☞ FormDef : Laboratory
 - ☞ Description
 - ☞ ItemGroupDef : Common
 - ☞ ItemGroupDef : Hematology
 - ☞ ItemGroupDef : Basic Metabolic Panel
 - ☞ Description
 - ☞ ItemDef : Glucose
 - ☞ ItemDef : Glucose Units
 - ☞ ItemDef : Glucose LOINC code
 - ☞ ItemDef : Blood Urea Nitrogen
 - ☞ ItemDef : Blood Urea Nitrogen Units
 - ☞ ItemDef : Blood Urea Nitrogen LOINC code
 - ☞ ItemDef : Creatinine [Mass/volume] in Blood
 - ☞ ItemDef : Creatinine Units
 - ☞ ItemDef : Creatinine LOINC code
 - ☞ ItemDef : Urea nitrogen/Creatinine [Mass Ratio] in
 - ☞ ItemDef : Urea nitrogen/Creatinine Units
 - ☞ ItemDef : BUN/Creatinine LOINC code
 - ☞ ItemDef : Calcium [Mass/volume] in Blood
 - ☞ ItemDef : Calcium Units
 - ☞ ItemDef : Calcium LOINC code
 - ☞ ItemDef : Sodium [Moles/volume] in Blood
 - ☞ ItemDef : Sodium Units
 - ☞ ItemDef : Calcium LOINC code
 - ☞ ItemDef : Potassium [Moles/volume] in Blood
 - ☞ ItemDef : Potassium Units
 - ☞ ItemDef : Potassium LOINC code
 - ☞ ItemDef : Chloride [Moles/volume] in Blood
 - ☞ ItemDef : Chloride Units
 - ☞ ItemDef : Chloride LOINC code
 - ☞ ItemDef : Carbon dioxide, total [Moles/volume] in E

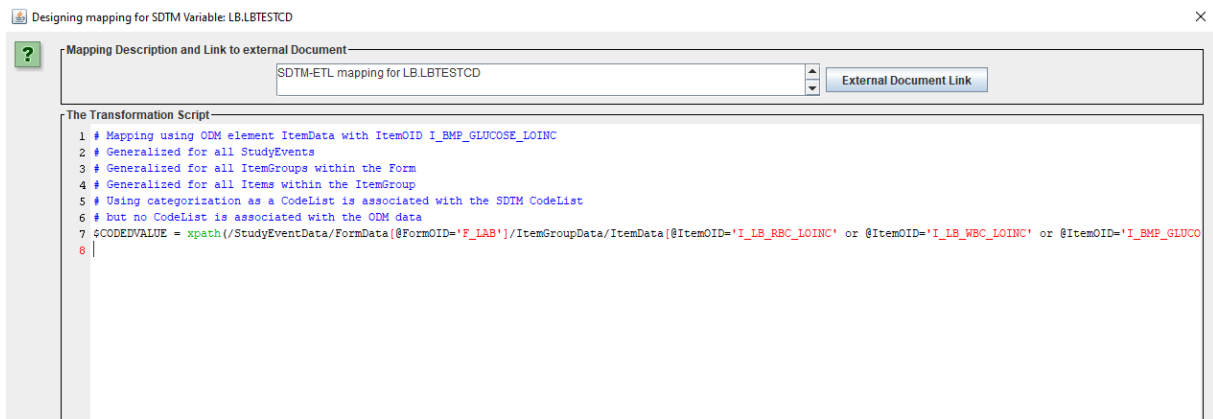
We see that there are several tests in the "Laboratory" form, each having a field for the result, for the unit, and for the LOINC code. When we then do a drag-and-drop of one of them to the LBTESTCD cell, we select that we want to have this for all the visits ("Generalize for all StudyEvents") and for both of the groups ("Generalize for all ItemGroups") "hematology" and "basic metabolic panel":



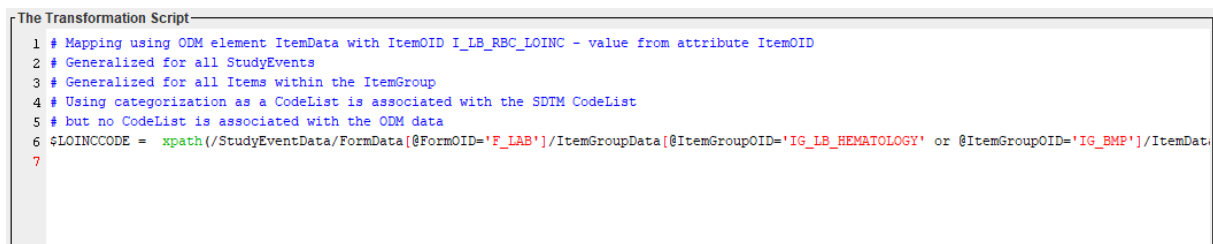
And by clicking "Only for", then select all the LOINC Code" items



leading to the following mapping script:



In total, selecting 11 items from the "Laboratory" form. We can now replace "\$CODEDVALUE" by e.g. "\$LOINCCODE" as that is more clear for what we want to do.

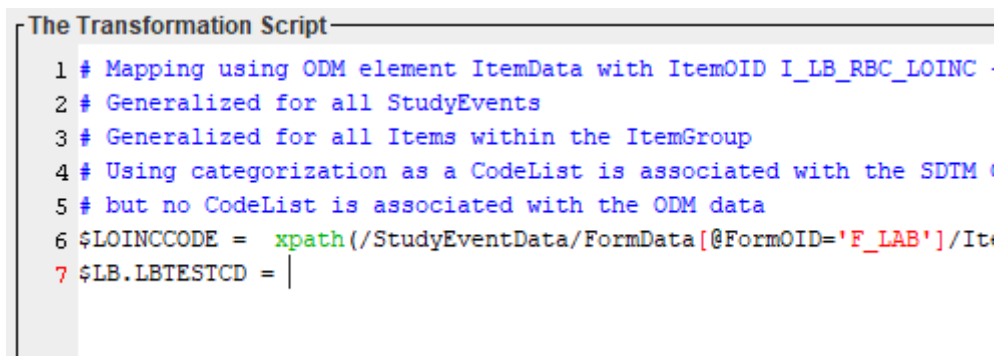


We now do exactly the same (or copy-paste the code) for the cells with LBLOINC, LBTEST, LBMETHOD, LBSPEC, LBMETHOD, as these are the SDTM variables we want to populate automatically starting from the LOINC code.

Let us now go back to the cell for LBTESTCD.

We can now use one of the new functions of SDTM-ETL that will use a RESTful web service¹ to use the mapping between the most popular LOINC coded and SDTM-LB published by CDISC.

We just type the first part of the equation for retrieving the value of LBTESTCD from LBLOINC:

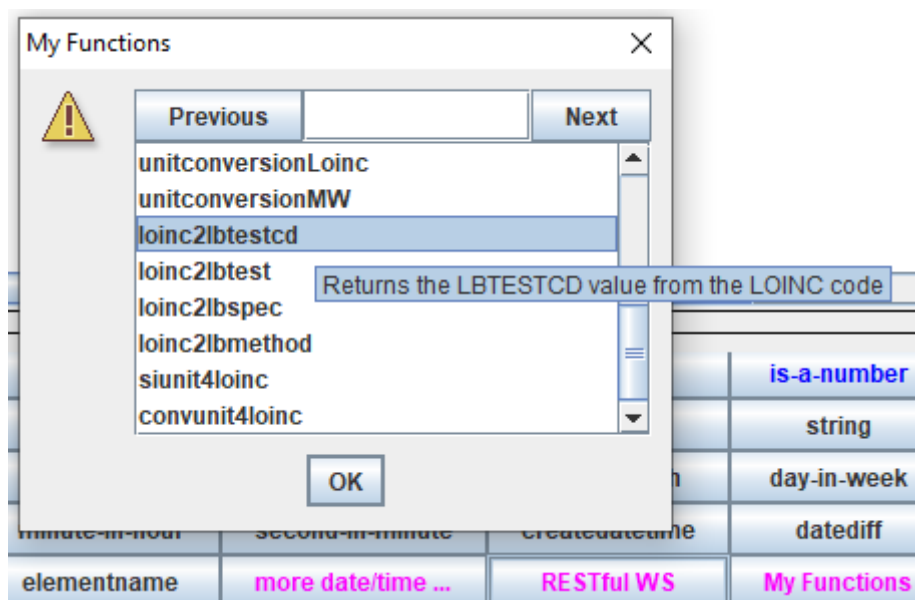


In the panel with functions, then look for the button "RESTful WS" (RESTful Web Services). You will find it near the bottom:

¹ The description of this RESTful web service can be found at: http://xml4pharmaserver.com/WebServices/LOINC2CDISC_webservices.html

second-in-minute	createdatetime	datediff	
more date/time ...	RESTful WS	My Functions	

then click the button and then look for the function "loinc2lbtestcd":



It takes just one parameter, which is the LOINC code.

Clicking the entry "loinc2lbtestcd" will then add it to the mapping code. One then only needs to add the variable \$LOINCCODE as the argument, e.g.:

```

The Transformation Script
1 # Mapping using ODM element ItemData with ItemOID IT.LOINCCODE
2 # Generalized for all StudyEvents
3 # Using categorization as a CodeList is associated with the SDTM C
4 # but no CodeList is associated with the ODM data
5 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/Item
6 $LB.LBTESTCD = rws:loinc2lbtestcd($LOINCCODE) ;

```

Very similar functions also exist for generating the mappings for LBTEST, LBSPEC and LBMETHOD starting from the LOINC code. These are:

- loinc2lbtest(string loinccode)
- loinc2lbspec(string loinccode)
- loinc2lbmethod(string loinccode)

It is advised to only use the methods "loinc2lbspec" and "loinc2lbmethod" when no explicit value for the specimen and/or analysis method is provided in the source data.

So, in our mapping scripts, we will then e.g. have:

```

The Transformation Script
1 # Mapping using ODM element ItemData with ItemOID IT.LOINCCODE
2 # Generalized for all StudyEvents
3 # Using categorization as a CodeList is associated with the SDTM
4 # but no CodeList is associated with the ODM data
5 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/It
6 $LB.LBTEST = rws:loinc2lbtest($LOINCCODE) ;

```

```

The Transformation Script
1 # Mapping using ODM element ItemData with ItemOID IT.LOINCCODE
2 # Generalized for all StudyEvents
3 # Using categorization as a CodeList is associated with the SDTM
4 # but no CodeList is associated with the ODM data
5 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/It
6 $SPECIMEN = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/It
7 if($SPECIMEN != '') {
8     $LB.LBSPEC = $SPECIMEN ;
9 } else {
10     $LB.LBSPEC = rws:loinc2lbspec($LOINCCODE) ;
11 }

```

```

The Transformation Script
1 # Mapping using ODM element ItemData with ItemOID IT.LOINCCODE
2 # Generalized for all StudyEvents
3 # Using categorization as a CodeList is associated with the SD
4 # but no CodeList is associated with the ODM data
5 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/
6 $LB.LBMETHOD = rws:loinc2lbmethod($LOINCCODE) ;

```

When we then execute the mappings on the clinical data, we get:

USUBJID	LB.LBTESTCD	LB.LBTEST	LB.LBCAT	LB.LBORRES	LB.LBORRESU	LB.LBSTRESN	LB.LBSTRESU	LB.LBLOINC	LB.LBSPEC
001	RBC	Erythrocytes		4.9	10 ⁶ /uL	4.9	10 ⁶ /uL	789-8	BLOOD
001	WBC	Leukocytes		6.2	10 ³ /uL	6.2	10 ³ /uL	6690-2	BLOOD
001	GLUC	Glucose		67.2	mg/dL	3.7301006	mmol/L	2339-0	BLOOD
001	UREAN	Urea Nitrogen		7.0	mg/dL	2.5	mmol/L	6299-2	BLOOD
001	CREAT	Creatinine		1.0	mg/dL	0.088403338	mmol/L	38483-4	BLOOD
001								44734-2	
001	CA	Calcium		8.75	mg/dL	2.1832427	mmol/L	49765-1	BLOOD
001	SODIUM	Sodium		140	mmol/L	140	mmol/L	2947-0	BLOOD
001	K	Potassium		4.2	mmol/L	4.2	mmol/L	6298-4	BLOOD
001	CL	Chloride		111	mmol/L	111	mmol/L	2069-3	BLOOD
001	CO2	Carbon Dioxide		26	mmol/L	26	mmol/L	20565-8	BLOOD
001	RBC	Erythrocytes		5.1	10 ⁶ /uL	5.1	10 ⁶ /uL	789-8	BLOOD
001	WBC	Leukocytes		6.4	10 ³ /uL	6.4	10 ³ /uL	6690-2	BLOOD
001	GLUC	Glucose		68.1	mg/dL	3.7800573	mmol/L	2339-0	BLOOD
001	UREAN	Urea Nitrogen		7.2	mg/dL	2.5714286	mmol/L	6299-2	BLOOD
001	CREAT	Creatinine		1.2	mg/dL	0.10608401	mmol/L	38483-4	BLOOD
001								44734-2	
001	CA	Calcium		8.9	mg/dL	2.2206697	mmol/L	49765-1	BLOOD
001	SODIUM	Sodium		137	mmol/L	137	mmol/L	2947-0	BLOOD
001	K	Potassium		4.0	mmol/L	4.0	mmol/L	6298-4	BLOOD

We see that all the values for LBTESTCD, LBTEST, LBSPEC and LBMETHOD (the latter not visible here) have automatically been populated, except for ... LOINC code 44734-2. The reason for this is that this LOINC code is NOT in the list of the 2000+ most popular test codes for which CDISC has published a mapping. This is also the reason that we decided to extend the mapping with several thousand of LOINC codes ourselves².

² We are currently working on this. All additional mappings will also incorporated into the RESTful web service. You can follow out progress at: http://xml4pharmaserver.com/WebServices/LOINC2CDISC_webservices.html

As the mapping for LOINC code 44734-2 cannot be fully automated, we need to do a little bit of manual work. First, we look up on the LOINC website what the LOINC code 44734-2 means. We find:

LOINC CODE: **44734-2**

LONG COMMON NAME: **Urea nitrogen/Creatinine [Mass Ratio] in Blood**

LOINC STATUS: **Active**

Part Descriptions

LP14492-0 Urea nitrogen
 Urea nitrogen is the waste product formed in the liver when proteins are broken down into their respective amino acids. Once proteins are metabolized in the liver, urea nitrogen is released from the liver to the kidneys via the bloodstream. The kidneys then filter the urea nitrogen out of the bloodstream into the urine. Although the kidneys filter out most of the urea nitrogen in the blood, small traces of stable urea nitrogen remain in circulation. To evaluate kidney function, a blood urea nitrogen test (BUN test) is conducted. High amounts of urea nitrogen in the BUN test indicate complications with kidney function or decreased blood flow to the kidneys. A decrease in urea nitrogen in the BUN test is indicative of severe liver disease or malnutrition. [Lab Tests Online:bun/tab/sample/]
 Source: Regenstrief LOINC

LP32035-5 Creatinine
 Creatinine or creatine anhydride, is a breakdown product of creatine phosphate in muscle. The loss of water molecule from creatine results in the formation of creatinine. It is transferred to the kidneys by blood plasma, whereupon it is eliminated by glomerular filtration and partial tubular excretion. Creatinine is usually produced at a fairly constant rate and measuring its serum level is a simple test. A rise in blood creatinine levels is observed only with marked damage to functioning nephrons; therefore this test is not suitable for detecting early kidney disease. Creatine and creatinine are metabolized in the kidneys, muscle, liver and pancreas.
 © Copyright ©2005-2009 Genome Alberta (Reference to original publication: Wishart DS, Knox C, Guo AC, et al. HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res. 2009 37(Database issue):D603-610.)
 Source: Human Metabolome Database, [Creatinine](#)

Ok, this looks to be a pretty easy one.

We then look up whether there is an entry for this in the CDISC Controlled Terminology. There are several ways to do it (e.g. by searching in the CDISC-CT Excel worksheet), but using the CDISC Library browser is of course more elegant:

cdisc LIBRARY Data Standards Browser

Search

Submission Value: LBTESTCD
 Definition: Terminology used for laboratory test codes of the CDISC Study Data Tabulation Model.
 Nci Preferred Term: CDISC SDTM Laboratory Test Code Terminology
 Synonyms: Laboratory Test Code

SDTM Controlled Terminology Package 42 Effective 2020-06-26

Urea/Creatinine

Exports

Term	Submission Value	Synonyms	Definition	Preferred Term
C96645	UREACRT	Urea/Creatinine	A relative measurement (ratio or percentage) of the urea to creatinine in a biological specimen.	Urea to Creatinine Ratio Measurement

Finding that the LBTESTCD for this is "UREACRT". Similarly, we find that the corresponding LBTEST is "Urea/Creatinine".

We can now incorporate this information in our mapping script (here e.g. for LBTESTCD) using an if-statement:


```

6 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/ItemGroupData[@ItemG:
7 # LOINC code 44734-2 is not covered by the LOINC to CDISC mapping
8 if($LOINCCODE = '44734-2') {
9     $TEMP = 'UREANCRT';
10 } else {
11     $TEMP = loinc2sdtmlb($LOINCCODE, 'LBTESTCD');
12 }
13 print 'LBTESTCD from LOINC code = ' $TEMP;
14 $LB.LBTESTCD = $TEMP;

```

and similar for LBSPEC:

```

6 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/ItemGroupData[@Ite
7 # LOINC code 44734-2 is not covered by the LOINC to CDISC mapping
8 if($LOINCCODE = '44734-2') {
9     $LB.LBSPEC = 'BLOOD';
10 } else {
11     $LB.LBSPEC = loinc2sdtmlb($LOINCCODE, 'LBSPEC');
12 }

```

Then executing the mappings again on the clinical data leads to:

USUBJID	LB.LBTESTCD	LB.LBTEST	LB.LBCAT	LB.LBORRES	LB.LBORRESU	LB.LBSTRESN	LB.LBSTRESU	LB.LBLOINC	LB.LBSPEC
001	RBC	Erythrocytes		4.9	10 ⁶ /uL	4.9	10 ⁶ /uL	789-8	BLOOD
001	WBC	Leukocytes		6.2	10 ³ /uL	6.2	10 ³ /uL	6690-2	BLOOD
001	GLUC	Glucose		67.2	mg/dL	3.7301006	mmol/L	2339-0	BLOOD
001	UREAN	Urea Nitrogen		7.0	mg/dL	2.5	mmol/L	6299-2	BLOOD
001	CREAT	Creatinine		1.0	mg/dL	0.088403338	mmol/L	38483-4	BLOOD
001	UREANCRT	Urea/Creatinine		9.6	g/g(creat)	9.6	g/g(creat)	44734-2	BLOOD
001	CA	Calcium		8.75	mg/dL	2.1832427	mmol/L	49765-1	BLOOD
001	SODIUM	Sodium		140	mmol/L	140	mmol/L	2947-0	BLOOD
001	K	Potassium		4.2	mmol/L	4.2	mmol/L	6298-4	BLOOD
001	CL	Chloride		111	mmol/L	111	mmol/L	2069-3	BLOOD
001	CO2	Carbon Dioxide		26	mmol/L	26	mmol/L	20565-8	BLOOD
001	RBC	Erythrocytes		5.1	10 ⁶ /uL	5.1	10 ⁶ /uL	789-8	BLOOD
001	WBC	Leukocytes		6.4	10 ³ /uL	6.4	10 ³ /uL	6690-2	BLOOD
001	GLUC	Glucose		68.1	mg/dL	3.7800573	mmol/L	2339-0	BLOOD
001	UREAN	Urea Nitrogen		7.2	mg/dL	2.5714286	mmol/L	6299-2	BLOOD
001	CREAT	Creatinine		1.2	mg/dL	0.10608401	mmol/L	38483-4	BLOOD
001	UREANCRT	Urea/Creatinine		9.3	g/g(creat)	9.3	g/g(creat)	44734-2	BLOOD
001	CA	Calcium		8.9	mg/dL	2.2206697	mmol/L	49765-1	BLOOD
001	SODIUM	Sodium		137	mmol/L	137	mmol/L	2947-0	BLOOD
001	K	Potassium		4.0	mmol/L	4.0	mmol/L	6298-4	BLOOD
001	CL	Chloride		119	mmol/L	119	mmol/L	2069-3	BLOOD
001	CO2	Carbon Dioxide		28	mmol/L	28	mmol/L	20565-8	BLOOD
001	RBC	Erythrocytes		5.4	10 ⁶ /uL	5.4	10 ⁶ /uL	789-8	BLOOD
001	WBC	Leukocytes		6.6	10 ³ /uL	6.6	10 ³ /uL	6690-2	BLOOD
001	GLUC	Glucose		73.2	mg/dL	4.0631453	mmol/L	2339-0	BLOOD
001	UREAN	Urea Nitrogen		7.3	mg/dL	2.6071429	mmol/L	6299-2	BLOOD
001	CREAT	Creatinine		0.9	mg/dL	0.079563005	mmol/L	38483-4	BLOOD
001	UREANCRT	Urea/Creatinine		8.2	g/g(creat)	8.2	g/g(creat)	44734-2	BLOOD
001	CA	Calcium		7.9	mg/dL	1.9711562	mmol/L	49765-1	BLOOD

Imagine that we did not have the LOINC code, and needed to puzzle out about suitable values of LBTESTCD, LBTEST, LBSPEC, LBMETHOD, ... ourselves ...! That is what I mean that LOINC is a blessing rather than a burden!

We also see that (surprise, surprise, ...) all values have been standardized to SI units in LBSTRESN/LBSTRESU. How this was fully automated is explained in the tutorial ["Performing Unit Conversions in SDTM-ETL"](#).

Generic mapping functions

One can also use the function "loinc2sdtmlb" that is taking two arguments:

rw:**loinc2sdtmlb**(string loinccode, string sdtmvarname)

where the second argument is the SDTM variable name (which can also be a Non-Standard variable, but then need to be prefixed by "SUPPLB.").

For example:

```
$LB.LBTESTCD = rws:loinc2sdtmlb('1751-7', 'LBTESTCD');
```

which is equivalent to using the function `rws:loinc2lbtested('1751-7')`.

The extra advantage of the current function "rws:loinc2sdtmlb" is that it can also be used for the new "standard non-standard" variables (NSVs) such as "SUPPLB.LBRSSCL" (result scale).

The second argument is always the SDTM variable name for which the value needs to be retrieved for. It also contains a number of "Supplemental Qualifier" variables. Here is a more complete list.

LBTESTCD (test code)	LBTEST (test name)	LBSPEC (location)	LBMETHOD (method)	LBTPT (time point)	LBANMETH (analysis method)
LBPOS (position)	LBLOC (location)	LBFAST (fasting status)	SUPPLB. LBRESTYP (result type)	SUPPLB. LBRSSCL (result scale)	SUPPLB. LBTSTOPO (test operational objective)
SUPPLB. LBLLOD (lower limit of detection)	SUPPLB. LBTSTCND (test condition)	SUPPLB. MTHSEN (method sensitivity)	SUPPLB. LBPTFL (point in time flag)	SUPPLB. LBP DUR (planned duration)	

Creating the mapping without using a RESTful web service (NOT IMPLEMENTED YET)

In the ultimate case that one cannot use the RESTful web service, there is a second function. It is using the CVS file that was published by CDISC and can be found in the folder "CDISC-CT".

Using this function works nicely as well, but has two disadvantages:

- it is considerably slower, as the whole file needs to be parsed at each function call (the RESTful web service is 5-10 times faster, as the underlying database is indexed)
- the mappings are limited to those published by CDISC, which is far from sufficient (only 2000+ LOINC codes). We are currently extending the mapping with several thousand LOINC codes, and these will also be available when using the RESTful web service.

The function is named "loinc2sdtmlb_local" and has the same two arguments as the function "loinc2sdtmlb". The results are the same as for the "loinc2sdtmlb" are identical, except for that the "extended mappings" are of course not covered.

For example, for LBMETHOD:

```
6 $LOINCCODE = xpath(/StudyEventData/FormData[@FormOID='F_LAB']/ItemGroupData[@ItemGroupC
7 $LB.LBMETHOD = loinc2sdtmlb_local($LOINCCODE, 'LBMETHOD');
```

Further work

We have also developed mappings for LOINC codes for COVID-19 tests to MB (Microbiology). A RESTful web service is also available for these mappings (see http://xml4pharmaserver.com/WebServices/LOINC2CDISC_webservices.html#loinccorona2mb) but we have not added a function for this mapping yet to SDTM-ETL. The same applies for a mapping using VSLOINC (LOINC code for vital signs) and the SDTM-VS variables. Other mappings that are still in the mapping phase are for EG (ECG Test Results) and QS (Questionnaires). More may follow.