cdisc

2022
CHINA
INTERCHANGE
29-30 JULY | VIRTUAL EVENT

Using LOINC in Regulatory Submissions

Jozef Aerts - XML4Pharma

Meet the Speaker

Jozef Aerts, Ph.D.

CEO XML4Pharma

CDISC Volunteer for more than 20 years

Retired Professor in Medical Informatics

Expert in the Use of EHRs for Medical Research
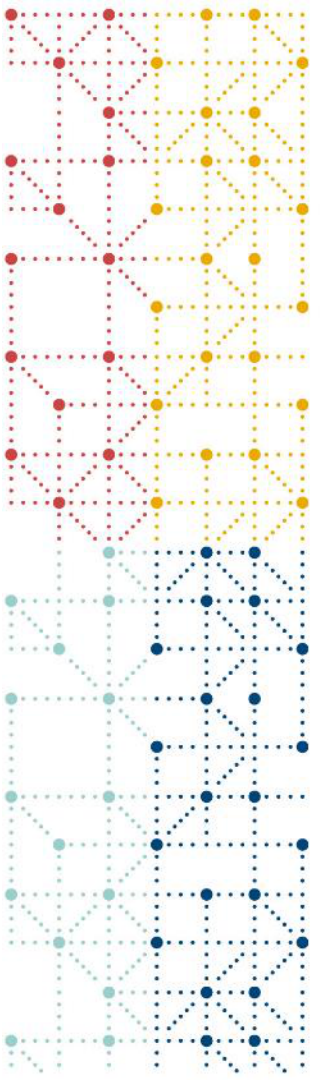
LOINC, UCUM, SNOMED-CT Specialist

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation
  are* solely *those of the author,
  and do not necessarily reflect the official policy or position of CDISC*

- *The author does have no real or apparent conflicts of interest to report.*

Agenda

1. What is LOINC?
2. LOINC and Regulatory Submissions
3. The LOINC CDISC Mapping and its Extensions
4. LOINC, SDTM, and SAS Transport format
5. Conclusions

# What is LOINC?

What is LOINC?

LOINC stands for Logical Observation Identifiers Names and Codes

- Logical: a system, not just a list
  - more logical than any CDISC controlled terminology list

- Observation
  - NOT only lab tests!
    - Also microbiology tests, vital signs, questionnaires, ...

- Identifiers
  - unique identifier for each test/observation
    (CDISC xxTESTCD/xxTEST does NOT at all uniquely identify a test or observation)

- Names and Codes
  - Each unique test has a code, a "long name" and a "short name"
  - The code is a 4-5 digit number + 1 check digit
  - The "short name" represents the 5/6 "dimensions" of the LOINC system

LOINC as a System
(not a list ...)

- One immediately recognizes some mapping opportunities

- But some aspects like "Property" and "Scale" are not present in SDTM at all

Serum or Plasma

Glucose

Mass Concentratation

System

Component

Property
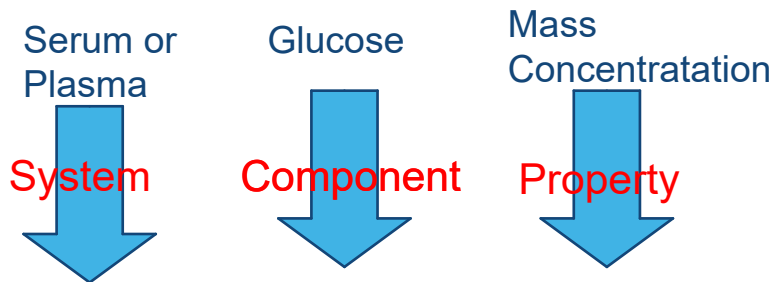
LOINC code: 2345-7
LOINC long name:
Glucose [mass/volume] in Serum or Plasma

Time

Scale

Method

Point in Time (Pt)

Quanti-tative

Empty (Optional)

Pre-coordinated versus Post-coordinated

LOINC is "pre-coordinated", CDISC-CT mostly "post-coordinated"

- Pre-coordinated is:
  - I plan carefully in advance
  - I know what type of results to expect
  - There is no need to (subjectively) categorize results, everything is already well pre-defined

- Post-coordinated is:
  - I do not plan well and carefully in advance
  - I do not know what type of results to expect
  - So I need to (subjectively) categorize the test results I obtain
    - Huge amount of work - error prone
    - There is no guarantee regulatory reviewers can compare results between studies and sponsors!
  - CDISC: categorization into xxTESTCD/xxTEST, xxSPEC, xxMETHOD, ...
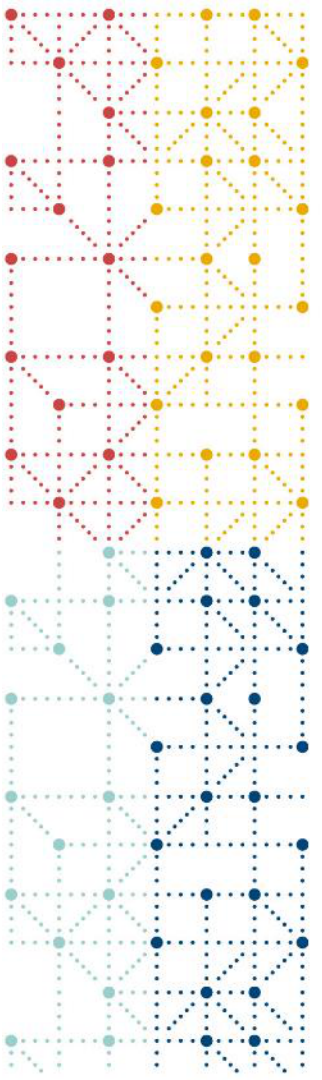
# LOINC and Regulatory Submissions

LOINC and the Regulatory Authorities

- 2017: FDA mandates the submission of the LOINC code for lab test results (SDTM LB domain) when the LOINC code can be obtained from the lab
  - LBLOINC should not be derived!
- Panic in the industry

- CDISC, FDA and LOINC (Regenstrief) start the development of a LOINC-CDISC mapping
  - limited to lab tests (excludes microbiology)
  - lmited to the 1,400 most popular LOINC lab test codes

- LOINC-CDISC mapping published in 2020
  - but only as Excel worksheet ...

Why the regulatory LOINC requirement?

- FDA cannot compare lab results between different studies from different sponsors using the CDISC-CT (combination of LBTESTCD/LBTEST, LBSPEC, LBMETHOD, ...)

- Only the LOINC code (LBLOINC), as the only real unique identifier, allows to compare lab results between different studies and sponsors

- Reason is that LBTESTCD/LBTEST, LBSPEC, LBMETHOD, etc, are (subjective) categorizations of the test
  - and some information is not present anyway (e.g. scale) or must be deduced from the results

- In EHRs, the LOINC code is the de-facto unique identifier for each observation (also for non-labs), in many countries even mandated by law
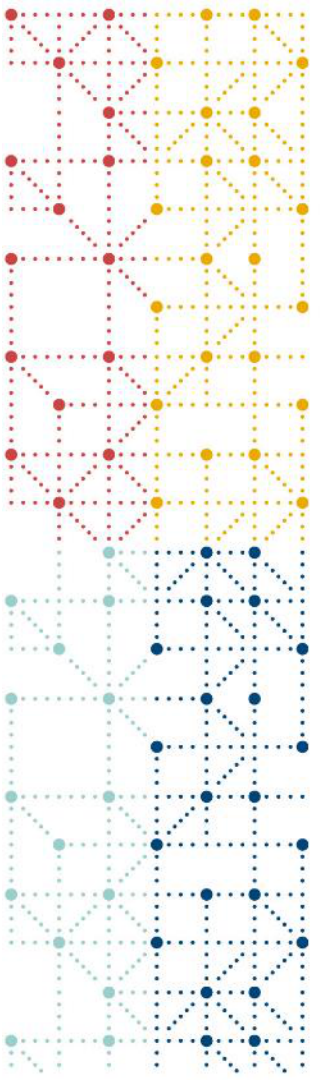
The LOINC-CDISC Mapping and its Extensions

The LOINC-CDISC Mapping

- Published by CDISC in 2020

- 2,298 mappings for 1,400 most popular LOINC codes
  - limited to lab tests, no microbiology tests, no COVID-19 tests, no vital signs, no ECG ...

- Published as an Excel worksheet
  - not really machine-executable

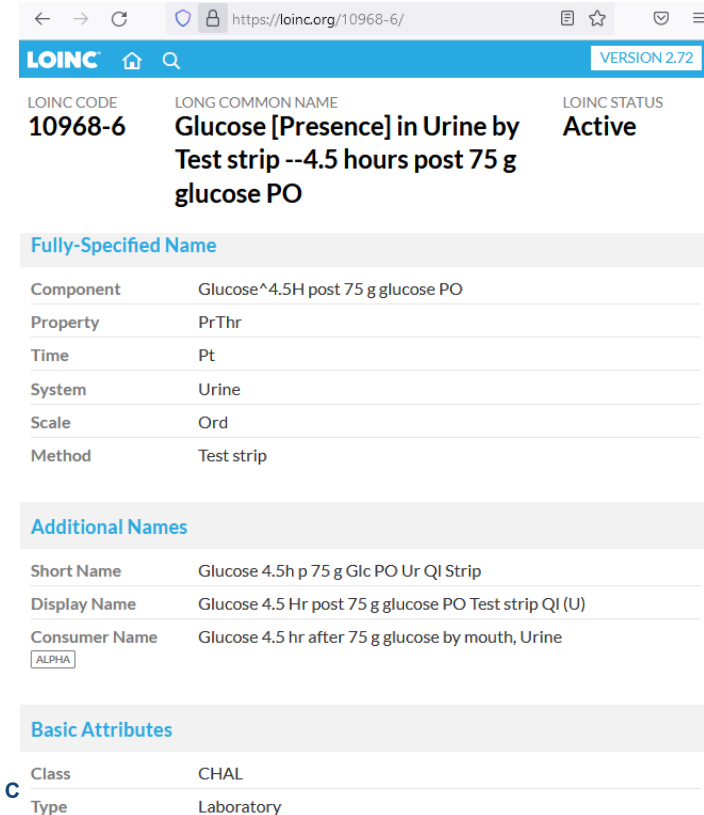|	A	B	C	D	E	F	G	H	I	J	K	L	M
	LOINC Class	LOINC Code	Component	Property	Time Aspect	System	Scale	Method	LOINC Short Name	CDISC LBTEST	CDISC LBTESTCD	CDISC LBTEST/TESTCD NCI C-code	CDISC LBSPEC
	BLDBK	1006-6	Direct antiglobulin test.IgG specific reagent	Imp	Pt	RBC	Nom		DAT IgG-Sp Reag RBC-Imp	Antiglobulin Test, Direct	ANGLOBDR	C81974	ERYTHROCYTES
	BLDBK	10331-7	Rh	Type	Pt	Bld	Nom		Rh Bld	Rh Factor	RH	C92948	BLOOD
	BLDBK	883-9	ABO group	Type	Pt	Bld	Nom		ABO Group Bld	ABO Blood Group	ABO	C125939	BLOOD
	CHAL	14771-0	Glucose^post CFst	SCnc	Pt	Ser/Plas	Qn		Glucose p fast SerPl-sCnc	Glucose	GLUC	C105585	SERUMOR PLASMA
	CHAL	14771-0	Glucose^post CFst	SCnc	Pt	Ser/Plas	Qn		Glucose p fast SerPl-sCnc	Glucose	GLUC	C105585	SERUM
	CHAL	14771-0	Glucose^post CFst	SCnc	Pt	Ser/Plas	Qn		Glucose p fast SerPl-sCnc	Glucose	GLUC	C105585	PLASMA
	CHAL	27873-9	Insulin^post CFst	ACnc	Pt	Ser/Plas	Qn		Insulin p fast SerPl-aCnc	Insulin	INSULIN	C74788	SERUMOR PLASMA
	CHAL	27873-9	Insulin^post CFst	ACnc	Pt	Ser/Plas	Qn		Insulin p fast SerPl-aCnc	Insulin	INSULIN	C74788	SERUM
	CHAL	27873-9	Insulin^post CFst	ACnc	Pt	Ser/Plas	Qn		Insulin p fast SerPl-aCnc	Insulin	INSULIN	C74788	PLASMA
	CHAL.ROUTINE	1558-6	Glucose^post CFst	MCnc	Pt	Ser/Plas	Qn		Glucose p fast SerPl-mCnc	Glucose	GLUC	C105585	SERUMOR PLASMA

The LOINC-CDISC Mapping

- Provides mappings between the LOINC code and (existing) SDTM variables
  - LBTESTCD / LBTEST
  - LBSPEC, LBMETHOD, LBANMETH, LBLOC, LBFAST
  - proposes values for LBORRESU (the rest of the medical world is using UCUM notation)
  - sets values for LBTPT where necessary

- Introduces new SDTM variables:
  - LBPDUR   (time aspect)
  - LBRESTYP (property that is measured, e.g. "mass concentration", "presence", ...)
  - LBRSLSCL (scale)
  - LBTSTOPO (Test Operational Objective, e.g. "screen", "confirm")
  - LBLLOD (Lower Limit of Detection)
  - LBTSTCND (Test Condition)
  - LBMTHSEN (Method Sensitivity, e.g. "high sensitivity")

The LOINC-CDISC Mapping and Real-World-Data

- Very soon, it became obvious that this LOINC-CDISC Mapping is insufficient for work with EHRs and Real-World-Data

    - Very limited in scope (only 1,400 LOINC codes)
    - No mappings for Microbiology, Vital Signs, ECG, Questionnaires, ...
    - No machine-executable implementation (Excel)

- For use with EHR and RWD, we extended the mapping:
    - for more than 10,000 LOINC codes
    - including mappings for COVID-19 tests (regularly extended)
    - including >600 Vital Signs test codes

    - and implemented all of this as a free RESTful Web Service

Extended Mapping: an Example

- LOINC code 10968-6: Glucose [Presence] in Urine by Test strip - 4.5 hours post 75 g glucose oral (PO)
  - A glucose challenge ...

- LBTESTCD = GLUC / LBTEST = GLUCOSE

- LBSPEC = URINE

- LBMETHOD = TEST STRIP

- LBTPT = POST DOSE

- LBELTM = PT4H30M

- LBTPTREF = 75 g glucose PO dose

- LBPTFL = Y

- LBRESTYP = THRESHOLD OR PRESENCE

- LBRESSCL = ORDINAL

LOINC CODE	LONG COMMON NAME	LOINC STATUS
10968-6	Glucose [Presence] in Urine by Test strip --4.5 hours post 75 g glucose PO	Active

Fully-Specified Name

Component	Glucose^4.5H post 75 g glucose PO
Property	PrThr
Time	Pt
System	Urine
Scale	Ord
Method	Test strip

Additional Names

Short Name	Glucose 4.5h p 75 g Glc PO Ur QI Strip
Display Name	Glucose 4.5 Hr post 75 g glucose PO Test strip QI (U)
Consumer Name ALPHA	Glucose 4.5 hr after 75 g glucose by mouth, Urine

Basic Attributes

Class	CHAL
Type	Laboratory

The Extended LOINC-CDISC Mapping: RESTful Web Services

- API description at:
http://xml4pharmaserver.com/WebServices/LOINC2CDISC_webservices.html

- Example:

```xml
<XML4PharmaServerWebServiceResponse ServerDateTime="2020-10-02T09:08:13">
    <WebServiceRequest>http://www.xml4pharmaserver.com:8080/CDISCCTService2/rest/LOINC2SDTMLB/34183-4</WebServiceRequest>
    <Response>
        <LOINC2SDTMMapping MappingSource="XML4Pharma">
            <LBTESTCD NCICode="C92294">GLOBBPT</LBTESTCD>
            <LBTEST NCICode="C92294">Beta Globulin/Total Protein</LBTEST>
            <LBORRESU_Example NCICode="C25613">%</LBORRESU_Example>
            <LBPOS/>
            <LBLOINC>34183-4</LBLOINC>
            <LBSPEC NCICode="C13283">URINE</LBSPEC>
            <LBLOC/>
            <LBMETHOD NCICode="C16536">ELECTROPHORESIS</LBMETHOD>
            <LBANMETH/>
            <LBFAST/>
            <LBTPT/>
            <SUPPLB.LBPTFL/>
            <SUPPLB.LBPDUR>PT24H</SUPPLB.LBPDUR>
            <SUPPLB.LBRESTYP>MASS FRACTION</SUPPLB.LBRESTYP>
            <SUPPLB.LBRSLSCL>QUANTITATIVE</SUPPLB.LBRSLSCL>
            <SUPPLB.LBTSTOPO/>
            <SUPPLB.LBLLOD/>
            <SUPPLB.LBTSTCND/>
            <SUPPLB.LBMTHSEN/>
            <Example_UCUM_Units>%</Example_UCUM_Units>
        </LOINC2SDTMMapping>
    </Response>
</XML4PharmaServerWebServiceResponse>
```

cdisc

The Extended LOINC-CDISC Mapping: RESTful Web Services

- The RESTful Web Service is very fast: about 20ms response time

- Is being used by many companies and organizations to fully automate
  - the creation of SDTM datasets for lab data starting from LOINC codes
  - the creation of SDTM datasets directly from EHRs, not only for LB datasets,
    but also for Vital Signs, COVID-19 Microbiology datasets etc..
    (also see HL7-Vulcan "Real World Data" project)

- However, as the LOINC code is the real unique identifier of the test,
  one starts thinking about whether it makes sense that one must still submit
  xxTESTCD/xxTEST, xxSPEC, xxMETHOD, etc., when the LOINC code is
  known.
  - Is this related to the primitiveness of the review tools at the regulatory authorities?

LOINC Codes and Regulatory Review Tools

- Most regulatory review tools are <u>not</u> "SDTM-savvy"

- A reviewer cannot know what a specific LOINC code means
  - and thus must rely on xxTESTCD/xxTEST, xxSPEC, xxMETHOD, xxTPT, etc..

- Therefore, we created the "Smart Submission Dataset Viewer" which "understands" SDTM, SEND and ADaM
  - freely available from:
    https://sourceforge.net/projects/smart-submission-dataset-viewe/

- The viewer works with CDISC Dataset-XML, Dataset-JSON and CSV files
  - SAS-XPT is admittedly not supported, as this is an outdated format
  - SAS-XPT files can however be converted "on the fly" to a modern format

The "Smart Submission Dataset Viewer": LOINC features

LOINC, SDTM, and SAS Transport format

LOINC, SDTM, and SAS Transport format

- SAS Transport Format 5 (XPT) is a catastrophic format, essentially blocking every innovation in our industry
  - One step beyond the punch card ...

- Moving to SAS Transport 8 will throw us back >20 years in time, blocking innovation for the next 10-20 years

- The new CDISC Dataset-JSON format is a much better, and very modern format for electronic submissions to regulatory authorities
  - considerably smaller file sizes that XPT
  - easy to import into SAS, R, or any other modern software (no SAS needed)
  - inherent support for non-ASCII characters (UTF-8 default encoding) - e.g "Chinese" characters (no "tricks" needed)
  - free (open source) viewers available
  - perfect for use with RESTful Web Services

- As the LOINC code is the real unique identifier, transposal with SAS Transport (5 or 8) will lead to problems (first character is a number ...)
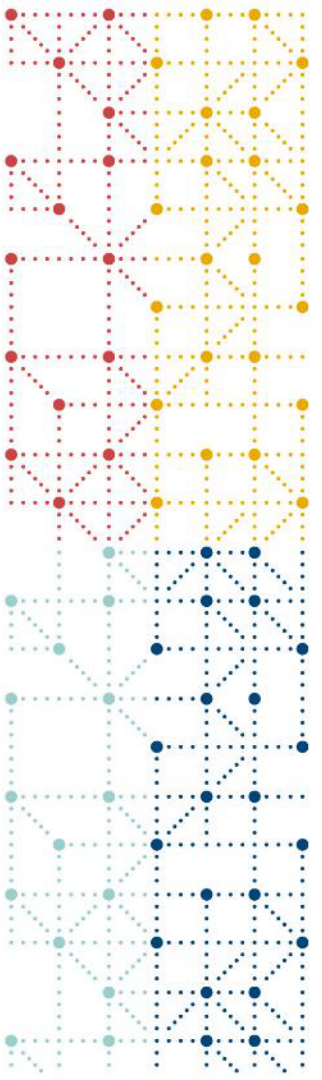
Conclusions

Conclusions

- Only the LOINC code is the real unique identifier of the test
  - this also applies to microbiology, vital signs, ECG, ...

- Regulatory authorities can only compare test results between studies and sponsors using the LOINC code
  - this especially applies to labs and microbiology

- Although the LOINC code is the real identifier, and CDISC still requires to submit xxTESTCD/xxTEST, xxSPEC, xxLOC, xxMETHOD, etc.,
a LOINC-SDTM-LB Mapping has been developed and published by CDISC

Conclusions

- This mapping has been considerably extended to over 10,000 LOINC codes including vital signs and COVID-19 microbiology tests

- A (free) RESTful Web Service has been installed to make these mappings usuable in applications, allowing automation

- Reviewers can use these and other RESTful web services in modern tools to work with LOINC codes
  (essentially making LBTESTCD/LBTESTC, etc. unnecessary)

- Regulatory authorities must move to a modern submission format (like CDISC Dataset-JSON) and not start using the catastrophic SAS Transport 8 format
    - This avoids problems when doing transposals with LOINC codes
    - and would once and for all solve the character encoding problem

Acknowledgements

- Special thanks are due to the CDISC-CT Team, especially *Erin Muhlbradt*, for the many useful discussions and timely processing of the many "new term requests"

- Also special thanks to a number of pharma companies, CROs and service providers in Scandinavia, France, Germany and the USA, for their continued encouragement and their involvement, essentially becoming "LOINC / UCUM evangelists"

cdisc

谢谢

Thank You!

Contact:

Jozef.Aerts@XML4Pharma.com

cdisc