



# Disclaimer and Disclosures

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

# What's Up with LOINC and UCUM? From EHR Records to LB Dataset in Just a Few Minutes

Presented by Jozef Aerts  
XML4Pharma  
Jozef.Aerts@XML4Pharma.com





# Agenda

1. Short Intro
2. **Demo:** From EHR Records to LB Dataset in Just a Few Minutes
3. How did we do this?
4. Limitations and future work
5. Conclusions

Jozef Aerts is a CDISC volunteer for about 18 years now, and contributor to several CDISC standards



# Introduction

## What's up with LOINC and UCUM at CDISC?

- **LOINC**: mandated in SDTM by FDA for laboratory data
  - "LOINC to CDISC-CT mapping" in public review
- **UCUM**: worldwide standard notation for units in science & engineering
  - mentioned by FDA in "Data Standards Catalog"
  - Implications for SDTM/SEND/ADaM submissions still unclear
  - Mandated/requested for SPL, drug establishment registration, drug listing
  - No intentions at CDISC to replace "UNIT" codelists by UCUM notation



# Demo: from FHIR EHR-records to SDTM in just a few minutes

## Challenge:

- Can we generate large SDTM-LB + DM datasets from Electronic Health Records (HL7-FHIR)?
- In just a few minutes?
- And standardize from US conventional units to SI units (or vice versa) fully automatically (LBSTRESC, LBSTRESN, LBSTRESU)?

# Yes, we can!

But only when using LOINC and UCUM

# Demo: from FHIR EHR-records to SDTM in just a few minutes

```
C:\FHIR2SDTM_Software>java -jar FHIR2SDTM.jar
# Log Appenders = 0
logging goes to: logs\FHIRLOINC2CDISC_LOG_2019_11_14_13-24-36.txt
changing logging level to "DEBUG"
# of LOINC codes from file = 41
# of LOINC codes from file = 41
Which FHIR-EHR System would you like to use?
0 - SyntheaStudy - server base: http://hapi.fhir.org/baseDstu3
1 - HAPIFHIR - server base: http://hapi.fhir.org/baseDstu3
2 - Vonk - server base: https://vonk.fire.ly
3 - SPARK Firely - server base: http://spark.furore.com/fhir
4 - Azure - server base: http://sqlonfhir-stu3.azurewebsites.net/fhir
5 - Pyro - server base: https://stu3.test.pyrohealth.net/fhir
0
Which Set of LOINC codes would you like to use?
0 - Urinalysis : 17 codes
1 - SI Urinalysis : 17 codes
2 - Synthea Lab codes : 41 codes
3 - Mini-mental Score Examination (MMSE) : 2 codes
0
2019-11-14 13:27:29,721 DEBUG # of FHIR records so far = 2716
2019-11-14 13:27:29,739 INFO Now sorting by USUBJID and LBDC
2019-11-14 13:27:29,754 INFO Now adding LBSEQ to records
Standardization for LBSTRESC, LBSTRESN, LBSTRESU ...
Following options are available:
0 : no standardization
1 : standardization to SI units
2 : standardization to US Conventional units
Please select your choice (0, 1, or 2)
1
```



How did we do it?

No, it is not "rocket science" ...

It just is **LOINC** and **UCUM**



Source: Wikipedia

# How does it work?

- Define set of lab tests by LOINC code
  - Best practice: LOINC codes are already in the protocol
- Query the EHR repository for the lab results (**FHIR "Observation"** resources) by the LOINC code
- Transform FHIR records to CDISC **LB** SDTM using
  - LOINC to CDISC mapping (RESTful web service)
  - LOINC RESTful web services
- Ask user whether standardization is desired

```
2019-11-14 13:27:29,721 DEBUG # of FHIR records so far = 2716
2019-11-14 13:27:29,739 INFO Now sorting by USUBJID and LBDC
2019-11-14 13:27:29,754 INFO Now adding LBSEQ to records
Standardization for LBSTRESC, LBSTRESN, LBSTRESU ...
Following options are available:
0 : no standardization
1 : standardization to SI units
2 : standardization to US Conventional units
Please select your choice (0, 1, or 2)
1
```

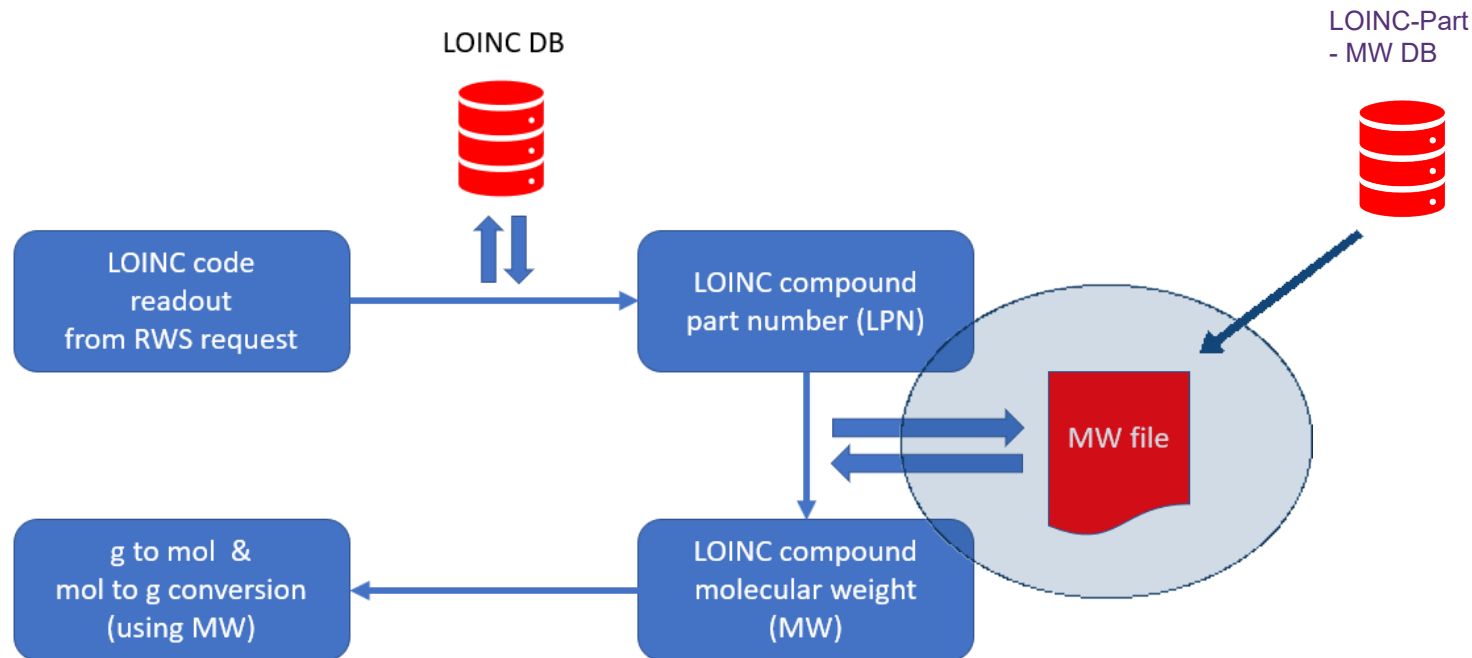


# How does it work?

- LOINC codes come in pairs
  - One with SI units ("substance concentration") and one with "conventional" ("mass concentration") UCUM units
- So, we can easily find the **"target" SI unit** from the "conventional" LOINC code
- Do UCUM conversion from **source** (original) units to **target** (standardized) units, using our **RESTful web service**  
(soon to be deployed at NLM servers)
- Store the generated **LB** SDTM records to file in CDISC Dataset-XML format
- Embed the original FHIR record into the SDTM record
  - Not possible using SAS-XPT ...
- Additionally, generate the **DM** SDTM dataset from the "Patient" resource (easy ...)

# The Conventional to SI unit conversion (and vice versa) RESTful web service

MW = Molecular Weight



The RESTful web service will soon be available by the NLM

# Example Conventional to SI unit conversion

## LOINC 21305-8 (Glucose [Mass/Vol] in 24 hours Urin)

- Value provided: 15 mg/dL
- LOINC code: **21305-8**
- Component / Analyte: Glucose => **LP14635-4** => MW: **180.2**
- Property: mass concentration (so ... "conventional")
- Corresponding molar concentration ("SI") LOINC code: **25916-8**
- Standard SI unit molar: **mmol/L**
- 15 mg/dL (conventional) =  $15 \cdot 10^{-3} \text{ g} / 10^{-2} \text{ L} = 15 \cdot 10^{-3} \cdot (180.2)^{-1} \text{ mol} / 10^{-2} \text{ L}$   
= **0.83 mmol/L**
- This kind of calculations can easily be automated

# Example Conventional to SI unit conversion LOINC 21305-8 (Glucose in 24 hours Urin)

```
▼ <UCUMWebServiceResponse ServerDateTime="2020-02-16T13:09:55">
  ▼ <WebServiceRequest>
    http://xml4pharmaserver.com:8080/UCUMService2/rest/ucumtransform/15/from/mg/dL/to/mmol/L/LOINC/21305-8
  </WebServiceRequest>
  ▼ <Response>
    <SourceQuantity>15.0</SourceQuantity>
    <SourceUnit>mg/dL</SourceUnit>
    <TargetUnit>mmol/L</TargetUnit>
    <MolecularWeightUsed>180.2</MolecularWeightUsed>
    <ResultQuantity>0.83240844</ResultQuantity>
  </Response>
</UCUMWebServiceResponse>
```

# Example Results

- 17 **Urinalysis** tests
- EHR record retrieval from Synthea EHR demo repository by RESTful Web Service (**FHIR-API**)
- 14,000 records retrieved and transformed to SDTM
  - Resulting in an SDTM-LB dataset of 15,000 records
- FHIR source "Observation" embedded within the LB record
  - Using CDISC Dataset-XML format
- Quantitative results standardized to **SI** units (LBSTRESC/LBSTRESN)
- For 32 unique subjects => SDTM-**DM** dataset





# Example Results

And all that ...

# In less than 5 minutes

# Example results

```
<ItemGroupData ItemGroupOID="IG.LB" data:ItemGroupDataSeq="3">
  <ItemData ItemOID="IT.STUDYID" Value="SyntheaStudy"/>
  <ItemData ItemOID="IT.DOMAIN" Value="LB"/>
  <ItemData ItemOID="IT.USUBJID" Value="04d8f766-99f8-4a71-9cac-17a730cf93f6"/>
  <ItemData ItemOID="IT.LBSEQ" Value="3"/>
  <ItemData ItemOID="IT.LBTESTCD" Value="BILI"/>
  <ItemData ItemOID="IT.LBTEST" Value="Bilirubin"/>
  <ItemData ItemOID="IT.LBCAT" Value="URINANALYSIS"/>
  <ItemData ItemOID="IT.LBORRES" Value="0.6546"/>
  <ItemData ItemOID="IT.LBORRESU" Value="mg/dL"/>
  <ItemData ItemOID="IT.LBSTRESC" Value="11.20"/>
  <ItemData ItemOID="IT.LBSTRESN" Value="11.20"/>
  <ItemData ItemOID="IT.LBSTRESU" Value="umol/L"/>
  <ItemData ItemOID="IT.LBLOINC" Value="20505-4"/>
  <ItemData ItemOID="IT.LBSPEC" Value="URINE"/>
  <ItemData ItemOID="IT.LBMETHOD" Value="TEST STRIP"/>
  <ItemData ItemOID="IT.LBDTC" Value="1975-11-12T16:29:46-05:00"/>
  <ItemData ItemOID="IT.LBDY" Value="1"/>
  <ItemData ItemOID="IT.LBEVLINT" Value="PT"/>
</ItemGroupData>
<Observation xmlns="http://hl7.org/fhir">
  <id value="4b51ff91-7892-4090-ae27-25984049d43c"/>
  <meta>
    <lastUpdated value="2019-04-09T12:04:23.039817+00:00"/>
    <versionId value="MTU1NDgxMTQ2MzAzOTgxNzAwMA"/>
  </meta>
  <fhir:text xmlns:fhir="http://hl7.org/fhir" xmlns:xhtml="http://www.w3.org/1999/xhtml">
    <fhir:status value="generated"/>
    <xhtml:div>
      <xhtml:p>Status: final</xhtml:p>
      <xhtml:p>Test: Bilirubin.total [Mass/volume] in Urine by Test strip (LOINC 20505-4)</xhtml:p>
      <xhtml:p>Subject: Patient/04d8f766-99f8-4a71-9cac-17a730cf93f6</xhtml:p>
      <xhtml:p>Encounter: Encounter/13b98546-4519-4d41-9386-320a83ee8bd6</xhtml:p>
    </xhtml:div>
  </fhir:text>
</Observation>
```

# Example results

STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBORRES	LBORRESU
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	1	COLOR	Color	MISC	Reddish color urine	
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	2	GLUC	Glucose	URINANALYSIS	Urine glucose test = ++ (finding)	
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	3	BILI	Bilirubin	URINANALYSIS	0.6546	mg/dL
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	4	BILI	Bilirubin	URINANALYSIS	Finding of bilirubin in urine (finding)	
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	5	KETONES	Ketones	URINANALYSIS	16.62	mg/dL

STUDYID	DOMAIN	USUBJID	LBSEQ	LBTESTCD	LBTEST	LBCAT	LBSTRESC	LBSTRESN	LBSTRESU	LBLOINC	LBSPEC	LBMETHOD	LBEVLINT
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	1	COLOR	Color	MISC	Reddish color urine			5778-6	URINE		PT
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	2	GLUC	Glucose	URINANALYSIS	Urine glucose test...			25428-4	URINE	TEST STRIP	PT
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	3	BILI	Bilirubin	URINANALYSIS	11.20	11.20	umol/L	20505-4	URINE	TEST STRIP	PT
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	4	BILI	Bilirubin	URINANALYSIS	Finding of bilirubi...			5770-3	URINE	TEST STRIP	PT
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	5	KETONES	Ketones	URINANALYSIS	284.3	284.3	mmol/L	5797-6	URINE	TEST STRIP	PT

Bilirubin: 0.6546 mg/dL (original) = 11.2 umol/L (standardized)

# Example results

STUDYID	DOMAIN	USUBJID ▲	LBSEQ	LBTESTCD	LBTEST	LBCAT
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	1	COLOR	Color	MISC
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	2	GLUC	Glucose	URINANAL...
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	3	BILI	Bilirubin	URINANAL...
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	4	BILI	Bilirubin	URINANAL...
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6 (USUBJID)				URINANAL...
SyntheaStudy	LB	<b>FHIR record:</b> Status: final Test: Bilirubin.total [Mass/volume] in Urine by Test strip (LOINC 20505-4) Subject: Patient/04d8f766-99f8-4a71-9cac-17a730cf93f6 Encounter: Encounter/13b98546-4519-4d41-9386-320a83ee8bd6 Date/Time of observation: 1975-11-12T16:29:46-05:00 <b>Result:</b> Result value: 0.6545735366462426 mg/dL (UCUM: mg/dL)				URINANAL...
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	5	GLUC	Glucose	URINANAL...
SyntheaStudy	LB	04d8f766-99f8-4a71-9cac-17a730cf93f6	6	BILI	Bilirubin	URINANAL...

FHIR source record embedded in SDTM-LB record  
 (visualized by the "Smart Submission Dataset Viewer")



## Future work



# Future work

- Implement SDTM-IG 3.3 / SDTM model v.1.7
- Use FHIR resources "**ResearchStudy**" and "**ResearchSubject**" for selecting the right subjects for the study
- Try out with new (preliminary) Corona Virus LOINC codes
- Extend to SDTM domains VS, MB, QS
  - LOINC contains a large number of tests for these domains but a mapping to SDTM is not available yet
- A similar approach would also work for MH and CM (no LOINC needed)



# Conclusions

# Conclusions

- We need **LOINC codes** for retrieving information from EHRs and transformation to SDTM records
  - And not only for lab tests
- Only the combination of LOINC with **UCUM notation** for units allows for automated Conventional ↔ SI unit conversion
  - Impossible using "CDISC units"
- Personal opinion: FDA and CDISC should start allowing UCUM notation for units in submissions
  - CDISC "UNIT" has too many limitations



# Thank You!

Contact: [Jozef.Aerts@XML4Pharma.com](mailto:Jozef.Aerts@XML4Pharma.com)

