

Creating CORE Rules from Biomedical Concepts

An abstract graphic on the right side of the slide. It features two hands cupping a bright, glowing yellow sphere. Surrounding the sphere are various geometric shapes: circles, squares, and hexagons, some of which are connected by thin lines, suggesting a network or data structure. The background is a dark blue gradient.

6th November 2023
phuse EU Connect, Birmingham
Johannes Ulander

Agenda

- Why do a Core rule from a BC?
- Process
- Comparison
CORE rule vs. Biomedical Concept
- Summary

What is a Systolic Blood Pressure?

Why do a Core rule from a BC?

- Proof of concept
- Different perspectives on the same data
- Core rule
 - Designed to check conformance, that the correct metadata is stored together on a row in a dataset
- Biomedical Concept
 - Identifiable: Has an identifier, unique
 - Complete: Everything is defined
 - Atomic: If it is split it loses meaning
 - Data Specification
 - Specification of the data, not how it is used with a particular technology

Why do a Core rule from a BC?

- Proof of concept
- Different perspectives on the same data
- Core rule
 - Designed to check conformance, that the correct metadata is stored together on a row in a dataset
- Biomedical Concept
 - Identifiable: Has an identifier, unique
 - **Complete: Everything is defined**
 - **Atomic: If it is split it loses meaning**
 - **Data Specification**
 - Specification of the data, not how it is used with a particular technology



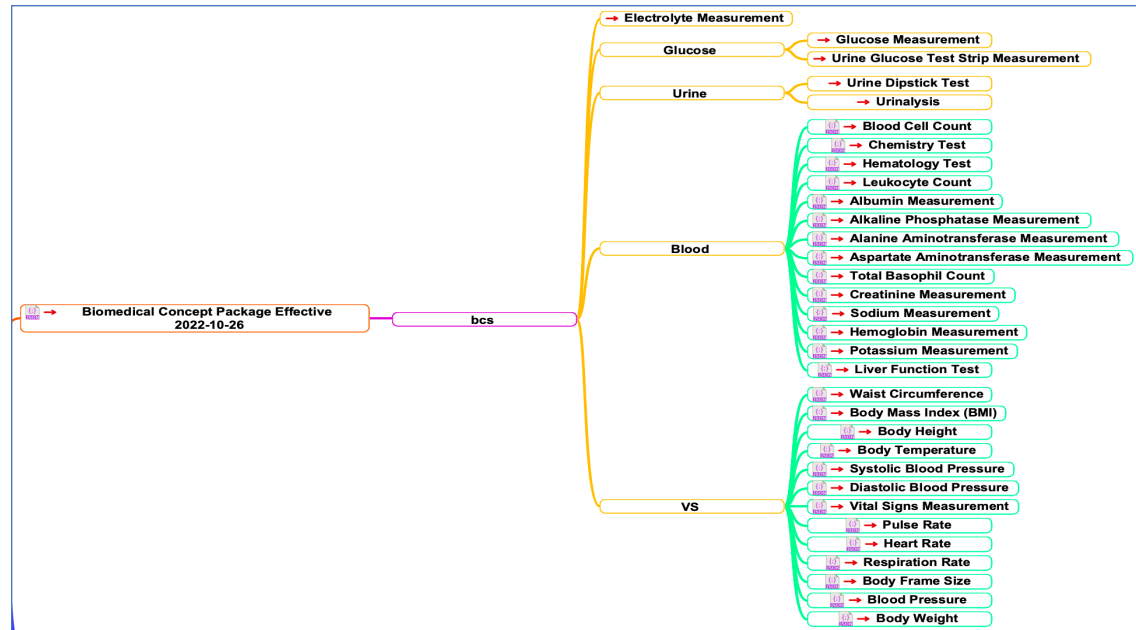
So the Core rule should be inside the BC?

Process

- 2022-10-26: First set of BC's available via CDISC Library API
 - Downloaded BC's as JSON from the CDISC Library API
 - Using python
- 2022-11-14: Attended Core Workshop at Phuse, Belfast
- On the bus trip from the conference to the airport (around 2 hours)
 - Converted the downloaded BC's to Core rules
 - Using Groovy/Java
 - Add some additional info linked from the BC (NCI concepts)
 - Created some test data
 - In csv format

When I got internet again...

- Uploaded Core rules and test data to Core test server
- I had working Core Rules for all BC's specifying --UNIT and --SPEC



Take a step back: Why do we need Core rules?

Example: Height – and the current state of definitions

Data

VSTESTCD	VSTEST	VSCAT	VSORRES	VSORRESU	VSPOS
HEIGHT	Height	BODY	70	in	

Specification

(<u>VSTESTCD</u>)	(<u>VSTEST</u>)	*	No terminology requirement	(<u>VSRESU</u>)	(<u>POSITION</u>)
ABSKNF BMI BMR BSA WEIGHT HEIGHT PULSE SYSBP DIABP ...	<i>Too long to show, but you know them</i>	Sponsor defined		m beats/min breaths/min C cm cmHg in mmHg Pascal ...	SITTING STANDING DECUBITUS FOWLERS LATERAL DECUBITUS LEFT LATERAL DECUBITUS ...

Systolic Blood Pressure

Core rule

vs

Biomedical Concept

```

---
Authority:
  Organization: "CDISC"
Check:
  all:
    - name: "VSTESTCD"
      operator: "equal_to"
      value: "SYSBP"
    - name: "VSORRESU"
      operator: "is_not_contained_by"
      value:
        - "cmHg"
        - "mmHG"
        - "Pascal"
Citations: ...
Core: ...
Description: "Systolic BP unit is incorrect"
Outcome:
  Message: "SYSBP unit is not allowed"
  Output Variables:
    - "VSTESTCD"
    - "VSORRESU"
References: ...
Rule Type: "Data Pattern and Format"
Scopes:
  Classes:
    Include:
      - "Findings"
  Domains:
    Include:
      - "VS"
  Standards: ...
Sensitivity: "Record"
Severity: "Warning"
  
```

VSTESTCD = SYSBP

VSORRESU is not:

- cmHg
- mmHg or
- Pascal

Output the error

(<u>VSTESTCD</u>)	(<u>VSTEST</u>)	*	No terminology requirement	(<u>VSRESU</u>)
ABSKNF BMI BMR BSA WEIGHT HEIGHT PULSE SYSBP DIABP ...	<i>Too long to show, but you know them</i>	Sponsor defined		m beats/min breaths/min C cm cmHg in mmHg Pascal ...

Systolic Blood Pressure

vs

Biomedical Concept

Core rule

```
---
Authority:
  Organization: "CDISC"
Check:
  all:
    - name: "VTESTCD"
      operator: "equal_to"
      value: "SYSBP"
    - name: "VSORRESU"
      operator: "is_not_contained_by"
      value:
        - "cmHg"
        - "mmHG"
        - "Pascal"
Citations: ***
Core: **
Description: "Systolic BP unit is incorrect"
Outcome:
  Message: "SYSBP unit is not allowed"
  Output Variables:
    - "VTESTCD"
    - "VSORRESU"
References: ***
Rule Type: "Data Pattern and Format"
Scopes:
  Classes:
    Include:
      - "Findings"
    Domains:
      Include:
        - "VS"
      Standards: ***
Sensitivity: "Record"
Severity: "Warning"
```

```
{
  "conceptid": "C25298",
  "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C25298",
  "category": [
    "Vital Signs"
  ],
  "shortName": "Systolic Blood Pressure",
  "synonym": [
    "SYSBP"
  ],
  "resultScale": "Quantitative",
  "definition": "The maximum pressure exerted into the systemic arterial circulation during the contraction of the left ventricle of the heart.",
  "coding": [
    {
      "code": "9480-6",
      "system": "http://hl7.org",
      "systemName": "LOINC"
    }
  ],
  "dataElementConcepts": [
    {
      "conceptid": "C173522",
      "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C173522",
      "shortName": "Vital Signs Result",
      "dataType": "string"
    },
    {
      "conceptid": "C49669",
      "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C49669",
      "shortName": "Unit of Pressure",
      "dataType": "string",
      "exampleSet": [
        "cmHg",
        "mmHg",
        "Pascal"
      ]
    },
    {
      "conceptid": "C83088",
      "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C83088",
      "shortName": "Vital Signs Location",
      "dataType": "string",
      "exampleSet": [
        "Arm",
        "Forearm",
        "Thigh",
        "Calf",
        "Artery",
        "Vein"
      ]
    },
    {
      "conceptid": "C123975",
      "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C123975",
      "shortName": "Vital Signs Laterality",
      "dataType": "string",
      "exampleSet": [
        "Left",
        "Right"
      ]
    },
    {
      "conceptid": "C82535",
      "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C82535",
      "shortName": "Vital Signs Method",
      "dataType": "string",
      "exampleSet": [
        "Auscultation automatic",
        "Auscultation manual",
        "Cuff manual palpated",
        "Doppler",
        "Occlusometry",
        "Atrial line",
        "Venous line",
        "Continuous noninvasive arterial pressure (CNAP)"
      ]
    },
    {
      "conceptid": "C83114",
      "href": "https://ncitthesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ncs=ncit&code=C83114",
      "shortName": "Vital Signs Position",
      "dataType": "string",
      "exampleSet": [
        "Decubitus",
        "Dowler",
        "Lateral Decubitus",
        "Prone",
        "Reverse Trendelenburg",
        "Semi-Fowler",
        "Semi-Recurrent",
        "Sitting",
        "Sitting Legs Dependent",
        "Lying",
        "Standing"
      ]
    }
  ]
}
```

Core rule

```
---
Authority:
  Organization: "CDISC"
Check:
  all:
    - name: "VSTESTCD"
      operator: "equal_to"
      value: "SYSBP"
    - name: "VSORRESU"
      operator: "is_not_contained_by"
      value:
        - "cmHg"
        - "mmHG"
        - "Pascal"
Citations: ...
Core: ...
Description: "Systolic BP unit is incorrect"
Outcome:
  Message: "SYSBP unit is not allowed"
  Output Variables:
    - "VSTESTCD"
    - "VSORRESU"
References: ...
Rule Type: "Data Pattern and Format"
Scopes:
  Classes:
    Include:
      - "Findings"
    Domains:
      Include:
        - "VS"
  Standards: ...
Sensitivity: "Record"
Severity: "Warning"
```

Systolic Blood Pressure

vs

Biomedical Concept

```
{
  "conceptId": "C25298",
  "href": "https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C25298",
  "category": [
    "Vital Signs"
  ],
  "shortName": "Systolic Blood Pressure",
  "synonym": [
    "SYSBP"
  ],
  "dataElementConcepts": [
    {
      "conceptId": "C173522",
      "href": "https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C173522",
      "shortName": "Vital Signs Result",
      "dataType": "integer"
    },
    {
      "conceptId": "C49669",
      "href": "https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C49669",
      "shortName": "Unit of Pressure",
      "dataType": "string",
      "exampleSet": [
        "cmHg",
        "mmHG",
        "Pascal"
      ]
    }
  ]
}
```

Identifier

SYSBP is a synonym identifier

List of variables

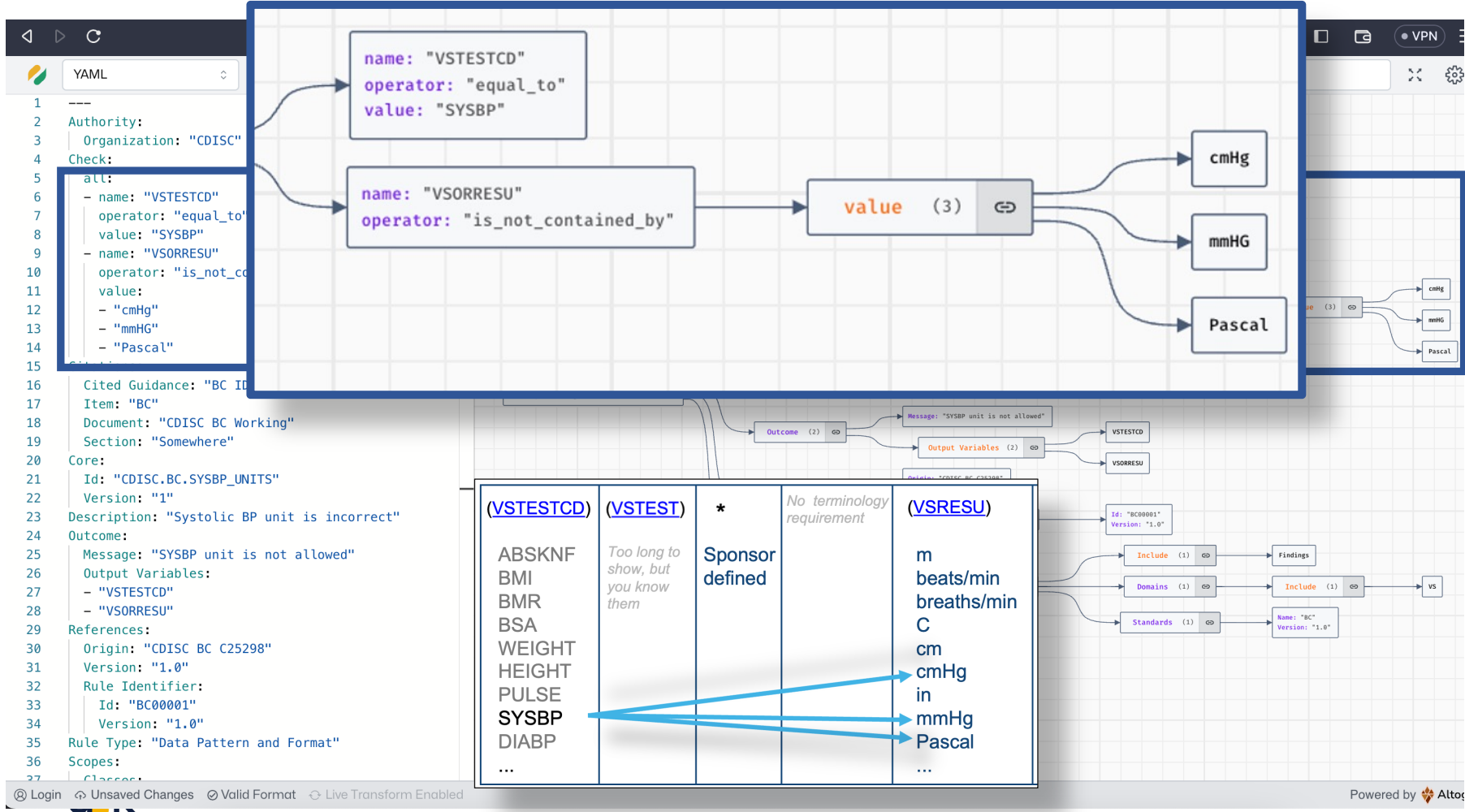
Vital Signs Result = VSORRES

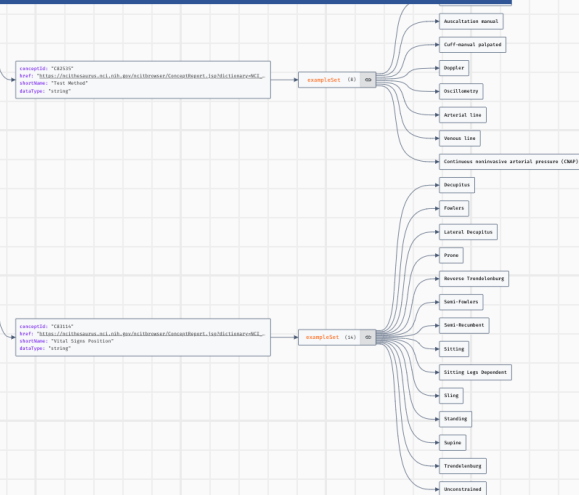
Unit of Pressure = VSORRESU

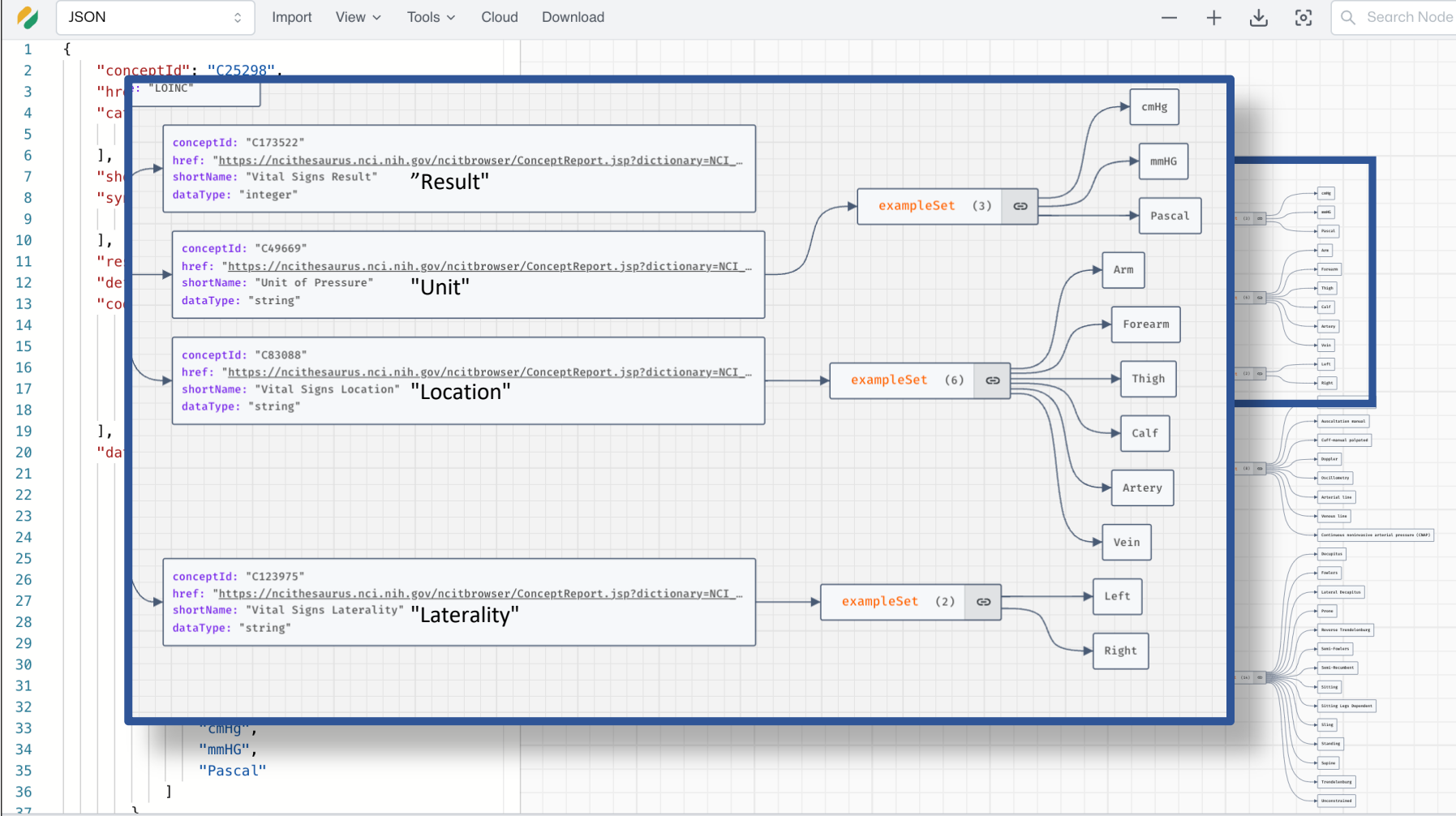
List of units for Unit of Pressure

Let's visualize differently

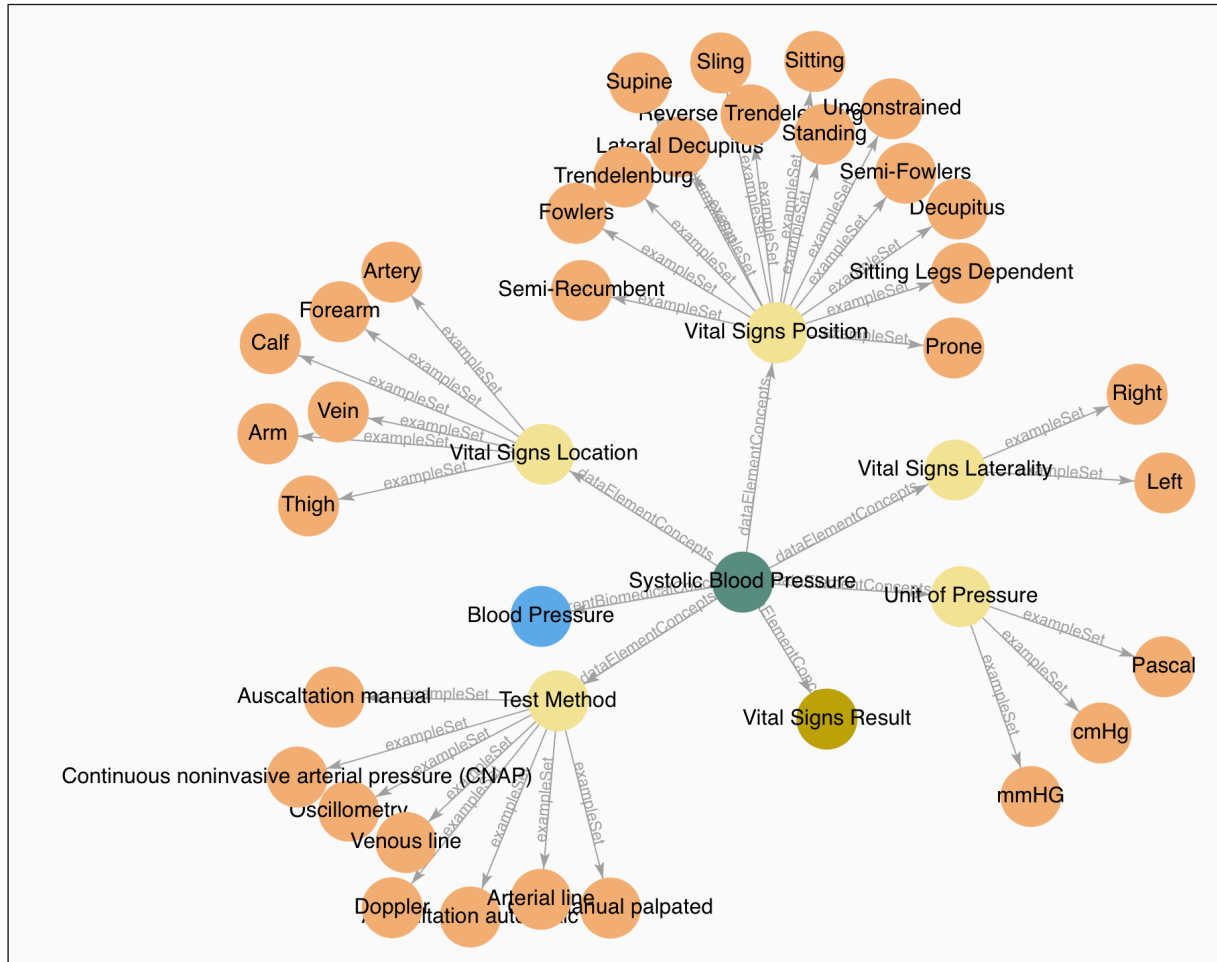
Core rule YAML







Same Biomedical Concept – graph view



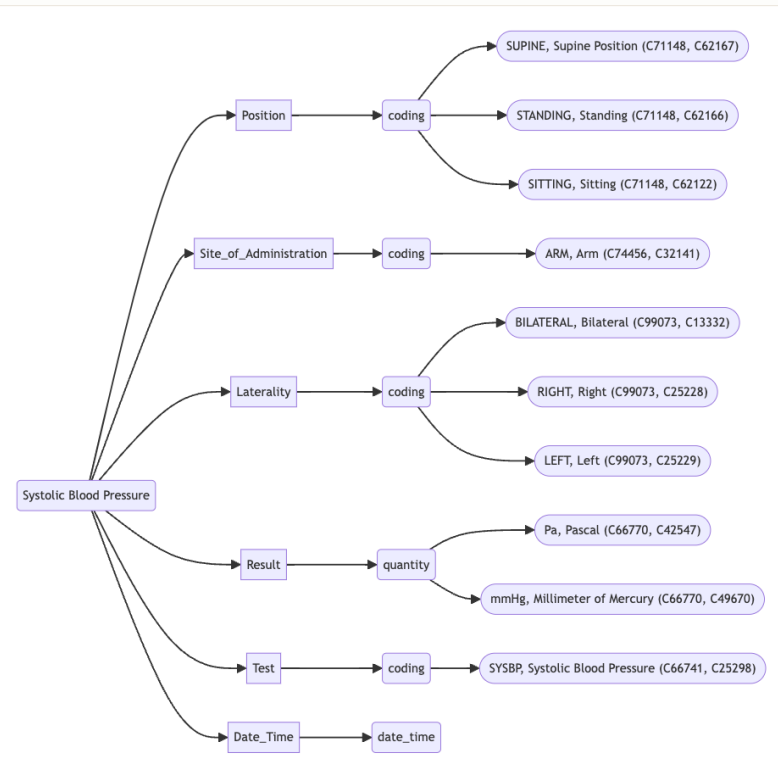
Subset of the same Biomedical Concept– graph and tabular view

SYSTOLIC BLOOD PRESSURE

Owner: d4k, Version: 1.0.0



Graphical View



Tabular View

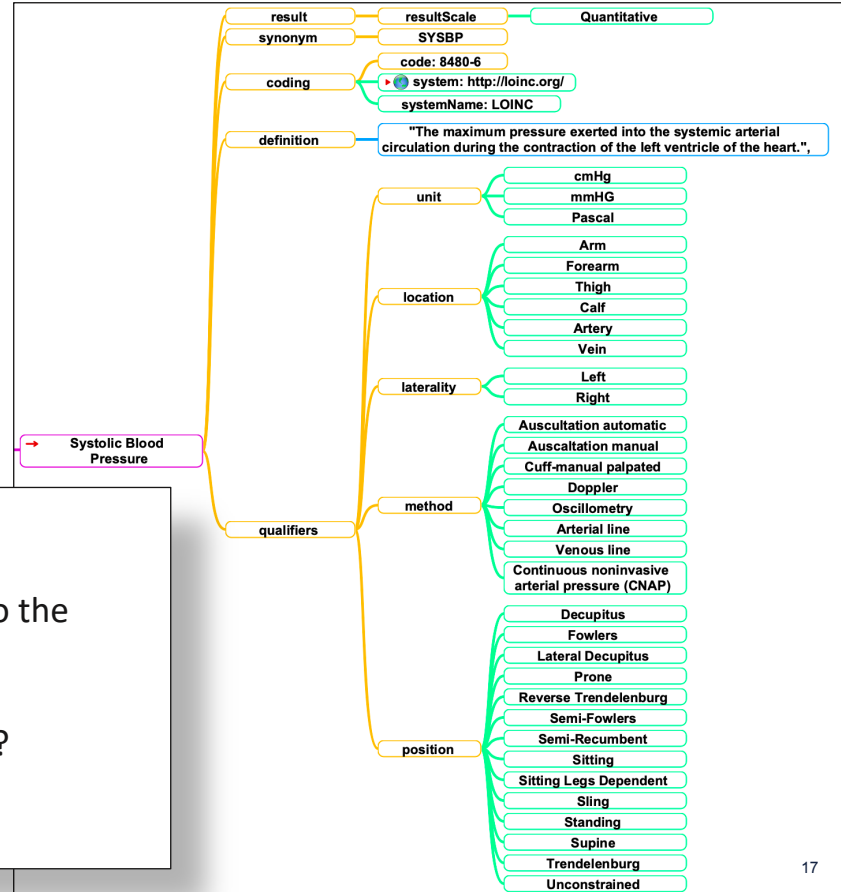
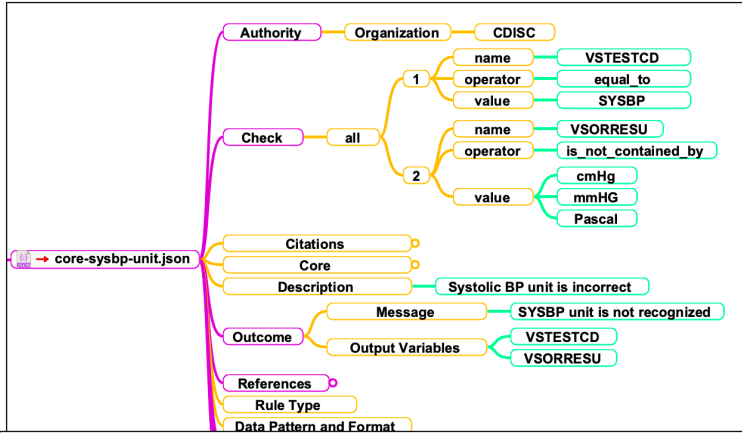
Item Name	Identifier	Data Type	Terms
Position		coding	SUPINE, Supine Position (C71148, C62167) STANDING, Standing (C71148, C62166) SITTING, Sitting (C71148, C62122)
Site of Administration		coding	ARM, Arm (C74456, C32141)
Laterality		coding	BILATERAL, Bilateral (C99073, C13332) RIGHT, Right (C99073, C25228) LEFT, Left (C99073, C25229)
Result		quantity	Pa, Pascal (C66770, C42547) mmHg, Millimeter of Mercury (C66770, C49670)
Test	✓	coding	SYSBP, Systolic Blood Pressure (C66741, C25298)
Date Time		date_time	

Systolic Blood Pressure

Core rule

vs

Biomedical Concept



The Biomedical Concept

- contains all qualifiers and associated terminology
- Core rules can be deduced from the relationships to the qualifiers and their associated terminology

Should we trigger the Core engine from the BC directly?

Can the BC also be displayed as a form?

Create a form from a BC and create SDTM

<https://github.com/jozza/simple-data-template>

The screenshot shows the Simple Data Template (SDTM) web application interface. The interface is divided into several sections:

- Form Section:** At the top, there is a dropdown menu for "Form" (currently set to "Form 1") and two buttons: "Render" and "Hide".
- View Section:** Below the form, there is a "View" section with a "Collected view" tab and a "Save Collected" button, and a "Data" section with a "References resolved" tab and a "Resolve" button.
- Empty Data Area:** A large empty box labeled "Empty" is present below the "References resolved" tab.
- SDTM view Section:** At the bottom, there is an "SDTM view" section with a "Make SDTM" button and a "Hide" button.

Callouts and annotations provide additional context:

- A callout box on the right side of the "Form" and "View" sections contains the following list:
 - Collect data
 - Present SDTM annotations from BC
- A callout box labeled "Transfer collected data" points to the "Make SDTM" button.
- A callout box labeled "Manage complex data" points to a JSON snippet in the "Data" section:

```
{
  "result": {
    "collected": [
      "C41260",
      "C41261"
    ],
    "specification": ""
  }
}
```
- A callout box labeled "Collected data put at its appropriate location in the BC" points to the "Make SDTM" button.
- A callout box labeled "View the collected data as a SDTM" points to the "Make SDTM" button.

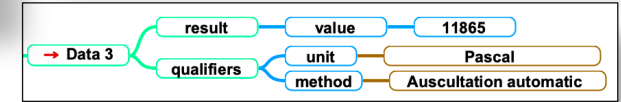
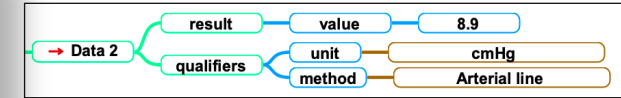
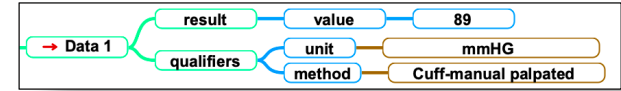
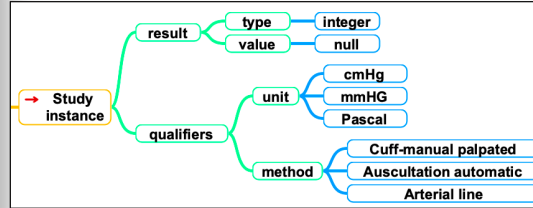
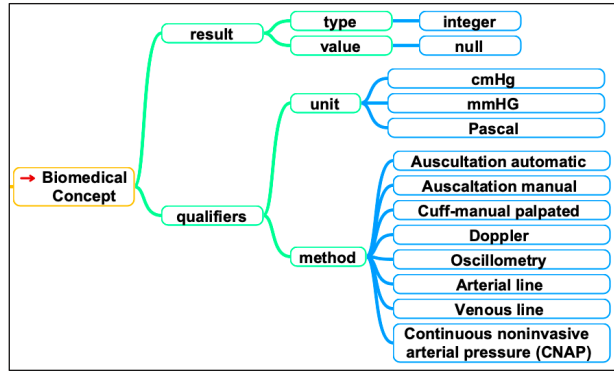
Biomedical Concept



Study instance



Data



- Identifiable: Has an identifier, unique
- Complete: Everything is defined
- Atomic: If it is split it loses meaning
- Data Specification
- **Does not have a result yet**
 - But describes the type

- A copy with its own identifier
- References back to the BC
- References back to the BC
- Specifies the needs of the study
- Qualifiers/terms a subset

- A copy with its own identifier
- References back to the study instance
- References back to the study instance
- Qualifiers/terms have the collected value
- **Result value is populated**

(Compare to Value Level Metadata in define.xml)

Summary

- Biomedical Concepts = Data centrality
 - Let's data speak by itself
 - It is not bound to a tool or system
- Provides specificity
 - Uniquely identified and can represent complexity of data
- Enables Automation
 - Core rules can be deduced
 - Can be presented as define-xml
 - Can be presented as a form for data collection
 - Including SDTM annotation
 - Add study instances to a protocol



Thank You for listening!

Johannes Ulander, data4knowledge ApS

ju@data4knowledge.dk

<https://www.data4knowledge.dk/>

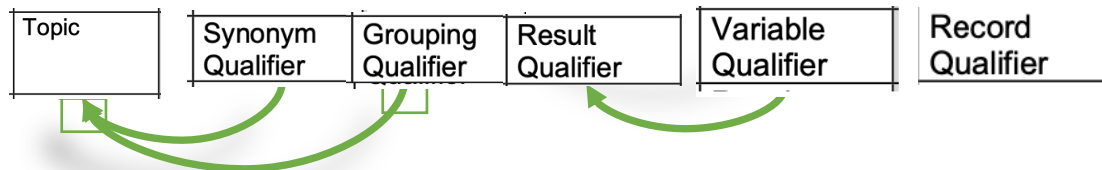
Backup

A step back: BC's kind of already in SDTM

Specification

VSTESTCD	VSTEST	VSCAT	VSORRES	VSORRESU	VSPOS
(<u>VSTESTCD</u>)	(<u>VSTEST</u>)	*	<i>No terminology requirement</i>	(<u>VSRESU</u>)	(<u>POSITION</u>)
ABSKNF BMI BMR BSA WEIGHT HEIGHT PULSE SYSBP DIABP ...	<i>Too long to show, but you know them</i>	* Sponsor defined		m beats/min breaths/min C cm cmHg in mmHg Pascal ...	SITTING STANDING DECUBITUS FOWLERS LATERAL DECUBITUS LEFT LATERAL DECUBITUS ...

SDTM IG
Variable Role



That was my alternative presentation

Example template Systolic Blood Pressure, Supine 100 Unit mmHg Position Supine Location Arm

Pulse 80 Unit beats/min Position Standing Location Left Arm

Race Subject self reports one or many

☐ American Indian or Alaska Native

☒ Asian

☐ Black or African American

☒ White

☐ Native Hawaiian or Other Pacific Islander

☐ Other

Collected view [Save Collected](#) [Hide](#)

id	result	location	position	specification
EXAMPLE-TEMPLATE	100	C32141		
PULSE	80	SPONSOR-LOC1	C62166	
Race	C41260,C41261			

References resolved [Resolve](#) [Hide](#)

"instruction": "Subject self reports one or many"

```
"PULSE": {  
  },  
  "location": "SPONSOR-LOC1",  
  "position": "C62166",  
  "protocol_text": "Pulse Rate"  
}
```

"protocol_text": "Self reported Race of Subjects"

SDTM view [Make SDTM](#) [Hide](#)

STUDYID	DOMAIN	USUBJID	VSSEQ	VSTESTCD	VSTEST	VSCAT	VSPOS	VSORRES	VSORRESU	VSLOC	VSLAT
PILOT01	VS		1	SYSBP	Systolic Blood Pressure	VITAL SIGNS	SUPINE	100	mmHg	ARM	
PILOT01	VS		2	PULSE	Pulse Rate	PULSE AND BLOOD PRESSURE	STANDING	80	beats/min	ARM	LEFT

STUDYID	DOMAIN	USUBJID	SUBJID	RFSTDTCD	RFENDTCD	INVID	INVNAM	BRTHDTC	AGE	AGEU	SEX	RACE	ETHNIC	RACE1	RACE2
PILOT01	DM											MULTIPLE		ASIAN	WHITE

```
"terms": {  
  "SPONSOR-LOC1": {  
    "display": "Left Arm",  
    "submission_value": {  
      "location": { "term": "C32141" },  
      "laterality": { "term": "C25229" }  
    },  
    "definition": "Sponsor specific"  
  },  
  "C25229": {  
    "display": "Left",  
    "submission_value": { "laterality": "LEFT" },  
    "definition": "Being or located on or directed toward the side of the body to the west"  
  },  
  "C32141": {  
    "display": "Arm",  
    "submission_value": { "location": "ARM" },  
    "definition": "The portion of the upper extremity between the shoulder and the elbow."  
  }  
}
```