

# COSA – BC & OpenStudyBuilder Workshop @ EU Interchange 2023

Introduction

Mikkel Traun, Novo Nordisk A/S

25th April 2023

# Workshop Description

## **COSA Biomedical Concept & OpenStudyBuilder Workshop**

At this workshop we will dive into what Biomedical Concepts (BC) is, and how they can be applied within a MDR data standards repository and a SDR study definitions repository – illustrated within the OpenStudyBuilder (OSB) solution.

We will as well relate to how BC's are defined within COSMoS, DDF, d4k and other models. There will be a shared introduction followed by 4 breakout sessions, leading to a shared reflection and discussion on how we can support and bring these initiatives forward. The 4 breakout sessions are currently defined as:

- Setup BC's in OSB SoA for a new study, run various queries to learn how BC's can be utilised
- Learn and understand the BC model in OSB versus the COSMoS, DDF, d4k and other models
- Create and curate OpenStudyBuilder BC content via the OSB Library and NeoDash reports and mining BC's from existing data sources like SDTM.

OpenStudyBuilder: A MDR and SDR open source project ([//novo-nordisk.gitlab.io/nn-public/openstudybuilder/project-description/](https://novo-nordisk.gitlab.io/nn-public/openstudybuilder/project-description/))

COSMoS: Conceptual and Operational Standards Metadata Services CDISC project ([//www.cdisc.org/cosmos](https://www.cdisc.org/cosmos))

DDF: TransCelerate Digital Data Flow project ([//transcelerate.github.io/ddf-home/](https://transcelerate.github.io/ddf-home/))

d4k: data4knowledge BC model ([//d4k.dk](https://d4k.dk))

# COSA – BC & OpenStudyBuilder Workshop @ EU Interchange 2023

9:00 – 9:30 Coffee and welcome

9:30 -11:00 Shared Introduction

11:00 – 12:00 Break-out part 1

12:00 – 13:00 Lunch

13:00 – 14:00 Break-out part 2

14:00 – 15:00 Sharing in plenum

## *Workshop drivers:*

- Anja Lundgreen
- Katja Glass
- Dave Ibersen-Hurst
- Lex Jansen
- Mikkel Traun
- Marius Conjeaud
- Kirsten Langendorf
- Chandrakant J S
- Nicolas de Saint Jorre
- Linda Lander

# Workshop Goals for the COSA BC & OpenStudyBuilder

## For **CDISC & COSA**

- Promote COSA as the Open Source community enabling use and sharing of Open Source projects within the CDISC community
- Promote collaboration on the COSMoS initiative – defining BC's in CDISC Library
- Promote COSA initiatives supporting COSMoS
- Promote CDISC as a standards organisation not only defining data standards, but also facilitating tool development and sharing of these

## For **Vendors**

- Explore and get insights into new business opportunities supporting open source initiatives with a focus on Biomedical Concepts as enabling end-2-end consistency and automation

## For **Participants**

- Be **Trained** and **Explore** how Biomedical Concepts can enable end-2-end consistency and automation
- **Get Insights** into possibilities in using the OpenStudyBuilder tool for defining and applying Biomedical Concepts
- Learn opportunities in **accessing** and **sharing** Biomedical Concepts

## For **Novo Nordisk**

- Learn and get feedback from our community
- Promote development of shared open source tools providing general value for pharma industry
- Get contributions from other pharma companies and technology providers to ensure better IT tools for Novo Nordisk

# Detailed Agenda – Shared Introduction

- Recap what is a BC (Dave)
- BC in OSB := Activity Concepts (Mikkel)
  - OSB definition, why,
  - Overview of OSB model
  - Demo in App
  - Demo in NeoDash report (Kirsten)
- OSB model versus other models
  - Link to COSMoS-BC Model + Terminology (Kirsten, Linda)
  - Link to DDF-BC Model (Dave)
  - Link to d4k Model (Dave)
  - Many representations exist, having different focus, context and purpose + discussion
- Present initial SWOT and Mind Map as input to break-out sessions
- Present break-outs, adjust participation as relevant

# Recap what is a BC (Dave)



# Biomedical Concepts

Dave Iberson-Hurst  
Partner, data4knowledge ApS & CDISC DDF Product Owner

25<sup>th</sup> April 2023





# Collections ...

## FDA'S STUDY DATA TECHNICAL CONFORMANCE GUIDE

Contains Nonbinding Recommendations

### Appendix: Data Standards and Interoperable Data Exchange

This appendix provides some of the guiding principles for the Agency's long-term study data standards management strategies. An important goal of standardizing study data submissions is to achieve an acceptable degree of *semantic interoperability* (discussed below). This appendix describes different types of interoperability and how data standards can support interoperable data exchange now and in the future.

At the most fundamental level, study data can be considered a collection of data elements and their relationships. A data element is the smallest (or *atomic*) piece of information that is useful for analysis (e.g., a systolic blood pressure measurement, a lab test result, a response to a question on a questionnaire).

A data value is by itself meaningless without additional information about the data (so called *metadata*). Metadata is often described as *data about data*. Metadata is structured information that describes, explains, or otherwise makes it easier to retrieve, use, or manage data.<sup>48</sup> For example, the number *44* itself is meaningless without an association with Hematocrit and the unit of measurement (e.g. "%"). Hematocrit in this example is metadata that further describes the data.

Just as it is important to standardize the representation of data (e.g., M and F for male and female, respectively), it is equally important to standardize the metadata. The expressions Hematocrit = 44; Hct = 44, or Hct Lab Test = 44 all convey the same information to a human, but an information system or analysis program will fail to recognize that they are equivalent because the metadata is not standardized. It is also important to standardize the definition of the metadata, so that the meaning of a Hematocrit value is constant across studies and submissions.

In addition to standardizing the data and metadata, it is important to capture and represent relationships (also called associations) between data elements in a standard way. Relationships between data elements are critical to understand or interpret the data. Consider the following information collected on the same day for one subject in a study:

Systolic Blood Pressure = 90 mmHg  
Position = standing  
Systolic Blood Pressure = 110 mmHg  
Time = 10:23 a.m.  
Time = 10:20 a.m.

<sup>48</sup> Metadata is said to "give meaning to data" or to put data "in context." Although the term is now frequently used to refer to XML (extensible markup language) tags, there is nothing new about the concept of metadata. Data about a library book such as author, type of book, and the Library of Congress number, are metadata and were once maintained on index cards. SAS labels and formats are a rudimentary form of metadata, although they have not historically been referred to as metadata.

December 2014



Contains Nonbinding Recommendations

### Appendix A: Data Standards and Interoperable Data Exchange

This appendix provides some of the guiding principles for the Agency's long-term study data standards management strategies. An important goal of standardizing study data submissions is to achieve an acceptable degree of *semantic interoperability* (discussed below). This appendix describes different types of interoperability and how data standards can support interoperable data exchange now and in the future.

At the most fundamental level, study data can be considered a collection of data elements and their relationships. A data element is the smallest (or *atomic*) piece of information that is useful for analysis (e.g., a systolic blood pressure measurement, a lab test result, a response to a question on a questionnaire).

A data value is by itself meaningless without additional information about the data (so called *metadata*). Metadata is often described as *data about data*. Metadata is structured information that describes, explains, or otherwise makes it easier to retrieve, use, or manage data.<sup>49</sup> For example, the number *44* itself is meaningless without an association with Hematocrit and the unit of measurement (e.g. "%"). Hematocrit in this example is metadata that further describes the data.

Just as it is important to standardize the representation of data (e.g., M and F for male and female, respectively), it is equally important to standardize the metadata. The expressions Hematocrit = 44; Hct = 44, or Hct Lab Test = 44 all convey the same information to a human, but an information system or analysis program will fail to recognize that they are equivalent because the metadata is not standardized. It is also important to standardize the definition of the metadata, so that the meaning of a hematocrit value is constant across studies and submissions.

In addition to standardizing the data and metadata, it is important to capture and represent relationships (also called associations) between data elements in a standard way. Relationships between data elements are critical to understand or interpret the data. Consider the following information collected on the same day for one subject in a study:

Systolic Blood Pressure = 90 mmHg  
Position = standing  
Systolic Blood Pressure = 110 mmHg  
Time = 10:23 a.m.  
Time = 10:20 a.m.  
Position = lying

<sup>49</sup> Metadata is said to "give meaning to data" or to put data "in context." Although the term is now frequently used to refer to XML (extensible markup language) tags, there is nothing new about the concept of metadata. Data about a library book such as author, type of book, and the Library of Congress number, are metadata and were once maintained on index cards. SAS labels and formats are a rudimentary form of metadata, although they have not historically been referred to as metadata.

March 2023

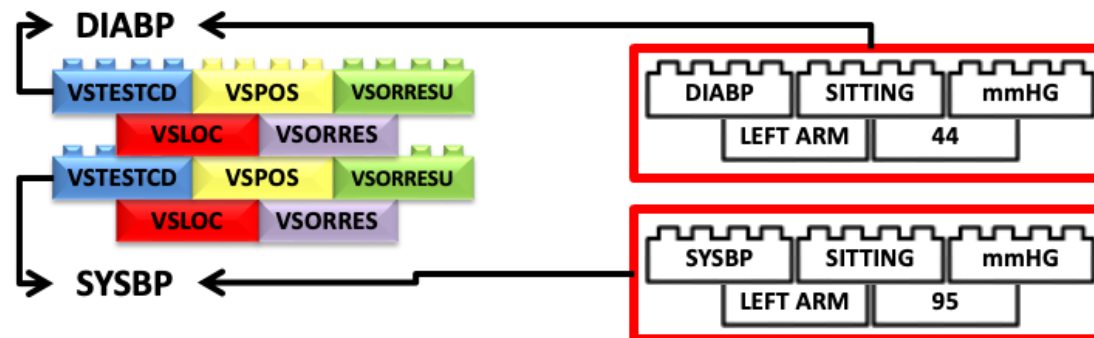


# Collections ...

... and the data ...

The data only make sense when we link it to the meta data

USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSLOC	VSORRES	VSORRESU
ABC-001-001	12	SYSBP	Systolic Blood Pressure	SITTING	LEFT ARM	95	mmHg
ABC-001-001	13	DIABP	Diastolic Blood Pressure	SITTING	LEFT ARM	44	mmHg



28

San Diego 2009

45th Annual Meeting

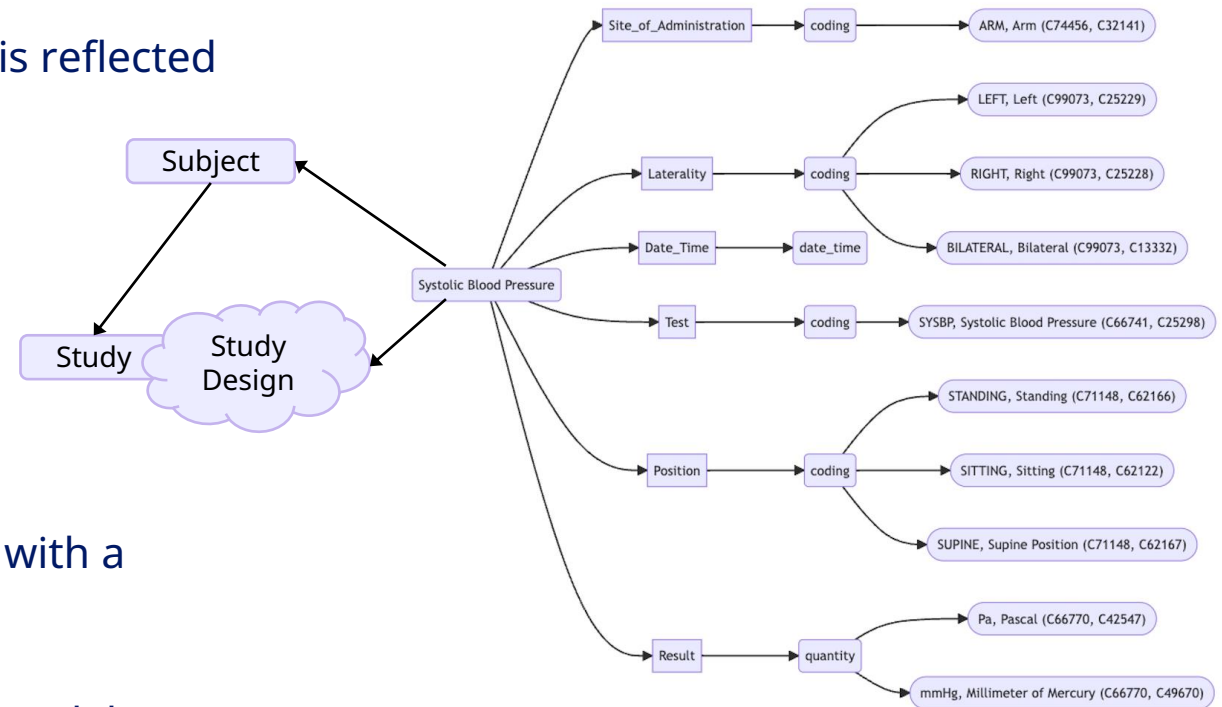
CDISC View of Goals of the HL7 CDISC Relationship and Projects

Dave Ibersen-Hurst  
VP Technical Strategy  
CDISC

CDISC  
DRUG INFORMATION ASSOCIATION

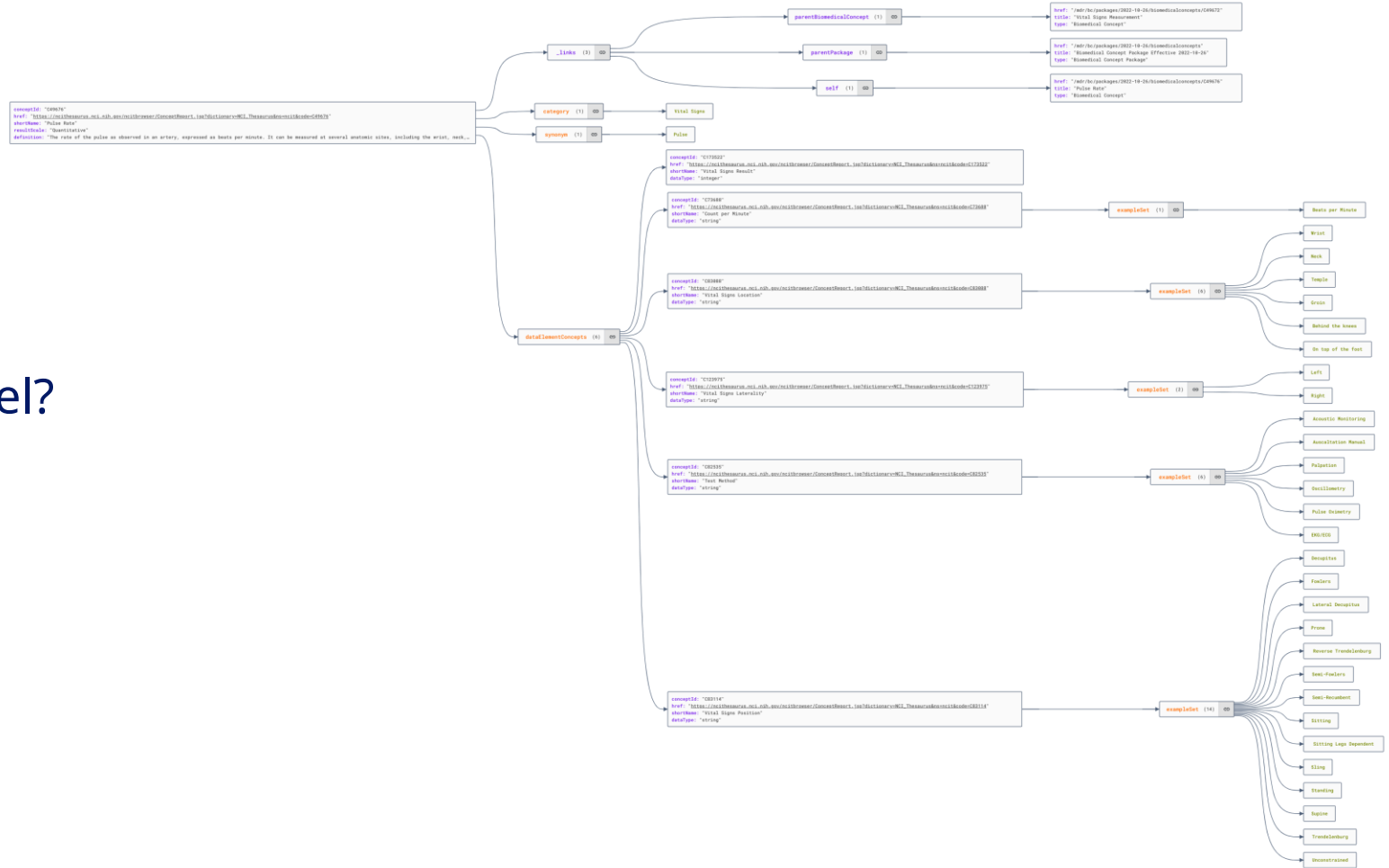
# A Biomedical Concept is ...

- A small model that defines a clinical concept in a standardized and reusable manner
- Atomic:
  - If it is split it loses meaning
  - Refers more to the data based on a BC but is reflected in the model
- Identifiable:
  - Has an identifier, unique
  - Find it, Reference it, Deploy it
- Complete:
  - Everything is defined
- Data Specification
  - Specification of the data, not how it is used with a particular technology
- Context:
  - A BC needs context, i.e the rest of the DDF model, a study, the encounters, activities, timing ...



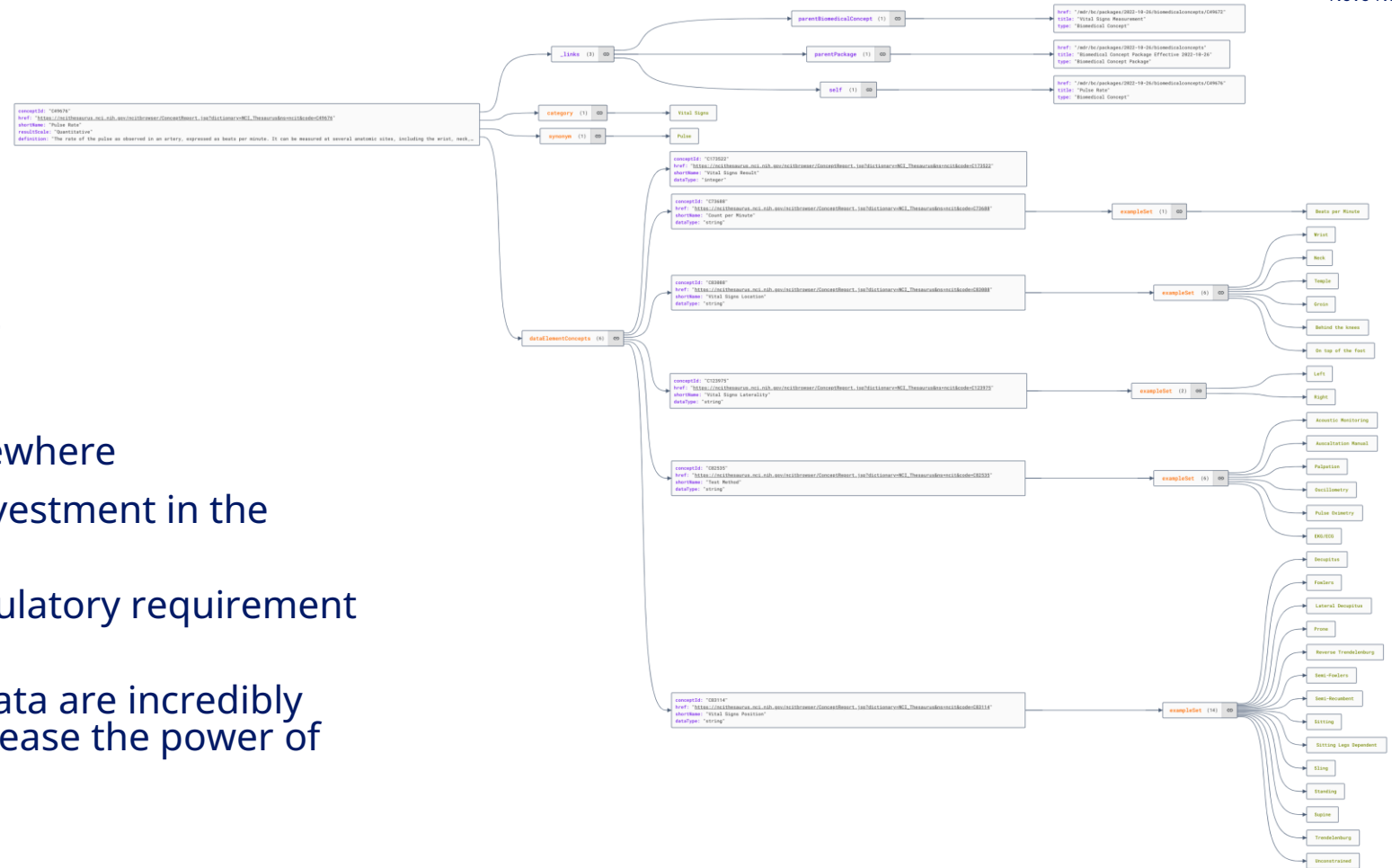
# BC Design

- Are CDISC BCs Unique?
- Are we reinventing the wheel?



# BC Design

- Are CDISC BCs Unique?
  - No
- Are we reinventing the wheel?
  - Yes and No
    - Clinical Models exist elsewhere
  - But ... Industry has a big investment in the current CDISC standards
    - SDTM Standard: Is a regulatory requirement for submission
    - SDTM Datasets: These data are incredibly valuable and BCs can release the power of all this "old" data
  - CT is a major challenge
    - SDTM based on CDISC CT
  - Healthcare and Research have different needs
  - In charge of your own destiny



# OHDSI

- Observational Health Data Sciences and Informatics
- Using SNOMED CT
- Hierarchy

**ATHENA** SEARCH DOWNLOAD LOGIN ?

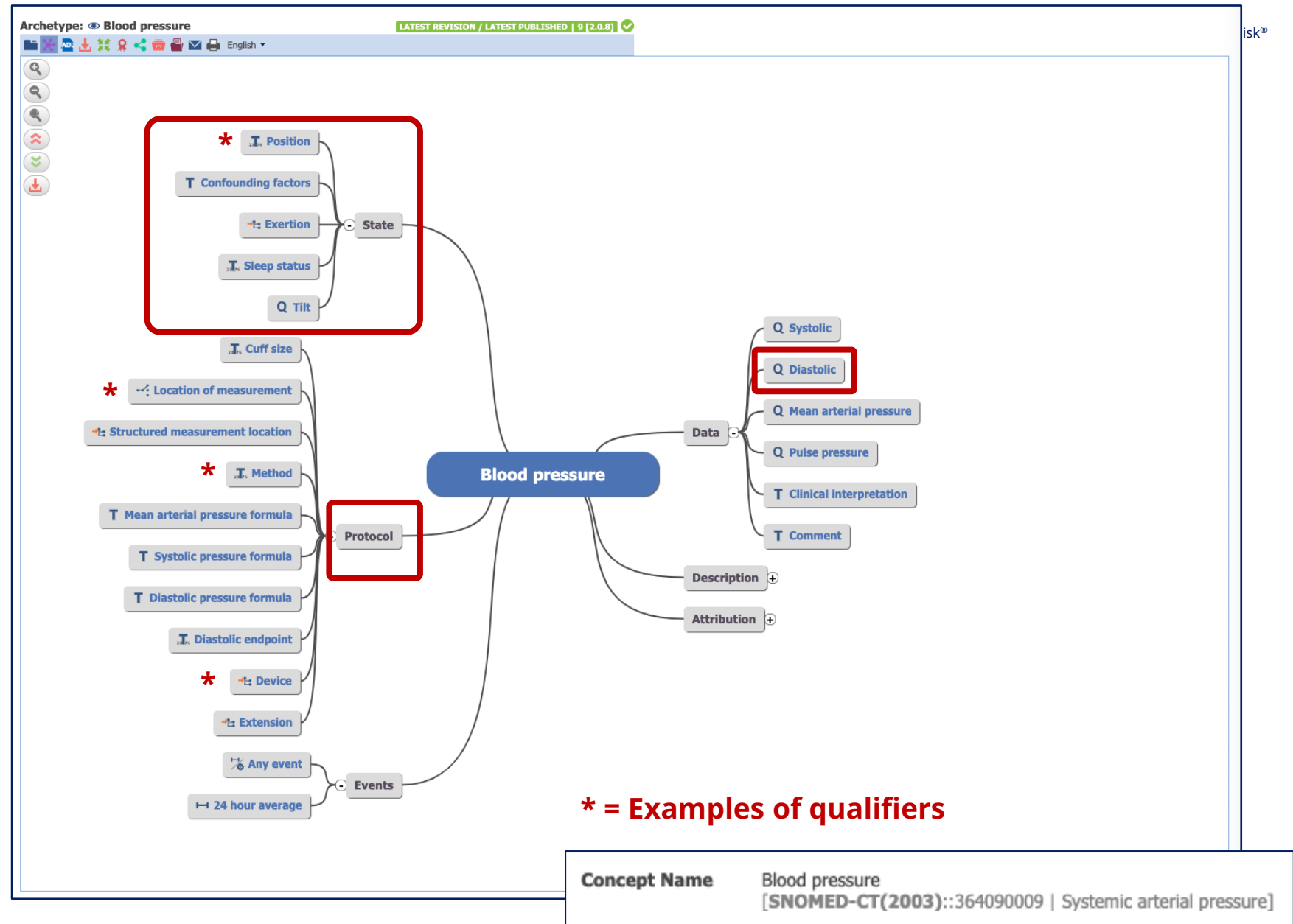
← Diastolic blood pressure

DETAILS	
Domain ID	Measurement
Concept Class ID	Observable Entity
Vocabulary ID	SNOMED <span>?</span>
Concept ID	4154790
Concept code	271650006
Validity	Valid
Concept	Standard
Synonyms	Diastolic blood pressure (observable entity) DAP - Diastolic arterial pressure Diastolic arterial pressure DBP - Diastolic blood pressure
Valid start	31-Jan-2002
Valid end	31-Dec-2099

TERM CONNECTIONS (32)				
		HIERARCHY	RELATED CONCEPTS	
RELATIONSHIP	RELATES TO	CONCEPT ID	VOCABULARY	
Active possibly_equivalent_to inactive (SNOMED)	Non-invasive diastolic arterial pressure	40454826	SNOMED	
Active same_as inactive (SNOMED)	Diastolic arterial pressure	40620464	SNOMED	
Characterizes (SNOMED)	Diastolic phase	36716961	SNOMED	
Has Module	SNOMED CT core	40642539	SNOMED	
Has property	Pressure	36716283	SNOMED	
Has scale type	Quantitative	4149267	SNOMED	
Has status	Defined	40642537	SNOMED	
Inheres in (SNOMED)	Structure of cardiovascular system	4014241	SNOMED	
Interprets of (SNOMED)	Abnormal diastolic arterial pressure	4177952	SNOMED	
	Decreased diastolic arterial pressure	4175340	SNOMED	
	Increased diastolic arterial pressure	4047613	SNOMED	
	Normal diastolic arterial pressure	4204973	SNOMED	
Is a	Blood pressure	4326744	SNOMED	
	Vascular measure	4092664	SNOMED	
Non-standard to Standard map (OMOP)	Diastolic blood pressure	4154790	SNOMED	
Standard to Non-standard map (OMOP)	Diastolic arterial pressure	3162060	Nebraska Lexicon	
	Diastolic arterial pressure	3464102	Nebraska Lexicon	
	Diastolic arterial pressure	40620464	SNOMED	
	Diastolic blood pressure	4154790	SNOMED	
	Diastolic blood pressure	45906295	CIEL	

# openEHR

- Using SNOMED CT
- Strong content



# HL7 FHIR Resource

- Using various CTs
  - LOINC
  - UCUM
  - SNOMED
  - Others ...
- Uses FHIR data types
  - Binding of result value and units
- Strong structure, less content focused

```

79   "component": [
80     {
118   },
119   {
120     "code": {
121       "coding": [
122         {
123           "system": "http://loinc.org",
124           "code": "8462-4",
125           "display": "Diastolic blood pressure"
126         }
127       ]
128     },
129     "valueQuantity": {
130       "value": 60,
131       "unit": "mmHg",
132       "system": "http://unitsofmeasure.org",
133       "code": "mm[Hg]"
134     },
135     "interpretation": [
136       {
137         "coding": [
138           {
139             "system": "http://terminology.hl7.org/CodeSystem",
140             "code": "L",
141             "display": "low"
142           }
143         ],
144         "text": "Below low normal"
145       }
146     ]
147   }
148 ]
149 }

```

```

1  {
2  "resourceType": "Observation",
3  "id": "blood-pressure",
4  "meta": {
5    "profile": [
6      "http://hl7.org/fhir/StructureDefinition/vitalsigns"
7    ]
8  },
9  "text": {
10   }
11 },
12 "identifier": [
13 ],
14 "basedOn": [
15 ],
16 "status": "final",
17 "category": [
18 ],
19 "code": {
20   "coding": [
21     {
22       "system": "http://loinc.org",
23       "code": "85354-9",
24       "display": "Blood pressure panel with all children optional"
25     }
26   ],
27   "text": "Blood pressure systolic & diastolic"
28 },
29 "subject": {
30   "effectiveDateTime": "2012-09-17",
31   "performer": [
32     ]
33   },
34   "interpretation": [
35     ]
36   },
37   "bodySite": {
38     "coding": [
39       {
40         "system": "http://snomed.info/sct",
41         "code": "368209003",
42         "display": "Right arm"
43       }
44     ]
45   },
46   "component": [
47     {
48       "code": {
49         "coding": [
50           {
51             "system": "http://loinc.org",
52             "code": "8462-4",
53             "display": "Diastolic blood pressure"
54           }
55         ]
56       },
57       "valueQuantity": {
58         "value": 60,
59         "unit": "mmHg",
60         "system": "http://unitsofmeasure.org",
61         "code": "mm[Hg]"
62       },
63       "interpretation": [
64         {
65           "coding": [
66             {
67               "system": "http://terminology.hl7.org/CodeSystem/v3-ObservationInterpretation",
68               "code": "L",
69               "display": "low"
70             }
71           ],
72           "text": "Below low normal"
73         }
74       ]
75     }
76   ]
77 }
78 ]
79 }
80 }

```



# LOINC

- Uses LOINC codes
- Precoordinated qualifiers
- Hierarchy

LOINC CODE	LONG COMMON NAME
<b>8454-1</b>	<b>Diastolic blood pressure--standing</b>

LOINC CODE	LONG COMMON NAME
<b>8453-3</b>	<b>Diastolic blood pressure--sitting</b>

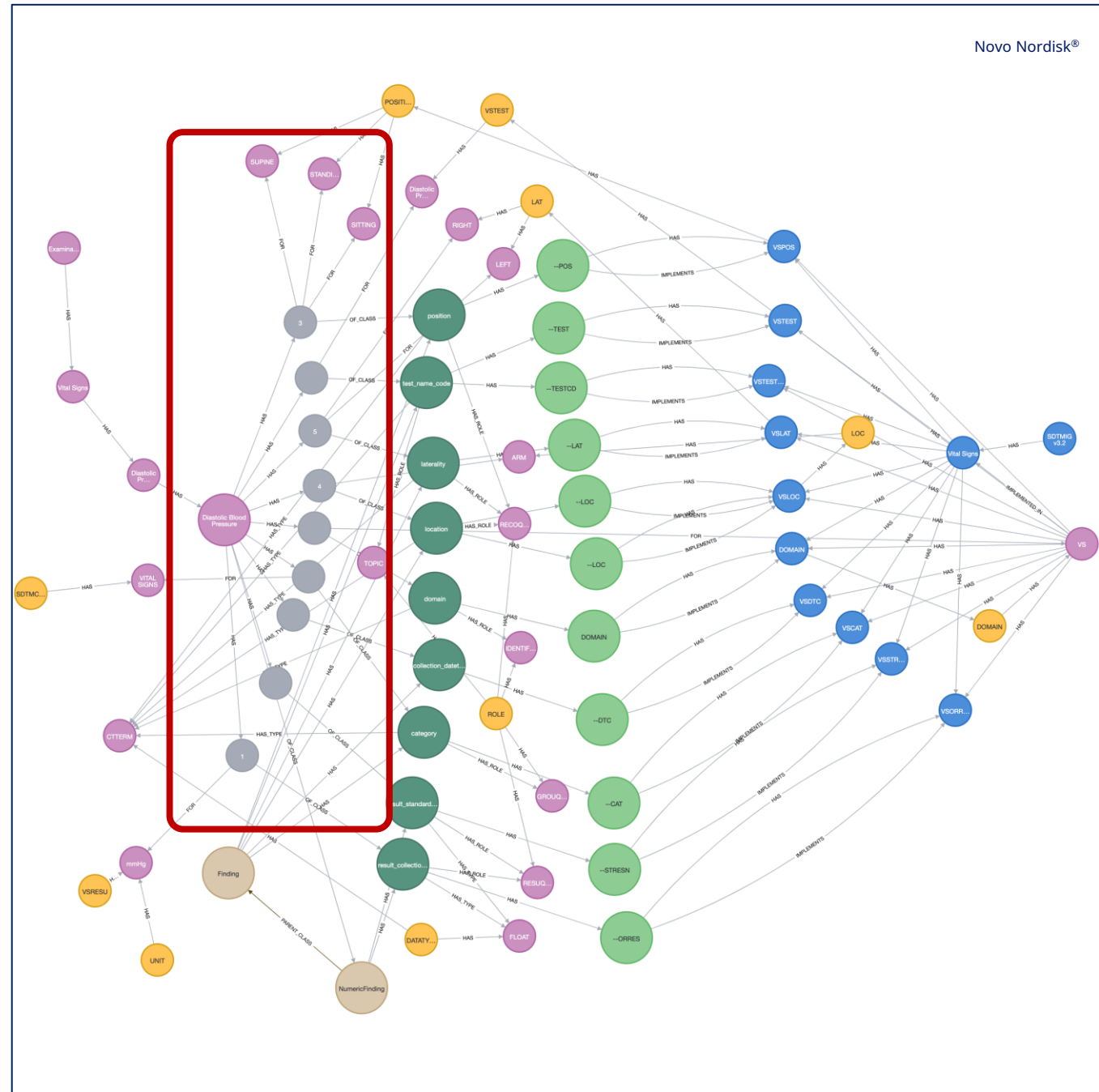
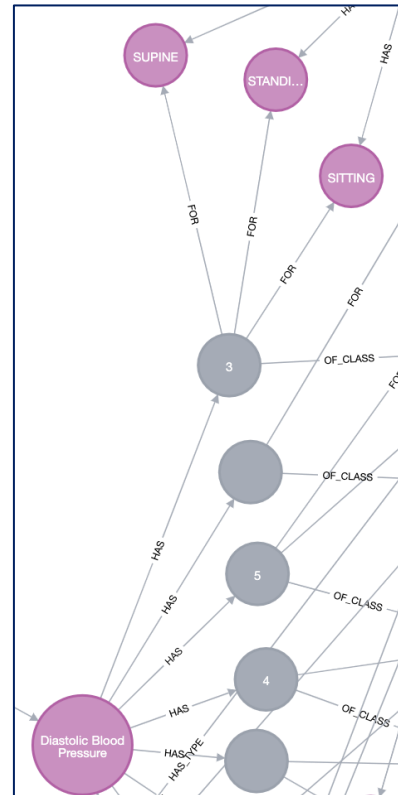
LOINC CODE	LONG COMMON NAME
<b>8455-8</b>	<b>Diastolic blood pressure--supine</b>

LOINC CODE	LONG COMMON NAME	LOINC STATUS		
<b>8462-4</b>	<b>Diastolic blood pressure</b>	<b>Active</b>		
<b>Fully-Specified Name</b>				
Component	Intravascular diastolic			
Property	Pres			
Time	Pt			
System	Arterial system			
Scale	Qn			
Method				
<b>Additional Names</b>				
Short Name	BP dias			
<b>Associated Observations</b>				
<p>This panel contains all of the "Special circumstances" LOINC codes that are used to report the specific context during which a particular measurement was taken, where that context may affect the measurement value. Circumstances that can affect measurement of certain variables include when the patient is asleep, in pain, crying, febrile, or during a period of apnea. This panel is attached to the generic LOINC code for each measurement so that the circumstance under which the measurement was taken can be reported along with the primary result.</p>				
LOINC	Name	R/O/C	Cardinality	Example UCUM Units
<b>89263-8</b>	<b>Special circumstances associated observations panel</b>			
10224-4	Hemodynamic method special circumstances			
55285-1	Glasgow coma score special circumstances			
55416-2	Oxymetry special circumstances			
8304-8	Body height special circumstances			
8337-8	Body weight special circumstances			
9278-3	Breath rate special circumstances			
9848-3	Body temperature special circumstances			
9855-8	Blood pressure special circumstances			
89299-2	Heart rate special circumstances			



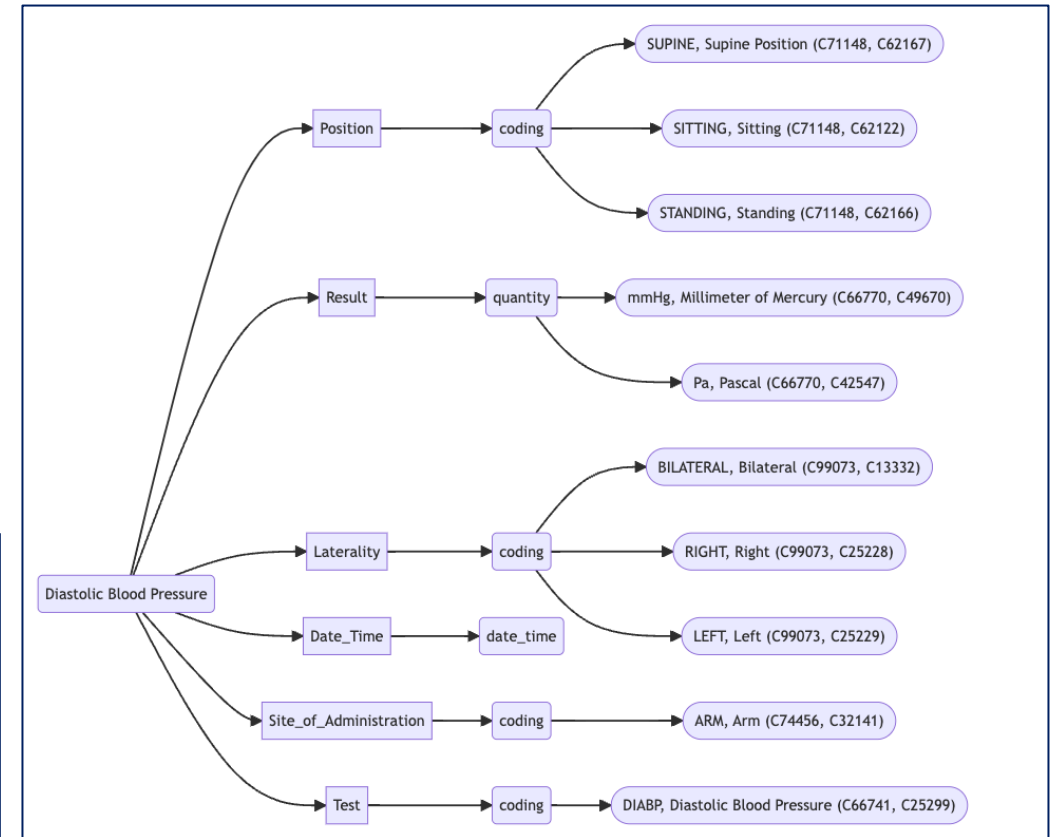
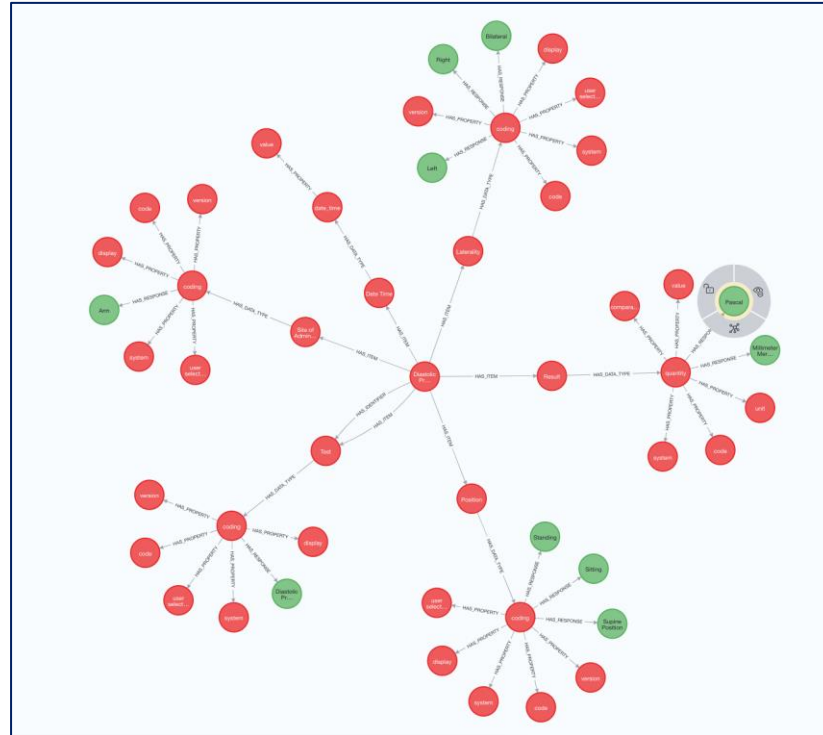
# Open Study Builder

- Unit of knowledge and properties
- Attached CT, using CDISC



# d4k

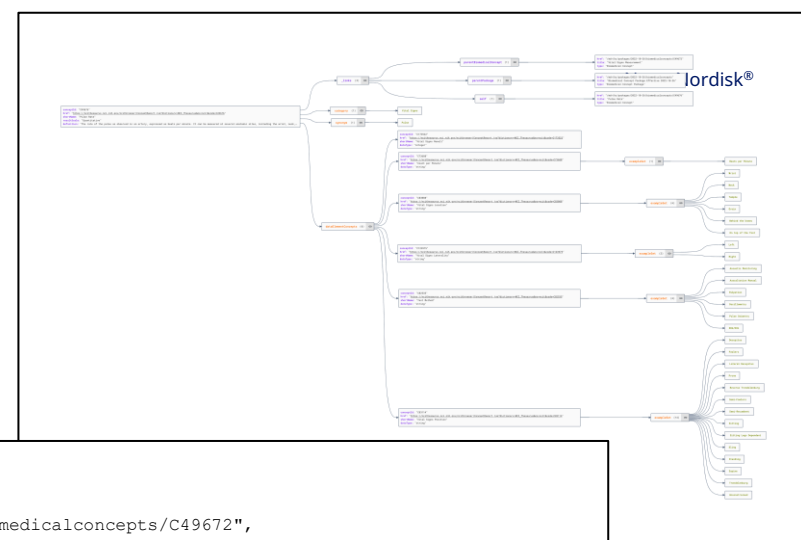
- Uses CDISC CT
- Uses FHIR data types



Item Name	Identifier	Data Type	Terms
Position		coding	SUPINE, Supine Position (C71148, C62167) SITTING, Sitting (C71148, C62122) STANDING, Standing (C71148, C62166)
Result		quantity	mmHg, Millimeter of Mercury (C66770, C49670) Pa, Pascal (C66770, C42547)
Laterality		coding	BILATERAL, Bilateral (C99073, C13332) RIGHT, Right (C99073, C25228) LEFT, Left (C99073, C25229)
Date Time		date_time	
Site of Administration		coding	ARM, Arm (C74456, C32141)
Test	✓	coding	DIABP, Diastolic Blood Pressure (C66741, C25299)

# So ...

- A central node, the root of the BC
- Identification
  - Unique (UUID, URI ...)
  - Resolvable would be nice
- Version managed
- A set of properties
  - Some must be there, e.g. result
- Controlled Terms defined
  - CDISC CT
- Complete
  - Everything we need defined
- Equivalence
  - Links to equivalent models in other systems
- Hierarchy
  - Subclass, membership type capabilities
- Configurable
  - Select properties and / or CT values



```
{
  "_links": {
    "parentBiomedicalConcept": {
      "href": "/mdr/bc/packages/2022-10-26/biomedicalconcepts/C49672",
      "title": "Vital Signs Measurement",
      "type": "Biomedical Concept"
    },
    "parentPackage": {
      "href": "/mdr/bc/packages/2022-10-26/biomedicalconcepts",
      "title": "Biomedical Concept Package Effective 2022-10-26",
      "type": "Biomedical Concept Package"
    },
    "self": {
      "href": "/mdr/bc/packages/2022-10-26/biomedicalconcepts/C49676",
      "title": "Pulse Rate",
      "type": "Biomedical Concept"
    }
  },
  "conceptId": "C49676",
  "href": "https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C49676",
  "category": [
    "Vital Signs"
  ],
  "shortName": "Pulse Rate",
  "synonym": [
    "Pulse"
  ],
  "resultScale": "Quantitative",
  "definition": "The rate of the pulse as observed in an artery, expressed as beats per minute. It can be measured at several anatomic sites, including the wrist, neck, temple, groin, behind the knees, or on top of the foot.",
  "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"
    },
    ...
  ]
}
```

## Definition

- A small model that defines a clinical concept in a standardized and reusable manner
- A unit of knowledge
- Independent of technology or means of collection
- A specification of the data

01

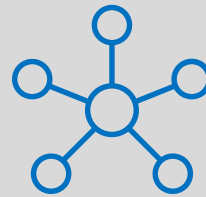
## Structure

- Central Node
- Properties
- Identification
- Version Managed
- Controlled Terms
- Complete
- Equivalence
- Hierarchy
- Configurable
- Data Types
- Templated

02

# Biomedical Concepts

*Perspectives for the road to automation*



## Adoption

- Don't initially seek perfection
- Iterate and learn
- Question every assumption
- Think data not presentation, a data-centric approach
- Consider linking to DDF USDM to bring ultimate power

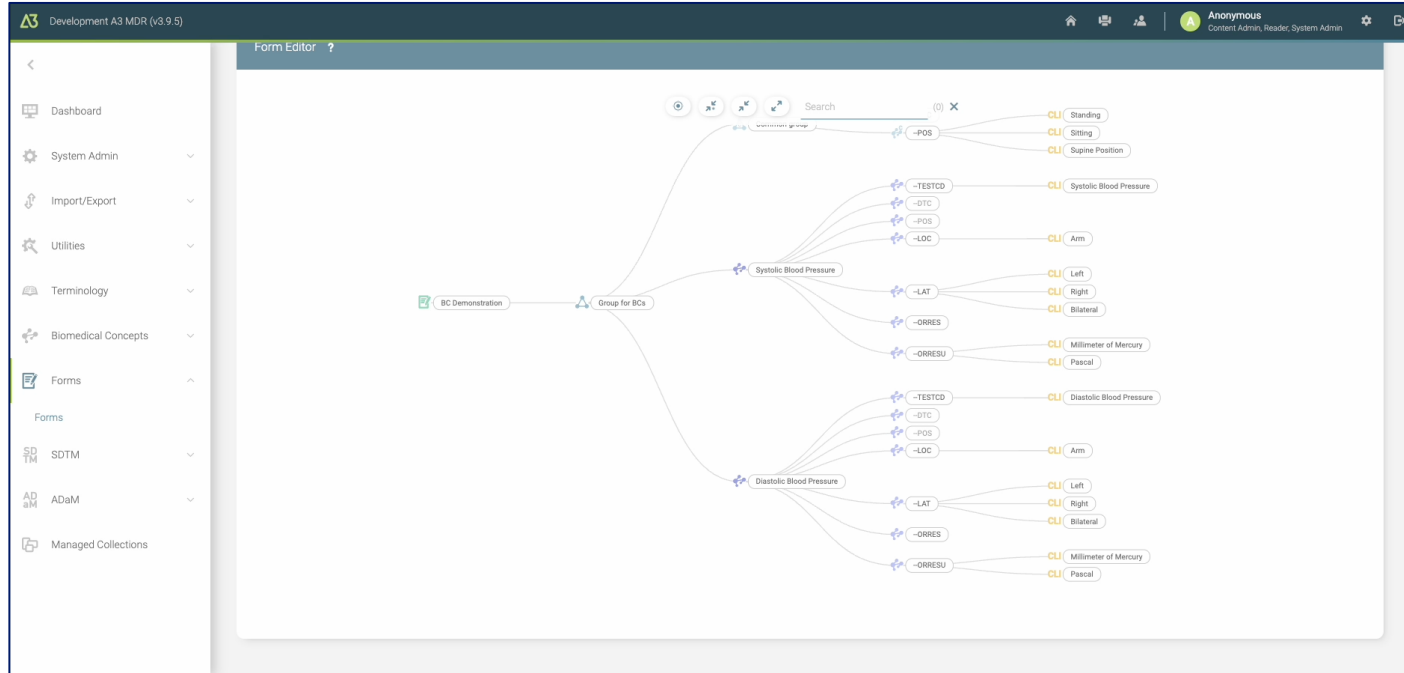
03

## Use Cases

- Form definition
- SDTM automation
- Define automation
- Impact analysis
  
- But, real power comes when combined with study design, e.g. USDM, for retrospective and prospective study builds

04

# Example Deployment



This screenshot shows the deployment view of the 'BC Demonstration' form. The form is titled 'VS=Vital Signs' and contains the following sections:

- Group for BCs
- Common group
- Date and time
  - VSDTC where VSTESTCD=DIABP
  - VSDTC where VSTESTCD=SYSBP
- Body Position
  - VSPOS where VSTESTCD=DIABP (Standing, Sitting, Supine Position)
  - VSPOS where VSTESTCD=SYSBP (Standing, Sitting, Supine Position)
- Systolic Blood Pressure
  - VSTESTCD (Systolic Blood Pressure)

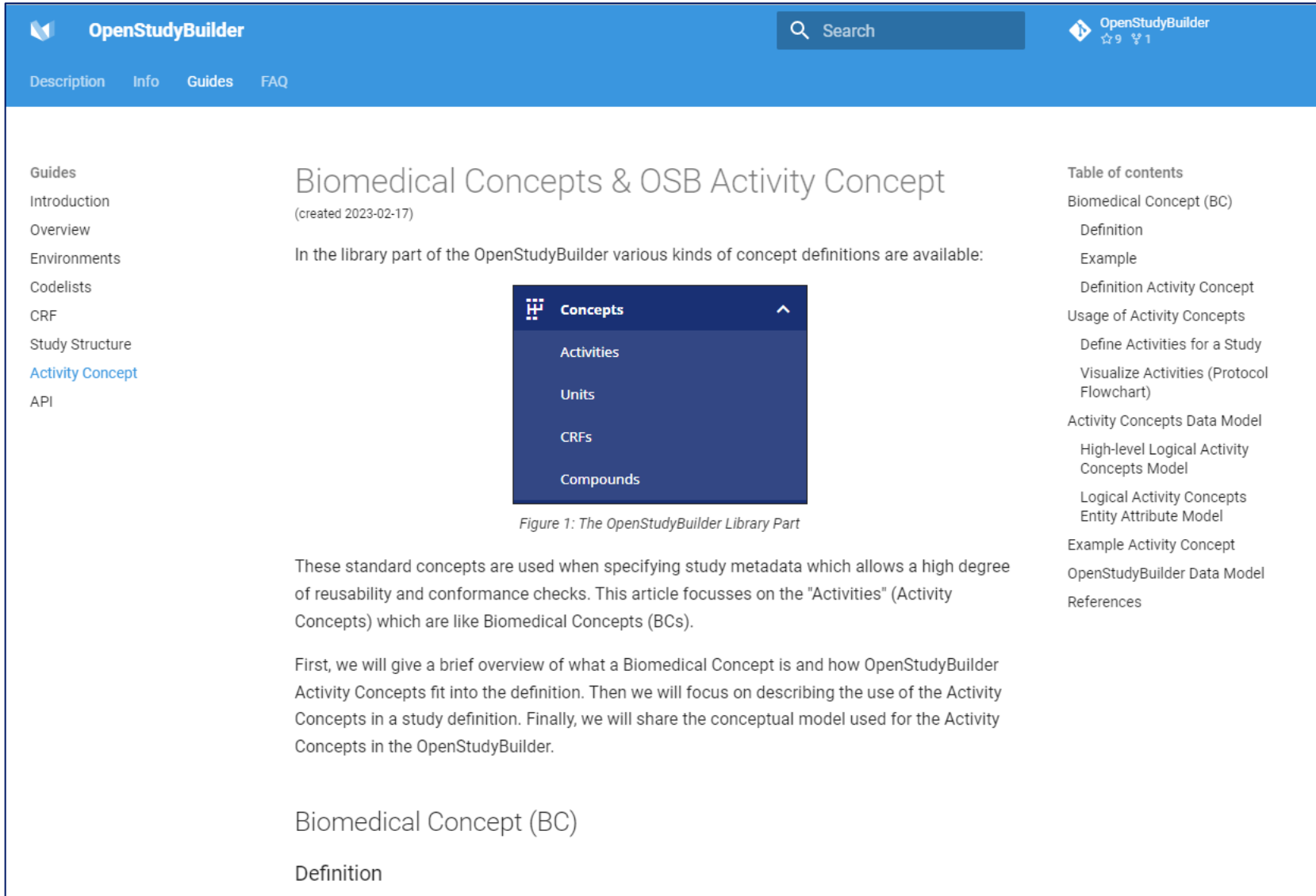


# BC in OpenStudyBuilder := Activity Concepts (Mikkel)

# BC in OSB := Activity Concepts

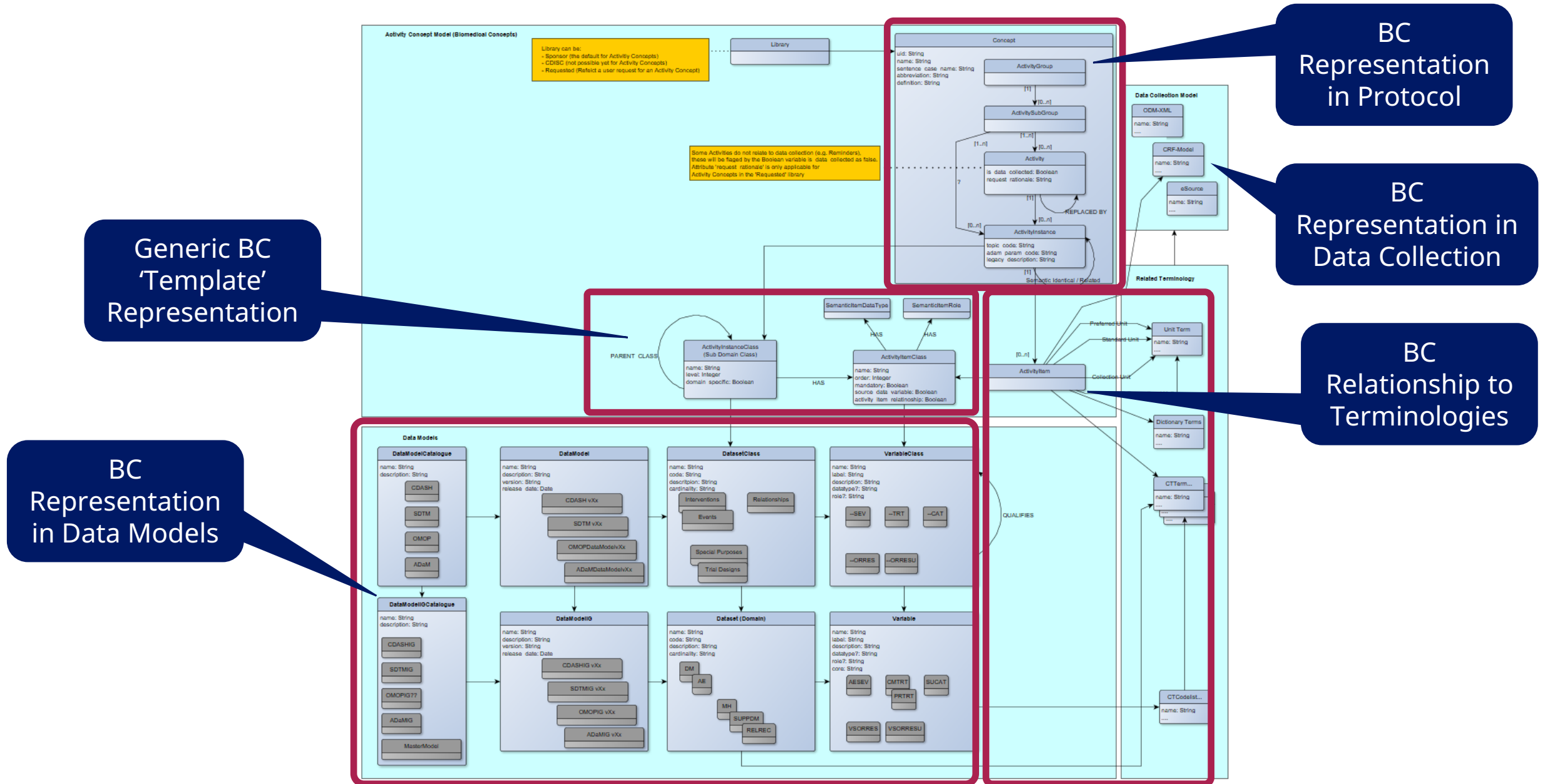
- OpenStudyBuilder is based on **Concept based Data Standards**
  - These are structures with more complex relationships
  - I.e. not only code-value pairs
  - They are applied for many different types of data, Activities (Clinical Procedures and Assessments), Compounds (linked to IDMP), Unit Definitions, Data Collection forms
- **Biomedical Concepts** (BC's) is generally defined as Activities (Clinical Procedures and Assessments)
- In OpenStudyBuilder we therefore use the general term **Concepts** and the specific term **Activity Concept := BC**

# Read more in our BC article on our GitLab site



The screenshot shows the OpenStudyBuilder website interface. At the top, there is a blue header with the OpenStudyBuilder logo, a search bar, and navigation links for Description, Info, Guides, and FAQ. The main content area is titled "Biomedical Concepts & OSB Activity Concept" and includes a sub-header "(created 2023-02-17)". Below the title, there is a paragraph stating: "In the library part of the OpenStudyBuilder various kinds of concept definitions are available:". This is followed by a dark blue dropdown menu labeled "Concepts" with an upward arrow, containing the following items: "Activities", "Units", "CRFs", and "Compounds". Below the menu is a caption: "Figure 1: The OpenStudyBuilder Library Part". The main text continues: "These standard concepts are used when specifying study metadata which allows a high degree of reusability and conformance checks. This article focusses on the 'Activities' (Activity Concepts) which are like Biomedical Concepts (BCs). First, we will give a brief overview of what a Biomedical Concept is and how OpenStudyBuilder Activity Concepts fit into the definition. Then we will focus on describing the use of the Activity Concepts in a study definition. Finally, we will share the conceptual model used for the Activity Concepts in the OpenStudyBuilder." To the right of the main text is a "Table of contents" section listing various topics such as "Biomedical Concept (BC)", "Definition", "Example", "Definition Activity Concept", "Usage of Activity Concepts", "Define Activities for a Study", "Visualize Activities (Protocol Flowchart)", "Activity Concepts Data Model", "High-level Logical Activity Concepts Model", "Logical Activity Concepts Entity Attribute Model", "Example Activity Concept", "OpenStudyBuilder Data Model", and "References".

# Discussion on BC data model in OpenStudyBuilder versus others



- ←
- About Library
- Process Overview
- Code Lists
- Dictionaries
- Concepts
- Activities
- Units
- CRFs
- Compounds
- Syntax Templates
- Template Instantiations
- Template Collections
- Data Exchange Standards
- List

Library / Concepts / CRFs

## CRFs (Case Report Forms) ?

CRF Templates Forms Item Groups Items CRF Tree ODM View Alias Extensions

RELOAD  

ODM version 1.3.2 with DoB

Annotated CRF [MSG2.0]

## Adverse Event

One AE should be reported per form. During conduct of the study, please transcribe data to EDC as soon as possible. The AE diagnosis, causality, seriousness and severity should be evaluated by the investigator or sub-investigator with physician background.

1: Any conditions / illnesses ?

MH (Medical History Domain)

[OID=G.MH.NS, Version=0.1]

Please state if there was any conditions / illnesses

2: Medical History item group

MH (Medical History Domain)

[OID=G.MH.CM, Version=0.1]

Please complete this Medical History item group before starting the treatment

Black label are Mandatory (otherwise Green)

Lock

\* Data Entry Required

Source Data Verification (SDV)

## Informed Consent and Demography

Please complete this Informed Consent and Demography form at the very beginning of the study General item design notes: Integration: A: Argus, Ax: Forms attached in Argus, C: CPR Dashboard, IW: IWRS, P: Impact, R: Reports, RT: RTSM General item design notes: Integration: A: Argus, Ax: rms attached in Argus, C: CPR Dashboard, IW: IWRS, P: Impact, R: Reports, RT: RTSM Oracle item des N notes: Key: [\*] = Item is required. Sex: Populated by IWRS. Item to trigger Childbearing potential form to appear if response = Female. Subject No.: Populated by IWRS and mapped from ENR to Inf Cons/DemogOracle item design notes: Key: [\*] = Item is required. Sex: Populated by IWRS. Item to trigger Childbearing potential form to appear if response = Female. Subject No.: Populated by IWRS and mapped from ENR to Inf Cons/Demog

1: Informed Consent item group

DM (Demographics Domain)

[OID=G.DM.IC, Version=0.1]

Please complete the Informed Consent item group before any other information



Study ID  
[OID=I.STUDYID, Version=0.1]

Although this field is not typically captured on a CRF, it should be displayed clearly on the CRF and/or the EDC system. This field can be included into the database or populated during SDTM-based dataset creation before submission.

 11 digit(s)

STUDYID



Time informed consent obtained  
[OID=I.RFICTIM, Version=0.1]

This will be the same information on informed consent used in the SDTM Disposition domain

 5 digit(s)

RFICDTC\_DSSTDTC

2: General Demographic item group

DM (Demographics Domain)

[OID=G.DM.DM, Version=0.1]

# Light OpenStudyBuilder demo

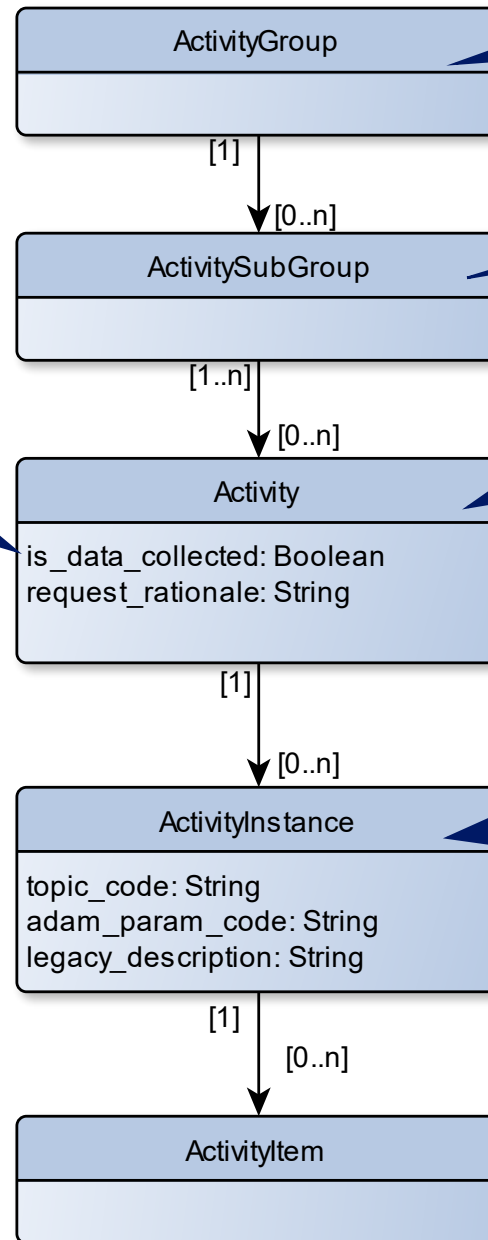
- Browse Activity Concept in Library -> Activity Concepts
  - Display details – will be CDISC COSMoS compatible
- Refer to Activity Concepts in Syntax Templates
- Apply as Endpoint Selection selecting Activity in Study Purpose
- Apply as Activity selection in SoA
- Bring to Protocol Document
  - Activity Concepts in endpoints based on syntax templates and SoA
- Drive metadata for SDTM
  - Both study design datasets as well as SDTM Define specification including value level metadata
- DDF SDR Compatibility

We generally use the term '**Activity**' to cover both Assessments based Activities as well as Activities without Assessments (like procedures, reminders, etc.)

Things in the flowchart related or not to data collection

Links to Generic Activity Instance Class model – as an Activity Instance Template

Links to Generic Activity Item Class model – as an Activity Item Template



Grouping of Activities, optionally only the grouping can be shown in the protocol SoA

The specific level in the hierarchy for protocol SoA. Independent on e.g. specimen, unit, SDTM Domain, ADaM PARAM, ...

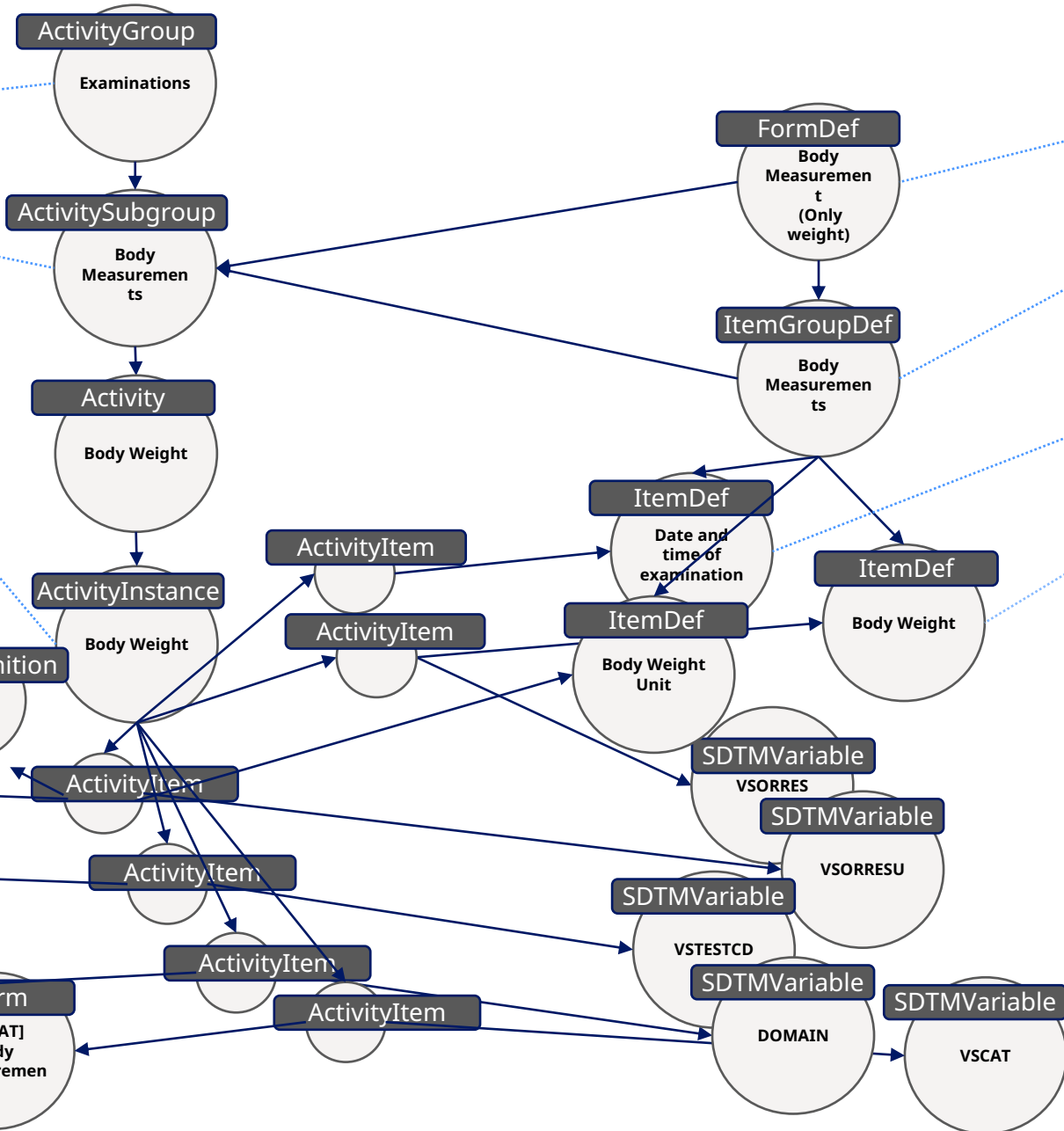
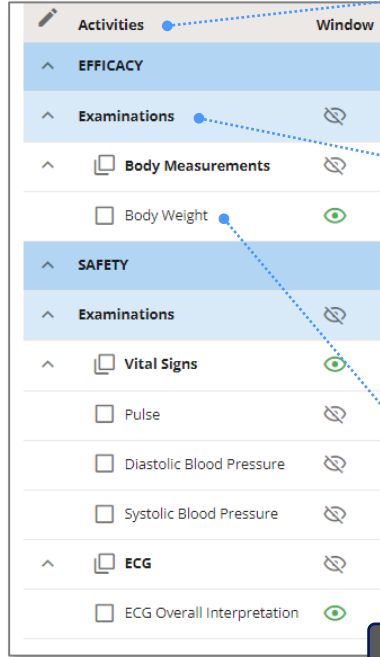
Correspond to our existing Topic Code, ADaM PARAM/PARAMCD. Specific to specimen, unit, SDTM Identify semantic observations

Links to terminology and cross data model variables



# Activity Concept data model sample – Body Weight

## Study Flowchart



## CRF

**Body Measurements**

**Body Measurements**

**Study: XX-XXXX**

Date and time of examination	Req/Unk <input type="checkbox"/> Req/Unk 24-hour clock <input type="checkbox"/>
Body weight	xxx.   o kg o lb

# Browse and curate BC content in OSB model

OpenStudyBuilder Activity Library Dashboard

Activity Library Content × Term impact assessment Activity in COSMOS format +

Number of Activities by groups

Group	Adverse Events	General	ECG	Body Measurements	Vital Signs	Physical Examination
Event	2	1	0	0	0	0
NumericFinding	24	28	22	0	0	0
CategoricalFinding	4	20	2	44	0	0
TextualFinding	2	0	0	1	0	0

Category: Activity Sub-Type Value: Number of Activities Group: Activity Sub-group

Activity Type: Finding

Activity Sub-type: Start typing...

Activity Group: Start typing...

Activity Sub-group: Body Measurements

List of Activities

Activity Type	Activity Sub-Type	Activity Group	Activity SubGroup	Activity	Activity Instance
Finding	NumericFinding	Examinations	Body Measurements	Estimated body weight	EWEIGHT
Finding	NumericFinding	Examinations	Body Measurements	Estimated height	HGEST
Finding	NumericFinding	Examinations	Body Measurements	Height	HEIGHT
Finding	NumericFinding	Examinations	Body Measurements	Height	HEIGHT3
Finding	NumericFinding	Examinations	Body Measurements	Height Average	MEANHGHT

OpenStudyBuilder Activity Library Dashboard

Activity Library Content Term impact assessment Activity in COSMOS format × +

Select Activity Instance

ActivityInstance adam\_param\_code

WEIGHT

Display Activity in CDISC Cosmos YAML format

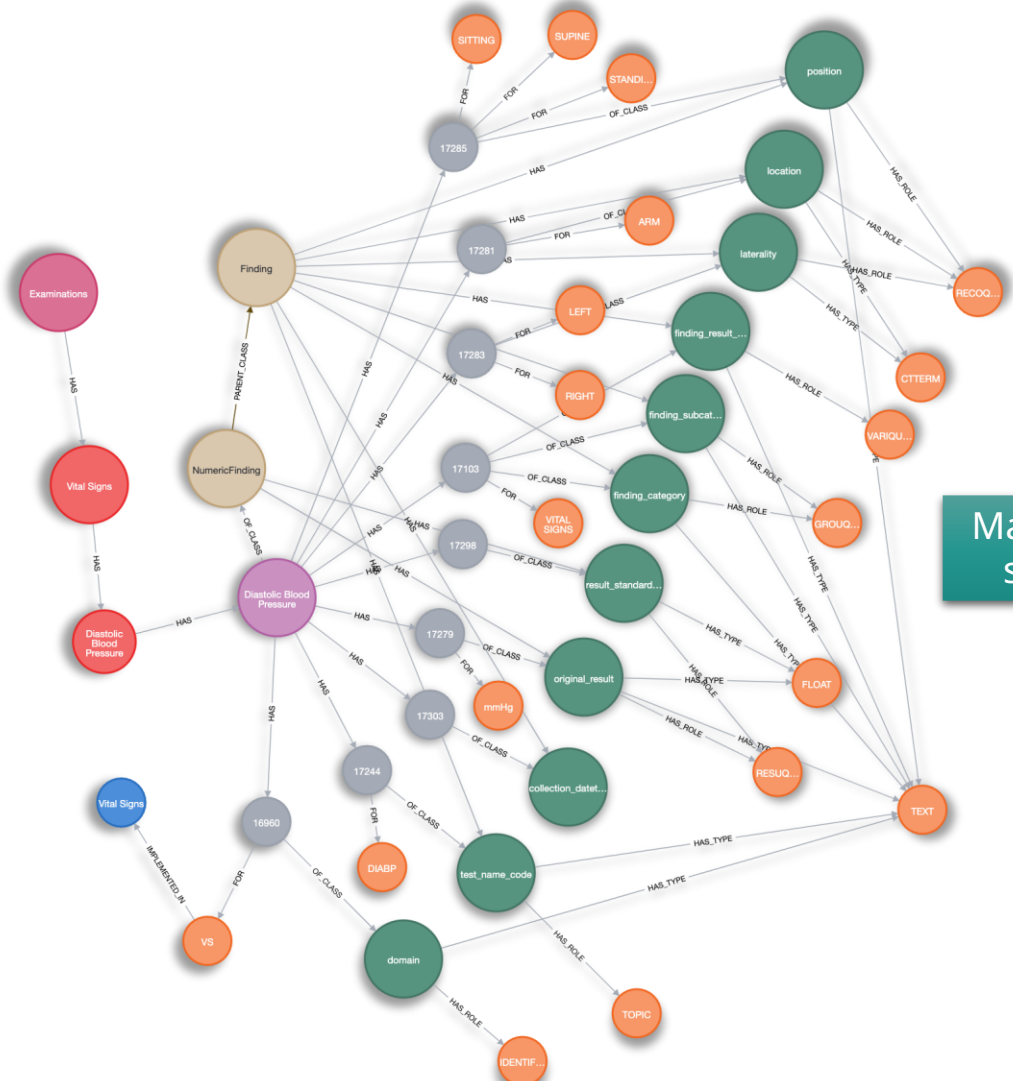
Variable	Variable Name	Value	Detail
packageDate		2022-11-13	
packageType		SDTM	
datasetSpecialisationID		BODY_WEIGHT	
domain		VS	
shortName		Body Weight	
source		VS.VSTESTCD	
sdtmigStartVersion		3.2	
sdtmigEndVersion			
biomedicalConceptId		C25208	
variables	VSORRES	name : VSORRES	
variables		codelist	(submissionValue=VSRESU, conceptId=C66770, href=https://ncithesaurus.nci.nih.gov/ncit)
variables		role : Result Qualifier	
variables		dataType : FLOAT	
variables		originType : Collected	
variables		mandatoryVariable : Yes	
variables	VSORRES	name : VSORRES	
variables		codelist	(submissionValue=UNIT, conceptId=C71620, href=https://ncithesaurus.nci.nih.gov/ncit)
variables		role : Result Qualifier	
variables		dataType : FLOAT	
variables		originType : Collected	
variables		mandatoryVariable : Yes	
variables	VSORRES	name : VSORRES	
variables		codelist	(submissionValue=PKUNIT, conceptId=C85494, href=https://ncithesaurus.nci.nih.gov/ncit)
variables		role : Result Qualifier	
variables		dataType : FLOAT	
variables		originType : Collected	

# OSB model versus other models



# OpenStudyBuilder Activity Model mapped to COSMoS-BC Model

+50TH SPECIALIZATIONS



Mapping steps

```
apoc.map.fromPairs([
  ['packageType', dmig.name],
  ['datasetSpecializationId', ai.topic_code],
  ['domain', ds.code],
  ['shortName', ai.name],
  ['source', ds.code+'.'+ds.code+'TESTCD'],
  ['sdmStartVersion', dmig.version],
  ['Variables', vars],
  ['category', asgrp.name],
  ['resultScale', aiclass]
]) as activity
```

Conversion

```
CASE when aic.name='NumericFinding' THEN 'Quantitative' ELSE
CASE WHEN aic.name='CategoricalFinding' THEN 'Ordinal' ELSE
null END END as aiclass
```

# OpenStudyBuilder Activity Model mapped to COSMoS-BC Model - finding terminology

Using NCI API

```
def nci_service_get_concept_code_from_term(term):
    url = 'https://api-eva-rest.nci.nih.gov/api/v1/concept/search?fromRecord=0&include=synonyms&pageSize=10&synonymSource=CDISC&term='+term+'&terminology=ncit&type=match'
    r = requests.get(url)
    concept_info = r.json()
    #avoid getting the CDASH one
    for concept in concept_info['concepts']:
        if 'CDASH' not in concept['name']:
            code = concept['code']
    return code

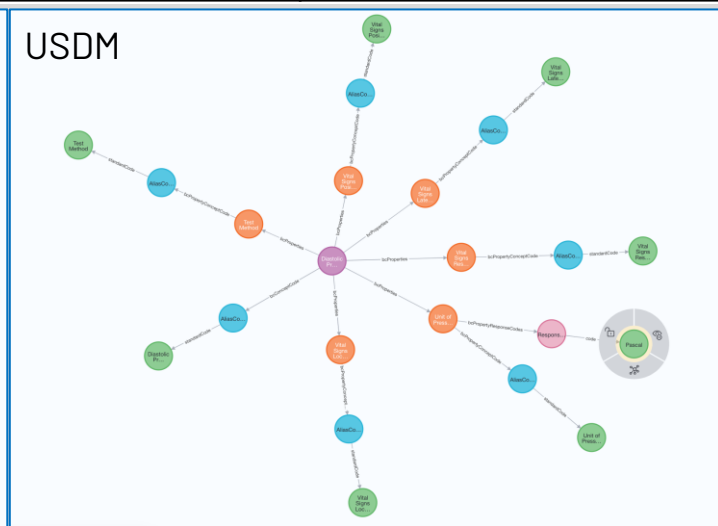
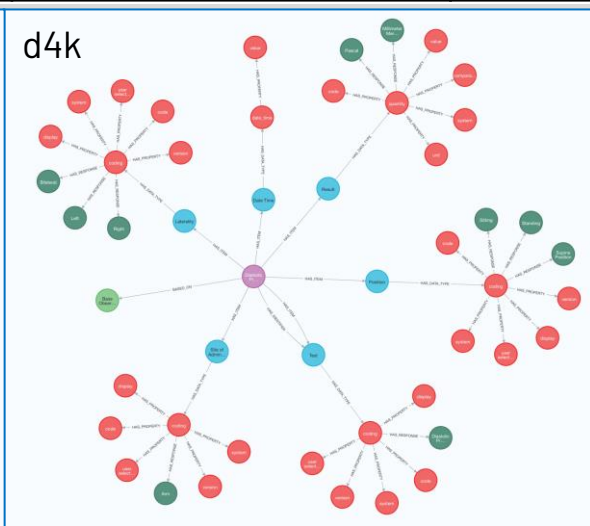
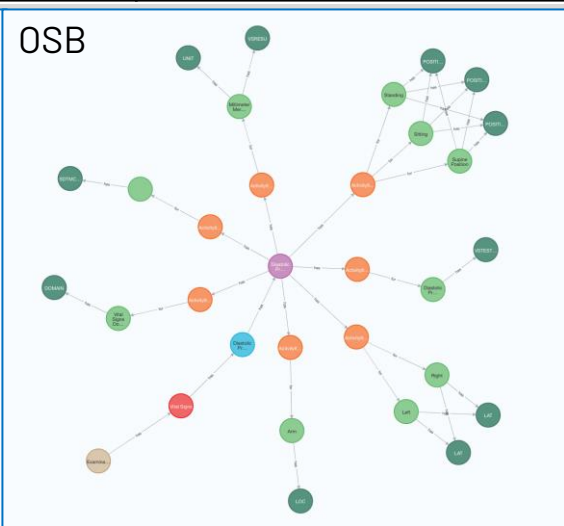
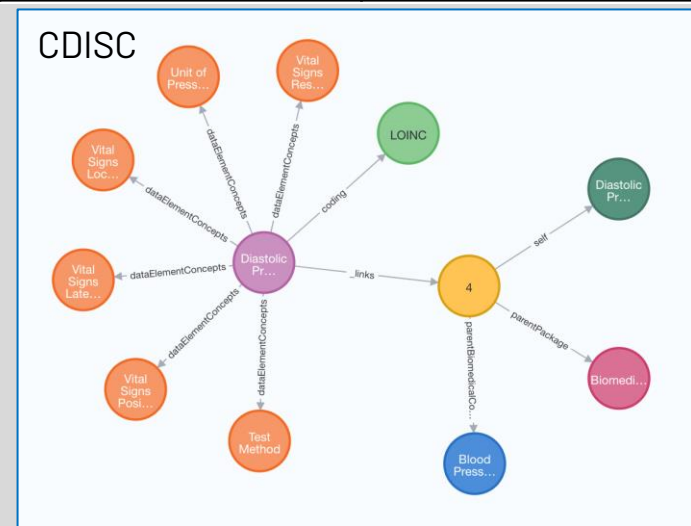
def get_testcd_subm_from_test_code(ccode):
    url = 'https://api-eva-rest.nci.nih.gov/api/v1/concept/search?fromRecord=0&include=summary%2Csynonyms&pageSize=10&synonymSource=CDISC&synonymTermType=PT&term='+ccode+'&terminology=ncit&type=match'
    r = requests.get(url)
    concept_info = r.json()

    #Get the SDTM TESTCD term based on test (i.e. get DIAP from 'Diastolic Blood Pressure' term)
    for concept in concept_info['concepts'][0]['synonyms']:
        if "code" in concept:
            if 'TESTCD' in concept['code'] and 'SDTM' in concept['code']:
                testcd = concept['name']
    return testcd
```

```
1 - packageDate: '2022-11-13'
2 packageType: bc
3 conceptID: C25299
4 href: https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C25299
5 parentConceptID: C54706
6 category:
7   - Vital Signs
8 shortName: Diastolic Blood Pressure
9 synonym:
10  - DIABP
11 resultScale: Quantitative
12 definition: The minimum pressure exerted into the systemic arterial circulation
13   during cardiac ventricular relaxation and filling.
14 coding:
15  - code: LOINC-CODE TO BE FOUND
16    system: http://loinc.org/
17    systemName: LOINC
18 dataElementConcepts:
19  - conceptID: C83466
20    href: https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C83466
21    shortName: VSTESTCD
22    dataType: CTERM
23 > - conceptID: C49672 ...
27 > - conceptID: C83454 ...
31 > - conceptID: C83452 ...
35 > - conceptID: C83114 ...
39 > - conceptID: C83088 ...
43 > - conceptID: C123975 ...
47 > - conceptID: C83108 ...
51 > - conceptID: C83108 ...
55 > - conceptID: C83464 ...
```



Perspective	A measure of the content of a BC. Inclusion or exclusion is not a mark of quality of fitness for purpose, the table is to show differences in approach	CDISC (Conceptual Layer)	OSB	d4k	USDM (Based on CDISC Model)
Central Node	Has a central node from which all BC information can be found				
Properties	Is the BC built up from a set of properties				
Identification	Does the BC have a unique identifier				
Version Managed	Is the BC explicitly version managed	Not currently explicit			Based on CDISC BC
Controlled Terms	Controlled terms defined as part of the BC and which CT used	CDISC CT	CDISC CT	CDISC CT	CDISC CT
Complete	Is the definition complete, everything needed for deployment	CT references			CT references
Equivalence	Does the BC allow for equivalence to other systems to be made		No?		
Hierarchy	Can the BCs be placed into a hierarchy		Yes (fixed)		
Configurable	Can the BCs be configured using attributes within the BC	Not designed to be	Planned		
Data Types	Do the BCs use complex data types in their design, if so which ones		Simple data types	FHIR	
Templated	Are the BC instances based on a template	No?	Yes, by Class concept		Based on CDISC BC



Present initial SWOT and Mind Map  
as input to break-out sessions



# Initial SWOT and Mind Map for next steps

- Use this framework to capture discussions and reflections during break-outs
- Present for all in last plenum session
  
- SWOT
  - How do we see this for BC's supporting our clinical data flows
- Mind Map for next steps
  - How can we contribute and support the adoption of BC's in tools and our use supporting digital data flows

# SWOT – Applying BC's in digital data flows

## Strengths

### BC's

- Generic representation independent of source and target data models
- Support end-2-end lineage across data standards
- Initial BC definitions shared and curated by CDISC

### Tools

- Hide complexity of BC's from end users
- Will support usage across skill areas
- Initial tools shared as open-source

## Opportunities

### BC's

- Influence future industry standards via BC adoption
- Consistency in how CDISC standards are applied cross pharma

### Tools

- Improved business insight through linking related elements via modern graph database allowing for intelligent dashboards and search functionality
- FAIR based data sharing through transparent API-based architecture

## Weaknesses

### BC's

- Perception of current state not reflecting actual status
- Pharma companies can manage with less
- Evolution of standards with BC's incompatibility

### Tools

- Higher expectations than what is realistic to deliver due to business process complexity

## Threats

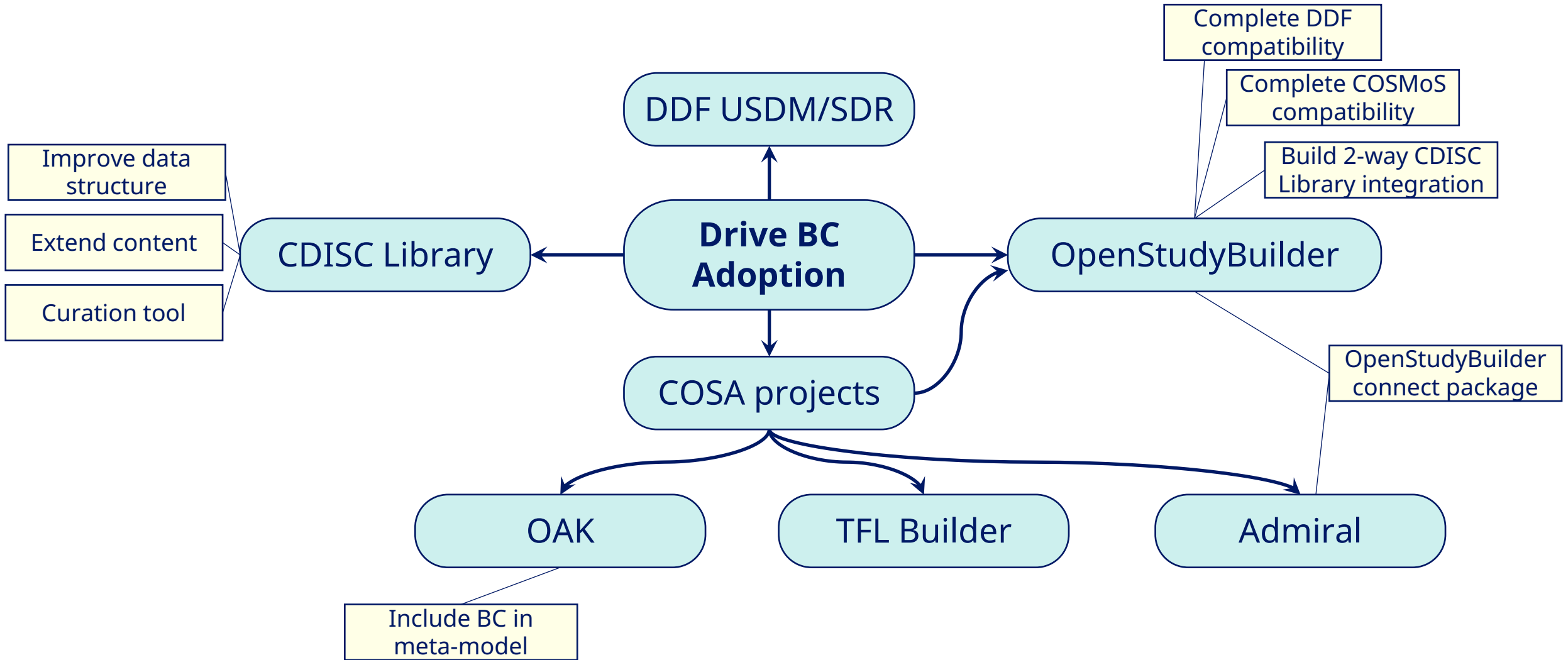
### BC's

- Too few SME resources and high dependency on few resources
- Limited sharing of BC definitions and curation of these
- Insufficient cross organisational/skill area allocation/commitment
- Dependency with other projects and initiatives

### Tools

- Lack of integration capabilities in consumer systems preventing realisation of business benefits
- Currently only custom solutions and not as a commercial system

# Mind Map for next steps in driving BC adoption



# Breakout sessions

11:00 – 12:00 Break-out part 1

12:00 – 13:00 Lunch

13:00 – 14:00 Break-out part 2

14:00 – 15:00 Sharing in plenum

# Breakout sessions

- Setup BC's in OSB SoA for a new study, run various queries to learn how BC's can be utilised
  - *BC beginners and those who want to see BCs applied in the OpenStudyBuilder*
- Learn and understand the BC model in OSB versus the COSMoS, DDF, d4k and other models
  - *BC engineers and for data modeling*
- Create and curate OpenStudyBuilder BC content via the OSB Library and NeoDash reports and mining BC's from existing data sources like SDTM.
  - *Modelling BCs from other sources and working with BCs, for example through dashboards*

- Room
  - Anja, Katja

- Room
  - Dave, Lex, Marius, Mikkel

- Room
  - Kirsten, Chandrakant, Nicolas, Linda

# Recap

Breakout 1 – BCs in OpenStudyBuilder  
Breakout 2 – BC Models  
Breakout 3 – Curation  
Final words

# Final Words

- Hope you enjoyed the workshop and learned a lot of BC
- Thanks for your input & discussions – all material will be shared
- Please give us feedback on the workshop – mail to Charles
- Looking forward to further collaborate
- OpenStudyBuilder Meet & Demo during Interchange (COSA Booth)

## Wednesday

10:30-11:00	Talk
12:00-13:00	Demo
13:00-13:30	Talk
13:30-14:00	Demo
17:00-18:00	Demo

## Thursday

9:00-10:00	Demo
12:00-13:00	Demo
13:00-13:30	Talk
13:30-14:00	Demo

# Final Words

- Links (also in COSA mail)

OpenStudyBuilder Homepage	<a href="https://novo-nordisk.gitlab.io/nn-public/openstudybuilder/project-description/">https://novo-nordisk.gitlab.io/nn-public/openstudybuilder/project-description/</a>
COSA Homepage	<a href="https://cosa.cdisc.org/">https://cosa.cdisc.org/</a>
CDISC Cosmos	<a href="https://www.cdisc.org/cdisc-biomedical-concepts">https://www.cdisc.org/cdisc-biomedical-concepts</a>
TransCelerate DDF	<a href="https://www.transceleratebiopharmainc.com/initiatives/digital-data-flow/">https://www.transceleratebiopharmainc.com/initiatives/digital-data-flow/</a>
D4K detailed BC paper	<a href="https://github.com/data4knowledge/biomedical_concepts/blob/main/docs/bc_treatise/Biomedical_Concepts_Treatise.pdf">https://github.com/data4knowledge/biomedical_concepts/blob/main/docs/bc_treatise/Biomedical_Concepts_Treatise.pdf</a>
<b>OpenStudyBuilder</b>	
Slack	Join OpenStudyBuilder <a href="#">Slack</a>
LinkedIn Newsletter	Subscribe to our newsletter <a href="#">here</a>
Guides	Various guides are available <a href="#">here</a>