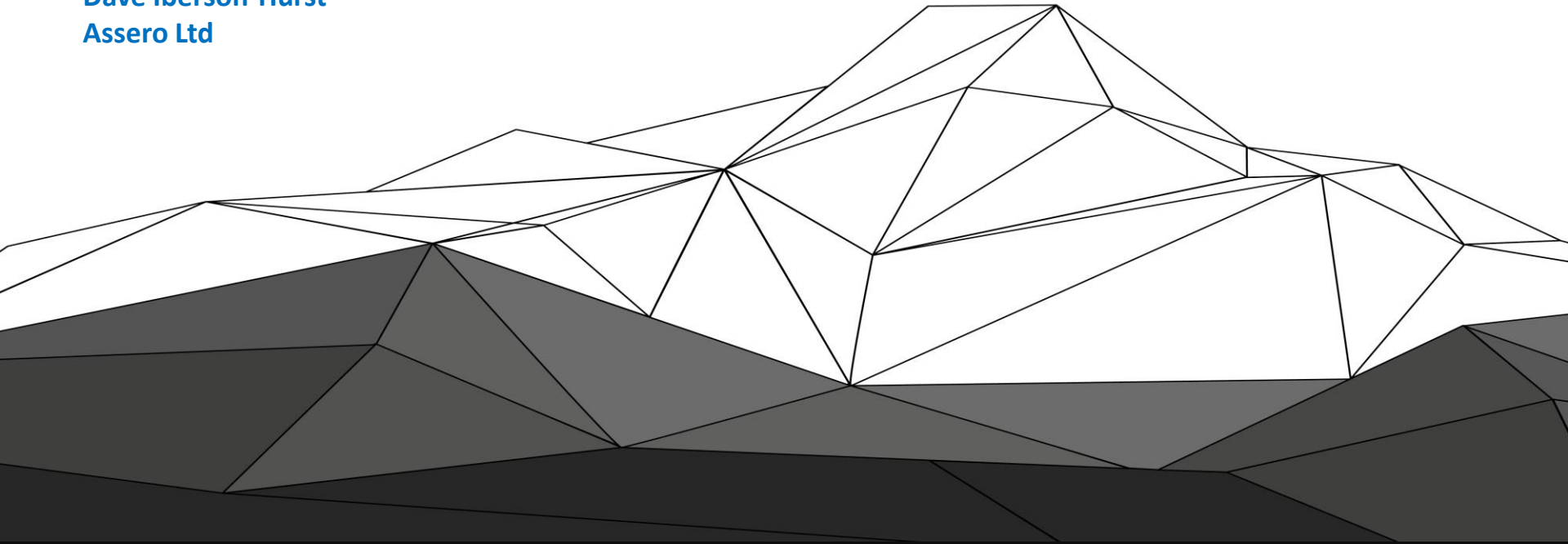
It’s Time To Change

CDISC ESUG Oxford

5th July 2017

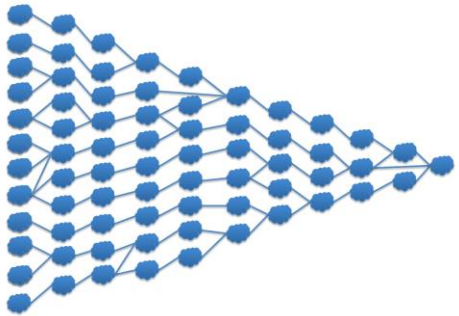
Dave Iberson-Hurst
Assero Ltd



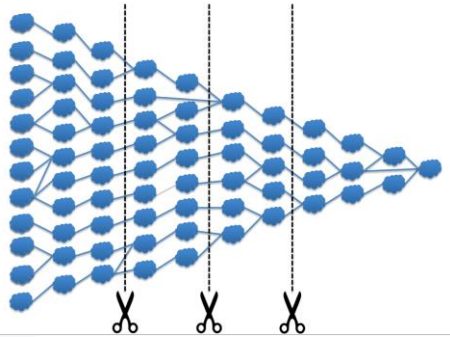
The Issues

A Conversation

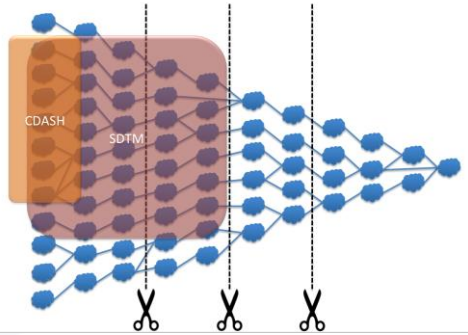
Our World Is a Very Big Graph



… Then We Make Life Hard

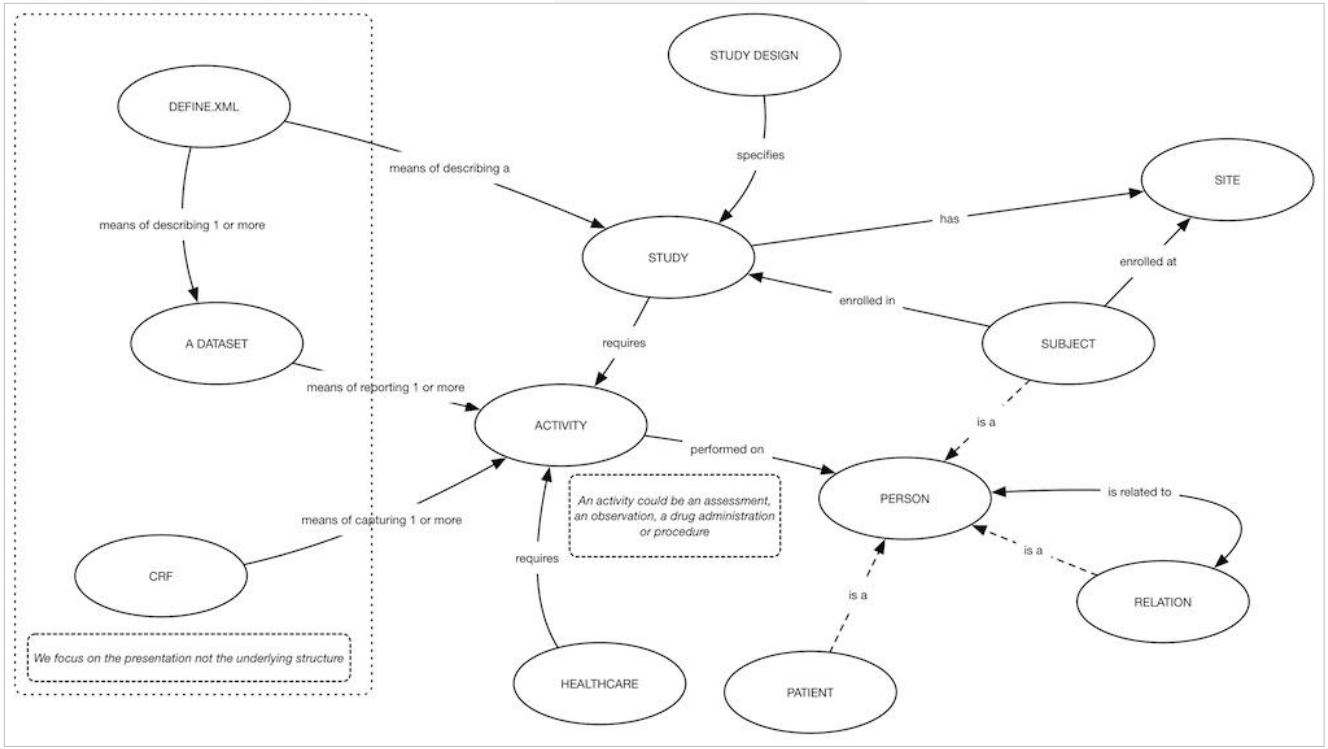


Our Standards Are Views

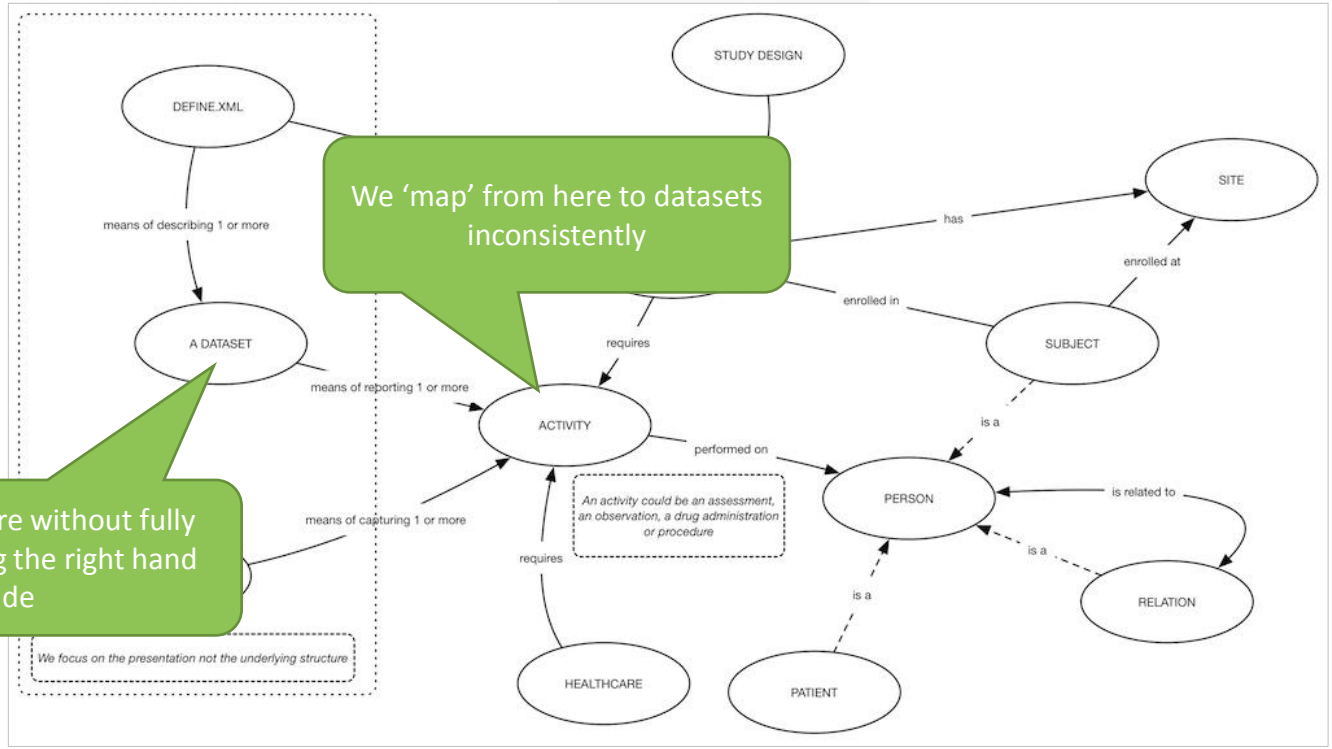


Left and Right Sides

Full Description: http://www.assero.co.uk/2017/a-left-side-and-a-right-side/



Left and Right Sides



FDA @ PhUSE CSS 2016

U.S. Food and Drug Administration
Protecting and Promoting Public Health
www.fda.gov

Define File

- Use Define.xml v2.0
  - Lacking a complete Define file greatly increases the amount of time reviewers spend understanding an application
- Include detailed description of data elements:
  - Detailed, reproducible computational algorithms for derived variables
  - Code lists that describe categories, subcategories, reference time-pts
  - Applicable value level metadata & description of SUPPQUAL domains
  - Explanations of sponsor-defined identifiers (e.g., –SPID, -GRPID)
- Provide separate unit code lists for each domain

Define File

6

U.S. Food and Drug Administration
Protecting and Promoting Public Health
www.fda.gov

Items in TCG (continued)

- Use CDISC controlled terminology variables when available (TCG Section 6)
  - Controlled terminology issues in 62% of applications
- Include Seriousness Criteria for all serious adverse events (TCG 4.1.1.3)
  - Missing (or with inconsistencies) in 50% of applications
  - Important to independently verify that AE was serious
- Include study day variable for all observational datasets (TCG 4.1.4.1)

Items in TCG

10

From a presentation by Mary Doi, M.D., M.S. (FDA CDER)

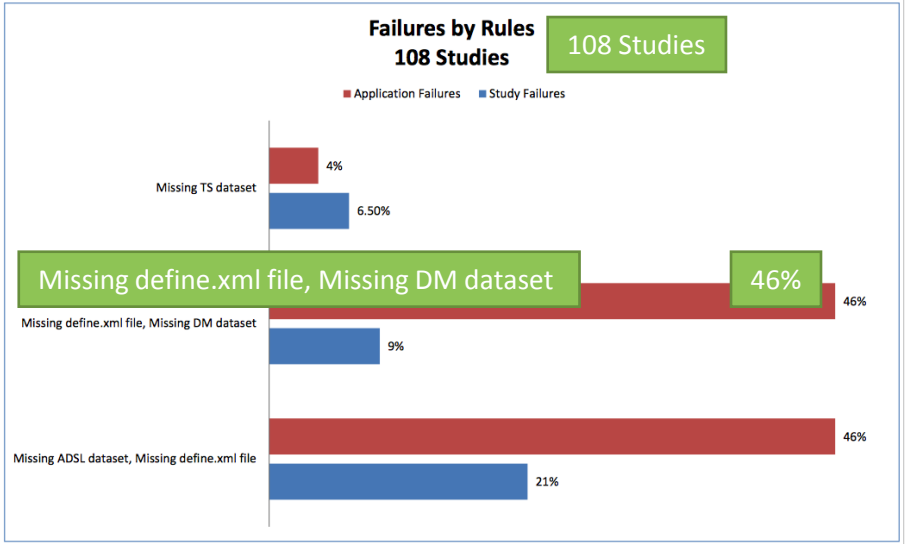
FDA @ PhUSE CSS 2017

eCTD Technical Validation of Study Data

High Level Rule #1	High Level Rule #2
Rule # 1734 : A Trial Summary (TS) dataset must be present for each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2	Rule #1736: DM dataset and define.xml must be submitted in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4. DM dataset, ADSL dataset, define.xml must be submitted in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2

https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM523539.pdf

11



13

From a presentation by Crystal Allard, Special Assistant to the Director Office of Computational Science

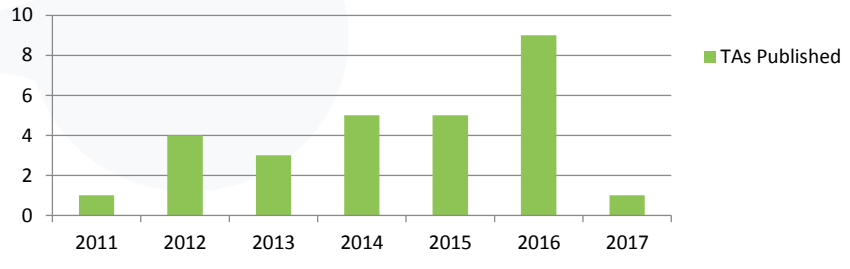
Rate of Change

Published TA User Guides

Project	Publication Date	Terminology	SDTM	CDASH	ADaM
Alzheimer's Disease v1	September 9, 2011	X	X		
Tuberculosis v1	June 29, 2012	X	X		
Pain v1	August 7, 2012	X	X		
Virology v1	December 6, 2012	X	X		
Parkinson's Disease v1	December 18, 2012	X	X		
Polycystic Kidney Disease v1	February 26, 2013	X	X		
Asthma v1	November 26, 2013	X	X		
Alzheimer's Disease v2	December 16, 2013	X	X		
Multiple Sclerosis v1	May 2, 2014	X	X		
Diabetes v1 (ADaM Supplement)	September 11, 2014 (December 18, 2015)	NA	X	X	X
Cardiovascular Endpoints v1	October 17, 2014	X	X		
Influenza v1	November 25, 2014	X	X		
QT Studies v1	December 12, 2014	X	X		
Chronic Hepatitis C Virus v1	May 8, 2015	X	X	X	
Schizophrenia v1	June 9, 2015	X	X		
Dyslipidemia v1	June 19, 2015	X	X	X	X
Virology v2	September 30, 2015	X	X		Partial
Traumatic Brain Injury v1	December 14, 2015	X	X	X	

Published TA User Guides

Project	Publication Date	Terminology	SDTM	CDASH	ADaM
COPD v1	January 26, 2016	X	X	X	X
Tuberculosis v2	February 26, 2016	X	X		
Breast Cancer v1	May 16, 2016	X	X	X	X
Rheumatoid Arthritis v1	November 14, 2016	X	X	X	X
Kidney Transplant	October 31, 2016	X	X	X	X
Major Depressive Disorder v1	December 5, 2016	X	X	X	X
Diabetic Kidney Disease v1	December 13, 2016	X	X	X	X
Pain v1.1 (update)	December 13, 2016	X	X		
Ebola v1	December 19, 2016	X	X	X	
Malaria v1	January 9, 2017	X	X	X	

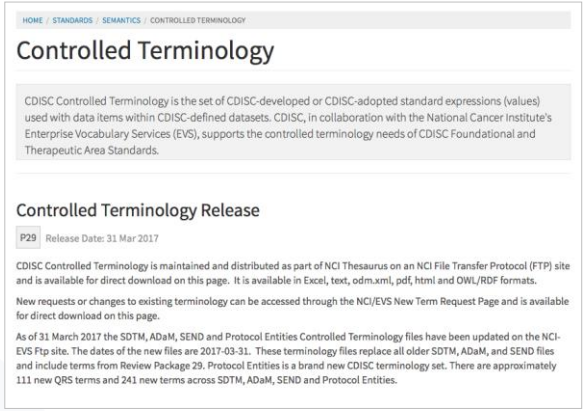


We Need …

Published TA User Guides

Project	Publication Date	Terminology	SDTM	CDASH	ADaM
Alzheimer's Disease v1	September 9, 2011	X	X		
Tuberculosis v1	June 29, 2012	X	X		
Pain v1	August 7, 2012	X	X		
Virology v1	December 6, 2012	X	X		
Parkinson's Disease v1	December 18, 2012	X	X		
Polycystic Kidney Disease v1	February 26, 2013	X	X		
Asthma v1	November 26, 2013	X	X		
Alzheimer's Disease v2	December 16, 2013	X	X		
Multiple Sclerosis v1	May 2, 2014	X	X		
Diabetes v1 (ADaM Supplement)	September 11, 2014 (December 18, 2015)	NA	X	X	X
Cardiovascular Endpoints v1	October 17, 2014	X	X		
Influenza v1	November 25, 2014	X	X		
QT Studies v1	December 12, 2014	X	X		
Chronic Hepatitis C Virus v1	May 8, 2015	X	X	X	
Schizophrenia v1	June 9, 2015	X	X		
Dyslipidemia v1	June 19, 2015	X	X	X	X
Virology v2	September 30, 2015	X	X		Partial
Traumatic Brain Injury v1	December 14, 2015	X	X	X	

- Control
  - Constant new versions
  - Rate-of-change of versions
- Precision
  - Which version am I using?
- Visibility
  - What changed?
  - When did it change
  - What is the impact of the change?
- Ease-of-use
  - Make it easier to use
  - Machine readable



HOME / STANDARDS / SEMANTICS / CONTROLLED TERMINOLOGY

Controlled Terminology

CDISC Controlled Terminology is the set of CDISC-developed or CDISC-adopted standard expressions (values) used with data items within CDISC-defined datasets. CDISC, in collaboration with the National Cancer Institute's Enterprise Vocabulary Services (EVS), supports the controlled terminology needs of CDISC Foundational and Therapeutic Area Standards.

Controlled Terminology Release

P29 Release Date: 31 Mar 2017

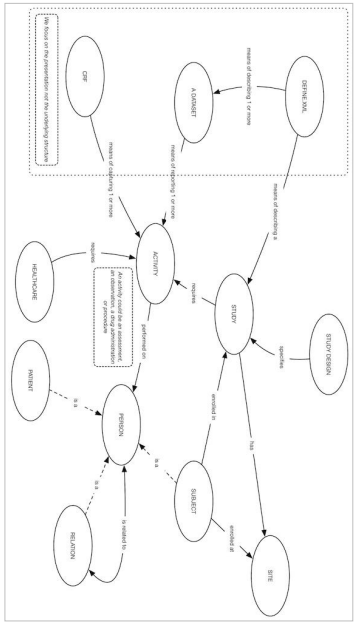
CDISC Controlled Terminology is maintained and distributed as part of NCI Thesaurus on an NCI File Transfer Protocol (FTP) site and is available for direct download on this page. It is available in Excel, text, odm.xml, pdf, html and OWL/RDF formats.

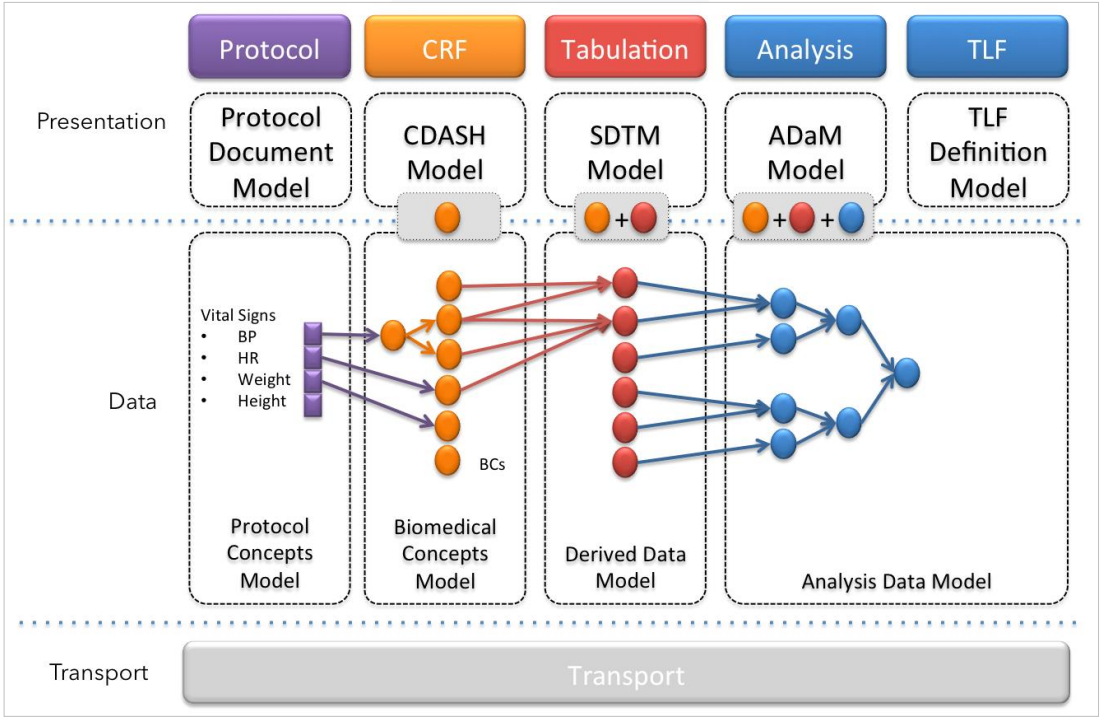
New requests or changes to existing terminology can be accessed through the NCI/EVS New Term Request Page and is available for direct download on this page.

As of 31 March 2017 the SDTM, ADaM, SEND and Protocol Entities Controlled Terminology files have been updated on the NCI-EVS Ftp site. The dates of the new files are 2017-03-31. These terminology files replace all older SDTM, ADaM, and SEND files and include terms from Review Package 29. Protocol Entities is a brand new CDISC terminology set. There are approximately 111 new QRS terms and 241 new terms across SDTM, ADaM, SEND and Protocol Entities.

Vision

Vision

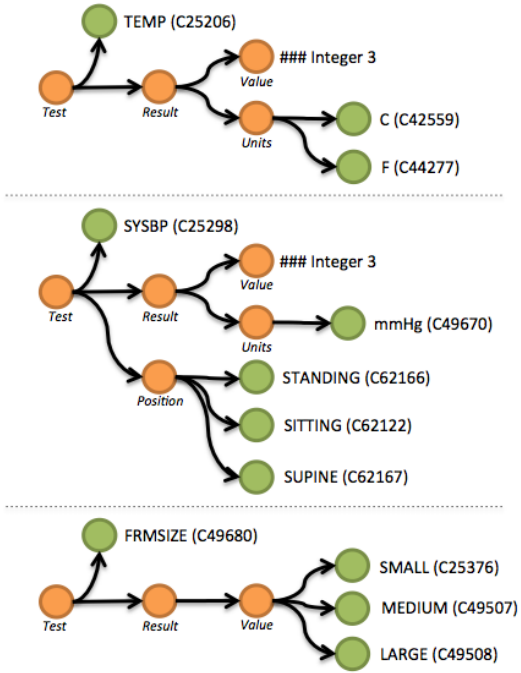




1. Utilize technology, notably semantic / graph methods
2. Build from the bottom
3. Iterate: small steps, learn, adjust

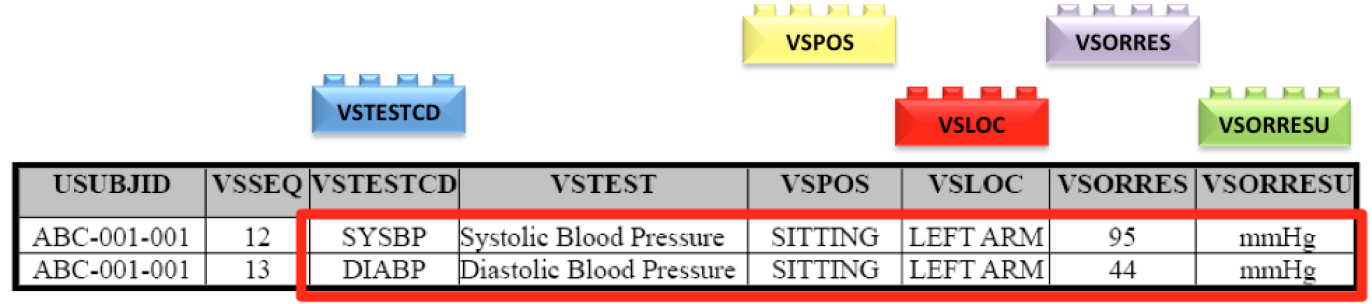
Biomedical Concepts

1. Better structure and relationships, not just individual variables 'floating'. A collection of variables with logical and consistent structure.
2. Note the definition of value level metadata
3. **CDISC have just released an updated SHARE model for review/comment containing BCs**
4. **BCs derived from concept maps as seen in TA User Guides**
5. Can exist independent of SDTM & CDASH

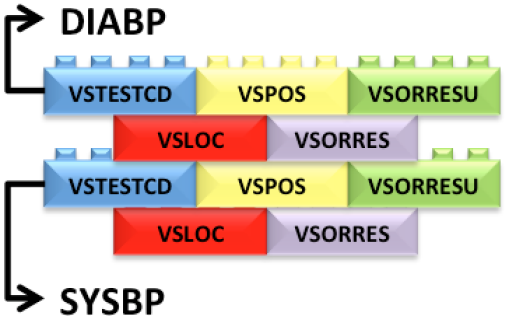


1. **Industry must share these definitions, they need to be standard.**
2. We know all of this information, it is currently within our data and our define.xml files

An Old Slide From 2009



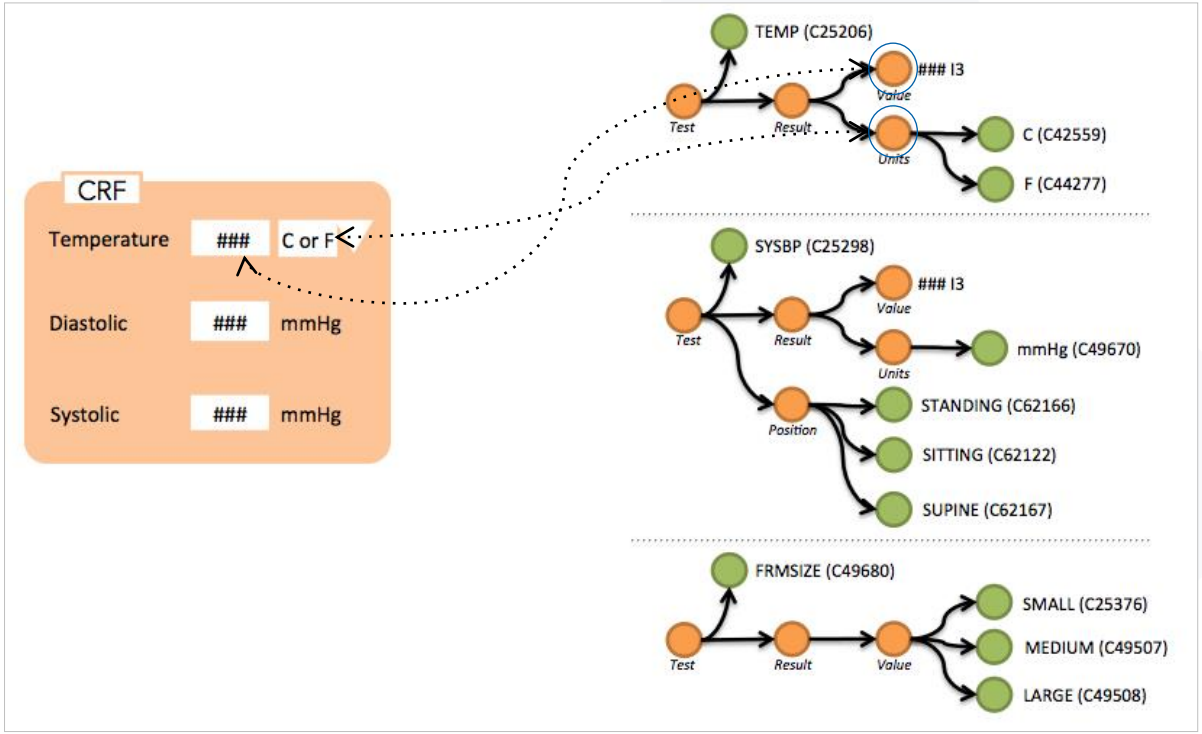
USUBJID	VSSEQ	VSTESTCD	VSTEST	VSPOS	VSLOC	VSORRES	VSORRESU
ABC-001-001	12	SYSBP	Systolic Blood Pressure	SITTING	LEFT ARM	95	mmHg
ABC-001-001	13	DIABP	Diastolic Blood Pressure	SITTING	LEFT ARM	44	mmHg



and blood pressure is the two together

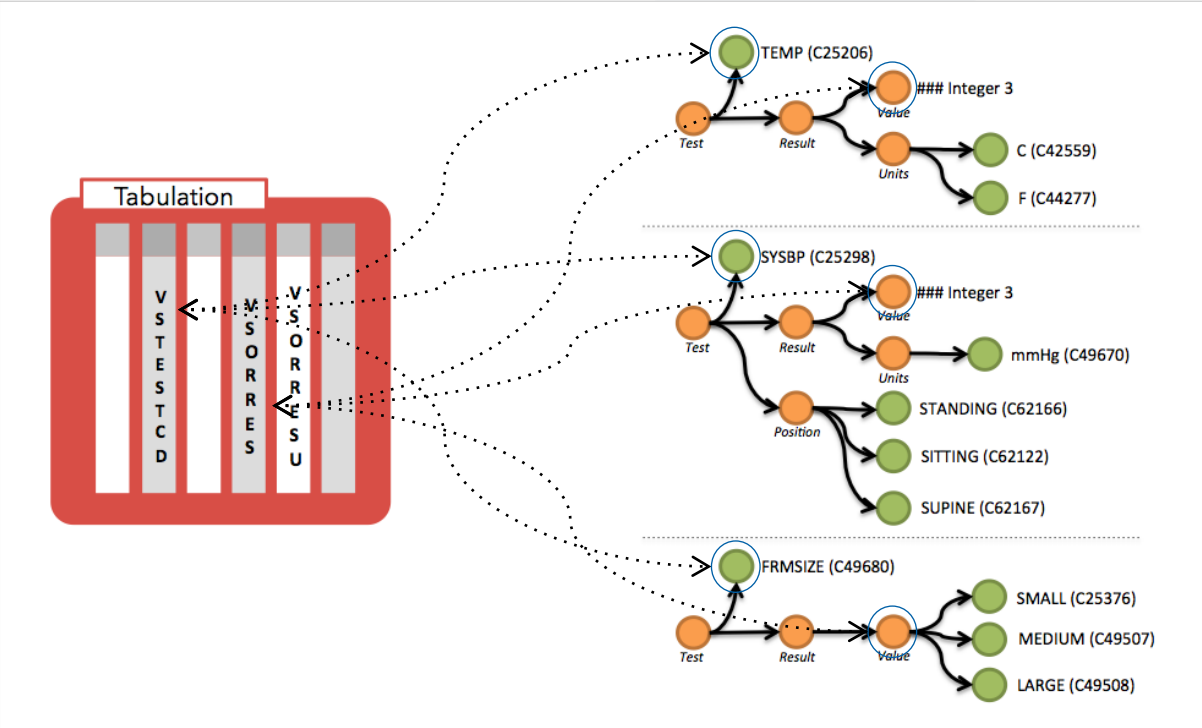
1. We are very good at 'losing' relationships in our data.
2. Our building block today is the variable.
3. It needs to be the next level up, a 'concept' that includes the relationships

Biomedical Concepts & Forms



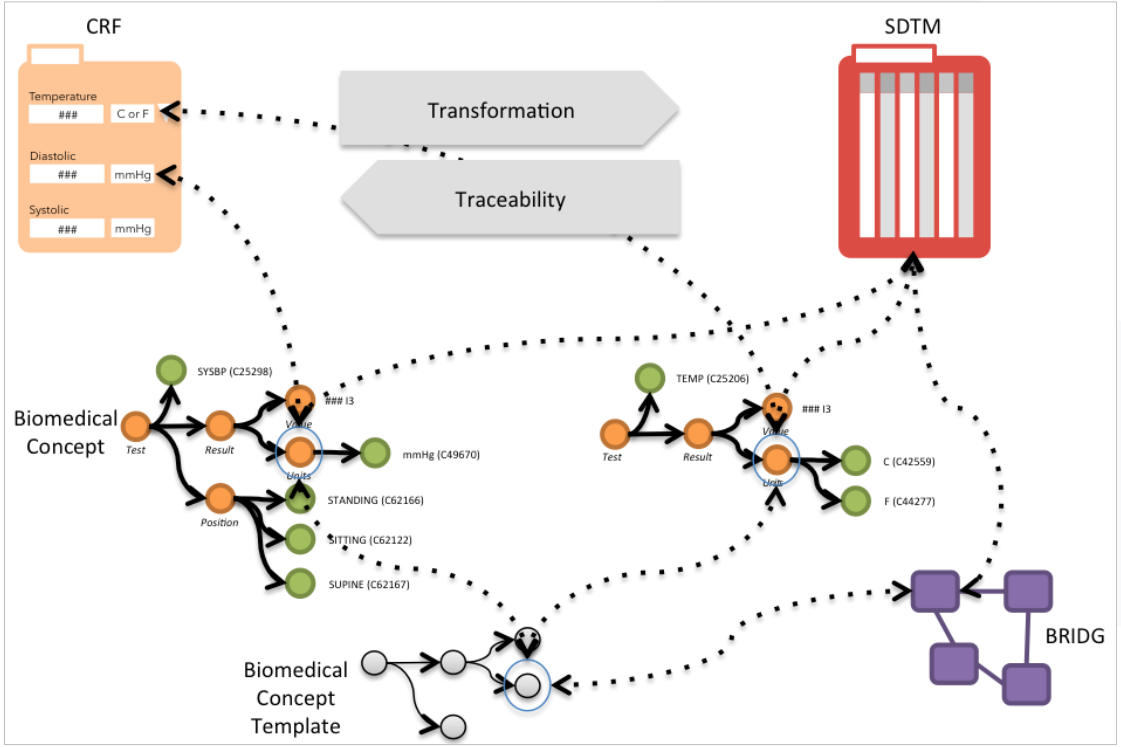
Note: Only a couple of relationships shown so as to convey the principle.

Biomedical Concepts & Domains



Note: Only a couple of relationships shown so as to convey the principle

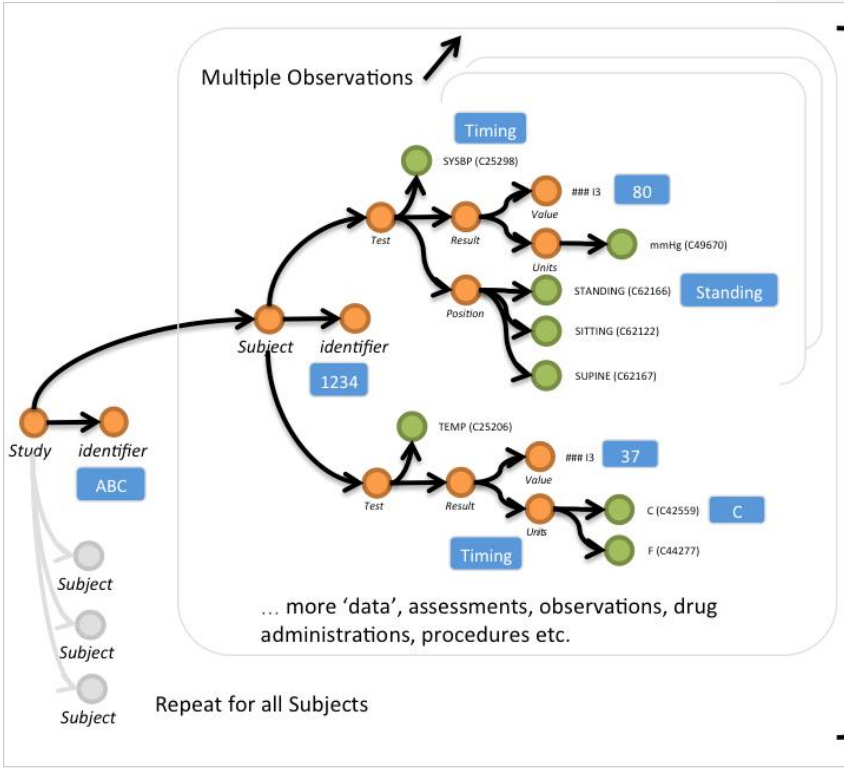
Using Biomedical Concepts



Notes:

1. A template is used to build BCs. There will be many templates.
2. The template is based on the BRIDG model.
3. BRIDG provides an invisible reference framework.
4. The dotted links provide machine capable automated processing and traceability.
5. Biomedical Concepts need only be associated with the target domain, the machine can link the individual pieces itself

SDTM Data I



1. The next step is to start generating domains automatically from data
2. Having a machine representation of the metadata (the study definition) allows for the data to be 'attached'
3. Then significant parts of a domain can be created via a query. SDTM is a view of the data
4. The data can be used for other purposes
5. Learn and iterate to allow for the other variables to be created automatically
6. **This is the subject of a FDA/PhUSE CSS Project**

Query would be something like:

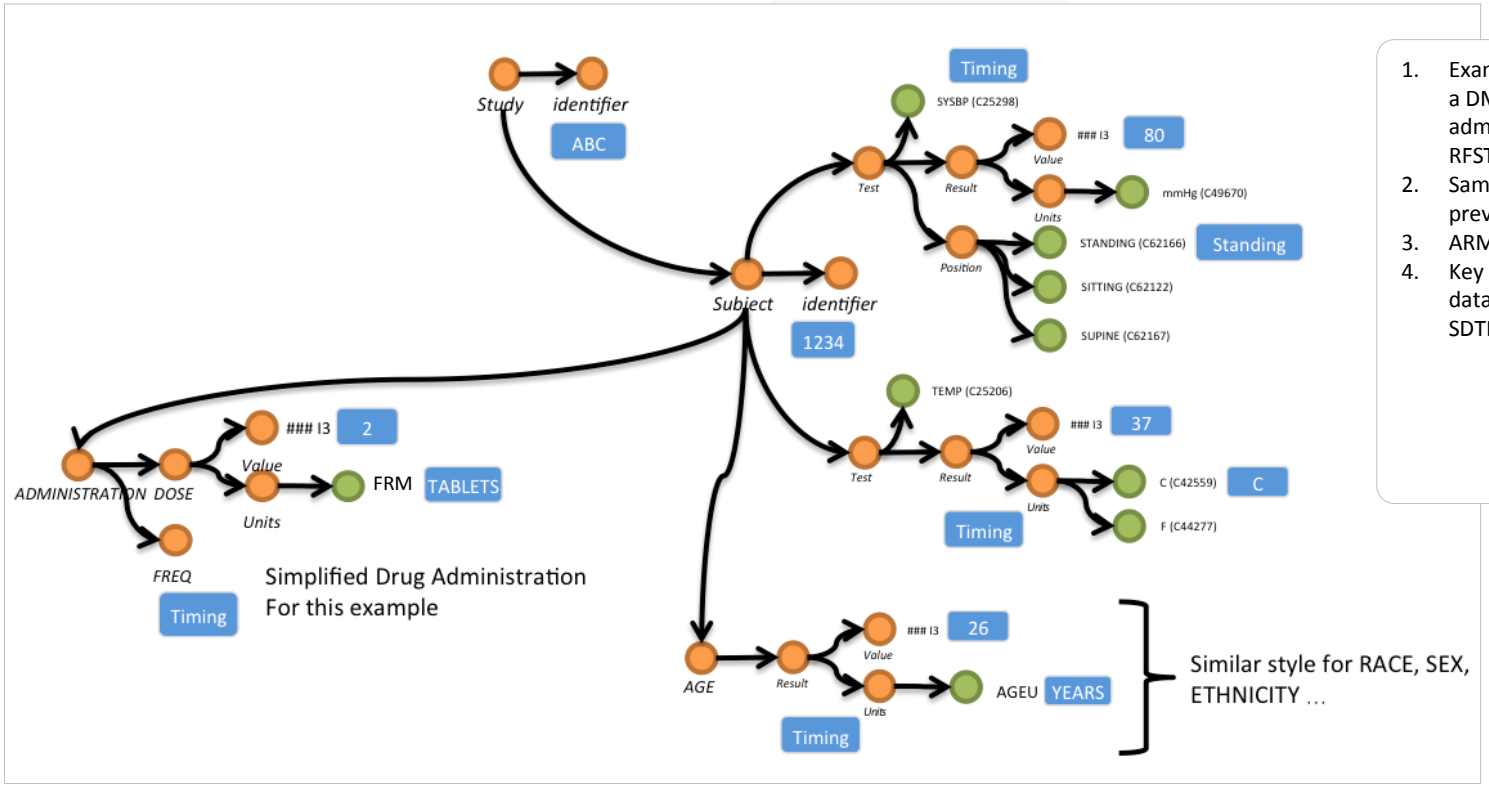
"For all VS tests start at study ABC, for all subjects get the identifier, get the test code, the result, additional properties and the date/time of the observation."

Query driven by the domain definition.

ABC	VS	ABC-1234	SYSBP	Standing	80	mmHg	…	Timing
ABC	VS	ABC-1234	TEMP		37	C	…	Timing

Standard Results & Units, Status, Baseline & Derived

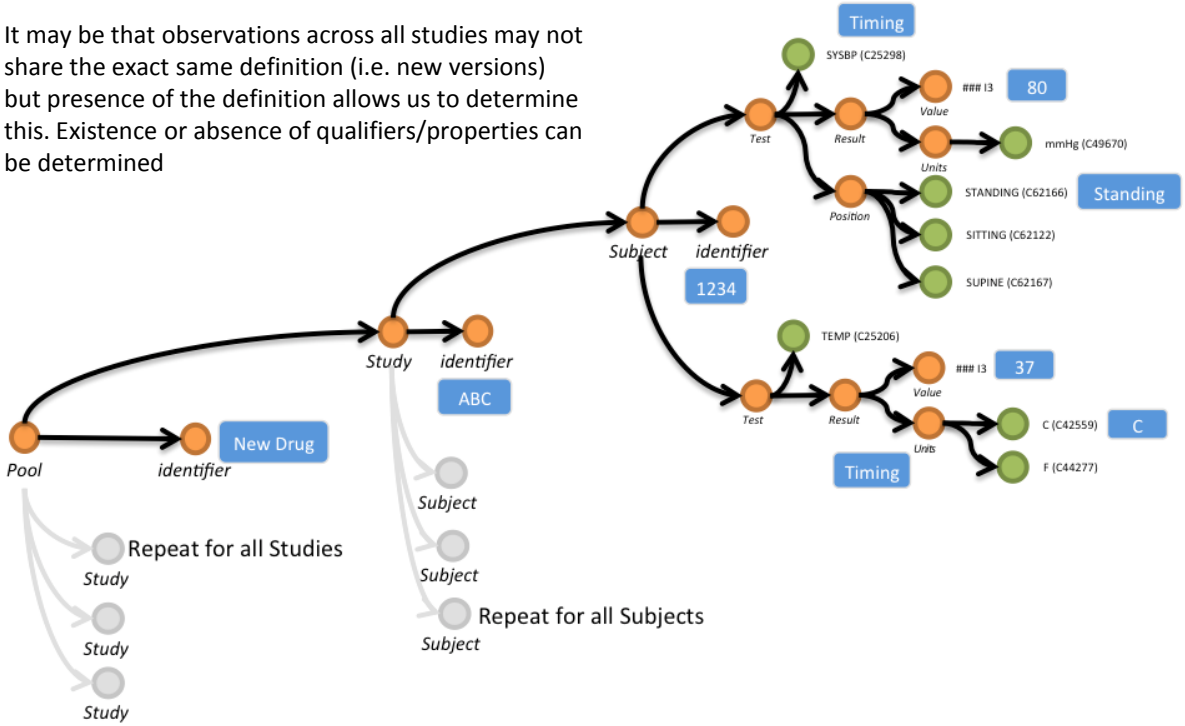
SDTM Data II



1. Example to give a flavour of generating a DM domain and the link to drug administration events to obtain RFSTDTC
2. Same query can be used for –DY for previous VS example.
3. ARM gets interesting!
4. Key high-level point is the raw captured data versus all of the 'derived' data in SDTM

Pooled Data

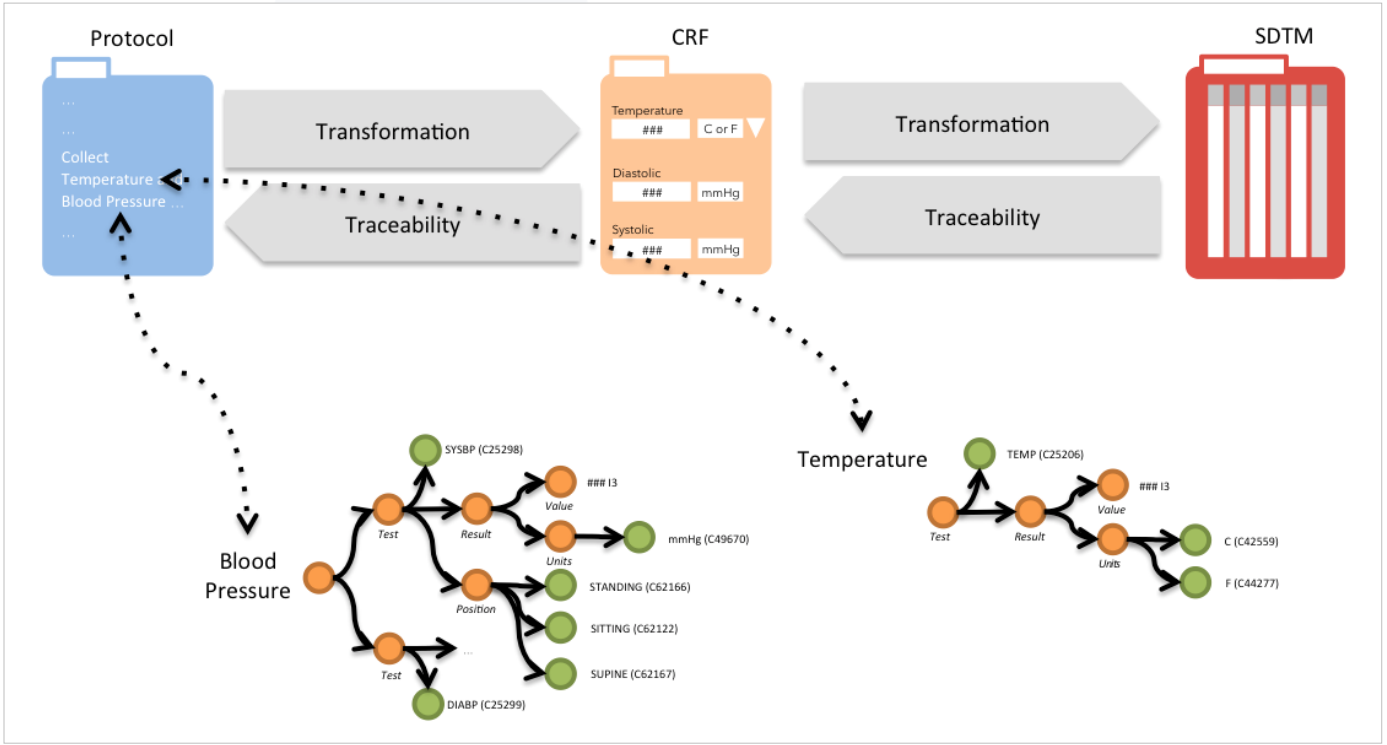
It may be that observations across all studies may not share the exact same definition (i.e. new versions) but presence of the definition allows us to determine this. Existence or absence of qualifiers/properties can be determined



1. Can get to larger pools of data
2. Query for a domain is the same, starting point is the Pool rather than the Study

Protocol

1. Link the names of the concepts back into the protocol document.
2. From a quick initial examination this looks possible with the TransCelerate protocol tool



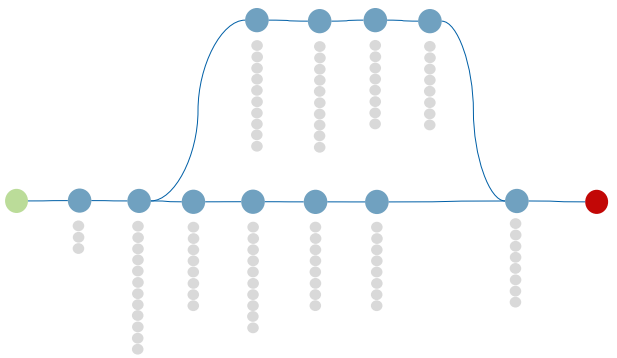
Protocol

Topic	Variable	Year						
		0	1		2	3	4	5/7	6/8
		Pre-V	V1a	V1b	V2	V3	V4	V5/V7	V6/V8
	Eligibility Form (Inclusion & Exclusion Criteria)	♦							
	Consent Form and Study Brochure	♦							
	Family Binder		♦	♦	♦	♦	♦	♦	♦
Kidney	Iohexol-based GFR		X		X		X		X
	Cystatin C		X		X	X	X	X	X
	Serum Creatinine		X		X	X	X	X	X
	Central Renal Panel[a]		X		X	X	X	X	X
	Central Uric Acid[b]		X		X	X	X	X	X
	Central Urine Creatinine and Protein		X		X	X	X	X	X
	Central Urine Albumin		X		X	X	X	X	X
	Local Complete Blood Count[c]		X		X	X	X	X	X
	Local Pregnancy Tests[d]		X		X	X	X	X	X
	Local Renal Panel[e]		X		X	X	X	X	X
	Local Urine Creatinine and Urine Protein[f]		X		X	X	X	X	X
Cardiovascular	Clinical Blood Pressure (centrally calibrated)		■	■	■	■	■	■	■
	Clinical Blood Pressure (locally measured)		■	■	■	■	■	■	■
	Lipid Profile				■		■		■
	Ambulatory Blood Pressure Monitoring				■		■		■
	Echocardiography				■		■		■
	Carotid Intima-Media Thickness[g]				■		■		■
	Pulse Wave Velocity[h]				■		■		■
	Cardiac Magnetic Resonance Imaging (MRI)[i]						■		
Neurocognitive	Pediatric Quality of Life			▲	▲	▲	▲	▲	▲
	Cognitive and Development Assessments			▲		▲		▲	
	Behavioral Assessments			▲		▲		▲	
Growth	Height/Length and Weight		●	●	●	●	●	●	●
	Head Circumference[j]		●	●	●	●	●	●	●
	Mid-Arm Circumference[j]		●	●	●	●	●	●	●
	Waist and Hip Circumferences[k]		●	●	●	●	●	●	●
	Tanner Stage		●		●	●	●	●	●
	Food Frequency Questionnaire (FFQ)			●	●	●	●	●	●
	Intact Parathyroid Hormone (iPTH)			●		●		●	
	High Sensitivity CRP (hsCRP)			●		●		●	
	Vitamin D			●		●		●	
	Fibroblast Growth Factor-23 (FGF-23)			●		●		●	
	6 Minute Walk Test (6MWT)					●		●	
	Grip Strength					●		●	

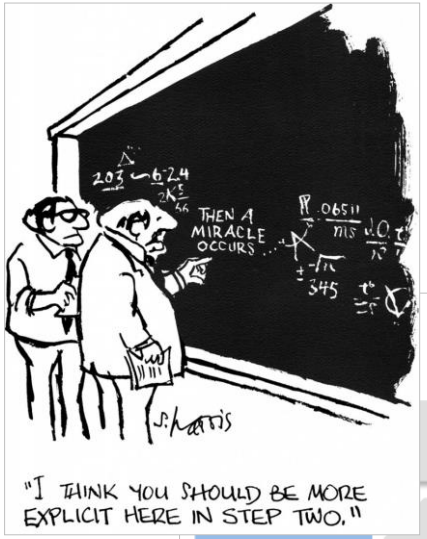
[a] **Central Renal Panel**: Blood drawn at each site and sent to Central Biochemistry Laboratory (CBL) where basic metabolic panel, phosphorous, and albumin are performed.
[b] **Cohort 2**: For Cohort 2, these tests will be measured at baseline and annual visits. For Cohort 1, the measurements of these tests were initiated at follow-up.
[c] **Local CBC**: The local laboratory at each clinical site will perform CBC tests.
[d] **Pregnancy Tests**: Pregnancy tests will be performed on females of child bearing potential. Childbearing potential occurs when the female has reached menarche.
[e] **Local Renal Panel**: Clinical sites will perform renal panel tests at their local laboratory for all participants in addition to the central renal panel tests that are sent to CBL.
[f] **Local Urine Creatinine and Urine Protein**: Clinical sites that require immediate results will perform urine creatinine and urine protein tests at their local laboratory in addition to the tests that are sent to CBL.
[g] **IMT**: At selected sites, sub-set of the entire cohort will have carotid IMT performed (N=100).
[h] **PWV**: At selected sites, sub-set of the entire cohort will have pulse wave velocity performed.
[i] **Cardiac MRI**: Sub-set of the entire cohort with a high probability of reaching ESRD will have cardiac MRI performed.
[j] **Head Circumference**: Head circumference will be measured at every study visit for children 3 years old and younger.
[k] **Mid-Arm Circumference**: Mid-arm circumference will be measured at every study visit for the entire cohort.
[l] **Waist and Hip Circumferences**: Waist and Hip circumference will be measured at every study visit for the entire cohort.

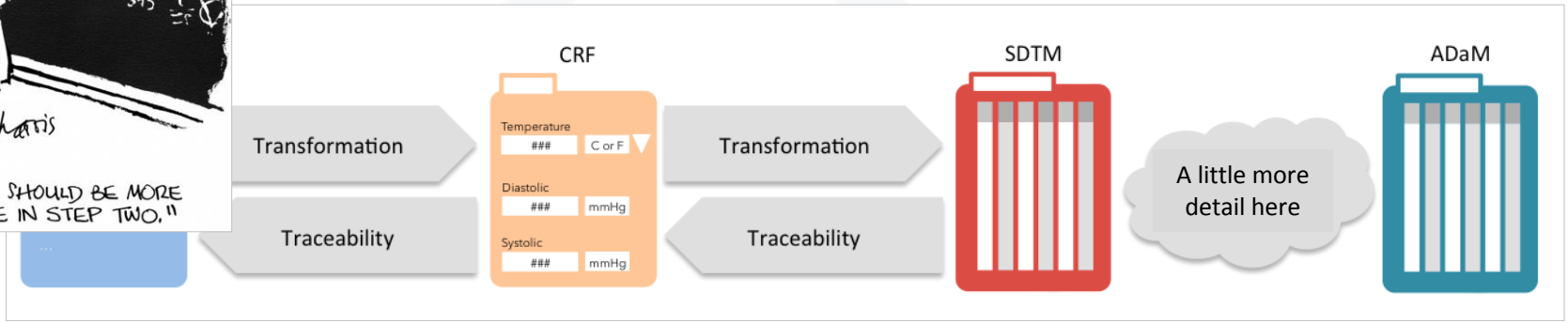
1. Is it possible to base protocols on the concepts to be collected rather than the form?
2. Create timelines rather than the rectangular ‘Schedule of Assessments’ we see today.
3. Link study designs in an explicit way to the data to be collected.
4. **Can still be presented in ‘traditional ways’.**
5. Important to remember to separate the data and its presentation.

Warning! ☺
Dangerous Thoughts

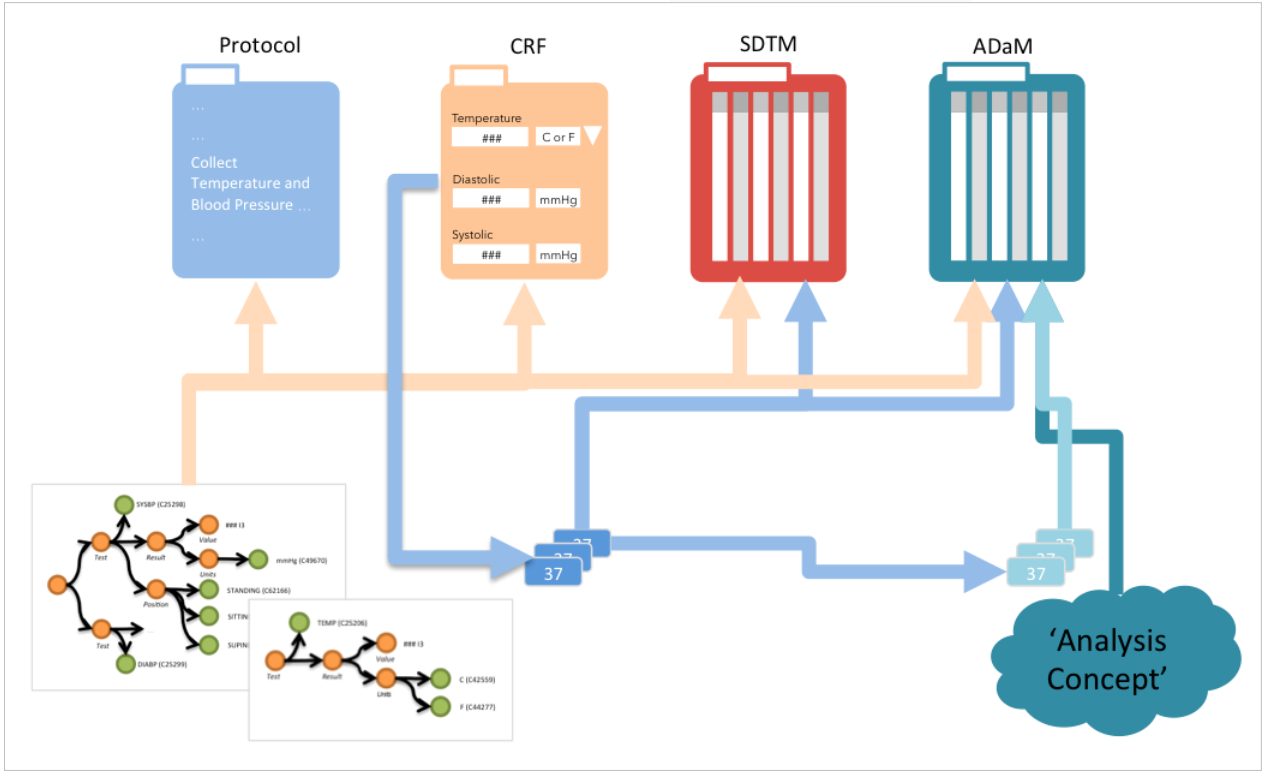


Analysis



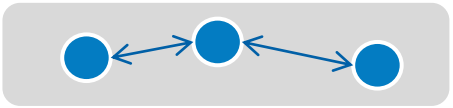


A Big Step



1. Results in an underlying model used across the life-cycle.
2. Requires 'analysis concepts', some thinking has been done in this area.
3. Large amount of work being done in sponsors re TLFs

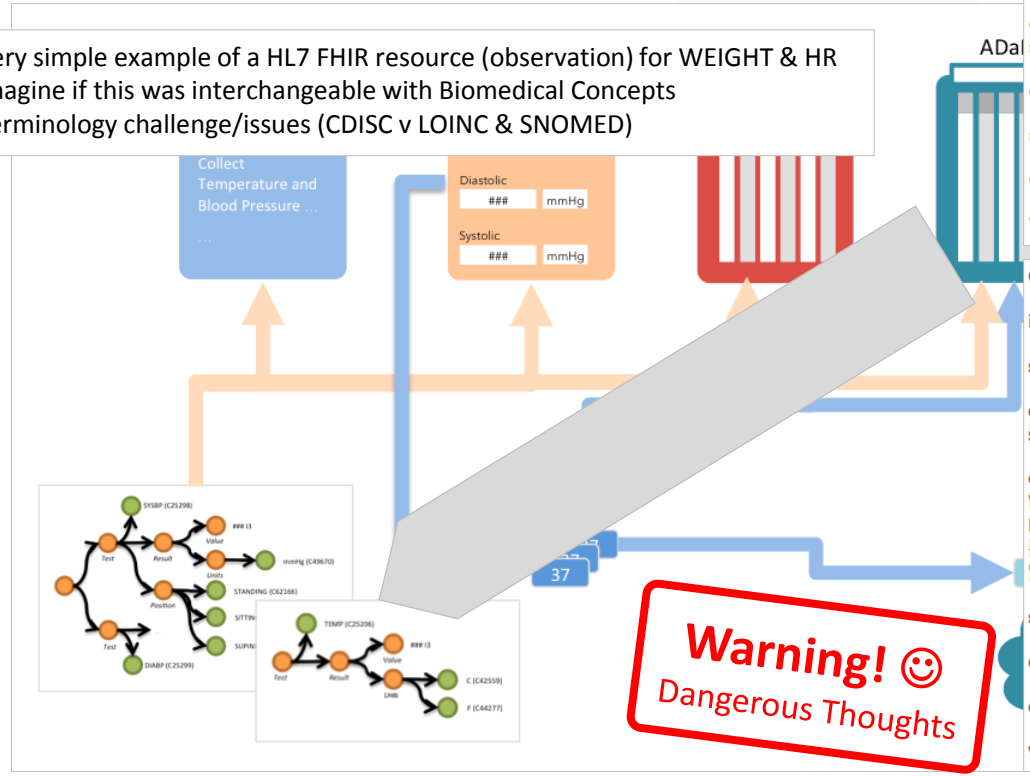
1. Forward and Reverse links through the entire lifecycle.
2. Traceability comes for free.



Warning! ☺
Dangerous Thoughts

Healthcare

1. Very simple example of a HL7 FHIR resource (observation) for WEIGHT & HR
2. Imagine if this was interchangeable with Biomedical Concepts
3. Terminology challenge/issues (CDISC v LOINC & SNOMED)



Generated Narrative with Details

id: heart-rate

meta:

status: final

category: Vital Signs (Details : {http://hl7.org/fhir/observation-category code 'vital-signs' = 'Vital Signs', given as 'Vital Signs'})

code: Heart rate (Details : {LOINC code '8867-4' = 'Heart rate', given as 'Heart rate'})

subject: Patient/example

effective: 02/07/1999

value: 44 beats/minute (Details: UCUM code /min = '/min')

Generated Narrative with Details

id: example

status: final

category: Vital Signs (Details : {http://hl7.org/fhir/observation-category code 'vital-signs' = 'Vital Signs', given as 'Vital Signs'})

code: Body Weight (Details : {LOINC code '29463-7' = 'Body weight', given as 'Body Weight'}; {LOINC code '3141-9' = 'Body weight Measured', given as 'Body weight Measured'}; {SNOMED CT code '27113001' = 'Body weight', given as 'Body weight'}; {http://acme.org/devices/clinical-codes code 'body-weight' = 'body-weight', given as 'Body Weight'})

subject: Patient/example

context: Encounter/example

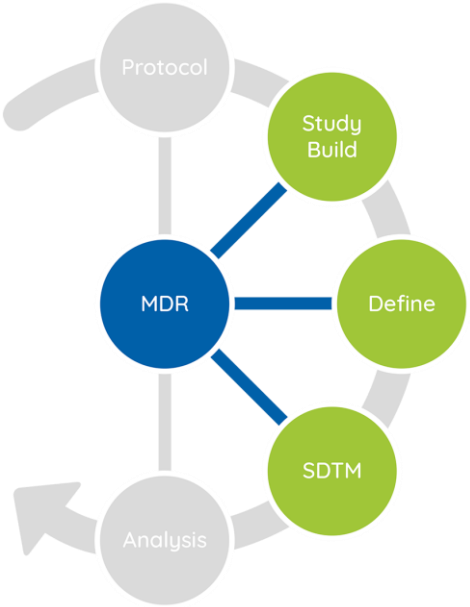
effective: 28/03/2016

value: 185 lbs (Details: UCUM code [lb_av] = 'lb_av')

Practical

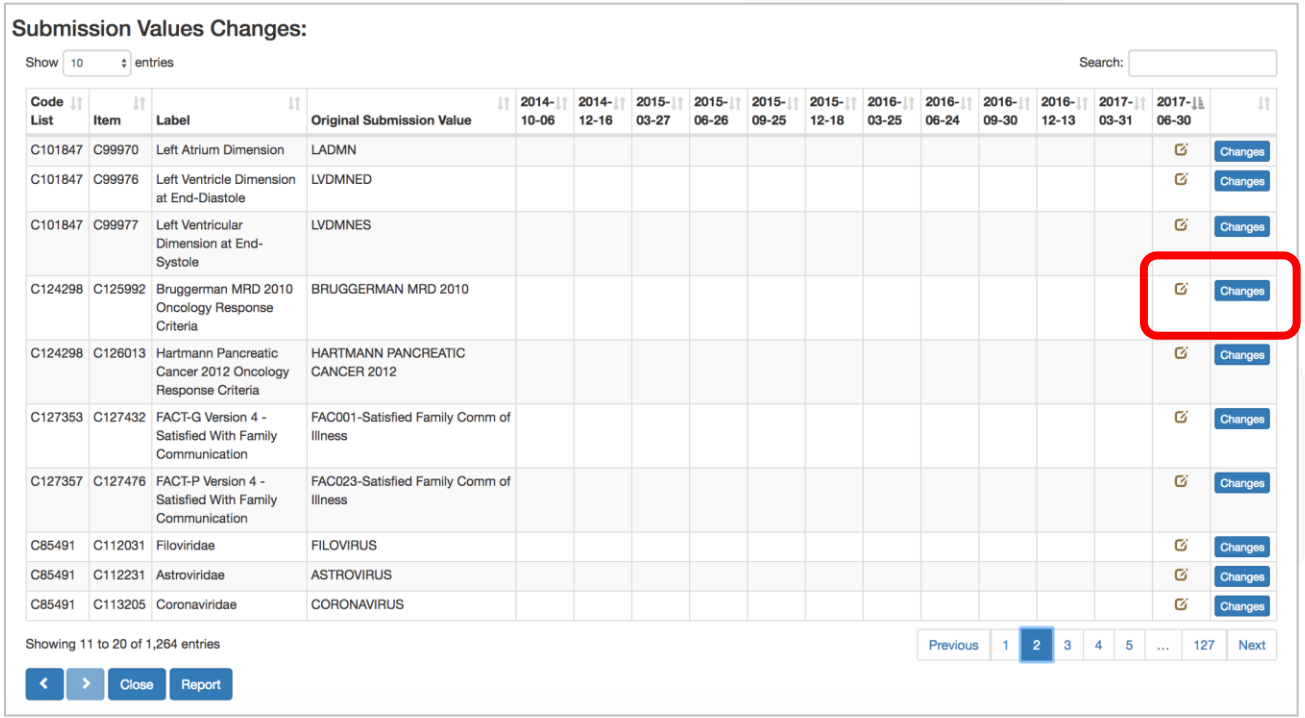
Today …

Tool	Description	Status
MDR	Metadata Repository handling CDISC Terminology, Biomedical Concepts, Forms and SDTM Model, Implementation Guides and Custom Domains	Production
Study Build	A tool that takes definitions from the MDR and allows studies to be constructed and then exported in ODM and ALS formats suitable for loading by EDC tools	Production
Define	Allows for a define.xml file to be built based on either a study definition, an existing file or without any information. Uses the MDR to aid the user into populating the define.xml	In Construction
SDTM	Takes captured data and facilitate the production of SDTM domains based on the study build information	Prototype



Assero
S-CUBED
A3 INFORMATICS

Something Simple



Submission Values Changes:

Show 10 entries — Search:

Code List	Item	Label	Original Submission Value	2014-10-06	2014-12-16	2015-03-27	2015-06-26	2015-09-25	2015-12-18	2016-03-25	2016-06-24	2016-09-30	2016-12-13	2017-03-31	2017-06-30	
C101847	C99970	Left Atrium Dimension	LADMN													Changes
C101847	C99976	Left Ventricle Dimension at End-Diastole	LVDMNED													Changes
C101847	C99977	Left Ventricular Dimension at End-Systole	LVDMNES													Changes
C124298	C125992	Bruggerman MRD 2010 Oncology Response Criteria	BRUGGERMAN MRD 2010													Changes
C124298	C126013	Hartmann Pancreatic Cancer 2012 Oncology Response Criteria	HARTMANN PANCREATIC CANCER 2012													Changes
C127353	C127432	FACT-G Version 4 - Satisfied With Family Communication	FAC001-Satisfied Family Comm of Illness													Changes
C127357	C127476	FACT-P Version 4 - Satisfied With Family Communication	FAC023-Satisfied Family Comm of Illness													Changes
C85491	C112031	Filoviridae	FILOVIRUS													Changes
C85491	C112231	Astroviridae	ASTROVIRUS													Changes
C85491	C113205	Coronaviridae	CORONAVIRUS													Changes

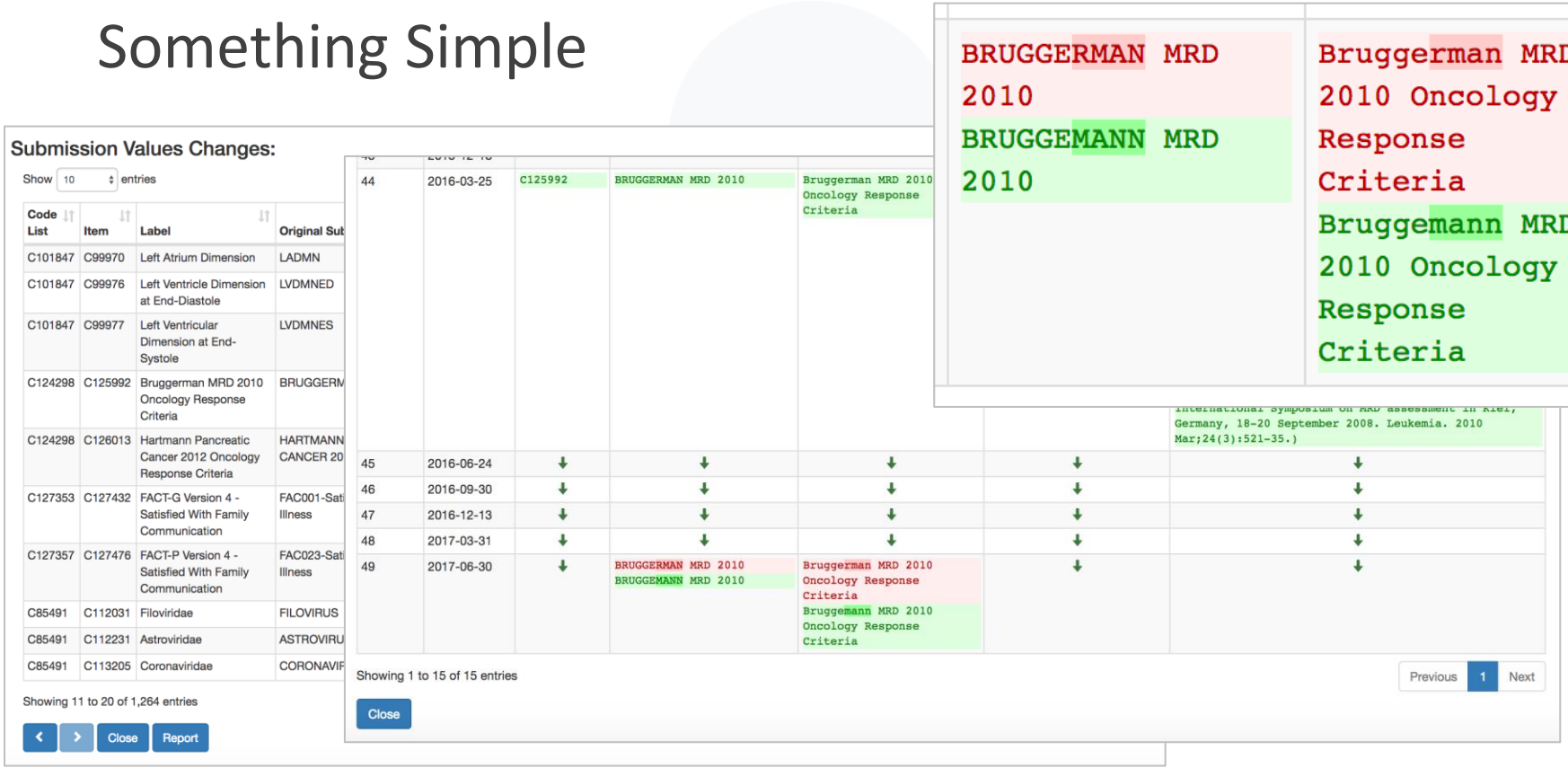
Showing 11 to 20 of 1,264 entries

Previous 1 2 3 4 5 ... 127 Next

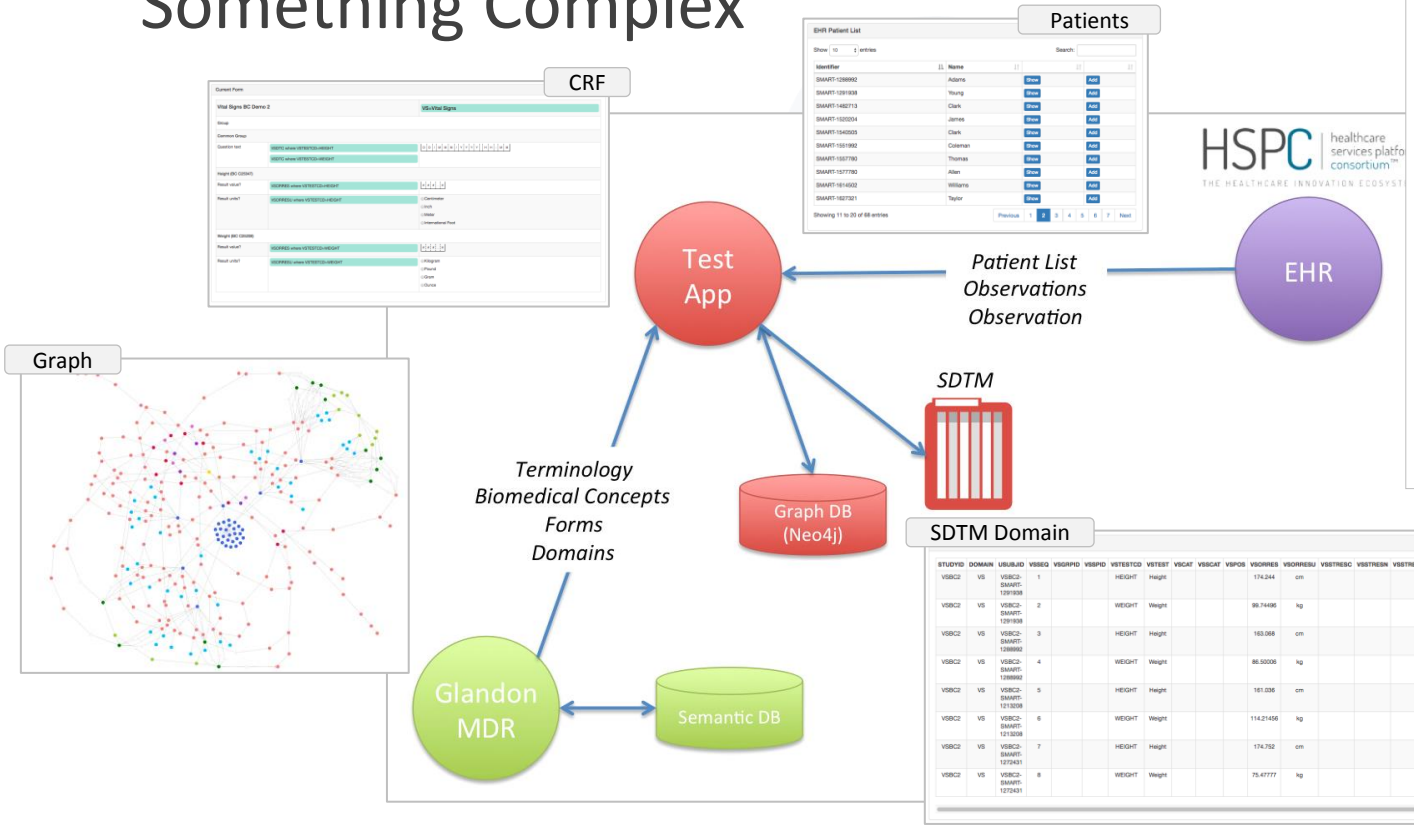
Close Report

1. Latest CDISC terminology release June 30th, 2017.
2. Loaded and differences available within minutes.
3. Can also be used within impact analysis, which items use the terminology, which items use those item, e.g. Terminology -> BCs -> Forms -> Domains.
4. Gives control.

Something Simple



Something Complex



1. Test application to test and demonstrate some of these ideas.
2. Use HL7 FHIR to obtain patient data (map LOINC/UCUM -> CDISC Terminology mapping).
3. Map to form selected from Glandon MDR built using Biomedical Concepts. Can build form on the fly and populate.
4. Put into graph (in effect a simple data warehouse) for multiple subjects.
5. Extract a presentation of the data (SDTM) using domain definition from Glandon MDR.

[Full Description: http://www.assero.co.uk/2017/all-the-toys-graphs-fhir-and-cdisc/](http://www.assero.co.uk/2017/all-the-toys-graphs-fhir-and-cdisc/)

Summary

Summary

- Build a linked view of the world and **break the silos down**.
- Consistent way to **generate the existing views** of that world.
- Do this in iterative fashion that **allows the old and the new to co-exist**.
- Solid foundation of linked high-quality definitions consistent across the life-cycle present opportunities for automation.

Contact Details

Dave Iberson-Hurst
dave.iberson-hurst@assero.co.uk
www.assero.co.uk