CDISC 360i
Connecting Standards from Design to Analysis to Submission

Peter Van Reusel

Sam Hume

Meet the Speaker

Peter Van Reusel

Title: Chief Standards Officer
Organization: CDISC

Peter Van Reusel provides executive leadership to the development and implementation of clinical standards in line with CDISC's strategy and operational plans, working closely with the President and CEO, as well as CDISC staff and stakeholders. He has over 20 years' experience in senior roles in pharma and at CROs, providing standards expertise and carrying out other standards work in various organizational settings. A long-time, CDISC-authorized instructor, Peter has helped significantly in developing CDISC training courses.

He previously served as CDISC's European Liaison, fostering relationships with key European regulatory, academic, and biopharma stakeholders. Peter is also an active PHUSE collaborator.

Meet the Speaker

Sam Hume

Title: Principal Consultant

Organization: CDISC

Sam Hume co-leads the CDISC Data Exchange Standards team, advises CDISC leadership on strategy and technical topics, and contributes to COSA, CORE, CDISC Library, and other CDISC projects. Sam formerly served as the CDISC VP of Data Science. During his 30 years in the biopharmaceutical industry, he has held several senior-level technology positions. Sam is an active PHUSE contributor. He holds a doctorate in Information Systems.

Agenda

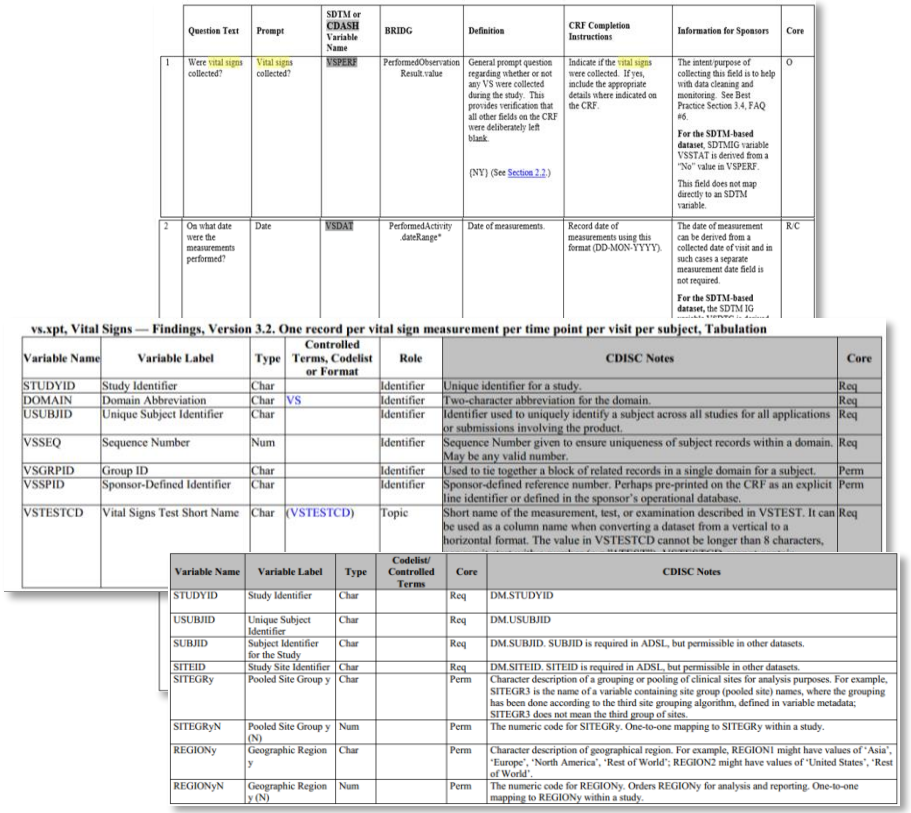
- CDISC 360 lessons learned & where are we now
- CDISC 360i: implementing the vision

CDISC 360 lessons learned

What were we trying to solve

Two dimensional standards

- CDISC Foundational models provide much needed structure
  - Normative Content
  - 2 dimensional (tables, columns)
  - Standard to represent data
- The Information itself is not defined
  - We do not need new structures
  - We need to define
    - Entities
    - Semantics (meaning)
    - Relationships between information
    - Rules in the data lifecycle



	Question Text	Prompt	SDTM or CDASH Variable Name	BRIDG	Definition	CRF Completion Instructions	Information for Sponsors	Core
1	Were vital signs collected?	Vital signs collected?	VSPERF	PerformedObservation Result.value	General prompt question regarding whether or not any VS were collected during the study. This provides verification that all other fields on the CRF were deliberately left blank. {NY} (See Section 2.2.)	Indicate if the vital signs were collected. If yes, include the appropriate details where indicated on the CRF.	The intent/purpose of collecting this field is to help with data cleaning and monitoring. See Best Practice Section 3.4, FAQ #6. **For the SDTM-based dataset**, SDTMIG variable VSSTAT is derived from a "No" value in VSPERF. This field does not map directly to an SDTM variable.	O
2	On what date were the measurements performed?	Date	VSDAT	PerformedActivity dateRange*	Date of measurements.	Record date of measurements using this format (DD-MON-YYYY).	The date of measurement can be derived from a collected date of visit and in such cases a separate measurement date field is not required. **For the SDTM-based dataset**, the SDTM IG	R/C

vs.xpt, Vital Signs — Findings, Version 3.2. One record per vital sign measurement per time point per visit per subject, Tabulation

Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role	CDISC Notes	Core
STUDYID	Study Identifier	Char		Identifier	Unique identifier for a study.	Req
DOMAIN	Domain Abbreviation	Char	VS	Identifier	Two-character abbreviation for the domain.	Req
USUBJID	Unique Subject Identifier	Char		Identifier	Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	Req
VSSEQ	Sequence Number	Num		Identifier	Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	Req
VSGRPID	Group ID	Char		Identifier	Used to tie together a block of related records in a single domain for a subject.	Perm
VSSPID	Sponsor-Defined Identifier	Char		Identifier	Sponsor-defined reference number. Perhaps pre-printed on the CRF as an explicit line identifier or defined in the sponsor's operational database.	Perm
VSTESTCD	Vital Signs Test Short Name	Char	(VSTESTCD)	Topic	Short name of the measurement, test, or examination described in VSTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in VSTESTCD cannot be longer than 8 characters,	Req

Variable Name	Variable Label	Type	Codelist/ Controlled Terms	Core	CDISC Notes
STUDYID	Study Identifier	Char		Req	DM.STUDYID
USUBJID	Unique Subject Identifier	Char		Req	DM.USUBJID
SUBJID	Subject Identifier for the Study	Char		Req	DM.SUBJID. SUBJID is required in ADSL, but permissible in other datasets.
SITEID	Study Site Identifier	Char		Req	DM.SITEID. SITEID is required in ADSL, but permissible in other datasets.
SITEGRy	Pooled Site Group y	Char		Perm	Character description of a grouping or pooling of clinical sites for analysis purposes. For example, SITEGR3 is the name of a variable containing site group (pooled site) names, where the grouping has been done according to the third site grouping algorithm, defined in variable metadata; SITEGR3 does not mean the third group of sites.
SITEGRyN	Pooled Site Group y (N)	Num		Perm	The numeric code for SITEGRy. One-to-one mapping to SITEGRy within a study.
REGIONy	Geographic Region y	Char		Perm	Character description of geographical region. For example, REGION1 might have values of 'Asia', 'Europe', 'North America', 'Rest of World'; REGION2 might have values of 'United States', 'Rest of World'.
REGIONyN	Geographic Region y (N)	Num		Perm	The numeric code for REGIONy. Orders REGIONy for analysis and reporting. One-to-one mapping to REGIONy within a study.

Biomedical Concepts

The CDISC 360 Proof of Concept: Adding a conceptual layer to standards

- Evolve from normative to informative standards
- Create and store standards as concepts which create meaning
- Electronically publish data standards as linked metadata
- Add computer executable process metadata which enables end to end automation
- Develop concept-based standard definitions, and test and demonstrate end-to-end automation of study specification, data processing, and analysis

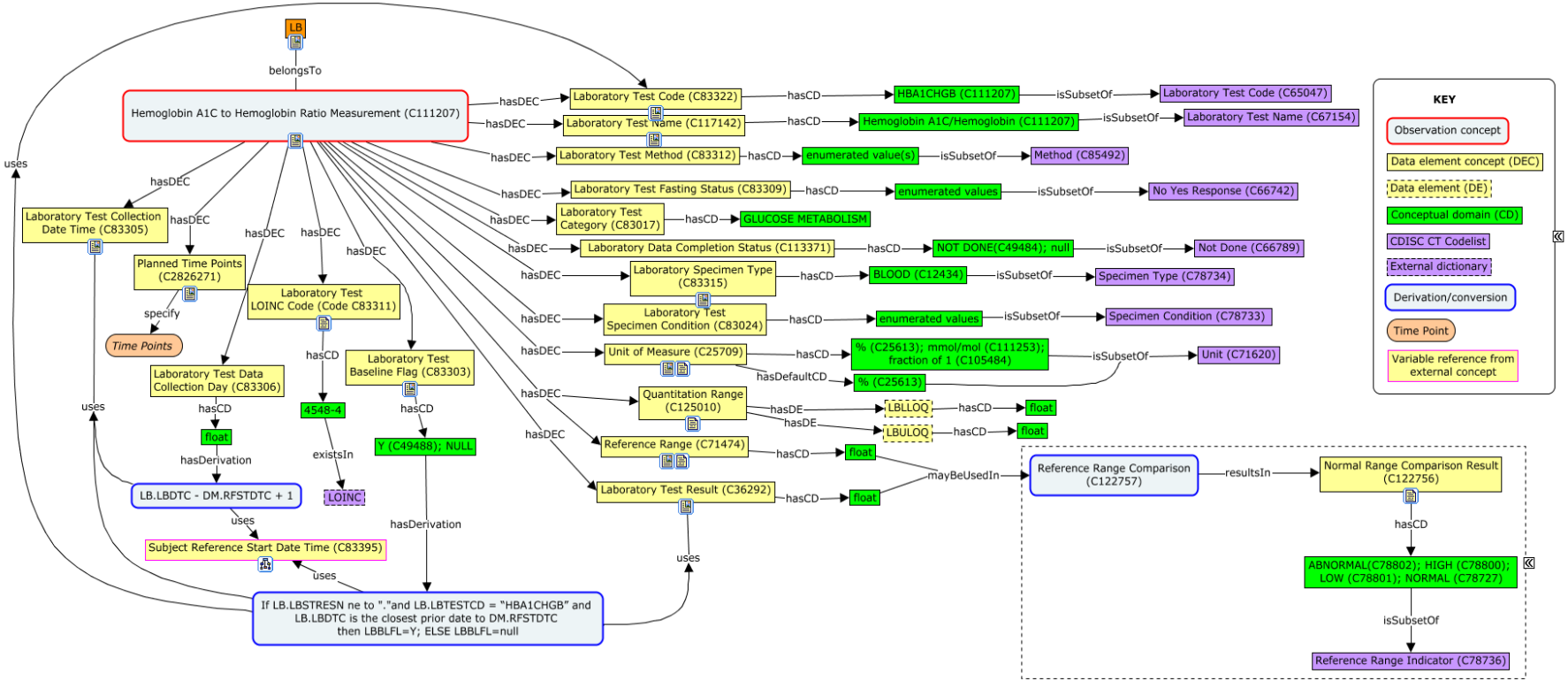
➔ *Test and demonstrate, but* ***not building software***

cdisc

CDISC 360 lessons learned

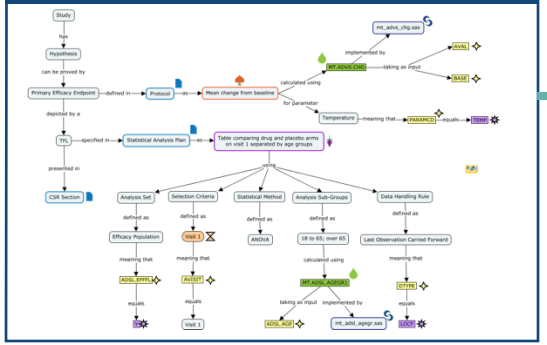
Project approach

Early ideation of a Biomedical Concept



CDISC 360 Workstreams

Workstream 1 - *ENHANCE STANDARDS*



Create concepts in knowledge graphs

Workstream 2 - *PUBLISH STANDARDS*

Load into library

Biomedical Concepts
Analysis Concepts
Foundational Standards

cdisc LIBRARY
formerly SHARE

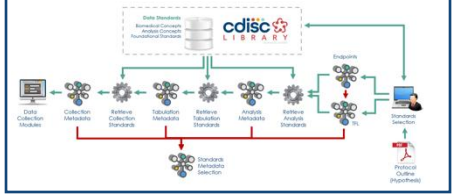
API

Extend API's

Transform concepts in machine readable form

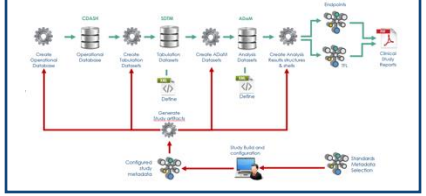
Study
Metadata
Library

Workstream 4 - *DEFINE*



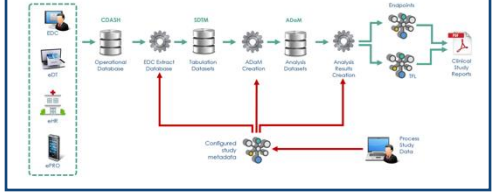
Identify and select standards specification (Use Case 1)

Workstream 5 - *BUILD*



Configure study specification and create artifacts (Use Case 2)

Workstream 6 - *EXECUTE*



Automatically process and transform data (Use Case 3)

CDISC 360 Workstreams

Enhance Standards

Bess LeRoy,
CDISC Head of Standards

Publish Standards

Sam Hume,
CDISC VP Data Science

Study Metadata Library

Industry

Define

Mikkel Traun,
Novo Nordisk

Build

Tianna Umann,
Microsoft

Execute

Bhavin Busa,
Vita Data Sciences

360 Participation Summary

Project Kickoff:

36 Resources specified

20 Organizations

Execution Phase:

107 Resources specified

38 Organizations

- Pharma-Biotech Sponsor: 20
- CRO: 6
- Technology Provider: 11
- Regulatory: 1


abbvie
Allergan
AMGEN
AstraZeneca
Bayer HealthCare
BeiGene
Biogen
BIOMARIN
C&R RESEARCH
Celgene
Cytel
STATISTICAL SOFTWARE & SERVICES
Deloitte.
dMed 绵脉
FDA
formedix
frontier science foundation
GILEAD
Inductive Quotient
Research . Bigdata . Analytics
Johnson & Johnson
Lilly
Mel Consulting
MERCK
Microsoft
novo nordisk
ORACLE
Otsuka
Pfizer
PharmaStat
PINNACLE 21
POINT CROSS
S-CUBED
SANOFI
sas
Syneos Health
TALENT MINE
VITA DATA SCIENCES
a division of SOFTWORLD
xclinical
Clinical Solutions Group

CDISC 360 lessons learned

Where are we today

Standards Development

- **Complete end to end** standards
  - ☑ ☒ Data Collection instruments
  - ☑ Analysis Results
  - ☒ Endpoint definitions
  - ☑ Digital protocol and study design
  - Trial Master File

- **Enrich** existing standards
  - ☑ ☒ Refine and test Biomedical Concepts
  - ☑ ☒ Link concepts to standard implementation
  - ☒ Understand Analysis Concepts and link to Biomedical Concepts
  - ☑ ☒ Add transformations and derivations content

➔ *Digitalize Therapeutic Area Guides*

cdisc

Standards Delivery

- Evolve **library** technology and schema
  - ☑ ☒ Refine and test the BC models
  - ☑ ☒ Refine and deploy BC software tools
  - ☑ Load BC content into the CDISC Library
  - ☑ Surface BC content via APIs and the Library Browser

➔ *Support Automation*

- Evolve toward collaborative **curation**
  - ☑ Develop and rollout governance process
  - ☒ Create CDISC Library curation tools
  - ☑ Develop standards curation training
  - ☒ Enhance CDISC Library to load community standards implementations

➔ *Crowdsource standards*

CDISC 360i: implementing the vision

CDISC 360i

- Define end to end standards
  - Digitalize information from protocol to reporting
  - Link concepts to representation standards
    - Forms definition, eDTs, DHT, SDTM specs, ADaM specs, TFL specs, …
  - Enrich with transformation & derivation logic

- Study design & build
  - Select concept and concept groups in digital Schedule of Activities
  - Automates study builds
    - Forms definition, SDTM specs, ADaM specs, TFL specs, …
  - Provides derivation & transformation algorithms

- Automate data flow
  - Demonstrate end to end automation
    - Starts with linking Schedule of Activities to Concepts (and Concept Groups)
  - Automates transformations & derivation between data states
    - Collection, tabulation, analysis, results

360i Study Package

Metadata	Data	Software	Demonstration
• Publish a complete study metadata package that covers the full study data pipeline from study design through TLFs.	• Publish the complete set of raw datasets for the test study to execute the automated study data pipeline.	• Publish a pre-configured set of open-source software tools that consume the study metadata and data to execute the study design using the test data.	• Execute the study build and data pipeline automation in a Connectathon event to demonstrate generating analysis results from a study design.

Description of the End State

360i has published a complete preconfigured study package with all the components defined in metadata from study design to submission, test data for the study, and software to execute the study data pipeline to generate analysis results

What's Changed since the CDISC 360 POC?

New standards:

- DDF USDM
- Analysis Results Standard
- Biomedical Concepts and Dataset Specializations
- sdtm.oak transformations
- Dataset-JSON for dataset exchange
- Open conformance rules

New software tools:

- Open Study Builder
- TLF Designer
- admiral
- CORE
- sdtm.oak
- Dataset-JSON conversion tools
- Other COSA tools
- Other Pharmaverse tools

These lists are examples and not intended to be comprehensive

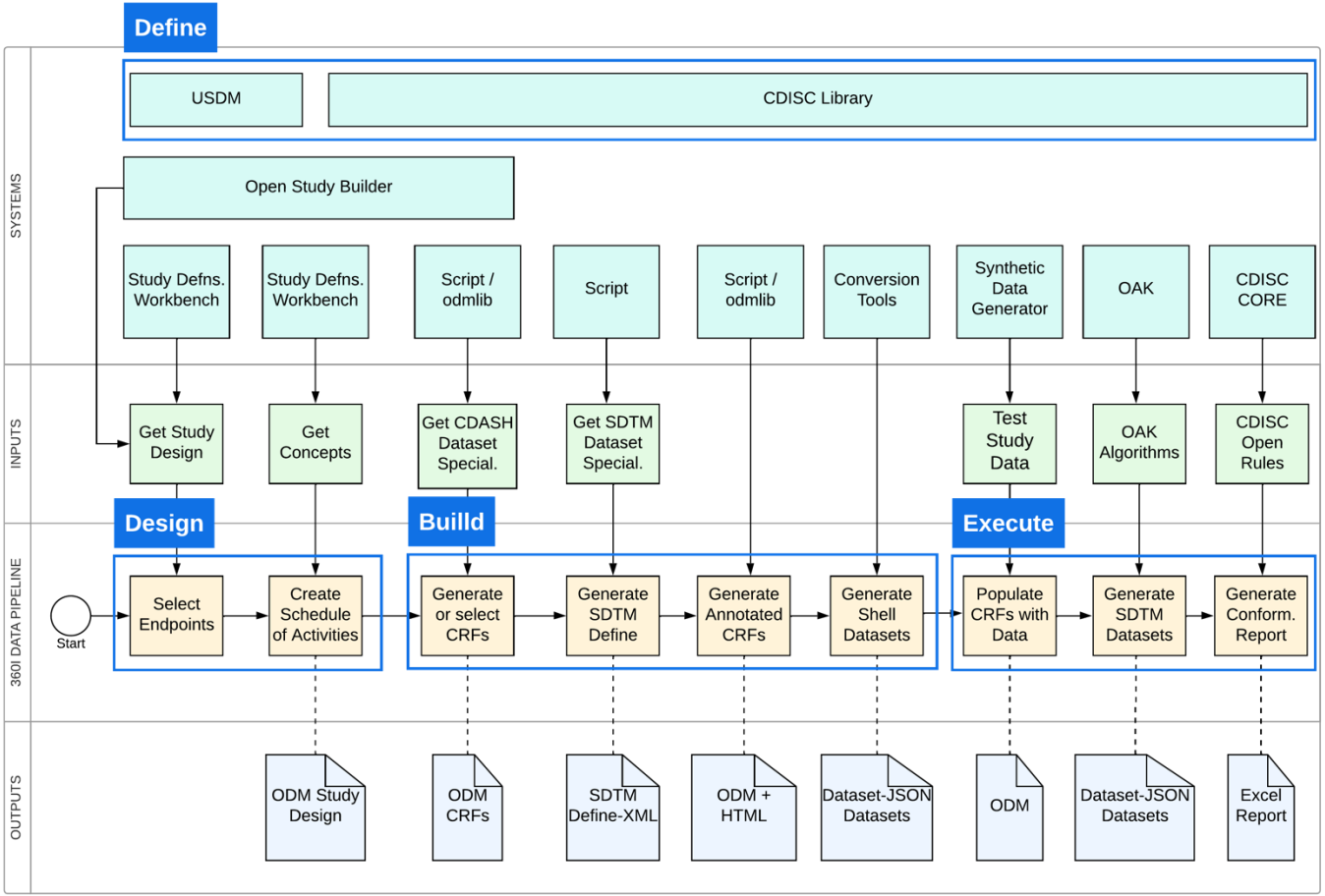
Software Tools since the CDISC 360 POC*

Study Design	Data Collection	SDTM	ADaM	Define-XML	ARS
Open Study Builder	odmlib	sdtm.oak	admiral	OpenCST	TLF Designer**
Study Definitions Workbench	ODM XML Stylesheet	Dataset-JSON Tools	Dataset-JSON Tools	defineR	cards
BC Browser		CORE	carver	odmlib	gtsummary
DDF SDR		Smart Dataset Viewer		Define XSL Stylesheet	JMP Clinical**

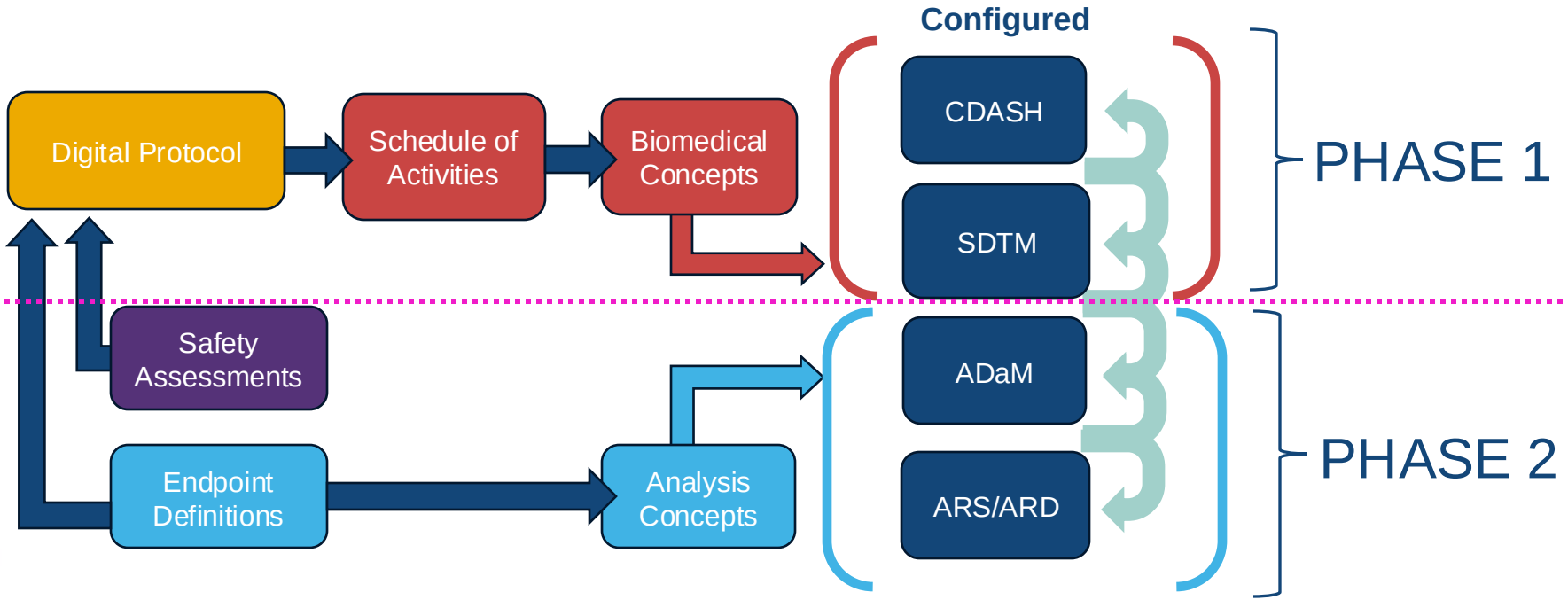
* Examples listed – not a comprehensive listing
** Open-source components expected to be available

CDISC 360i Phase 1: Generating SDTM

Technical Process



360i Phase 2: Generating Analysis Results



Phase 2 pre-work includes defining Analysis Concepts

Standards driven automation: Faster Innovation at Reduced Costs

Community development

End to End standards provide faster innovation

Open-source software

Path of least resistance

Thank You!

We Want Your Feedback!