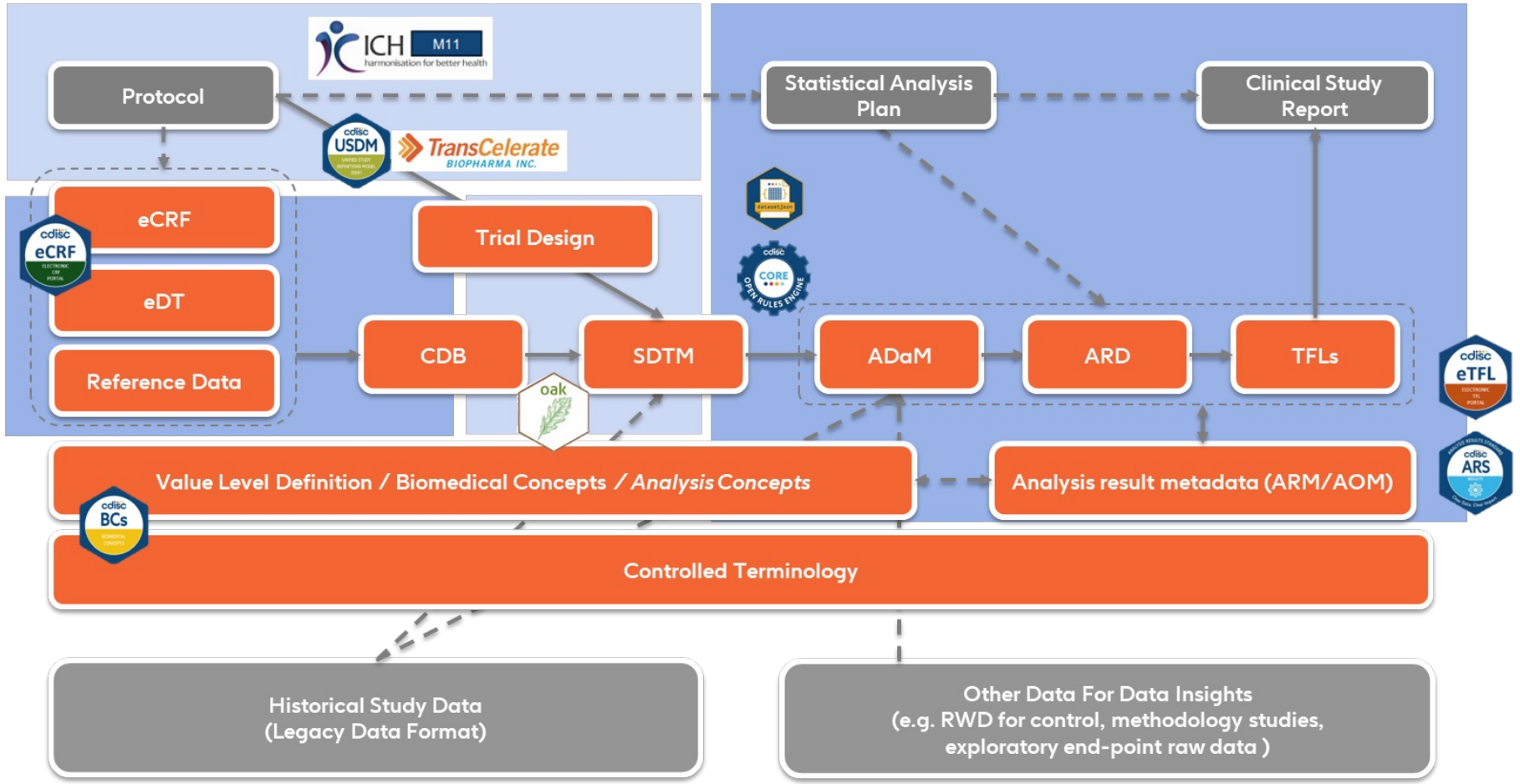
Biomedical Concept/Analysis Concept Use Cases
- Protocol, Collection, Analysis, Mapping

Warwick Benger, Igor Klaver

Linked/Connected Metadata for Clinical Trials

Enabling automation & reuse



Some key industry elements may help to unlock enhanced value, especially;

CDISC USDM
CDISC BCs
? HL7 FHIR

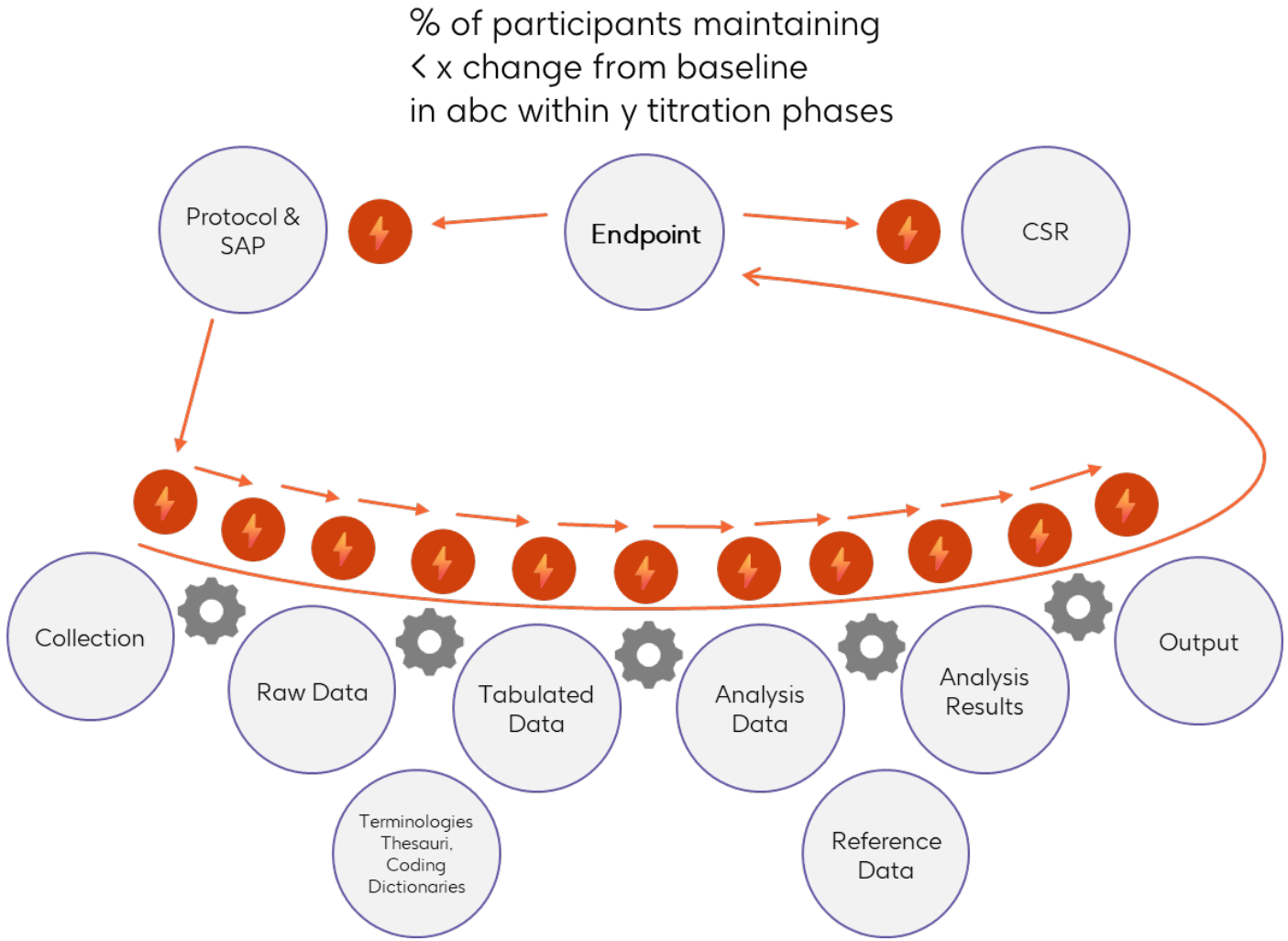
VULCAN HL7 FHIR

DHT

DIGITAL MEDICINE SOCIETY

Interoperability is critical to automation and data value

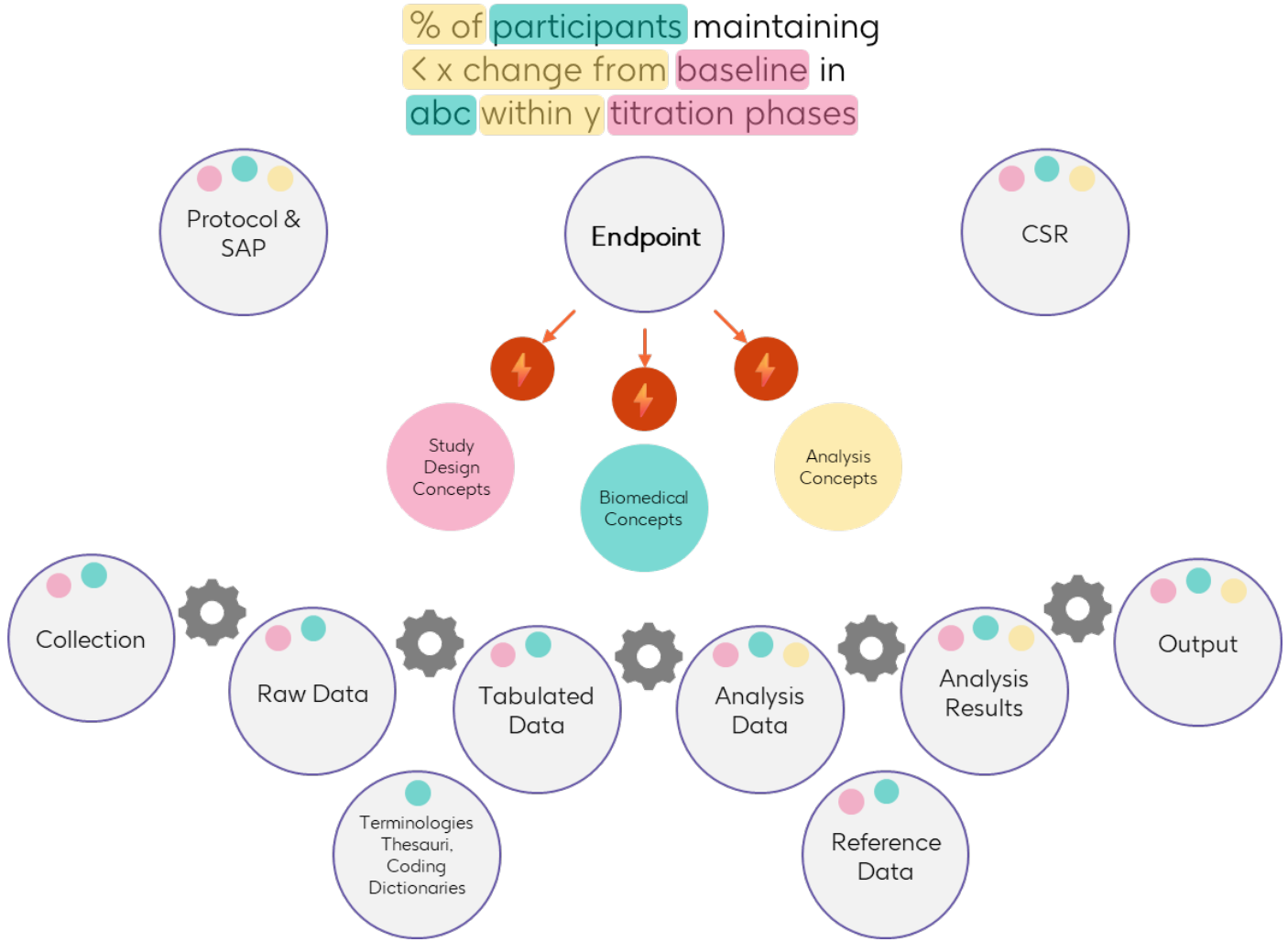
Traditional flow, requiring many user interpretations



- Study design articulated through documents
- Sequential interpretation of design along the delivery (even if definitions are linked)
- Issues
  - Poor and inefficient design decisions
  - Full digitisation of study design does not occur until after go-live
  - Issues with dependency and impact management (even if definitions are 'linked')
  - Poor forward alignment – too much or little data, or not what is needed
  - Definitions used (e.g. CDISC) are transient, so establishing common meaning for re-use is hard
  - Low acuity and ease of re-use of definitions

Interoperability is critical to automation and data value

Accelerated flow, driven by digital, semantically-linked study definition



- Study design articulated digitally
- Stable reference for complex linking of definitions, enabling automation & reuse
- Benefits
  - Data-driven protocol design/definition
  - Optimised data collection and use
  - Tied meaningfully and permanently to a centralised definition
    - → common, accurate meaning across diverse data
    - → high acuity and ease of re-use of definitions and data
  - Enhanced dynamic reporting
  - Improved quality and consistency

CONCEPTS at work

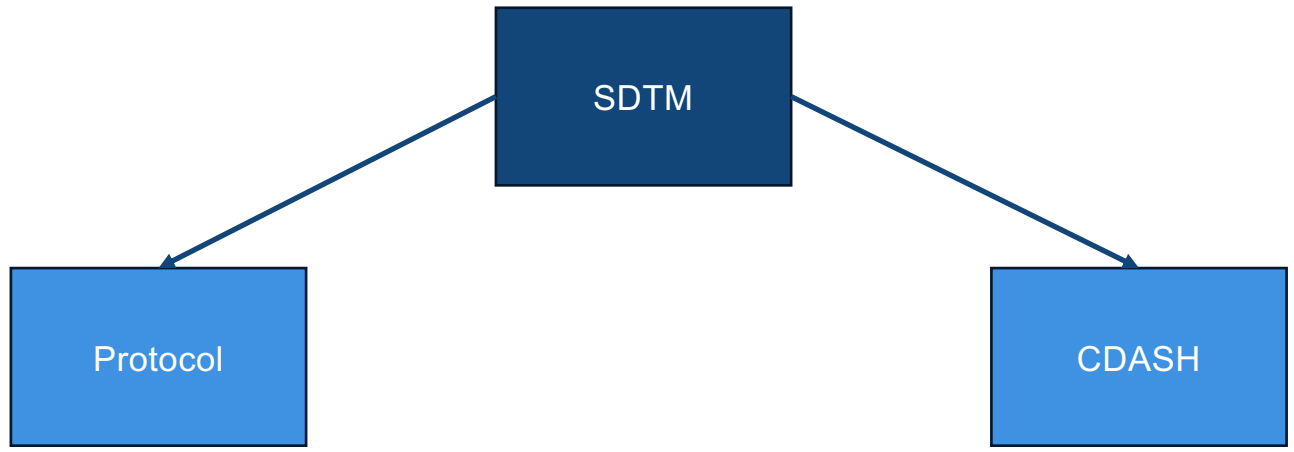
	Data Collection			Data Handling	Analysis
Submission deliverables	**SoA**	**eCRF & external data**	**SDTM**	**ADaM**	**Analysis Results**
Concepts	**Biomedical Concepts** *Meaning of collected data*			*[extend to analysis entities]*	
	Study design concepts				
				Derivation Concepts *Handling of data*	
					Analysis concepts *Analysis of data*
eDocuments	**eProtocol (USDM)**			**eSAP**	

Biomedical Concepts: Use Case

Oncology – Terminology Consistency from Protocol to SDTM

package_date	short_name	bc_id	ncit_code	parent_bc_id	bc_categories	definition	example_set
2023-07-06	Matted Tumor Mass Present	C94525	C94525	C82547	Evaluation Criteria in Solid Tumors Version 1.1;Tumor Identifier Test;Tumor Identification;RECIST 1.1;Merged	A finding indicating that two or more tumors have merged to create a single cancerous mass.	TARGET
2023-07-06	Response in Target Lesion	C94534	C94534	C50995	Response Evaluation Criteria in Solid Tumors;Response Evaluation Criteria in Solid Tumors Version 1.1;Disease Response Assessment Test;Disease Response;RECIST 1.1;Target	A qualitative or quantitative measurement of the response of a target lesion(s) to the therapy.	SD;PR;CR;PD
2023-07-06	Response in Non-Target Lesion	C94535	C94535	C50995	Response Evaluation Criteria in Solid Tumors;Response Evaluation Criteria in Solid Tumors Version 1.1;Disease Response Assessment Test;Disease Response;RECIST 1.1;Non-Target	A qualitative or quantitative measurement of the response of a non-target lesion(s) to the therapy.	CR; PR; SD; PD; NA; NE; NED
2023-07-06	Overall Response	C96613	C96613	C50995	Response Evaluation Criteria in Solid Tumors;Response Evaluation Criteria in Solid Tumors Version 1.1;Disease Response Assessment Test;Disease Response;RECIST 1.1;Overall	An assessment of the overall response of the disease to the therapy.	CR; PR; SD; PD; NA; NE; NED
2023-07-06	Tumor Fragmentation	C96642	C96642	C82547	Response Evaluation Criteria in Solid Tumors;Response Evaluation Criteria in Solid Tumors Version 1.1;Tumor Identifier Test;Tumor Identification;RECIST 1.1;Split	A finding indicating that a tumor mass has been divided into two or more tumors.	TARGET
2023-07-06	Tumor Status	C96643	C96643	C171082		ndition or state of the tumor at a particular time.	PRESENT;ABSENT;UNEQUIVOCAL;EQUIVOCAL
2023-07-06	Longest Diameter	C96684	C96684	C25285		gest possible length of a straight line passing through the of a circular or spheroid object that connects two points on cumference.	12;15;17;TOO SMALL TO MEASURE
2023-07-06	Longest Perpendicular	C96685	C96685			plane through a body or figure given line or plane.	12;15;17

Specializations of BC C96613



Protocol Specialization

bc_id	doma	vlm_source	vlm_group_id	short_name	sdtm_variable	codelis	ion_value	value_list	assigned_tern	assigned_value
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSTESTCD	C96782	ONCRTSCD		C96613	OVRLRESP
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSTEST	C96781	ONCRTS		C96613	Overall Response
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSCAT	C124298	ONCRSCAT		C124415	RECIST 1.1
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSORRES	C96785	ONCRSR	CR;NE;PD;PR;SD;NED;NON-CR/NON-PD		
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSDTC					

Table 36 presents the overall response at an individual time point for all possible combinations of tumor responses per RECIST 1.1.

Table 36 Evaluation of overall response

Use only the first 8 rows for studies requiring measurable/target lesions at baseline. Use only the last 3 rows for studies without a requirement of evidence of disease at baseline.

TLs	NTLs	New Lesions	Overall Response
CR	CR or NA	No	CR
CR	Non-CR/Non-PD or NE	No	PR
PR	Non-PD or NA or NE	No	PR
SD	Non-PD or NA or NE	No	SD
NE	Non-PD or NA or NE	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD
NA	CR	No	CR
NA	Non-CR/non-PD	No	Non-CR/non-PD[a]
NA	NE	No	NE
NA	Unequivocal PD	Yes or No	PD
NA	Any	Yes	PD
NA	NA	No	NED
NA	NA	NE	NE
NA	NA	Yes	PD

CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease; NA = not applicable; NE = not evaluable; NED = no evidence of disease; NTL = non-target lesion; TL = target lesion.

CDASH Specialization

bc_id	doma	vlm_source	vlm_group_id	short_name	sdtm_variable	codelis	ion_value	value_list	assigned_tern	assigned_value
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSTESTCD	C96782	ONCRTSCD		C96613	OVRLRESP
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSTEST	C96781	ONCRTS		C96613	Overall Response
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSCAT	C124298	ONCRSCAT		C124415	RECIST 1.1
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSORRES	C96785	ONCRSR	CR;NE;PD;PR;SD;NED;NON-CR/NON-PD		
C96613	RS	RS.RSTESTCD	OVRLRESP	Overall Response (RECIST 1.1)	RSDTC					

Forms

OID	Name	Description	Repeating	Aliases	Therapeutic Area	Clinical Stage
RS	Form RS - Disease Response	RS - Disease Response [en]	No			

Section

OID	Name	Repeating	Description	Order No.	Mandatory	Aliases	Condition	IsReferenceData	Repeating Information	SASDatasetName	Domain
CDASH_2-1_IG_33	RS - Disease Response	No	RS - Disease Response [en]	1	Yes						RS

Questions

OID	Name	bc_id	Text	DataType	Order No.	Mandatory	Terminology	Length	Significant Digits	Units	Description
IT.TRGRESP_RSORRES	TRGRESP_RSORRES	C94534	Target Response [en]	text	6	No	Target Response	15			
IT.NTRGRESP_RSORRES	NTRGRESP_RSORRES	C94535	Non-Target Response [en]	text	7	No	Non-Target Response	15			
IT.OVRLRESP_RSORRES	OVRLRESP_RSORRES	C96613	Overall Response [en]	text	8	No	Overall Response	15			

Terminologies

OID	Name	DataType	SASFormatName	Terminology Aliases	Coded Value	Text	Order No.
Overall Response	Overall Response	text			CR	Complete Response (CR)	
Overall Response	Overall Response	text			PR	Partial Response (PR)	
Overall Response	Overall Response	text			SD	Stable Disease (SD)	
Overall Response	Overall Response	text			NON-CR/NON-PD	Non Complete Response/ Disease (NON-CR/NON	
Overall Response	Overall Response	text			PD	Progressive Disease (P	
Overall Response	Overall Response	text			NE	Not Evaluable (NE)	
Overall Response	Overall Response	text			NED	No Evidence of Disease	

CDASH Specialization

Form RS - Disease Response

1 RS - Disease Response

1.8	Overall Response	○ [CR] Complete Response (CR) ○ [PR] Partial Response (PR) ○ [SD] Stable Disease (SD) ○ [NON-CR/NON-PD] Non Complete Response/Non Progressive Disease (NON-CR/NON-PD) ○ [PD] Progressive Disease (PD) ○ [NE] Not Evaluable (NE) ○ [NED] No Evidence of Disease (NED)

OVRLRESP_RSORRES

bc_id = C96613

GSK's Value Level Definition (VLD)

- GSK's VLDs are similar with CDISC Biomedical Concept (BC)/SDTM Specialization.
- We believe VLD/BCs will fill gaps in the current standards by adding semantics, variable relationships, and the detailed metadata needed to generate CRFs or Define-XML.

VLDsource	vld_group	WhereV	CODELIS	COMPARAT	Value	VLM_TARG	CCR_Category	Data_Typ	Origin	Length	Forma	Significant_Dig	Mandator	DASsource_name	target_name
LB LBTESTCD	LABCHEMGLUCPL	LBCAT	LBCAT	EQ	CHEMISTRY		LBTESTCD							External Datasets.LB_CENTR	Findings.LB.LBCAT
LB LBTESTCD	LABCHEMGLUCPL	LBFAST	NY		<define at study level>		LBTESTCD								
LB LBTESTCD	LABCHEMGLUCPL	LBORRES			<define at study level>	True	LBTESTCD	text	eDT	200			No	External Datasets.LB_CENTR	Findings.LB.LBORRES
LB LBTESTCD	LABCHEMGLUCPL	LBSPEC	SPECTYPE	EQ	PLASMA		LBTESTCD								
LB LBTESTCD	LABCHEMGLUCPL	LBTEST	LBTEST_CH		Glucose		LBTESTCD								
LB LBTESTCD	LABCHEMGLUCPL	LBTESTCD	LBTESTCD_	EQ	GLUC		LBTESTCD							External Datasets.LB_CENTR	Findings.LB.LBTESTCD
LB LBTESTCD	URINDIPGLUC	LBCAT	LBCAT	EQ	URINALYSIS		LBTESTCD							External Datasets.LB_CENTR	Findings.LB.LBCAT
LB LBTESTCD	URINDIPGLUC	LBMETHOD	METHOD	EQ	DIPSTICK		LBTESTCD							External Datasets.LB_CENTR	Findings.LB.LBMETHOD
LB LBTESTCD	URINDIPGLUC	LBORRES			<define at study level>	True	LBTESTCD	text	eDT	200			No	External Datasets.LB_CENTR	Findings.LB.LBORRES
LB LBTESTCD	URINDIPGLUC	LBSPEC	SPECTYPE	EQ	URINE		LBTESTCD								
LB LBTESTCD	URINDIPGLUC	LBTEST	LBTEST		Glucose		LBTESTCD								
LB LBTESTCD	URINDIPGLUC	LBTESTCD	LBTESTCD	EQ	GLUC		LBTESTCD								

Glucose measurement for Chemistry Panel

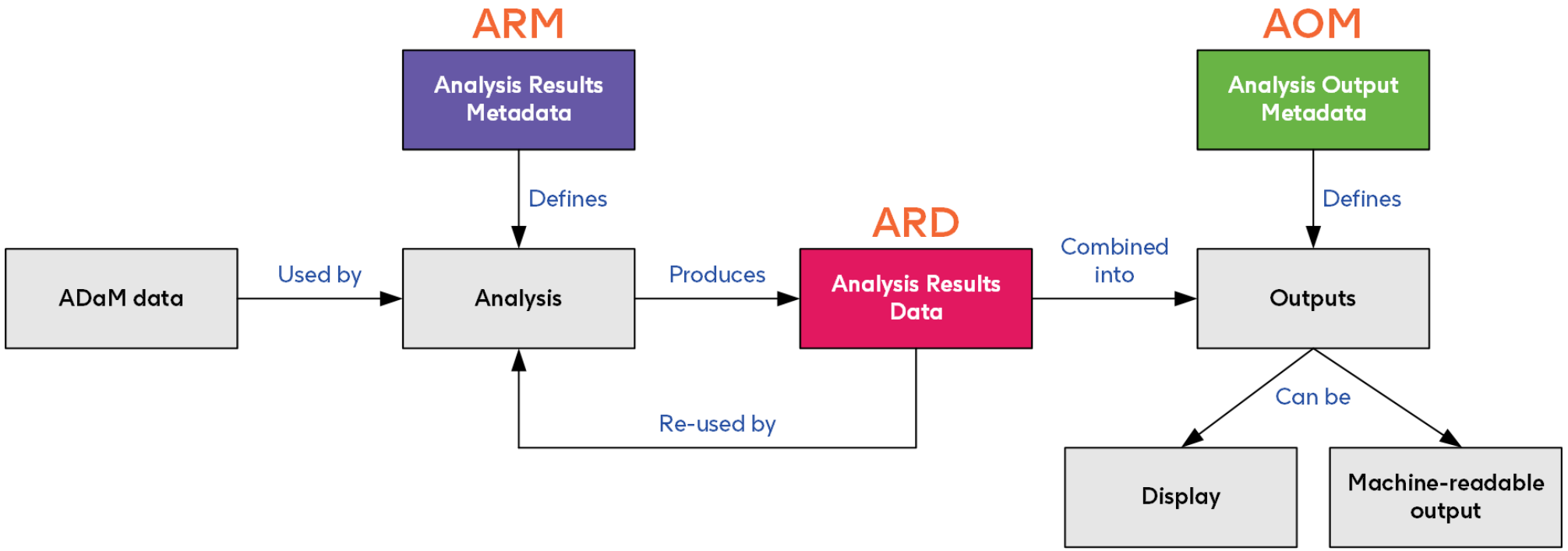
Glucose measurement for Urine Dipstick Panel

Summary

- SDTM specializations can be used to develop upstream standards using a metadata driven approach:
  - Protocol
  - CDASH
  - Review models
  - External data
- Incorporating BCs into e2e standards:
  - Ensures consistency
  - Accelerates timelines
  - Reduces conformance errors
  - Allows powerful impact assessments

Analysis Metadata & Concepts

ARMADA - our vision for analysis result and reporting automation



Benefits and principle

Increased

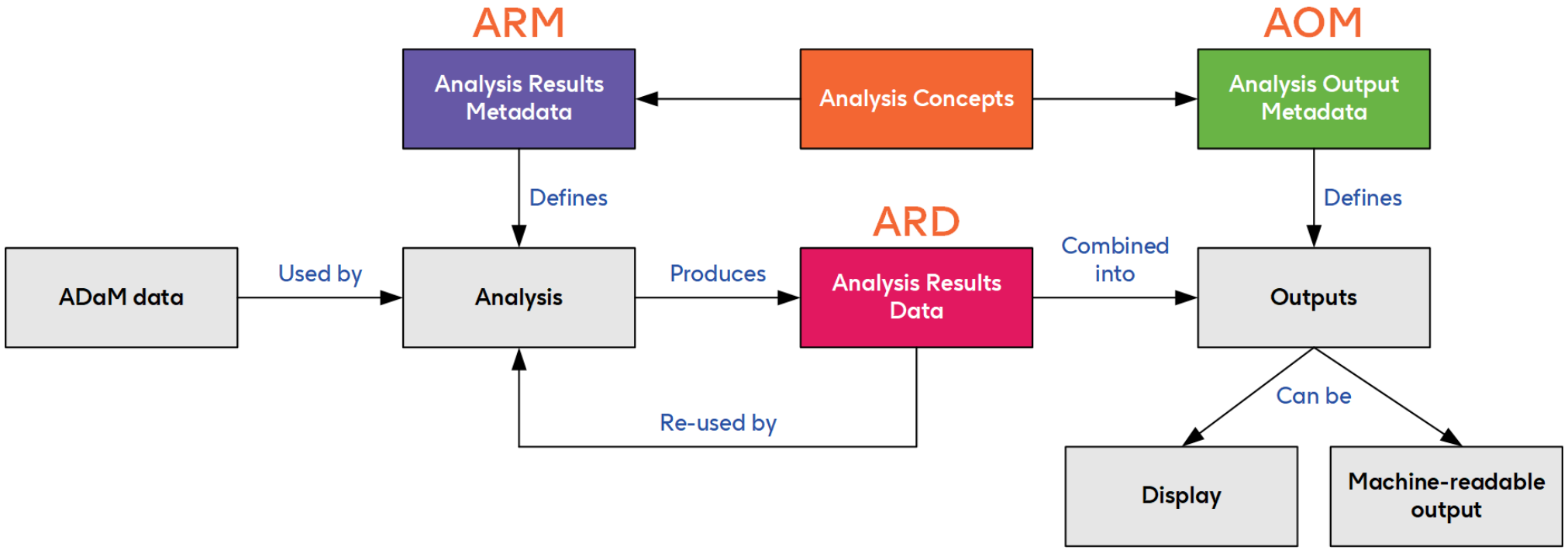
- Traceability
- Transparency
- Automation
- Consistency
- Flexibility

WORM

Write Once, Read Many

- Any analysis defined once
- Any analysis executed once
- Any analysis validated once
- Re-use analyses across outputs
- Re-use analyses across analyses

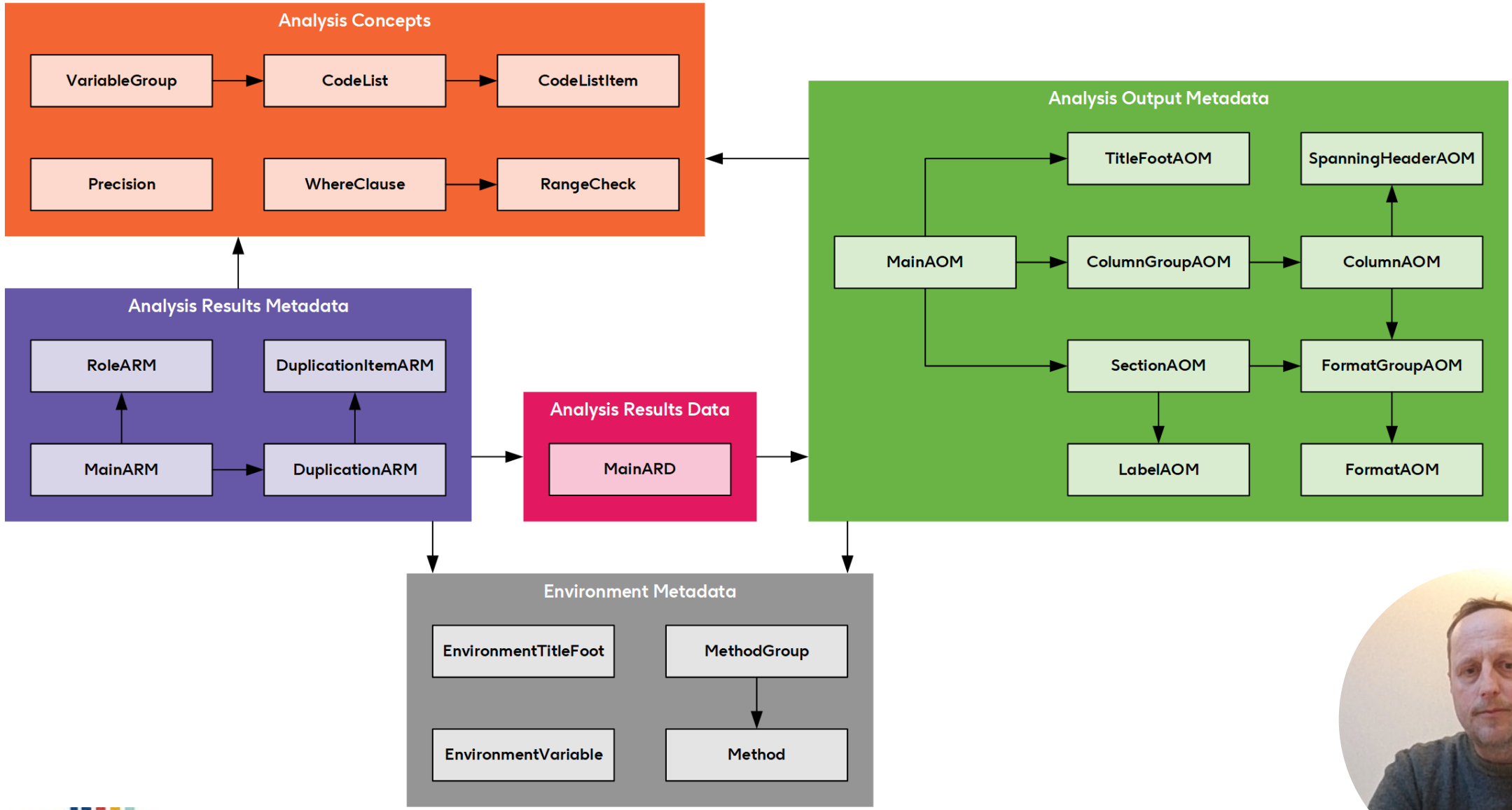
Where analysis concepts come in



Enables us to ensure analysis entities are reusable

Limits direct references to ADaM data sets and variables to the underlying reusable biomedical and analysis concepts

ARMADA data model



The role of analysis concepts

VariableGroup

- Explicitly links variables and codelists (e.g. PARAMN, PARAMCD and PARAM and their codelist)
- Defines where a full matrix is produced during the analysis (e.g. total treatment column)

CodeList and its child CodeListItem

- Explicitly links the triplicate of numeric, code and decode
- Includes values not present in ADaM (e.g. aggregate values like total treatment)

WhereClause and its child RangeCheck

- Defines re-usable where clauses (series of meaningful additive range checks)

Precision

- Defines precision of numeric input

Analysis results: concepts in practice (at GSK)

Analysis (defined in **ARM**): mean change from baseline of the lab parameter ALT by treatment, visit and timepoint in the safety analysis set

- *mean*: the analysis method (an **Analysis Concept**)
- *change from baseline*: the analysis variable CHG (a **Derivation Concept**)
- *of the (result of) the lab parameter for ALT*: (AVAL for) subset of ADLB defined by the where clause PARAMCD EQ "ALT" (a **Biomedical Concept**) and its input precision (an **Analysis Concept**)
- *by*: by variables (**Study Design Concepts** and/or **Biomedical Concepts**) combined into a variable group (an **Analysis Concept**)
  - *treatment, visit and timepoint*
- *in the safety population*: analysis set defined by the where clause SAFFL EQ "Y" and its label "Safety" (an **Analysis Concept**)

Analysis concepts discussion

Analysis concepts at GSK are part of our **operational model** and not a **conceptual model**. Looking at it conceptually, what should we as an industry define as analysis concepts?

- Is the analysis method part of an analysis concept?
- Is the analysis set part of the analysis concept? Or is it perhaps a separate analysis concept? Or is it a subset or child of an overarching analysis concept?
- Are the by variables part of an analysis concept? Or are they separate concepts in a list of concepts to pick and choose from?
- How should we manage the distinction between “analysis concept” and “derivation concept”

Thank you

warwick.j.benger@gsk.com

igor.x.klaver@gsk.com