



In Partnership with



INDIA CDISC DAY

PHUSE APAC Connect Hyderabad India 20 February 2026

Analysis Concepts Initial Perspectives from the CDISC Working Group

Stuart Malcolm, Head of Standards, Efficiency and Automation, Veramed.

Meet the speaker



Stuart Malcolm

Head of Standards,
Efficiency and Automation

Veramed

- Stuart has over 30 years' experience in regulated industries including healthcare, telecommunications, finance and media.
- He is co-chair of the PHUSE Emerging Trends and Innovation working group, was co-lead of the Dataset-JSON project and a member of the CDISC 360i Analysis Concept project

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the authors and do not necessarily reflect the official policy or position of CDISC or Veramed.*
- *The authors have no real or apparent conflicts of interest to report.*



Agenda

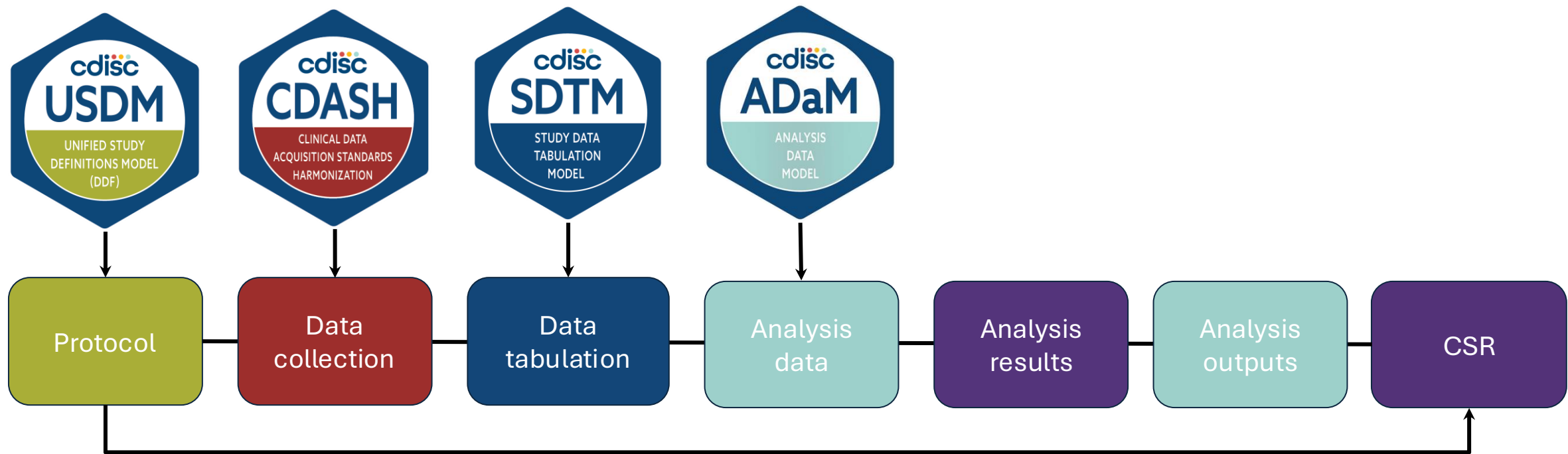
1. Problem
2. Solution
3. Next Steps
4. Questions



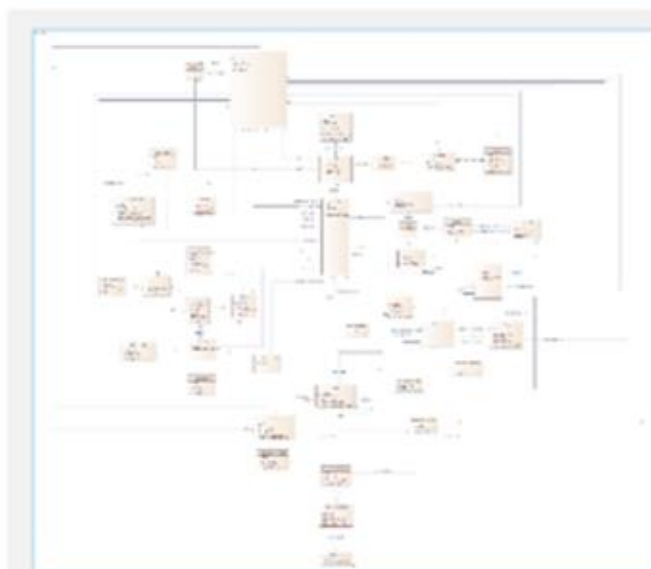
Problem Statement

Data Standards Are Evolving...

CDISC Standards – November 2025



USDM Content



Controlled
Terms

Study, Identifiers,
Amendments

Estimands

Unstructured Content

Populations

Study Designs,
Arms, Epochs

Inclusion &
Exclusion

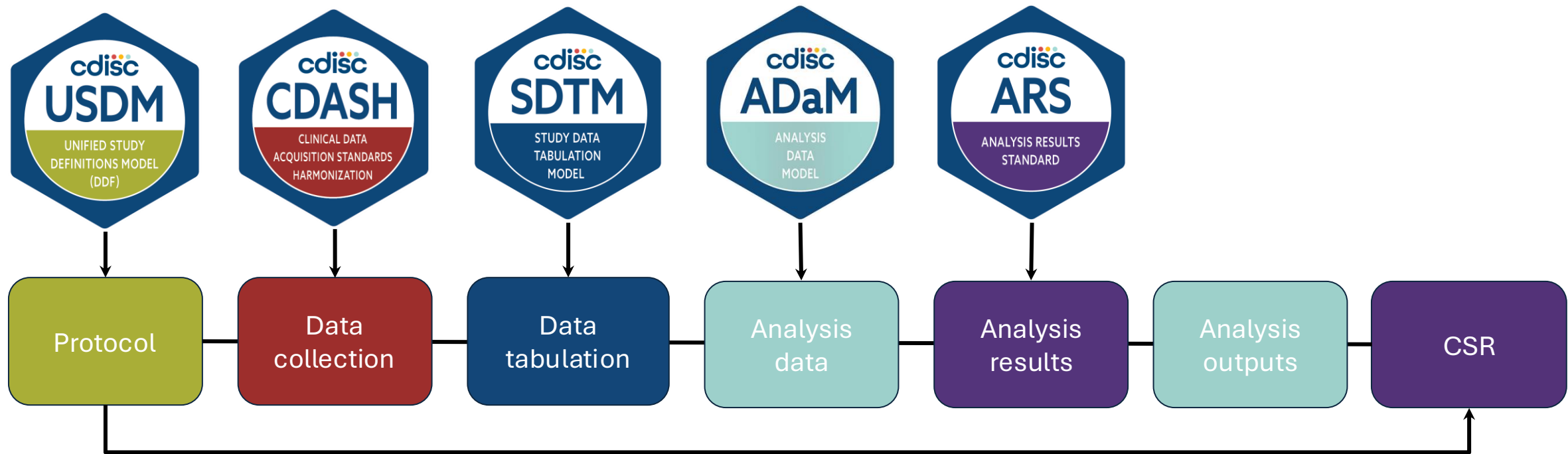
Interventions &
Indications

Detailed Study Logic,
Encounters

Objectives &
Endpoints

Procedures, Biomedical Concepts

CDISC Standards – November 2025



CDISC Analysis Result Standards – Released 2024



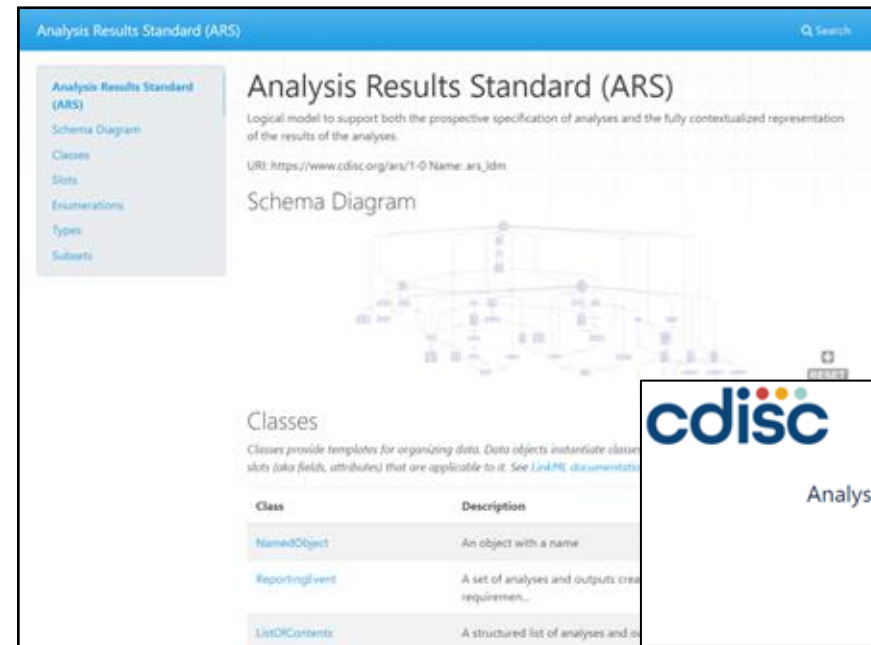
Large trials generate many analysis results in the form of tables, figures, and written reports, yet these results are rarely output in a form that is machine-readable. Previously, there has been no standard way of describing and organizing these results, making it difficult to automate their generation, make them reproducible, trace their origin, or enable them to be reused in other outputs.

To address these inefficiencies, CDISC has developed the [Analysis Results Standard \(ARS\)](#), which aim to facilitate automation, reproducibility, reusability, and traceability of analysis results data.

Features of ARS v1.0

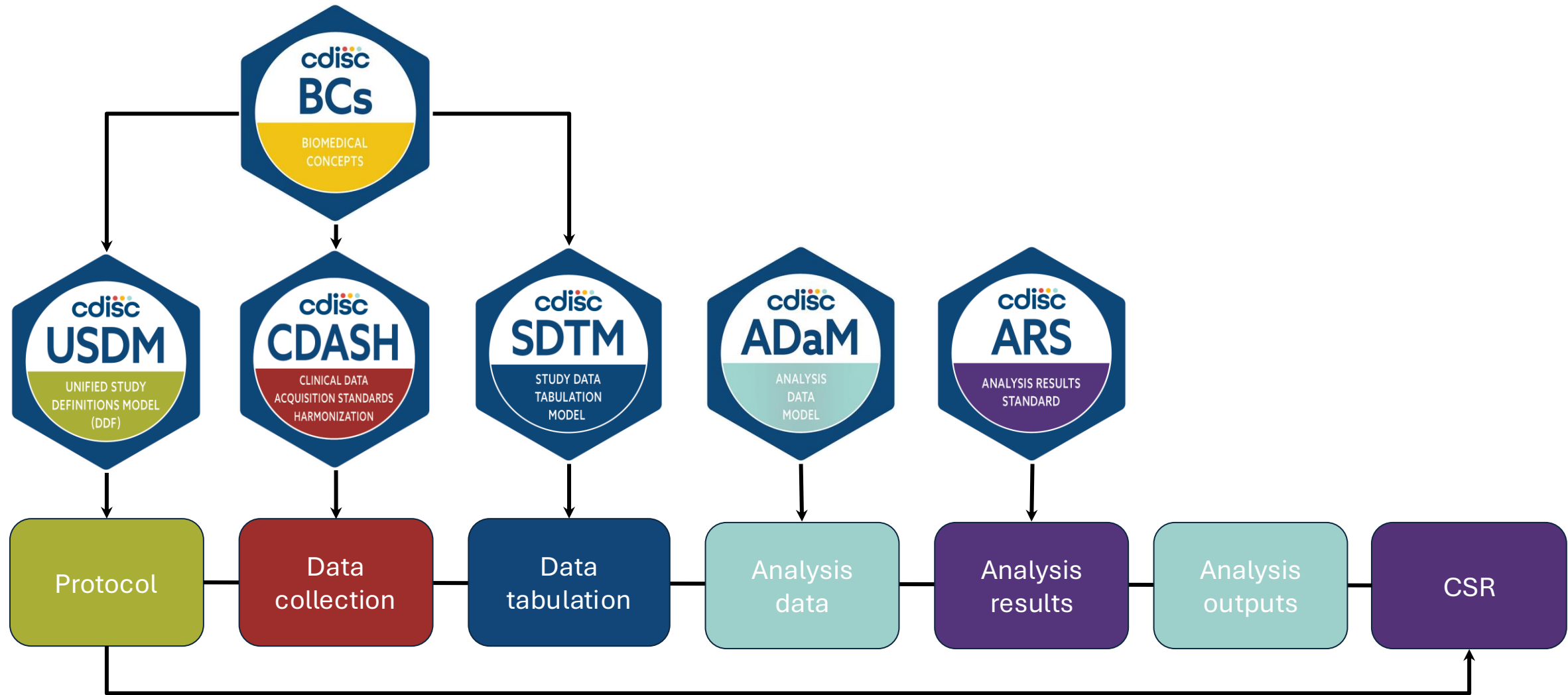
- A Logical Data Model that describes analysis results and associated metadata.
- A User Guide to illustrate and exercise the model with common safety displays.

<https://cdisc-org.github.io/analysis-results-standard/>



<https://wiki.cdisc.org/display/ARSP/Analysis+Results+Standard+User+Guide+v1.0>

CDISC Standards – November 2025



CDISC Biomedical Concepts

Home

Search by Word

BC Categories

BC Short Name

NCIt Code

Biomedical Concepts (BCs)

BC Short Name

Synonyms

More Dropdowns

Home

Search by word

SDTM Short Name

BC Short Name

Domain

SDTM Dataset Specializations

Domain

SDTM Short Name

SDTM Variable

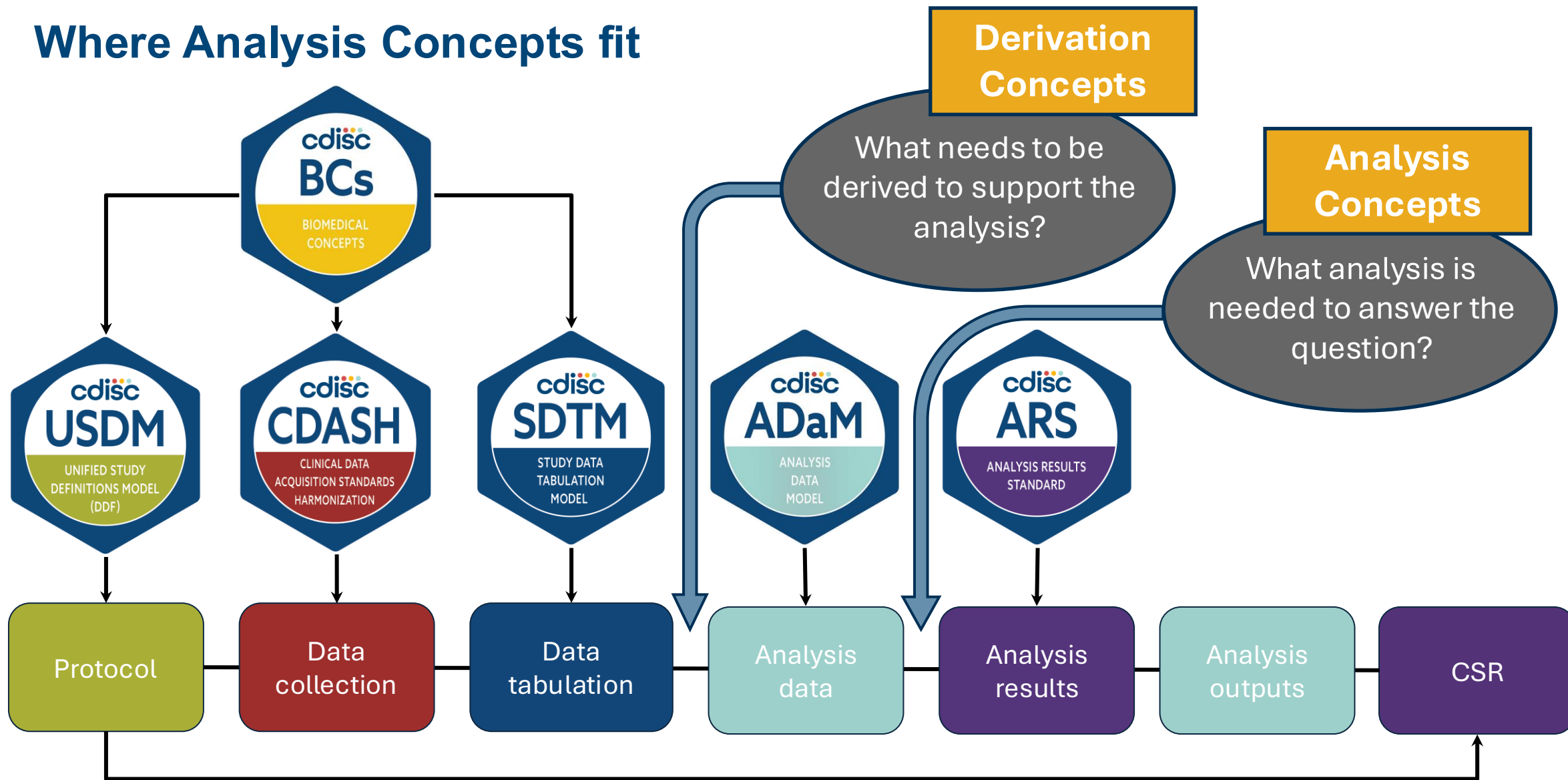
More Dropdowns

Note: When applying one filter, it applies to the other dropdowns. Use ctrl + click to do multiple selections in the same filter.

Package Date	BC Short Name
2025-09-23	6MWT - Distance at 6 Minutes
2025-09-23	6MWT - Distance at 6 Minutes
2025-09-23	6MWT - Distance at 6 Minutes
2025-09-23	6MWT - Distance at 6 Minutes
2025-09-23	6MWT - Distance at 6 Minutes
2025-09-23	Allergen-induced Antibody Measurement
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgA Antibody Measurem
2025-09-23	Allergen-induced IgE Antibody Measurem
2025-09-23	Allergen-induced IgE Antibody Measurem

Package Date	BC ID	SDTMIG Start Version	SDTMIG End Version	Domain	VLM Source	VLM Group ID	SDTM Short Name	SDTM Variat
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTCAT
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTDTC
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTORRES
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTREASND
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTSCAT
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTSTAT
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTSTRESC
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTSTRESN
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTTEST
2025-09-23	C100177	3-2		FT	FT.FTTESTCD	ADCRL	CDISC ADAS-Cog - Word Recall Average Score	FTTESTCD

Where Analysis Concepts fit



Standards Use-Cases

Use Case 1: Accelerating Study Start-Up

As a study statistician,
I want to select from a library of Analysis Concepts
so that I can quickly draft the digital Statistical Analysis Plan (SAP) and identify needed datasets and variables.

Use Case 2: Enabling Metadata-Driven Automation

As a clinical data standards expert, **I want to** build a structured repository of Analysis Concepts
so that downstream automation tools (e.g., TFL code generators, validation scripts) can use them as inputs.

Use Case 3: Automating Standard Analyses Across Studies

As a statistical programmer, **I want to** use pre-defined Analysis Concepts
so that I can automate the generation of standard TFLs across studies without recreating logic each time.



The AC/DC Framework

Work in progress

Today's specification of analysis

- Written in plain text
 - Not machine readable
- Some details are left out
 - Often (at best) retrievable from the analysis executable program/define.xml
 - Reason why regulators require programs to be submitted
- Specifications are not always software agnostic
 - SAS versus R syntax
 - Different set of options

Primary analysis from CDISC pilot study

The primary analysis of the ADAS-Cog (11) at Week 24 will use the efficacy population with LOCF imputation for any missing values at Week 24.

An ANCOVA model will be used with the baseline score, site, and treatment included as independent variables.

Treatment will be included as a continuous variable, and results for a test of dose response will be produced.

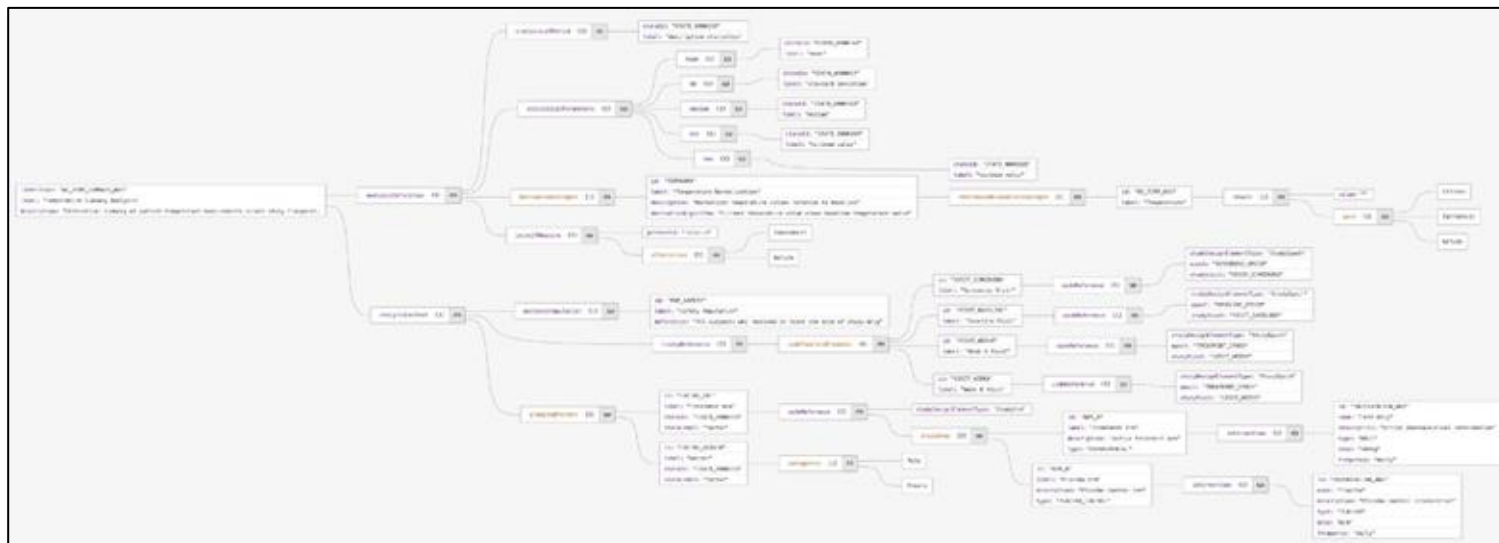
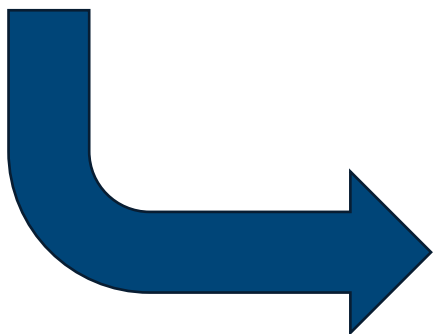
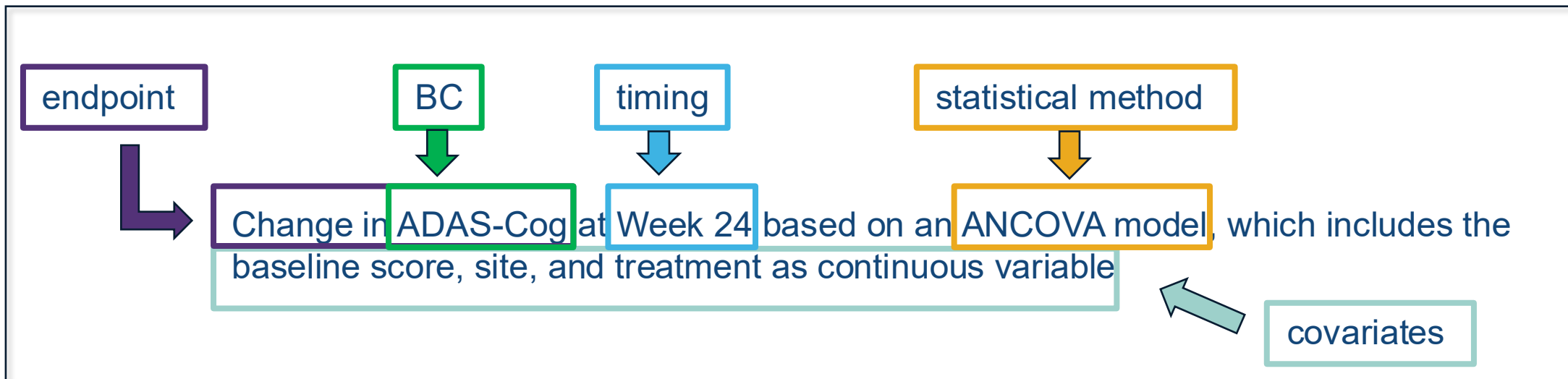
Interaction terms will not be investigated.

If the test for dose response is statistically significant, pairwise comparisons among the 3 groups will be performed and evaluated at a significance level of 0.05.

Proposed 5-Layer Architecture

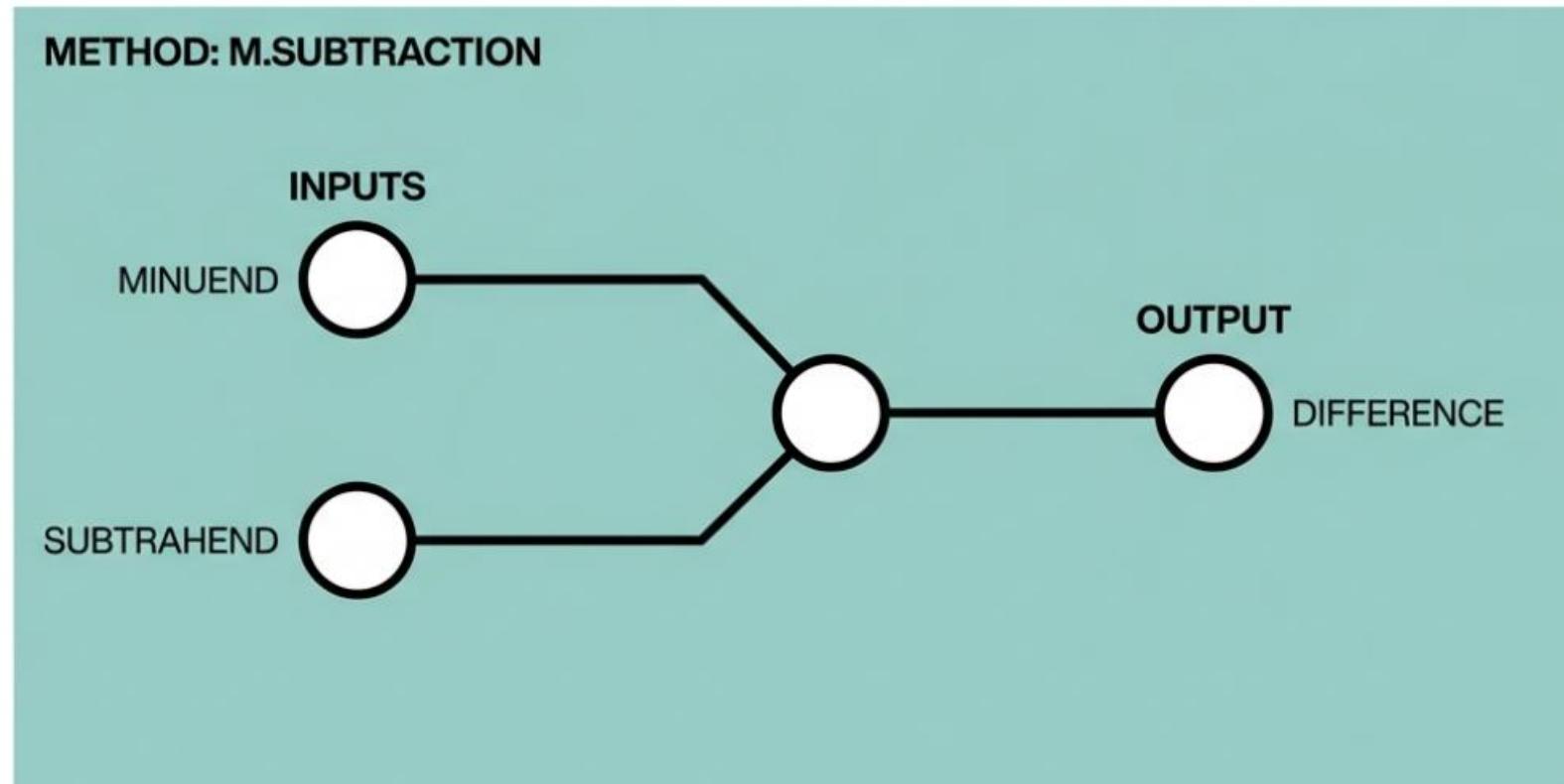


Layer 1: Human-readable analysis descriptions

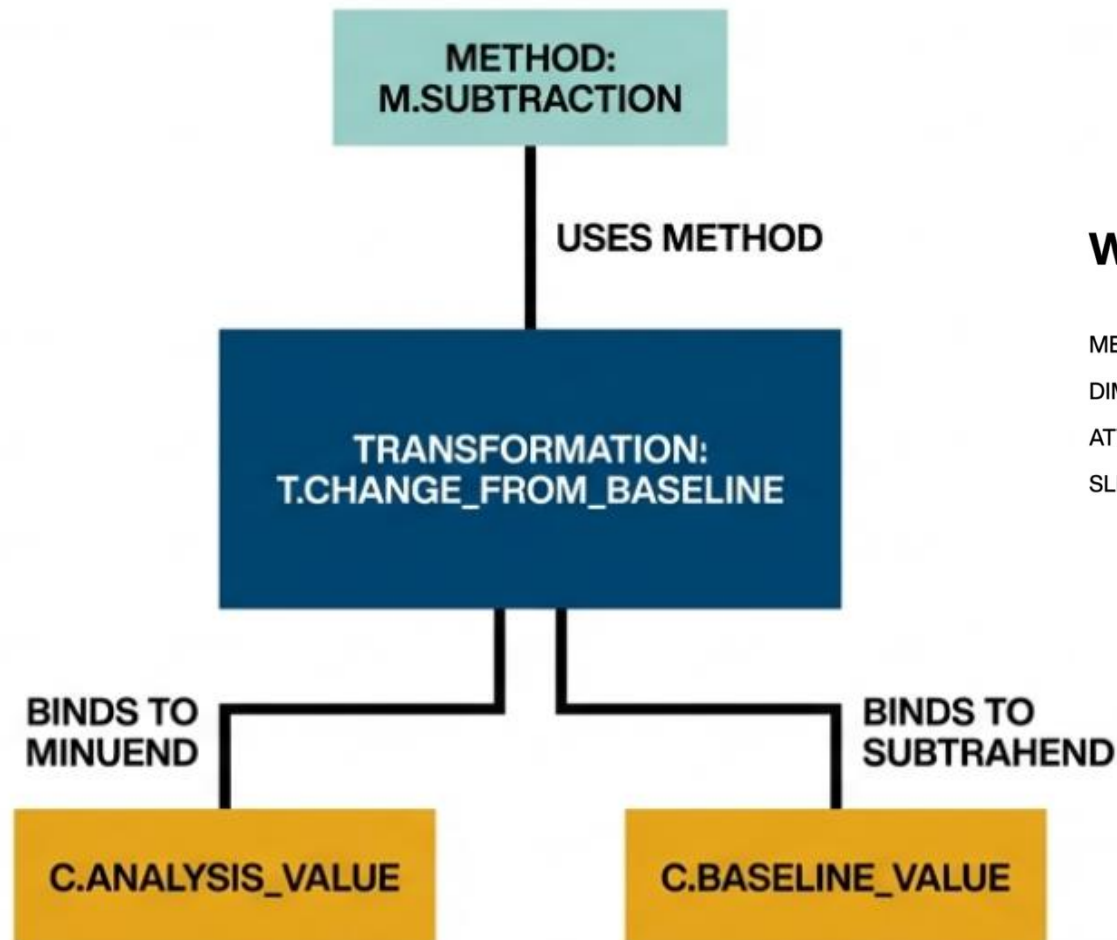


Layer 2: Abstract statistical logic

Pure operations. Standards-Agnostic. Knows nothing of Concepts

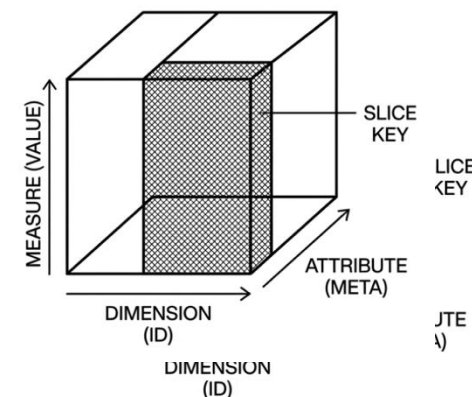


Layer 3: Describing the analysis data structure



W3C DATA CUBE ALIGNMENT

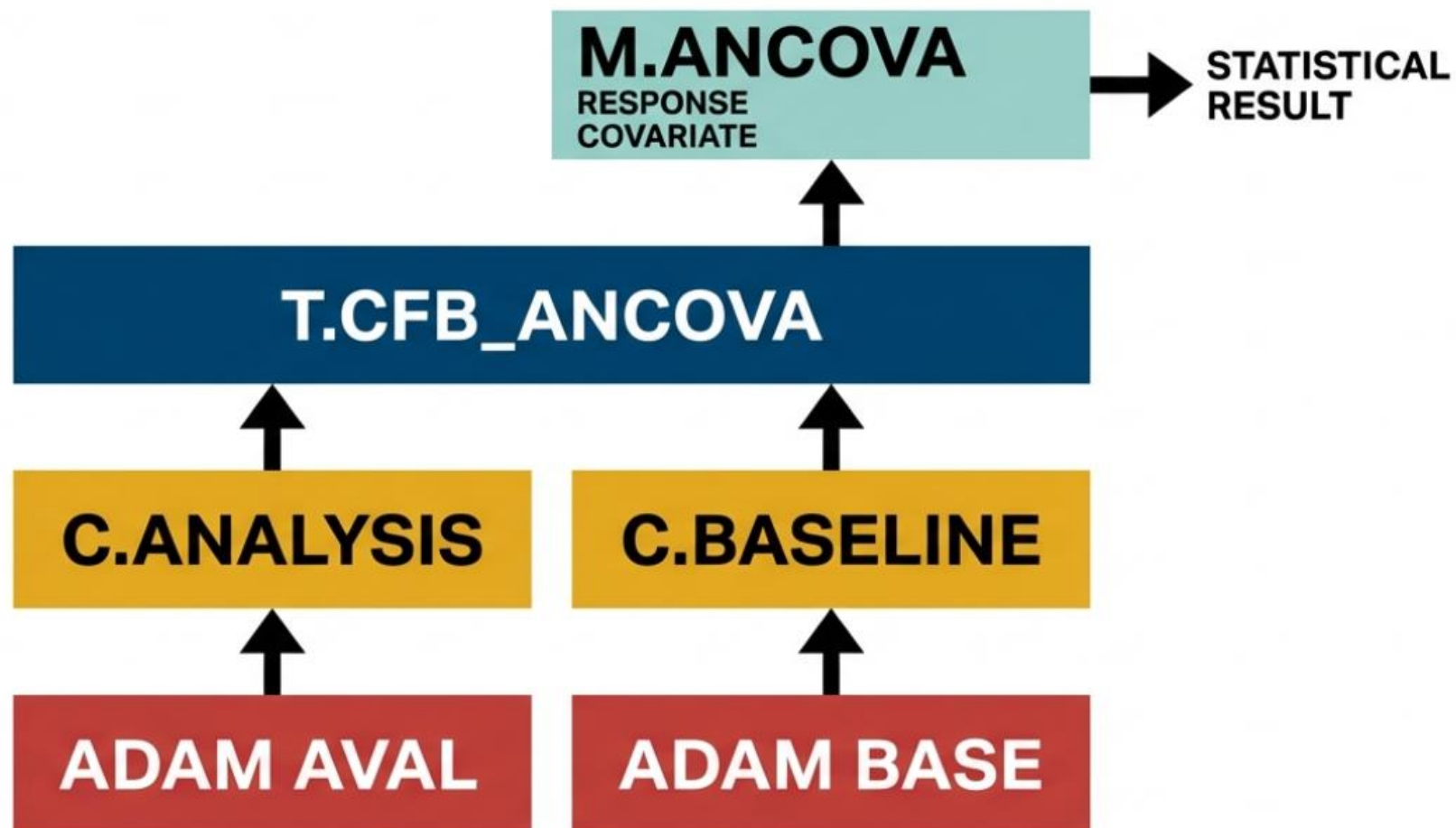
MEASURE: OBSERVED VALUE
DIMENSION: IDENTIFIERS
ATTRIBUTE: METADATA
SLICE MECHANISM: DEFINES SUBSETS



Layer 4: Standards-Independent semantic definitions

	Analysis Concept (AC)	Derivation Concept (DC)
Purpose	Examines existing data or information to draw conclusions, identify patterns, or test hypotheses	Generates new derived values from raw or derived data
Direction	Typically works with completed measurements or observations to extract meaning	Transforms or processes data to create new representations
Process	Involves applying statistical methods, critical thinking, and interpretative frameworks to understand data	Uses mathematical operations, formulas, or algorithms to calculate new quantities
Outcome	Produces insights, conclusions, or evaluations based on the data - aggregated data (not subject-level)	Creates derived data (subject level) that serve as inputs for subsequent analysis or derivations
Example 1	The p-value (from Type III Sums of Squares for treatment dose), based on linear model analysis of CHG for dose response; using randomized dose and site group in model.	CHG: Change from Baseline to Week 24 in ADAS Cog (11). Use LOCF is missing value at week 24.
Example 2	Mean value of CHG by visit	CHG: Change from BASELINE in ADAS Cog (11) by visit BASELINE = 'Y' if ADAS Cog (11) at visit 2

Example (ANCOVA)



Use cases

- as a statistician, I can define the nuances of my analyses so that a human or AI programmer can accurately implement the analyses
- as a statistician, I can search for and reuse Analysis Methods
- as a statistician, I can connect methods that have different implementations, different implementation contexts
- as an academic, I can publish structured analysis design to journals
- as a programmer, I can understand the impact analysis of changes to SAP
- as a regulatory reviewer, I can understand unambiguously what analysis was done
- as a statistician, I can trace from the analysis conducted to the data point which contributed to the analysis

360i AC/DC Automation Benefits

- Interchange
 - Standardized structure allows **interchange** of analysis specifications between systems and organizations
- Analysis automation and reuse
 - **Machine-readable** analysis specifications can directly link to statistical programming code reducing transcription errors, allowing automated validation and reuse
- Reduced ambiguity
 - Standardized structure **enforces precision** in specifying analysis setting and assumptions
- Increased traceability
 - Provides **clear linkage** from results all the way back to objectives and endpoints in USDM and vice versa
- Streamlined collaboration
 - Provides **common language** between statisticians, clinicians, data managers, programmers and other stakeholders



Next Steps

Technical and Model Development



Community & Industry Engagement



Strategic Opportunities & Challenges





Questions



Thank You!

