



e2e Metadata Driven Model for Efficient Reporting

Unveiling a Revolutionary Method for Defining Metadata and Reporting

Priyanka Sarode

Agenda

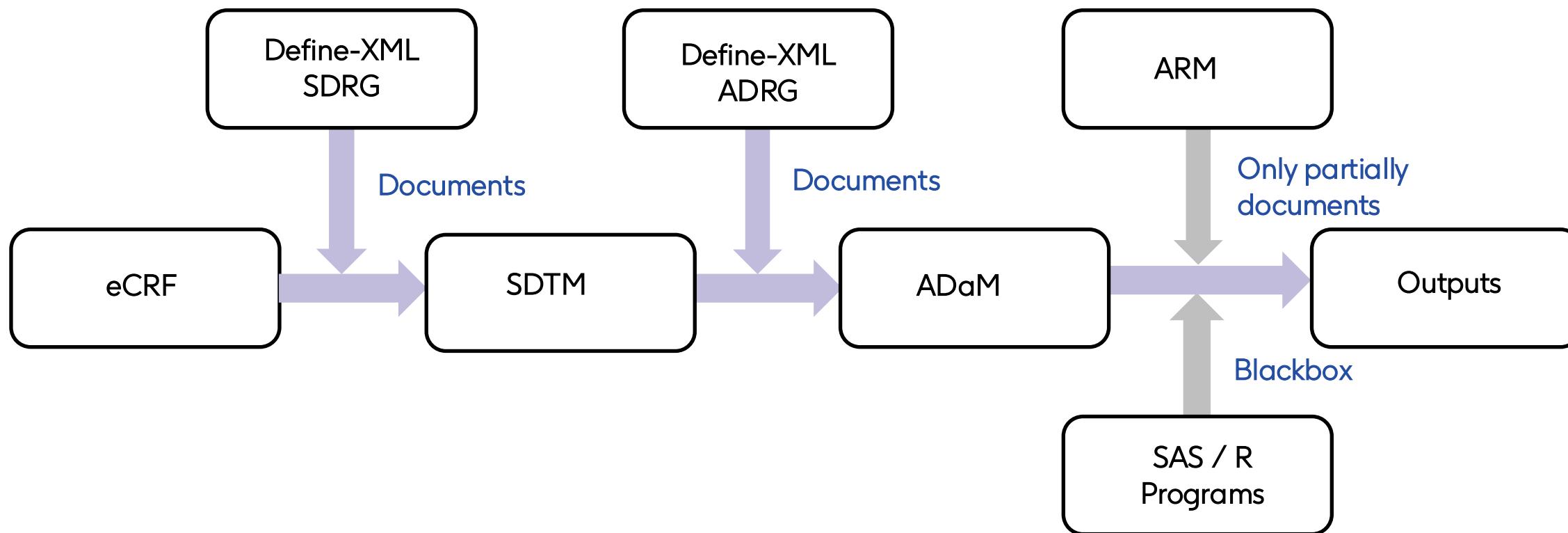
1. Introduction
2. Problem Statement & Solution: Driving Efficiency Through Metadata
3. ARMADA Model – e2e Metadata Driven Model for Efficient Reporting
4. Metadata-Driven Automation Workflow
5. Summary & Key Highlights
6. Conclusion
7. Q&A

Introduction

- CDISC standards (SDTM, ADaM) and documentation (Define-XML, aCRF) establish a foundation for data traceability and transparency.
- Traceability is limited for analyses not explicitly defined in ARM (Analysis Results Metadata) within Define-XML.
- ARM (Analysis Result Metadata) is often created retrospectively and isn't designed to drive automation, leading to traceability gaps.
- While SAS/R programs offer some traceability, heavy use of macros or packages can render them "black boxes," obscuring analysis logic.
- ARMADA (Analysis Metadata and Data Automation) Model is developed to enhance traceability, transparency, automation, and consistency from ADaM data through to final outputs.

Introduction

General Reporting Workflow



Problem Statement & Solution: Driving Efficiency Through Metadata

Let's Analyse a basic Demography table, where 4 distinct analysis can be identified

An analysis is defined as a single analysis on one variable using one method.

Protocol: ABC123456

Page 1 of 1

Analysis Set: Safety

(Data as of: ddMMMyyyy)

Table X: Summary of Demographic and Baseline Characteristics

	Treatment A (N=xxx)	Treatment B (N=xxx)	Total (N=xxx)
Sex			
n	xxx	xxx	xxx
F	xx (xx%)	xx (xx%)	xx (xx%)
M	xx (xx%)	xx (xx%)	xx (xx%)
Age (YEARS) at First Dose [1]			
n	xxx	xxx	xxx
Mean	xx.x	xx.x	xx.x
SD	x.xx	x.xx	x.xx
Median	xx.x	xx.x	xx.x
Min.	xx	xx	xx
Max.	xx	xx	xx

[1] Age is imputed when full date of birth is not provided.

/Directory/program.R ddMMMyyyy HH:mm

Frequency counts by treatment

Frequency counts by treatment for subjects with non-missing sex

Frequency counts by treatment and sex

Summary statistics on age by treatment

Problem Statement & Solution: Driving Efficiency Through Metadata

Problem Statement

Repeated generation of the same analysis (e.g., big N counts) for use as column headers or denominators leads to inefficiency and potential inconsistencies.

Key Question

How can we ensure these values are generated once, validated, and reused across analyses?

Solution

The ARMADA (Analysis Results Metadata & Data Automation) model, based on the WORM (Write Once, Read Many) principle, addresses this challenge by:

- Generating and validating numbers once
- Reusing them across multiple analyses

ARMADA Model – e2e Metadata Driven Model for Efficient Reporting

Key Components for Metadata-Driven Automation

The e2e Metadata Driven Model for Efficient Reporting introduces a groundbreaking approach to streamline the analysis and reporting of clinical trial data.

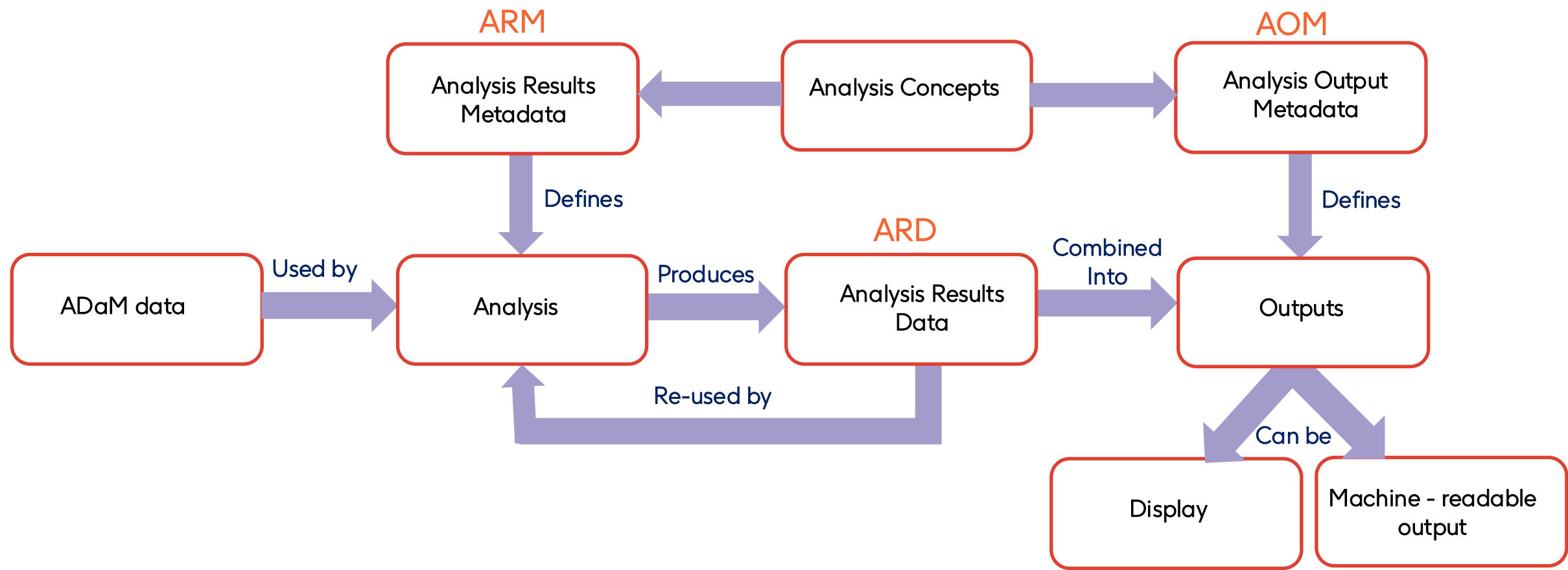
Goal: Enhance reusability, consistency, transparency, and automation.

Key Components of ARMADA:

- **Analysis Results Metadata (ARM):** Defines single analysis protocols, ensuring traceability and auditability across trials. Standardizes metadata frameworks for clarity.
- **Analysis Results Data (ARD):** Stores unformatted analysis data for reuse across outputs. Enables integration with for automated compliance reporting.
- **Analysis Output Metadata (AOM):** Specifies output construction rules, allowing flexible formatting without reanalysis. Supports diverse, tailored reports efficiently.

ARMADA Model – e2e Metadata Driven Model for Efficient Reporting

Metadata Driven Workflow



Defining Analysis Metadata

Protocol: 12345
Analysis Set: Safety

Page 1 of 1
Data as of : 01NOV2025

Table 1: Summary of Demographic characteristics

	TRT A	TRT B	Total
Sex			
n	xx	xx	xx
F	xx (xx%)	xx (xx%)	xx (xx%)
M	xx (xx%)	xx (xx%)	xx (xx%)
Age (YEARS) [1]			
n	xxx	xxx	xxx
Mean	xx.x	xx.x	xx.x
SD	x.xx	x.xx	x.xx
Median	xx.x	xx.x	xx.x
Min	xx	xx	xx
Max	xx	xx	xx

[1] Age is imputed when full date of birth is not provided.

A
O
M

A
O
M

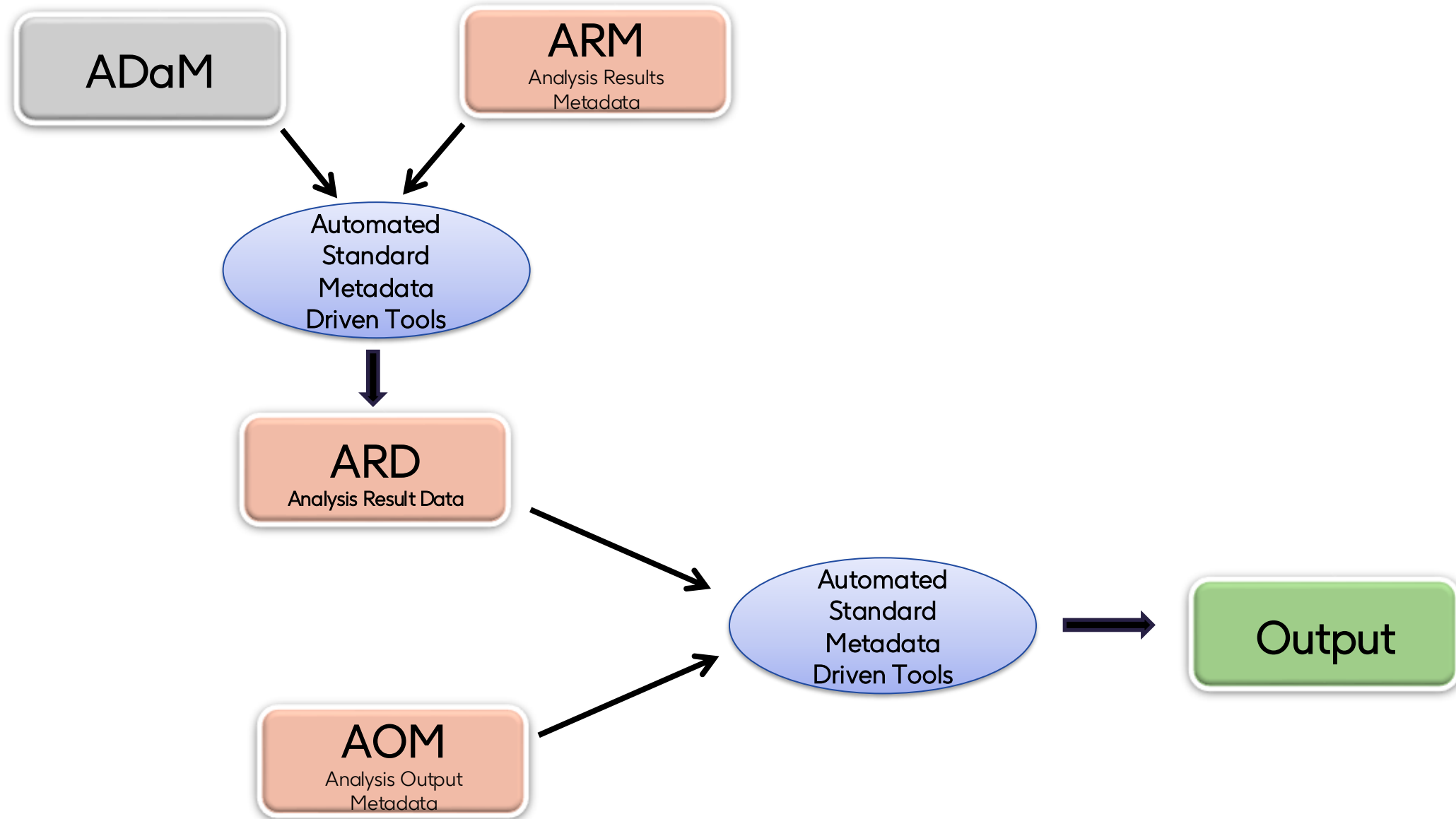
A
O
M

A
R
D

A
R
M

A
R
M

Metadata-Driven Automation Workflow



Summary & Key Highlights

The ARMADA framework introduces an **e2e metadata-driven approach** to streamline clinical reporting. It addresses inefficiencies caused by repeated analysis generation and ensures **traceability, consistency, and automation** across outputs.

Key Highlights:

WORM Principle (Write Once, Read Many): Enables reuse of validated analysis results across multiple outputs.

Three Metadata Layers:

- **ARM (Analysis Results Metadata):** Defines analysis protocols for traceability.
- **ARD (Analysis Results Data):** Stores reusable, unformatted analysis data.
- **AOM (Analysis Output Metadata):** Specifies output construction rules for flexible formatting.

Benefits:

- Eliminates redundant computations
- Improves compliance and auditability
- Enhances efficiency and transparency in reporting

Conclusion

The ARMADA model transforms clinical reporting by introducing an e2e metadata-driven approach that ensures:

- Reusability
- Traceability across all analyses
- Consistency in outputs
- Automation for efficiency

Metadata is not just documentation - it is the engine for scalable, transparent, and efficient reporting.

Q&A

GSK

GSK