

APAC Connect 2026

18–21 February
Novotel Hyderabad International Convention Centre

**AD11: E2E Clinical Metadata Journey:
‘Single Source of Truth’ to Accelerate
Clinical Research – From Collection
and Tabulation to Analysis and
Reporting**

Indumathi Narasetty
AstraZeneca

Disclaimer

- The views and opinions expressed in this presentation are those of speakers and do not reflect opinions of any other person or organization.

AGENDA

- Introduction
- Problems We Observed
- MetaData Repository (MDR) Implementation Journey
- Agile Framework & Scrum Dots..
- Study Instance Metadata (SIM)
- Study Instance Metadata (SIM) flow
- SIM life cycle and versioning
- Key outcomes
- Benefits
- Conclusion



Introduction

- **Challenge:** Demand for faster, higher-quality clinical delivery is rising, but organizations face fragmented metadata, inconsistent standards, limited automation, and integration gaps.
- **Solution:** An integrated MetaData Repository (MDR) + Study Instance Metadata (SIM) framework.
- **MDR** acts as a single source of truth for governed standards and metadata assets - driving consistency, traceability, and reuse.
- **SIM** translates MDR standards into study-specific, machine-readable metadata to enable downstream automation.
- This presentation covers the **MDR + SIM architecture** and its end-to-end impact on standards-driven clinical delivery.



Problems We Observed

- **Fragmented Metadata**

Metadata dispersed across protocol tools, EDC libraries, standards repositories, and mapping/programming/submission platforms, resulting in duplication and shadow repositories.
- **Siloed stakeholder workflows**

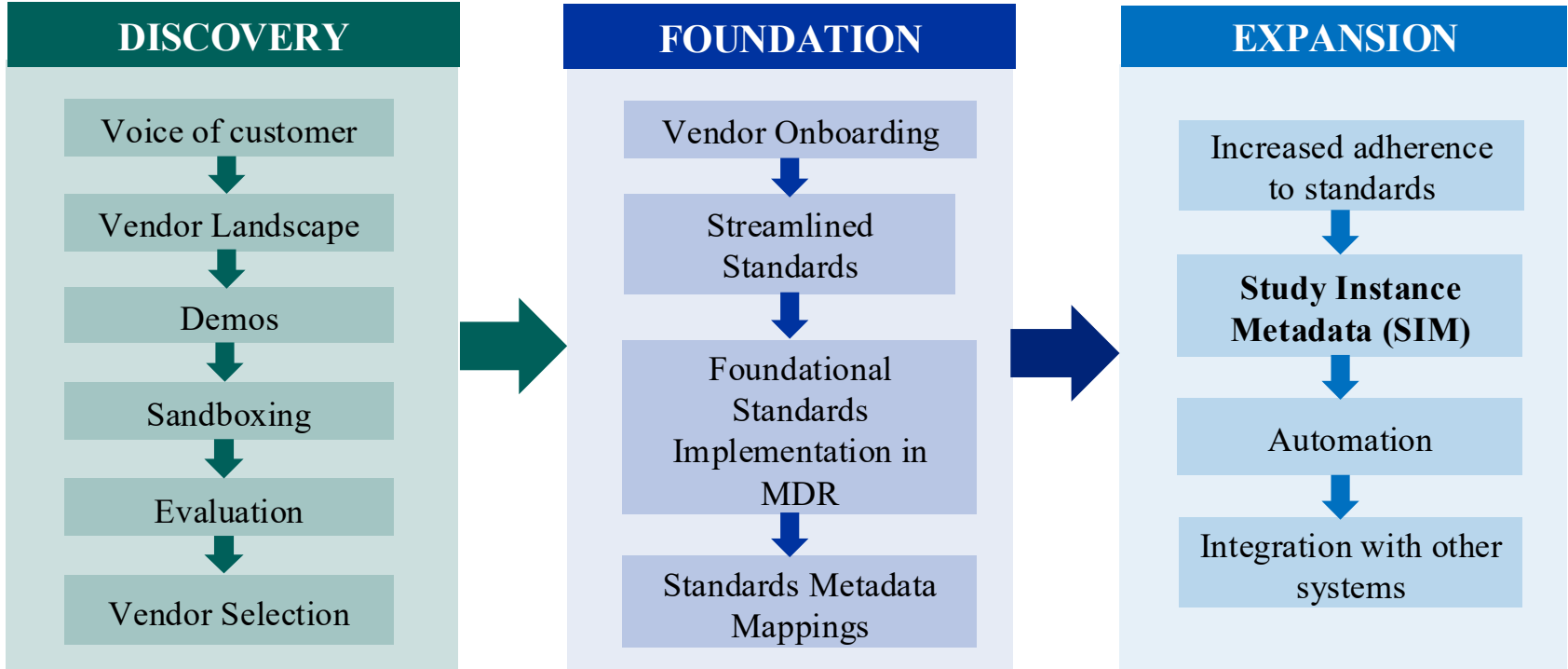
Limited real-time visibility across Data Management, Statistical Programming, Biostatistics, and Standards; changes don't reliably propagate, causing rework and misalignment.
- **Lack of machine-readable metadata** limits automations for study build, data cleaning, tabulation & analysis

Problems We Observed Contd...

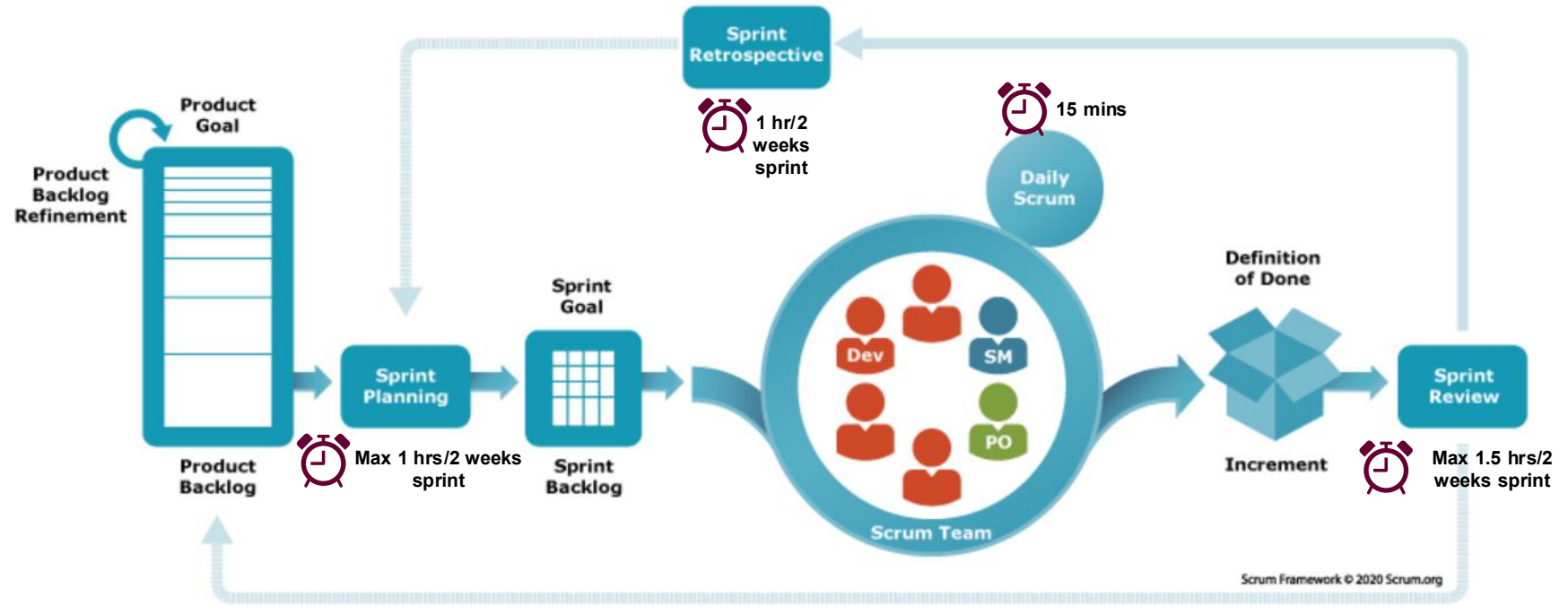
- **Missing end-to-end traceability**
No traceability from protocol → CRF → SDTM → ADaM → TLFs, undermining impact analysis, change control, and audit readiness
- **Integration challenges with downstream platforms**
- **Accessing Standards and consistency gaps**
Difficulty accessing the latest approved standards for accurate data collection, lacking standardization across studies



MDR Implementation Journey



Agile Framework



Scrum Framework © 2020 Scrum.org



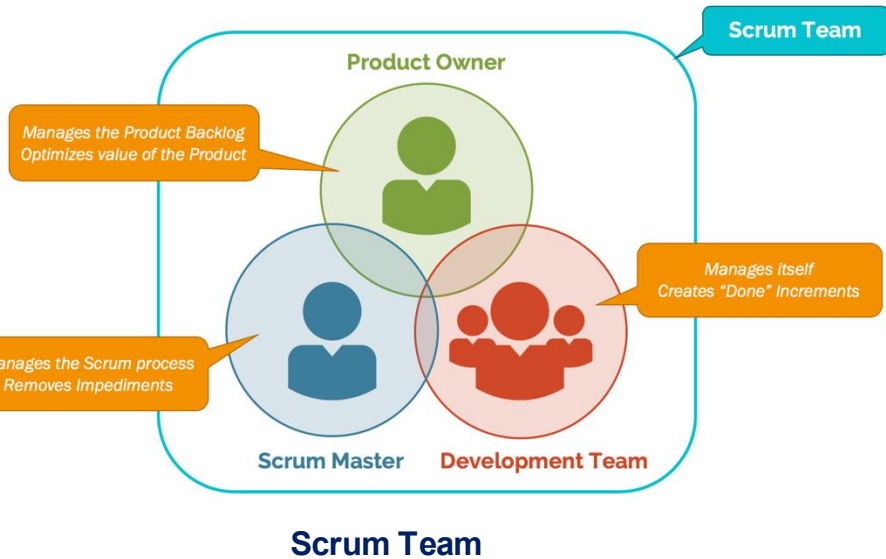
All events are time-boxed



Transparency & inspection



Scrum Dots..



Product Backlog
Sprint Backlog
Increment

Scrum Artifacts



Product Goal
Sprint Goal
Definition of Done

Commitments



Release Highlights

Please find the detailed MDR release information below; all listed releases are GxP-validated.

Releases	Success Criteria	Dataset	Time taken for creation of Dataset (s)	Improvement
Release 1.0	Minimum Viable Product (MVP) for the product	-	-	MVP
Release 1.1	Bug Fixes, Integration Improvements	-	-	Testing Phase
Release 1.2	General Product Enhancements, Bug Fixes	-	-	Testing Phase
Release 2.0	RAW - SDTM Mapping functionality deployment	-	-	Testing Phase
Release 2.1	Mapping Bug fixes	AE ADAE	AE: 1.5 - 2 hours ADAE: 6 - 7 hours	-
Release 2.2	Plugin Updates, Bug fixes, Domain restructuring, Performance Improvement	AE ADAE	AE: 1.2 - 1.6 hours ADAE: 5 - 7 hours	~ 25%
Release 2.3	RAW Plugin Updates - Visualization enhancements, ADaM Plugin Updates	AE ADAE	AE: 1.2 - 1.6 hours ADAE: 5 - 7 hours	No performance improvement

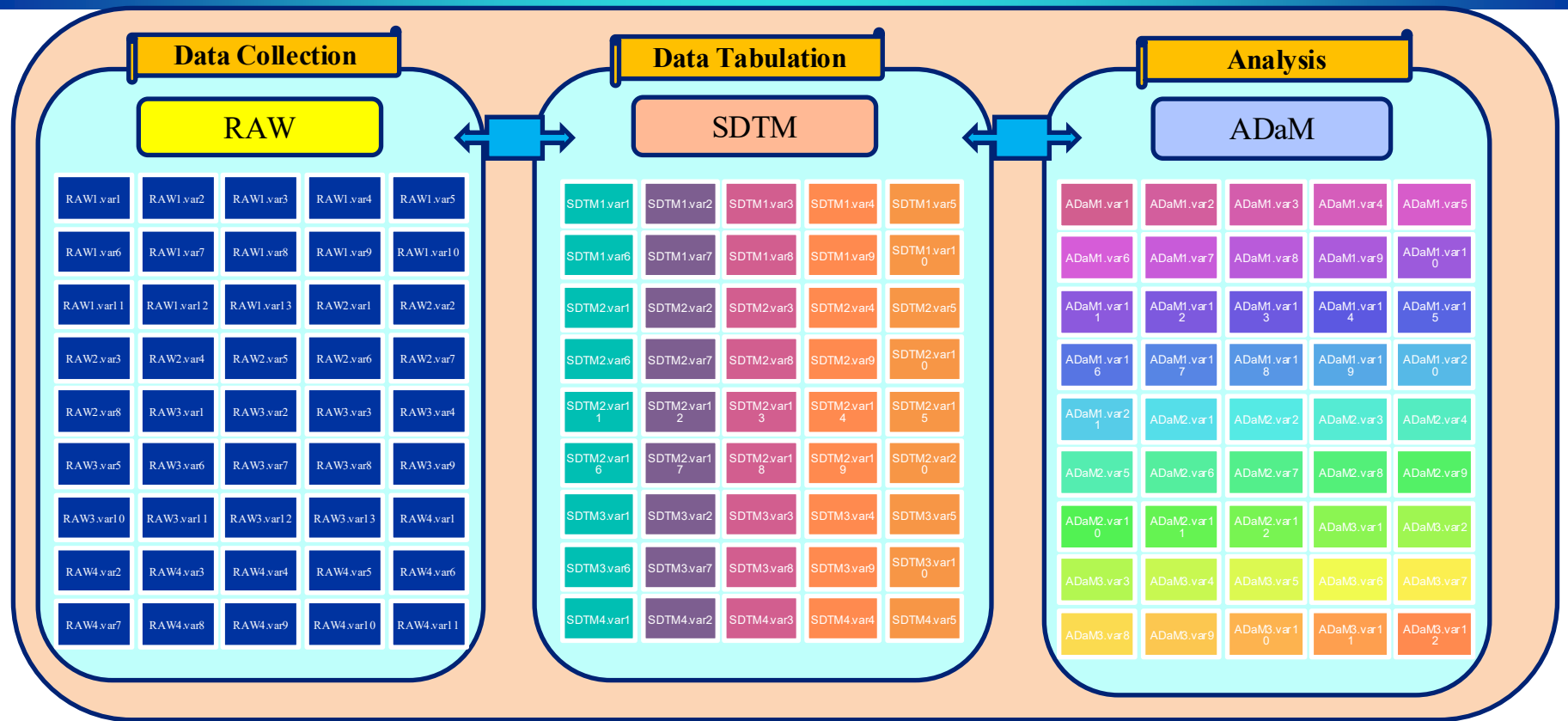


Releases Highlights Contd..

Releases	Success Criteria	Dataset	Time taken for creation of dataset(s)	Improvement
Release 2.4	Mapper Plugin Updates related to mapping functionality enhancements	AE ADAE	AE: 1.2 - 1.6 hours ADAE: 5 - 7 hours	No performance improvement
Release 3.0	SIM Technical Release, Performance Improvement, Integrations with few downstream systems	AE ADAE	AE : 0.8 - 1.2 hour ADAE : 4 - 6 hours	~ 30% - 40%
Release 4.0	Performance Improvement, Integrations with few other downstream systems, Study Compare reports, Mapping enhancements, SDTM & ADaM Plugin updates for compatibility of define-xml submission readiness	AE ADAE	AE : 0.6 - 1 hour ADAE : 3 - 5 hours	~ 25%
Release 4.1	ADaM Plugin updates, ADaM Value Level Metadata content enhancement, Change Log integration enhancement etc.,	AE ADAE	AE : 0.6 - 1 hour ADAE : 3 - 5 hours	No performance improvements in this release



Study Instance Metadata (SIM) – SIM Bank



Data Collection

RAW

RAW1.var1	RAW1.var2	RAW1.var3	RAW1.var4	RAW1.var5
RAW1.var6	RAW1.var7	RAW1.var8	RAW1.var9	RAW1.var10
RAW1.var11	RAW1.var12	RAW1.var13	RAW2.var1	RAW2.var2
RAW2.var3	RAW2.var4	RAW2.var5	RAW2.var6	RAW2.var7
RAW2.var8	RAW3.var1	RAW3.var2	RAW3.var3	RAW3.var4
RAW3.var5	RAW3.var6	RAW3.var7	RAW3.var8	RAW3.var9
RAW3.var10	RAW3.var11	RAW3.var12	RAW3.var13	RAW4.var1
RAW4.var2	RAW4.var3	RAW4.var4	RAW4.var5	RAW4.var6
RAW4.var7	RAW4.var8	RAW4.var9	RAW4.var10	RAW4.var11

Data Tabulation

SDTM

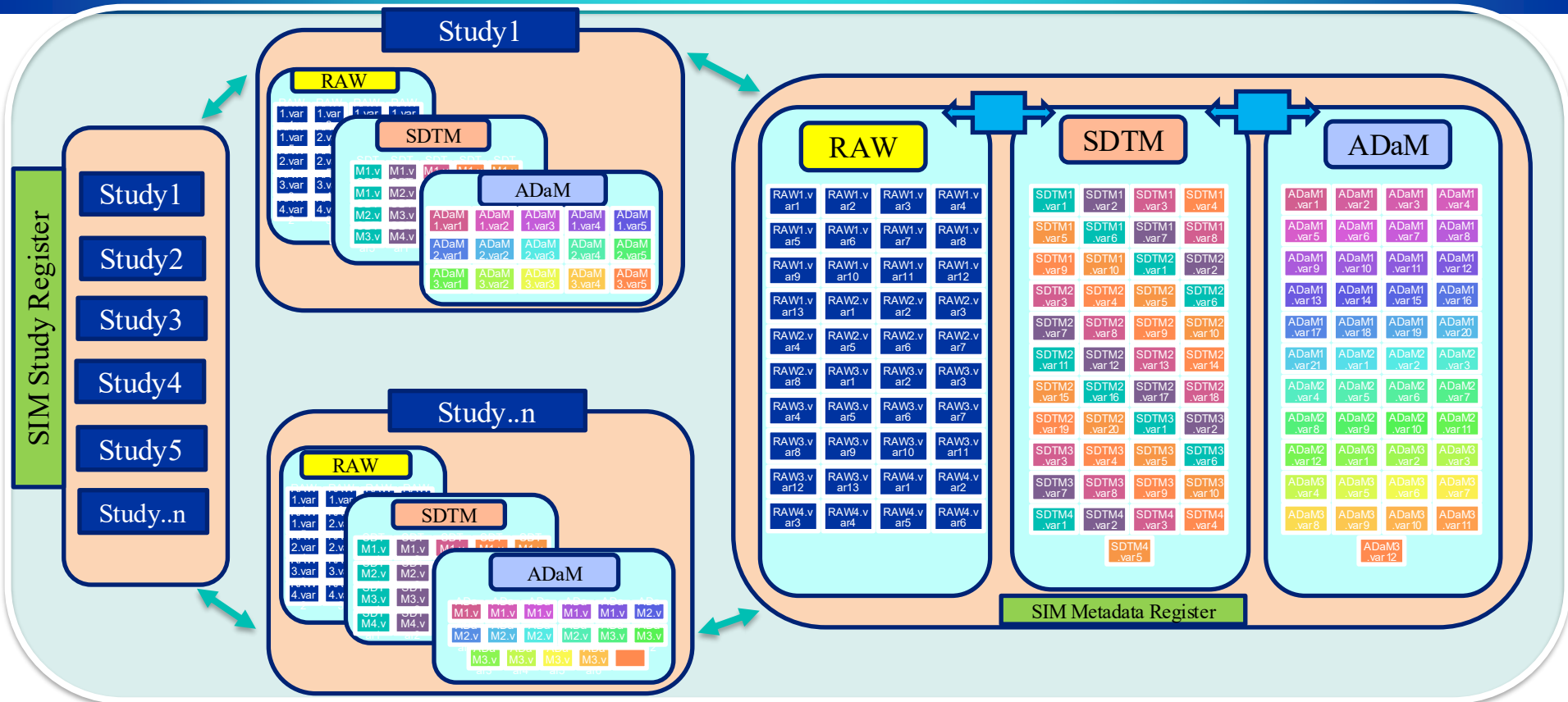
SDTM1.var1	SDTM1.var2	SDTM1.var3	SDTM1.var4	SDTM1.var5
SDTM1.var6	SDTM1.var7	SDTM1.var8	SDTM1.var9	SDTM1.var10
SDTM2.var1	SDTM2.var2	SDTM2.var3	SDTM2.var4	SDTM2.var5
SDTM2.var6	SDTM2.var7	SDTM2.var8	SDTM2.var9	SDTM2.var10
SDTM2.var11	SDTM2.var12	SDTM2.var13	SDTM2.var14	SDTM2.var15
SDTM2.var16	SDTM2.var17	SDTM2.var18	SDTM2.var19	SDTM2.var20
SDTM3.var1	SDTM3.var2	SDTM3.var3	SDTM3.var4	SDTM3.var5
SDTM3.var6	SDTM3.var7	SDTM3.var8	SDTM3.var9	SDTM3.var10
SDTM4.var1	SDTM4.var2	SDTM4.var3	SDTM4.var4	SDTM4.var5

Analysis

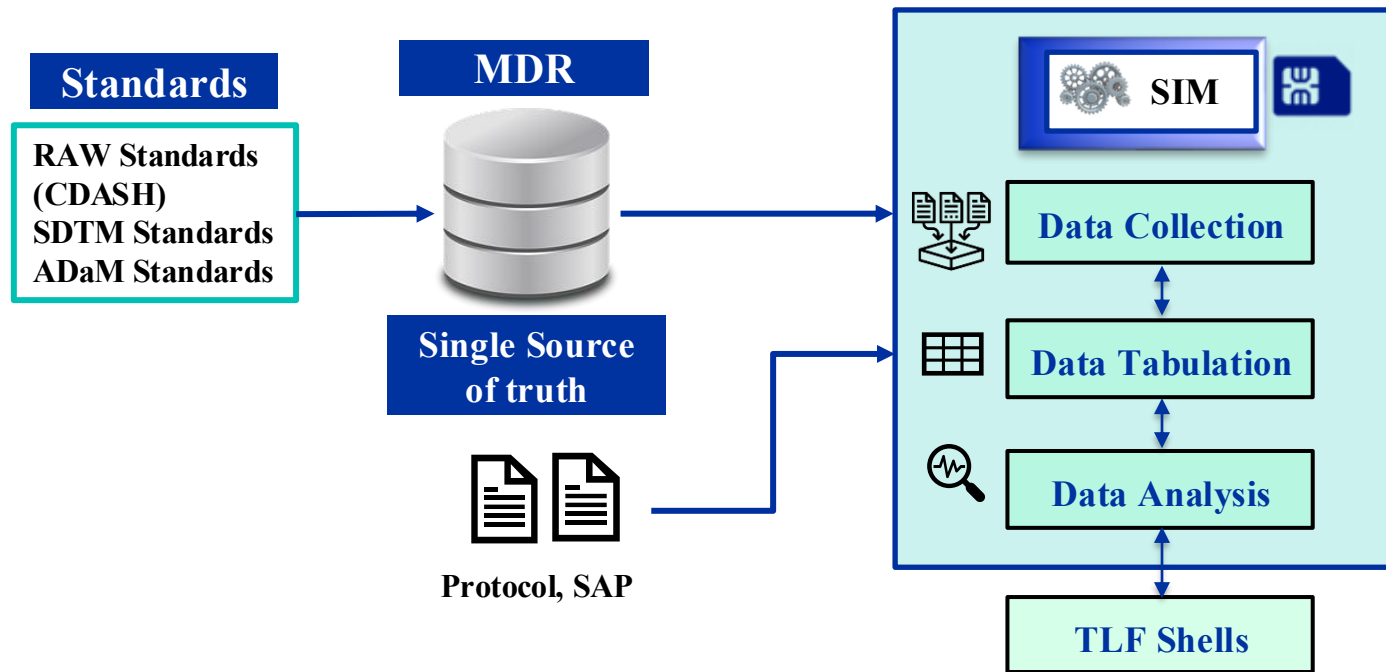
ADaM

ADaM1.var1	ADaM1.var2	ADaM1.var3	ADaM1.var4	ADaM1.var5
ADaM1.var6	ADaM1.var7	ADaM1.var8	ADaM1.var9	ADaM1.var10
ADaM1.var11	ADaM1.var12	ADaM1.var13	ADaM1.var14	ADaM1.var15
ADaM1.var16	ADaM1.var17	ADaM1.var18	ADaM1.var19	ADaM1.var20
ADaM2.var1	ADaM2.var2	ADaM2.var3	ADaM2.var4	ADaM2.var5
ADaM2.var6	ADaM2.var7	ADaM2.var8	ADaM2.var9	ADaM2.var10
ADaM2.var11	ADaM2.var12	ADaM2.var13	ADaM2.var14	ADaM2.var15
ADaM2.var16	ADaM2.var17	ADaM2.var18	ADaM2.var19	ADaM2.var20
ADaM3.var1	ADaM3.var2	ADaM3.var3	ADaM3.var4	ADaM3.var5
ADaM3.var6	ADaM3.var7	ADaM3.var8	ADaM3.var9	ADaM3.var10
ADaM3.var11	ADaM3.var12	ADaM3.var13	ADaM3.var14	ADaM3.var15
ADaM3.var16	ADaM3.var17	ADaM3.var18	ADaM3.var19	ADaM3.var20

Study Instance Metadata (SIM)



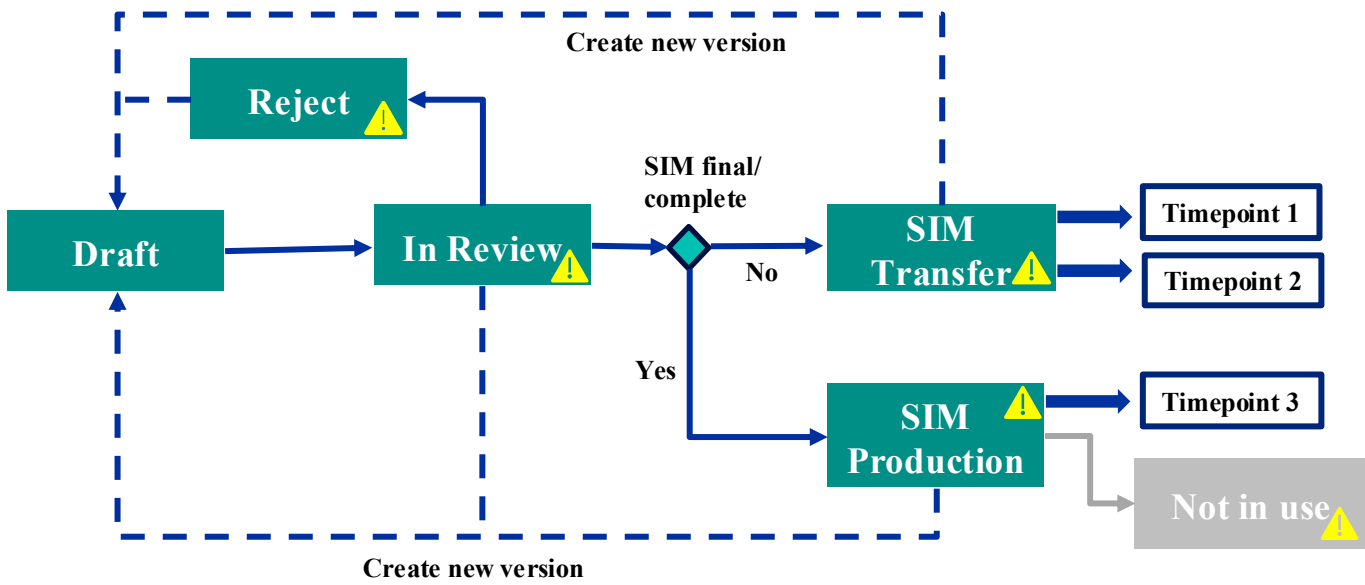
Study Instance Metadata (SIM) Flow



MDR - MetaData Repository; SAP - Statistical Analysis Plan; SIM - Study Instance Metadata; CDASH - Clinical Data Acquisition Standards Harmonization; TLF - Tables, Listings, Figures; SDTM - Study Data Tabulation Model; ADaM - Analysis Data Model;



SIM Life Cycle & Versioning



Timepoint 1: SIM version 1 containing Final RAW, initial SDTM, initial ADaM transferred to downstream

Timepoint 2: SIM version 2 containing Final RAW, final SDTM, initial ADaM transferred to downstream

Timepoint 3: SIM version 3 containing Final RAW, final SDTM, final ADaM transferred to downstream

SIM not editable

Key Outcomes



Generation of RAW
Specification



define-xml ready Study
Metadata Specifications -
SDTM, ADaM, TLF Metadata



RAW – SDTM Mapping
Specification (Facilitating the
Traceability from RAW –
SDTM)



Generate compliant **define-xml** for submissions.



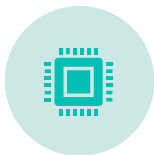
Study Metadata - For RAW,
SDTM, ADaM, TLF Metadata

Benefits



Faster study builds

Up to **30-40%** reduction in setup time via reusable, standardized metadata



Stronger standards governance

Improved maintenance of CDASH, SDTM, and ADaM libraries



Streamlined setup

Efficient **EDC build** and smoother **Analysis & Reporting** (SDTM, ADaM specs, **define-xml**)



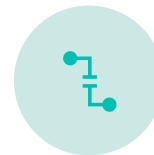
Machine-readable SIM

Structured metadata enabling automation, consistency, and reuse



Accelerated sister-study setup

Leverages the **SIM study register** to quickly replicate similar studies within the same compound



API-centric integration

Seamless connectivity with downstream systems



End-to-end traceability

Clear lineage from Protocol → eCRFs → SDTM → ADaM → Submission



Easier CDISC upgrades

Faster adoption of newer **CDISC** versions

Conclusion

- **Operational Impact:**

- Our MDR journey demonstrates how a centralized, governed metadata foundation can transform clinical delivery.
- By unifying standards, enabling automation, and integrating systems, **MDR + SIM** offers a resilient, reusable framework.

- **Clinical Outcomes:**

Accelerates study build, enhances traceability, and strengthens compliance, ultimately driving smarter, faster, and higher-quality clinical research.

References

- Transforming Biometrics: The Future of Statistical Programming Through AI-Powered Automation and Agile Scrum
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2025/SDE/APAC/Bengaluru/PRE_Bengaluru08.pdf
- Scrum.org



Indumathi Narasetty

AstraZeneca India Pvt Ltd

+91-9581229439




indumathi.narasetty@astrazeneca.com

www.linkedin.com/in/indumathi-narasetty-01a92483



The Global Healthcare Data Science Community

Contact Channels

-  UK +44 1843 609600
-  office@phuse.global
-  phuse.global

Social Media

-  [@phusetwitta](https://twitter.com/phusetwitta)
-  [/phusebook](https://facebook.com/phusebook)
-  [/phusetube](https://youtube.com/phusetube)
-  [/company/phuse](https://linkedin.com/company/phuse)