



connect·share·advance

Single Day Events

TRIVANDRUM, INDIA
Saturday 20 June 2026

phuse.global





connect·share·advance

Standards First Automation for AI Enabled Clinical Data Activities

GOWTHAMI PADMA KURRA
(Clinical Database Developer)

&

VEDALA PRASANNA KUMAR
(Expert Clinical Database Developer)

Novartis India

phuse.global

Outline

- **Why This Topic Now?**
- **Core Challenges in Current Practices**
- **Standards-First Approach**
- **Role of CDISC Across the Lifecycle**
- **Regulatory Expectations**
- **Operationalizing Standards**
- **Impact on Build, Validation, and AI**
- **Governance & Sustainability**
- **End-to-End Benefits & Key Takeaways**



Why This Topic Now?

Rapid Transformation in Clinical Trials

- Automation, AI, and ML are accelerating change in clinical trials, managing complex data and decentralized sources.
- Trial complexity (decentralized trials, adaptive designs) increasing exponentially
- Data volume growing across EDC, ePRO, labs, wearables

Regulatory Expectations

- FDA and EMA demand higher standards of data traceability, consistency, and integrity in AI-assisted processes.

Importance of Data Standards

- Sustainable AI value depends on robust, standardized clinical data aligned with regulatory requirements.



The Core Challenge in AI-Enabled Clinical Trials

Fragmented Clinical Data Practices

- Organizations face inconsistent data capture due to study-specific CRF designs across therapeutic areas and vendors

Rule Proliferation and Limited Reuse

- Custom logic and edit checks lead to uncontrolled proliferation of rules, limiting reuse across studies.

Metadata Underutilization

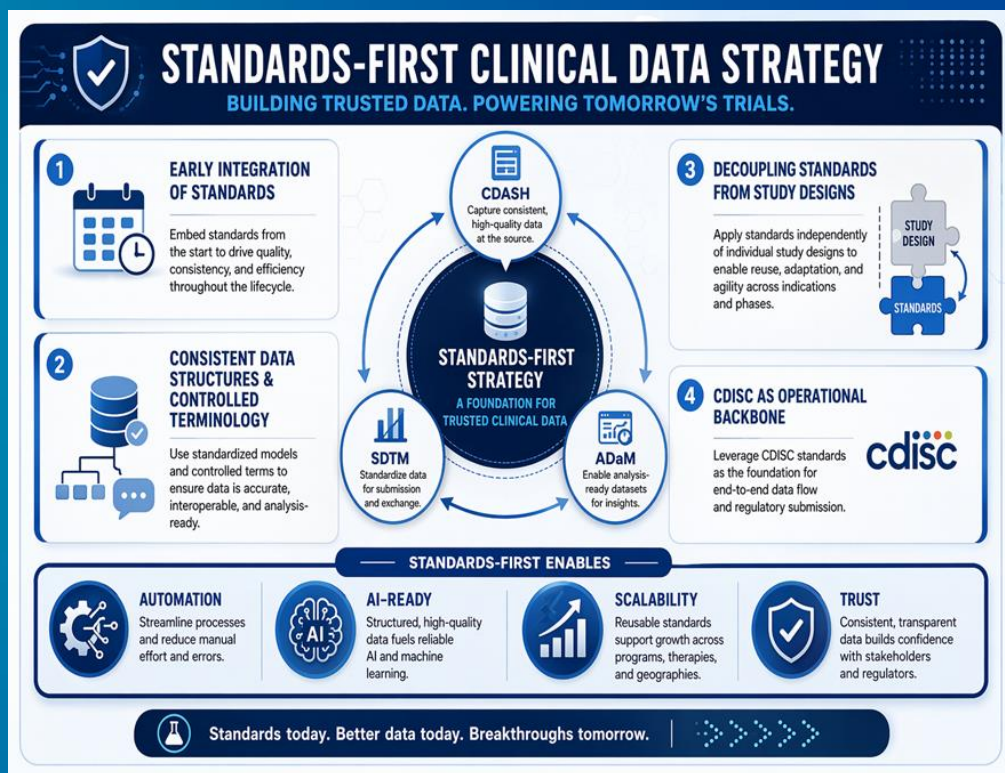
- Validated specifications and lessons from prior studies are rarely reused due to poor metadata management.

Impact on AI and ML Success

- AI failures often stem from data variability , affecting model explainability and regulatory acceptance.



What “Standards First” Really Means



Early Integration of Standards

- Embedding standards from the start ensures consistent data structures and controlled terminology throughout.
- Decoupling standards from individual study designs

CDISC as Operational Backbone

- CDISC standards like CDASH, SDTM, and ADaM guide decisions across the clinical data lifecycle effectively.

Enabler of Innovation

- Standards-first approach creates a stable environment that supports automation, AI, scalability, and trust.



Role of CDISC Across the Clinical Data Lifecycle

Standardized Data Capture with CDASH

- CDASH standardizes data capture at the source by aligning case report forms to common elements, reducing ambiguity.

Data Transformation Using SDTM

- SDTM transforms collected data into regulator-ready tables that ensure traceability and review efficiency.

Analysis-ready Datasets via ADaM

- ADaM provides analysis-ready datasets supporting reproducible statistics and clear derivation logic.

Continuous Data Lineage

- Together, CDASH, SDTM, and ADaM create continuous data lineage, enhancing data integrity and simplifying processes.

Regulatory Expectations from FDA and EMA

Data Traceability & Consistency

- Data lineage from CRF → SDTM → ADaM with minimum variability
- Explainability of derived variables and transformations
- Consistency across studies within a program



AI & Automation

- Auditability of automated AI processes
- Traceability of AI decisions
- Early error detection
- Shift from reactive remediation to proactive design



Standards-First Approach

- CDISC standards as foundation
- Embedded regulatory principles
- Reusable & scalable frameworks
- Built-in inspection readiness



Practical Enablement: From Standards to Execution

CDASH-Aligned CRF Design

CDASH-aligned CRF library with predefined modules (AE, CM, LB)

Metadata-Driven Reuse

Metadata repository for fields, variables, and controlled terminology

Governed Rule Libraries

Governed rule libraries for edit checks enable validated logic reuse, enhancing consistency across trials.

Version-controlled reuse across studies

Efficiency and Compliance

Operationalizing standards drives efficiency by reducing build time and variability while maintaining quality and compliance.

Refs: SAS Proceedings – Metadata-Driven Data Management





Impact on Database Build and Validation

Standardization Benefits

- Standardized CRFs, metadata, and rules speed up database builds and reduce customization needs.
- Parallelization of build and validation activities.

Reusable Edit Checks & Custom Functions

- Validated reusable edit checks, CF's lower error rates and minimize repeated testing efforts.
- Reuseable UAT Scenarios across programs.

Scalable Validation

- Consistent structures enable scalable and efficient validation across multiple programs.

Improved Quality and Speed

- Reuse of standard libraries reduces build time (approx. 30-40%) significantly without compromising quality.





Standards as an Enabler for AI and ML

Consistent Training Data

Standards ensure high-quality, consistent datasets that enhance AI and ML model reliability and effectiveness. Ex: AI Model trained on lab data; AI identifies Protocol deviations.

Cross-study Model Reuse

Conforming to standards like CDISC enables AI models to be reused across studies, improving scalability and learning.

Automated Review and Validation

Standards support intelligent validation and anomaly detection, helping automated review processes identify outliers effectively.

Safe Deployment in Regulated Environments

Standards create conditions for safe, compliant AI deployment, balancing innovation with regulatory requirements.





Debunking the “Standards Limit Innovation” Myth

Misconception: Standards restrict flexibility

Reality:

Standards Enhance Clarity

- Standards reduce noise and variability, allowing true signals and insights to emerge more clearly in innovation processes.

Explainability and Auditability

- In AI and ML, standards support explainability and auditability, which are essential for regulatory acceptance and trust.

Stable Baseline Enables Focus

- Standards provide a stable baseline, enabling organizations to focus innovation on advanced analytics and predictive insights.



Governance Models That Make It Sustainable

Clear Ownership and Stewardship

- Defined ownership of standards ensures accountability and drives continuous improvement in automation governance.

Change Control Processes

- Controlled change processes for metadata and rules prevent divergence while allowing necessary evolution with version traceability.

Oversight of AI and Automation

- Oversight mechanisms align AI and automation with regulations and ethical standards to ensure responsible use. Ex: AI model lifecycle (train, validate, deploy)

Centralized Metadata Repository

- A centralized metadata repository governed by cross-functional teams (data management, biostats, and IT) supports scalable and sustainable automation.





End-to-End Lifecycle Benefits

Accelerated Study Startup

- Reuse of standardized CRFs and rules speeds up study startup, reducing initial delays significantly.

Efficient Trial Conduct

- Consistent data enables efficient monitoring, review, and swift issue resolution during trial conduct.

Faster Study Closeout

- Reduced rework and clearer traceability accelerate the study closeout phase effectively.

Portfolio Scalability

- Organizations achieve scalability with consistent standards across multiple programs and therapeutic areas.
- Portfolio-level performance optimization.



Key Takeaways

- **Standards First is Essential** - Foundational for scalable automation, AI, and compliance
- **Consistency Enables Intelligence** - High-quality, standardized data drives reliable AI outcomes
- **CDISC Ensures Traceability** - End-to-end lineage from CRF → SDTM → ADaM is critical
- **Reuse Drives Efficiency** - CRFs, edit checks, and metadata libraries reduce build time and variability
- **Governance Sustains Scale** - Controlled processes ensure consistency, auditability, and continuous improvement
- **AI + Standards = Trust & Innovation** - Standards enable explainable, compliant, and scalable AI adoption

connect·share·advance



THANK YOU!

connect·share·advance



The Global Healthcare Data Science Community

Contact Channels

✉ office@phuse.global
🌐 phuse.global

Social Media

🌐 [/company/phuse](https://www.linkedin.com/company/phuse)
▶ [/phusetube](https://www.youtube.com/channel/UCphuse)
✕ [/phuse_global](https://twitter.com/phuse_global)
f [/phusebook](https://www.facebook.com/phusebook)