Paper SM13

Turning Data into Deliverables: Automating Biometrics from Metadata to TFLs

Inderjeet Arora, GenInvo, Toronto, Canada
Daragh Ryan, GenInvo, Horsham, UK

ABSTRACT

The clinical trial industry has long relied on well-defined macros and tools for generating Tables, Listings, and Figures (TLFs). However, the broader biometrics workflow i.e., from transforming raw/source data to SDTM, then to ADaM, and ultimately to TLFs remains complex and involves multiple manual steps. Despite automation tools, significant effort is still required to manage data mappings, transformations, and validations across these stages. A unified platform that automates the end-to-end process i.e., mapping raw data to SDTM, SDTM to ADaM, and ADaM to TLFs - can significantly streamline this workflow. By incorporating AI algorithms, such a system can intelligently manage data transformations and TLF generation, while built-in audit trails and governance features ensure controlled access, transparency, and traceability. leading to reduction in repeated communication, and manual execution. In this paper, we will explore how automation can be effectively applied to simplify and accelerate the journey from raw data to TLF generation.

INTRODUCTION

The generation of TLFs forms a critical component of the clinical trial data analysis and reporting process. Over the years, the pharmaceutical and clinical research industries have relied on standardized macros and limited rule based tools to support TLF generation, ensuring compliance, reproducibility, and consistency across studies. However, while these solutions have streamlined parts of the reporting workflow, the broader biometrics process—from transforming raw or source data to Study Data Tabulation Model (SDTM), then to Analysis Data Model (ADaM), and ultimately to TLFs—remains complex and fragmented.

This multi-stage pipeline typically involves several manual interventions for data mapping, transformation, and validation. Each stage requires significant coordination among data managers, statisticians, and programmers, often leading to redundant communication cycles, prolonged timelines, and increased potential for human error. Despite the introduction of automation tools for specific sub-tasks, a holistic solution that seamlessly integrates the entire data flow is still lacking in most organizations.

To address these challenges, a unified automation platform capable of managing end-to-end data transformation—from raw data to SDTM, SDTM to ADaM, and ADaM to TLFs—offers a promising direction. By incorporating artificial intelligence (AI) and machine learning techniques, such a system can intelligently handle data mappings and TLF generation with minimal human intervention. Additionally, the integration of audit trails, access control, and governance mechanisms can ensure transparency, traceability, and regulatory compliance.

This paper explores how automation and AI-driven approaches can be effectively applied to simplify and accelerate the journey from raw data to TLF generation. It examines the potential benefits of a unified platform, discusses the associated implementation challenges, and outlines how such an approach can enhance efficiency, quality, and collaboration within clinical trial data processing workflows.

RAW DATA

Raw data in clinical trials represent the foundational layer of clinical information collected during the execution of a study. These data encompass the initial, unprocessed observations, measurements, and entries that are directly obtained from study participants or associated clinical systems. They capture the factual record of what was observed, measured, or reported at the point of care or during scheduled study visits—before any modification, cleaning, or standardization is applied.

Typically, raw data includes a wide range of clinical and operational information such as demographic details, laboratory test results, vital signs, electrocardiogram (ECG) readings, imaging data, patient-reported outcomes, and adverse event reports. These data may originate from multiple sources, including electronic case report forms (eCRFs), laboratory information systems, medical devices, electronic health records (EHRs), or third-party vendor systems.

In essence, raw data forms the starting point of the clinical data lifecycle, providing the evidence base from which all subsequent analyses, statistical models, and clinical interpretations are derived.

CLINICAL DATA INTERCHANGE STANDARDS CONSORTIUM

The Clinical Data Interchange Standards Consortium (CDISC) provides a globally recognized framework of data standards that ensure consistency, traceability, and regulatory compliance in clinical research. These standards

define how clinical data should be structured, formatted, and exchanged across the different stages of the trial data lifecycle—from raw data capture to final analysis and reporting.

STUDY DATA TABULATION MODEL

The Study Data Tabulation Model (SDTM) defines a standardized structure for organizing and submitting clinical trial data to regulatory agencies such as the FDA or PMDA. Its purpose is to present collected data in a clear, consistent, and machine-readable format that facilitates review and validation.

SDTM datasets are typically derived from cleaned raw data and are organized into domains, each representing a specific type of data—such as Demographics (DM), Adverse Events (AE), Laboratory Tests (LB), Vital Signs (VS), or Exposure (EX). Each domain follows a predefined structure with variable names, labels, and controlled terminologies defined by CDISC standards.

By mapping raw data into SDTM domains, sponsors ensure that regulators can efficiently navigate and interpret study data using a familiar structure.

ANALYSIS DATA MODEL

The ADaM standard focuses on data prepared for statistical analysis. It builds upon SDTM datasets by organizing variables and derived measures in a way that facilitates reproducible and traceable statistical computations.

ADaM datasets include both observed and derived variables (e.g., change from baseline, treatment-emergent flags) and are typically structured around the analytical requirements defined in the study's Statistical Analysis Plan (SAP). The most common ADaM dataset types include:

- ADSL (Subject-Level Analysis Dataset) – Contains one record per subject with key demographics and treatment information.
- ADAE (Adverse Events Analysis Dataset) – Used for safety analyses.
- ADLB (Laboratory Analysis Dataset) – Supports clinical laboratory analyses.

Each ADaM dataset is designed to ensure traceability back to SDTM and, ultimately, to the original raw data, maintaining a clear lineage of data transformations.

METADATA REPOSITORY

A metadata repository serves as a centralized knowledge hub that defines, manages, and governs metadata across all stages of the clinical data lifecycle—from Raw Data to SDTM and ADaM. It enables a structured, traceable, and automated approach to data standardization and transformation, ensuring consistency, compliance, and efficiency across clinical studies.

The metadata repository acts as the single source of truth for all data definitions, mappings, derivations, and relationships across datasets. It maintains the structure and semantics of variables, datasets, and transformation logic that guide how raw clinical data are standardized into CDISC-compliant SDTM domains and then transformed into ADaM datasets for analysis.

By providing a unified framework, the repository supports end-to-end data automation, traceability, and governance — key enablers for reliable and compliant data processing.

STATISTICAL ANALYSIS PLAN

A Statistical Analysis Plan (SAP) is a formal document that outlines the detailed statistical methodologies and procedures to be applied in a clinical trial. It translates the study's objectives, hypotheses, and endpoints into concrete analysis steps, defining how raw data will be transformed into SDTM and ADaM datasets for statistical evaluation. The SAP specifies analysis populations, such as intent-to-treat, per-protocol, or safety sets, and documents rules for handling missing data, data derivations, and transformations. By pre-specifying the analysis approach, including primary, secondary, and exploratory endpoints, the SAP ensures that all results are reproducible, traceable, and aligned with regulatory expectations.

A key feature of SAP is its emphasis on traceability. Every analysis variable and dataset is linked back to its SDTM source and, ultimately, to the original raw data. This linkage ensures transparency and allows regulatory reviewers or auditors to validate the derivation of results. The SAP also defines output specifications for Tables, Listings, and Figures (TLFs), often including mock tables or shells, and provides derivation logic for analysis variables, making the analytical workflow systematic and reproducible. The Biometrics team relies on details and instructions provided in SAP to generate essential outputs such as TLFs. Hence, it's imperative to draft SAP in a manner that facilitates understanding and adherence, ensuring seamless data handling and analysis.

MOCK SHELLS

In clinical trials, a mock shell (or mock table/listing/figure shell) is a template or blueprint that defines the expected structure, layout, and content of statistical outputs such as TLFs — even before actual data are available. Mock shells are typically created as part of the SAP development process and serve as a bridge between statistical programming, medical writing, and clinical interpretation teams.

Mock shells clearly specify what each output will display, how it will be formatted, and which analysis variables or subgroups will be included. They define titles, footnotes, column headers, population definitions, and any derivations

or calculation rules. By predefining these elements, mock shells ensure alignment between statisticians, programmers, and reviewers on how data will ultimately be presented for regulatory submission or clinical reporting. Mock shells are essential for quality, consistency, and compliance in clinical reporting. They:

- Ensure a clear understanding of output expectations before programming begins.
- Serve as documentation for auditors and regulators to verify pre-specified analyses.
- Facilitate programming automation, especially in metadata-driven environments where shells can be linked to specifications in SDTM and ADaM.
- Support traceability from SAP-defined objectives to final outputs submitted in the Clinical Study Report (CSR).

TABLES, LISTINGS, AND FIGURES

TLFs are the primary outputs generated during the statistical analysis and reporting phase of a clinical trial. They present the analyzed data in a structured, interpretable, and regulatory-compliant format, forming the foundation of the Clinical Study Report (CSR) submitted to health authorities such as the FDA or EMA. Each TLF provides a unique perspective on trial results — tables summarize data numerically, listings display subject-level information, and figures visualize trends and comparisons.

Tables typically provide aggregated summaries of key variables, such as demographics, efficacy outcomes, or safety parameters. For example, a table might display the mean change in blood pressure by treatment arm or the frequency of adverse events by severity grade. Listings, on the other hand, present participant-level data — showing raw or derived values for each subject — and are often used for review, quality control, or audit purposes. Figures (such as Kaplan-Meier survival curves, forest plots, or bar charts) are used to graphically illustrate patterns, relationships, or trends in data that may not be easily captured in tabular form.

TLFs are driven by the Statistical Analysis Plan (SAP), which defines the analysis methods, populations, and variables to be included. The generation of TLFs relies on standardized analysis datasets (typically in the ADaM format), ensuring traceability from raw data through SDTM and ADaM to final outputs.

TRADITIONAL APPROACH

In the traditional clinical data flow, the process begins with collecting raw data from multiple sources such as EDC systems, laboratories, ePRO tools, or vendor files. This data represents the unprocessed records of patient demographics, adverse events, lab results, and other trial observations. The next step is to transform this raw data into SDTM datasets based on CDISC standards, ensuring a consistent, structured format suitable for regulatory submission. Each SDTM domain (like DM for demographics or AE for adverse events) adheres to controlled terminology, variable naming conventions, and documentation in Define.xml. The SDTM layer provides a standardized and traceable view of collected data, enabling reviewers at agencies such as the FDA or EMA to easily navigate and understand study results.

STANDARDS-COMPLIANT DATA TRANSFORMATION

The traditional process of transforming clinical data into standards-compliant formats (e.g., CDISC-SDTM) typically involves manual, rule-based procedures that are time-consuming, error-prone, and require specialized expertise. It often necessitates developing customized scripts to extract data from multiple sources, manually mapping raw clinical data to standard domains and elements, and performing repeated validations against compliance rules.

Common challenges associated with this manual approach include:

- Labor-intensive and time-consuming activities
- Manual handling of codelist mapping, unit conversion, and similar repetitive tasks
- High risk of human error affecting accuracy and completeness
- Inconsistent mappings across studies
- Requirement for in-depth understanding of raw data and metadata before mapping

To address this challenge, our internal solution leverages advanced Machine Learning (ML) techniques to automate the data transformation and mapping process, thereby reducing time, manual effort, and the need for repetitive programming while ensuring the creation of high-quality, standards-compliant datasets (such as SDTM, ADaM) for regulatory submission.

Once SDTM is finalized, the data is further processed into ADaM datasets, which are specifically designed to support statistical analysis. These datasets include derived variables such as baseline values, change from baseline, treatment flags, or analysis groups, maintaining full traceability back to SDTM. Using ADaM data, programmers and statisticians generate TFLs—the final outputs that summarize and visualize study results in the CSR. Tables display numerical summaries, figures visualize trends, and listings provide subject-level details. This structured approach ensures data integrity, regulatory compliance, and reproducibility, making it the cornerstone of modern clinical trial data management.

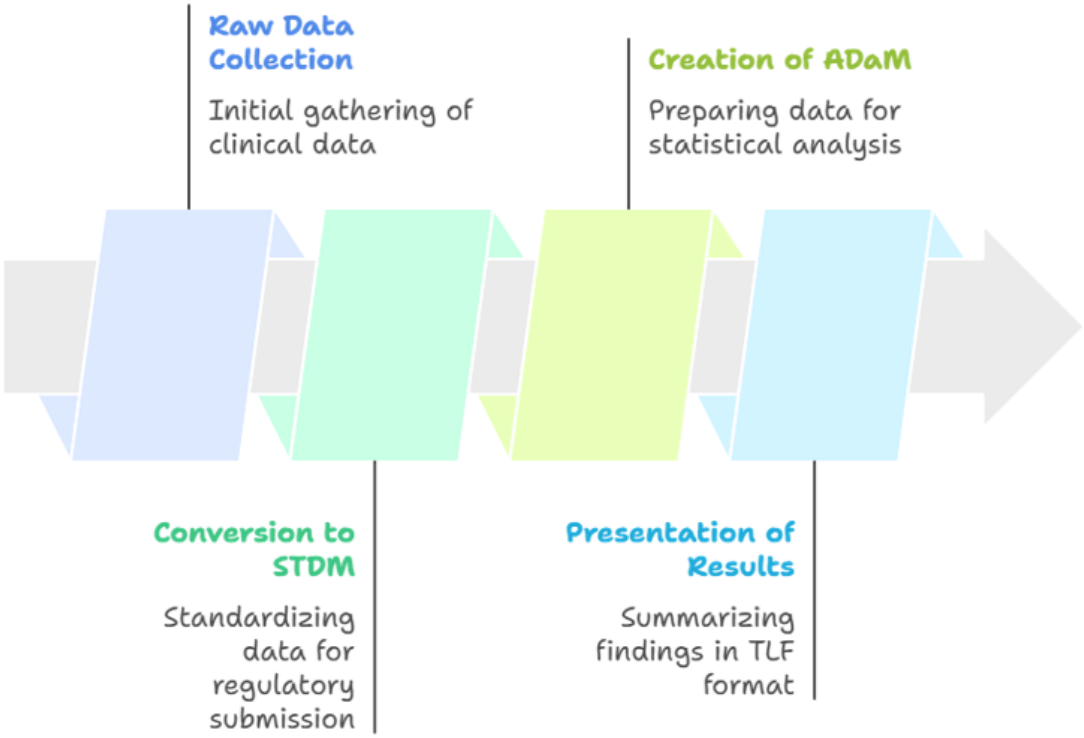
CREATING TABLES, LISTINGS, AND FIGURES

After the ADaM datasets are prepared, statistical programmers manually develop and execute SAS programs to generate each TLF according to the SAP and mock shell specifications. Every output—whether a summary table, subject listing, or visual figure—is coded individually, validated, and quality-checked to ensure accuracy and

compliance. The process involves several review cycles between programmers, biostatisticians, and clinical teams to confirm that the outputs are consistent with the SAP and mock shells.
While this traditional approach ensures traceability, precision, and regulatory compliance, it also has several significant drawbacks:

- High Manual Effort: Each TLF requires individual programming, which increases the workload for statistical programmers, especially in large or complex studies.
- Time-Consuming Process: Multiple layers of programming, validation, and review extend study timelines.
- Limited Reusability: TLF programs are often study-specific, making it difficult to reuse or standardize across trials.
- High Risk of Human Error: Manual coding and updates can introduce inconsistencies and inaccuracies.
- Resource Intensive: Requires extensive involvement from programmers, statisticians, and reviewers at each stage.
- Challenging to Maintain Traceability: Ensuring linkage from protocol to SAP to ADaM and final TLFs demands meticulous documentation and verification.
- Difficulty Managing Changes: Any change in SAP or data structure may require reprogramming multiple outputs, adding to the effort and risk.

Figure 1: Traditional Flow – Raw to TLFs



OUR APPROACH

METADATA REPOSITORY

The MDR module functions as a centralized hub that captures, manages, and governs the metadata associated with clinical studies. It supports various components such as source data definitions, SDTM domain metadata, mapping specifications, transformation rules, and derivation logic. By consolidating all these elements in one environment, organizations can streamline their data standardization and transformation processes, improve data integrity, and reduce operational inefficiencies.
Key advantages of our internal MDR include:

Centralized governance of metadata, ensuring all study teams work from a single, validated source of truth.

Automated mapping between study-level and standard-level metadata, significantly reducing manual effort and potential errors.

Enhanced traceability and auditability, allowing every change, version, and transformation to be tracked for compliance and transparency.

Improved collaboration and consistency across cross-functional teams through a shared metadata framework.

Scalability and reusability, allowing new studies to be set up quickly by leveraging existing metadata templates and configurations.

By maintaining a centralized and traceable metadata structure, the MDR simplifies data validation, accelerates regulatory submission readiness, and promotes transparency throughout the clinical data lifecycle. It ensures that organizations can efficiently manage metadata changes while maintaining compliance with standards such as CDISC SDTM and ADaM.

Key Features of our internal MDR Module:

Workflow Management – Automates metadata workflows, from the creation of new standards to the maintenance of existing ones, ensuring consistency and efficiency.

Automated Study Setup – Supports importing of both legacy and CDISC-compliant metadata with auto-recognition of standards, and automatically maps study metadata to the central library.

Metadata Management – Facilitates management of all metadata components, including domains, elements, and controlled terminology, with full edit and update capabilities.

Governance Framework – Incorporates built-in governance models to support metadata review, approval, and change-control processes.

Impact Analysis – Evaluates the downstream and upstream effects of metadata changes before implementation, providing complete visibility into potential impacts on standards and datasets.

Change Control – Enables users to log and manage change requests through a defined workflow that tracks approval, implementation, and closure, maintaining an end-to-end audit trail.

Versioning and Traceability – Maintains historical versions of all metadata objects with complete lineage information for transparency and regulatory compliance.

Reporting and Dashboards – Offers intuitive dashboards and customizable reports to monitor key performance indicators, track updates, and maintain visibility across metadata assets.

Export Functionality – Allows exporting of metadata and mappings for integration with other systems or for documentation purposes.

User Management and Access Control – Supports configurable user roles and permissions aligned with organizational and regulatory requirements.

Regulatory Compliance – Fully compliant with 21 CFR Part 11 standards, ensuring secure, auditable, and validated metadata management.

AUTOMATING DATA TRANSFORMATION

By combining Machine Learning, and Human-in-the-Loop (HIL) review, our solution automates and optimizes data transformation to produce standards-compliant datasets (such as CDISC-SDTM or ADaM) with greater accuracy and efficiency. The process includes the following key steps:

Mapping Automation:

Modern ML and pattern recognition algorithms enable automated variable mapping, significantly reducing manual intervention.

Model Training: The ML model is trained using existing standards-compliant datasets (e.g., CDISC-SDTM or ADaM) and their metadata. Through training, the model learns relationships between raw clinical trial data and their corresponding standard domains, variables, and values—allowing it to predict how raw data should align with compliant structures.

Input Data Processing: The trained model accepts raw data in various formats, such as spreadsheets, relational databases, or other structured data sources.

Automated Mapping Output: Using learned patterns, the model autonomously maps raw variables to their relevant standard domains and variables (such as SDTM or ADaM), generating a complete standards-compliant dataset. This automation minimizes manual effort and reduces human error, making it particularly effective for large or frequently updated studies.

Although automation greatly enhances efficiency, expert validation remains critical. The Human-in-the-Loop (HIL) process ensures that domain experts review and confirm the mappings, fine-tuning results as needed based on therapeutic and study-specific requirements.

DATA VALIDATION AND STANDARDIZATION

Ensuring accuracy, consistency, and compliance is central to generating standards-compliant data. AI-driven automation enhances these steps by identifying and resolving data quality issues efficiently.

Error Identification: Machine learning algorithms analyze SDTM datasets to detect missing values, incorrect formats, and other inconsistencies quickly.

Inconsistency Resolution: The model can identify contradictory data entries across domains (e.g., discrepancies in death information across AE, DM, and DD domains), flagging them for review to maintain data integrity.

Missing Data Handling: By analyzing data patterns, the model detects missing values and supports proactive query generation—helping data managers and programmers accelerate the data cleaning process.

Data Standardization: AI algorithms align non-standard data representations with clinical data standards, ensuring structural and semantic consistency. This includes mapping, conversion, and harmonization of variables and values to meet CDISC requirements.

AUTOMATING MOCK SHELLS GENERATION

This sub-module enables automated generation of mock shells by intelligently extracting and interpreting information directly from the Protocol and Statistical Analysis Plan (SAP) documents. Traditionally, mock shells are created manually, requiring extensive review of these documents to identify endpoints, variables, and layouts for each output. This automation module streamlines that process by leveraging AI and machine learning algorithms to read, understand, and translate the specifications into ready-to-use mock shells.

The workflow for mock shell generation involves the following key steps:

Step 1: Select TLF Type

Users can choose from a wide range of predefined TLF categories, such as Summary Tables, Demographics Tables, Vital Signs Tables, Adverse Event Listings, Laboratory Listings, Box Plot Figures, and many others. This selection guides the system in identifying the relevant data points and layout structures required for that specific type of output.

Step 2: Upload Protocol and SAP Documents

The user uploads or selects the corresponding Protocol and SAP documents from the integrated repository. These documents serve as the primary source of truth for study objectives, analysis requirements, and output definitions.

Step 3: Automated AI/ML Extraction and Shell Design

The TLF module then applies a smart AI/ML-driven algorithm to parse the Protocol and SAP, identifying key sections, tables, endpoints, and variable mappings relevant to the chosen TLF type. The system automatically extracts this contextual information and generates corresponding mock shells, including layout, titles, footnotes, and placeholders. This process replaces the manual effort traditionally required to design shells, saving significant time and reducing the risk of human oversight.

Additionally, the tool includes a Mock Shell Repository, which securely stores all generated mock shells along with version control and audit tracking features. This ensures full traceability of updates, facilitates collaboration among statisticians and medical writers, and maintains a consistent historical record of all changes.

By automating mock shell creation, this module drastically reduces manual workload, enhances standardization, and accelerates the preparation of analysis-ready templates—laying the groundwork for efficient and compliant TLF generation in clinical trials.

AUTOMATING TABLE, LISTING AND FIGURES

Creating TLFs is traditionally a labor-intensive and iterative process that demands continuous quality checks to verify the accuracy and completeness of information. Throughout this process, medical writers and statisticians often request modifications to the outputs, requiring programmers to rewrite or adjust the underlying code that generates the TLFs. Each change triggers new rounds of validation and quality control, extending timelines and increasing the risk of inconsistencies. These repetitive cycles contribute to significant delays in finalizing the Clinical Study Report (CSR) for regulatory submission, which can ultimately slow the delivery of new therapies and life-saving treatments to patients.

To overcome these challenges, there is a clear and growing need for automation in the creation and management of TLFs within clinical trials. Automation introduces efficiency, consistency, and accuracy throughout the TLF development lifecycle—reducing manual effort while ensuring compliance with regulatory standards. Below are key benefits that highlight the importance of adopting automation for TLF generation:

Efficiency and Time Savings: Automated processes dramatically reduce the time and effort required to create, modify, and validate TLFs. This is particularly beneficial when dealing with large datasets, frequent SAP updates, or multiple study deliverables. Automation accelerates the end-to-end process, enabling faster turnaround from data availability to final report generation.

Consistency and Standardization: Automation ensures that TLFs adhere to predefined templates, metadata structures, and formatting standards. Consistency across mock shells, outputs, and supporting documentation not

only enhances clarity but also simplifies regulatory review by maintaining uniform presentation across all study outputs.

Improved Accuracy: By minimizing manual intervention, automation reduces the likelihood of human errors in data processing, calculations, and code updates. This is critical in clinical trials, where even small inaccuracies can affect interpretation and regulatory outcomes. Automated generation ensures precision and reproducibility across multiple outputs.

Enhanced Data Integrity: Automation helps maintain accuracy and consistency of data throughout the TLF lifecycle. Built-in validation checks, data lineage tracking, and automated audit trails ensure transparency and reliability of outputs—key aspects of regulatory compliance and clinical data integrity.

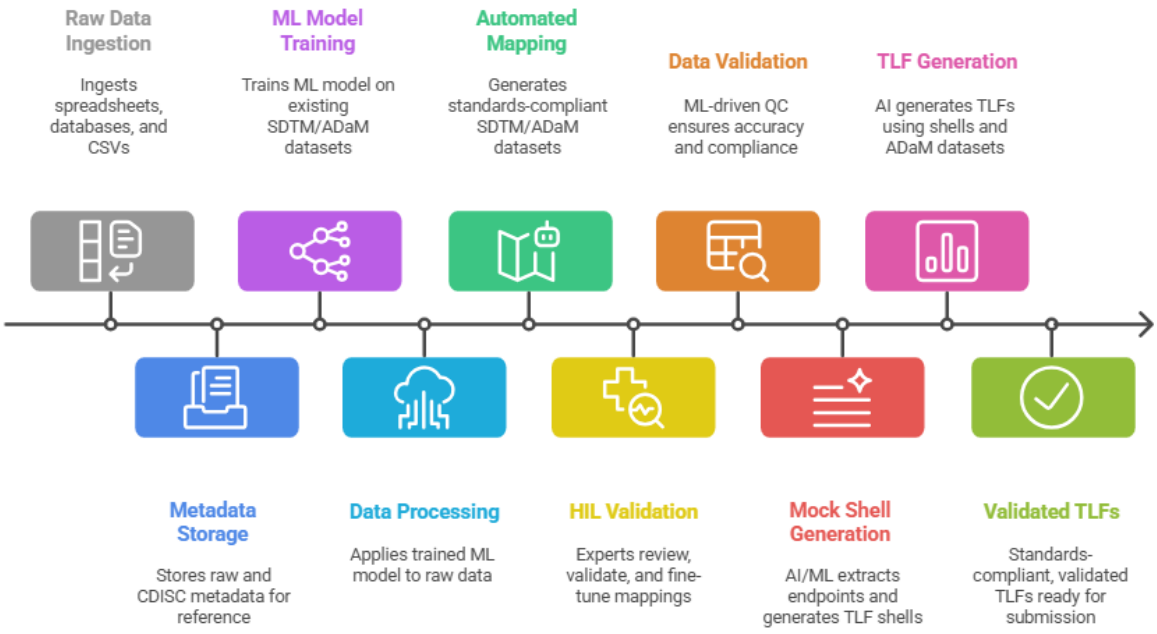
Reduced Workload and Optimized Resources: By automating repetitive and time-consuming tasks such as mock shell generation, layout alignment, data extraction, and display programming, teams can focus their efforts on more strategic tasks—such as interpretation, clinical insights, and decision-making. This leads to optimized resource allocation and improves overall productivity.

Regulatory Compliance and Traceability: Automated systems maintain detailed logs and audit trails, ensuring every step in the TLF generation process is traceable and compliant with regulatory requirements (e.g., FDA, EMA, PMDA). Standardized and automated documentation strengthens submission readiness.

Improved Collaboration and Transparency: With automated workflows, TLFs become more accessible and standardized across teams. Researchers, biostatisticians, and medical writers can collaborate more effectively, review outputs in real-time, and make informed decisions with greater efficiency and alignment.

To meet these evolving needs, our internal solution leverages machine learning and automation technologies to streamline the entire TLF generation process—from mock shell creation to final output production. The system intelligently processes multiple data layers, applies predefined templates and rules, and generates TLFs that meet clinical and regulatory expectations with minimal manual intervention. This approach not only ensures accuracy and consistency but also significantly shortens timelines, helping organizations accelerate the delivery of high-quality CSRs and ultimately bring new treatments to patients faster.

Figure 2: Automated Biometrics Workflow from Metadata to TFLs



USE OF MEANINGFUL SYNTHETIC DATA

To further enhance and accelerate the generation of TLFs from metadata, we utilize meaningful synthetic data to design and validate a robust end-to-end automation pipeline. This approach allows us to simulate real study conditions and test the entire workflow — from data transformation to TLF creation — well before the actual clinical data becomes available. By doing so, we ensure that all mappings, derivations, and processing logic are pre-validated and fully optimized.

Once the real study data arrives, the established pipeline can be executed seamlessly and without any modifications to mappings or processes. This proactive setup not only shortens development timelines but also ensures a smoother transition from test to production environments. As a result, the team can achieve faster turnaround times, improved consistency, and a higher degree of confidence in the quality and compliance of the final TLF outputs.

CONCLUSION

The evolution of clinical trial data processing—from raw data collection to the generation of TLFs—has reached a pivotal moment. Traditional, manual, and rule-based methods, while foundational in ensuring compliance and accuracy, no longer meet the growing demands for speed, scalability, and standardization in today's data-driven research environment. The integration of automation, artificial intelligence (AI), and metadata-driven frameworks represents a transformative shift toward a more efficient, transparent, and intelligent biometrics ecosystem.

By leveraging advanced machine learning algorithms, metadata repositories, and the use of meaningful synthetic data, organizations can now establish fully automated, end-to-end pipelines that span from raw data transformation to TLF generation. This holistic automation not only eliminates redundant manual tasks but also enhances data quality, consistency, and regulatory compliance. Furthermore, features like AI-driven mock shell generation, real-time validation, and audit-ready traceability empower teams to collaborate more effectively, reduce turnaround times, and maintain full control over the data lifecycle.

Ultimately, automation in clinical data workflows is not just an operational enhancement—it is a strategic enabler of innovation. It allows researchers and statisticians to focus on scientific insights rather than repetitive programming, accelerates clinical reporting timelines, and supports faster delivery of critical therapies to patients. As the industry continues to evolve, adopting intelligent automation solutions will be key to achieving sustainable efficiency, regulatory excellence, and a future-ready clinical trial ecosystem.

REFERENCES

GenInvo. (2025). Let's Talk Future: Automated Solution for Standard-Compliant Data, Mock Shells and TFL Generation Using Open-Source Platforms with AI
https://phuse.s3.eu-central-1.amazonaws.com/Archive/2025/Connect/US/Orlando/PAP_SM10.pdf)

CDISC. (2023). Study Data Tabulation Model (SDTM) Implementation
https://www.cdisc.org/standards/foundational/sdtm

CDISC. (2023). Analysis Data Model (ADaM) Implementation Guide
https://www.cdisc.org/standards/foundational/adam

ACKNOWLEDGMENTS

We would like to extend our heartfelt gratitude to all individuals that contributed to the completion of this paper. We appreciate our colleagues for their valuable feedback and support.

RECOMMENDED READING

GenInvo. (2023). Importance of "Table, Listing and Figures" Automation in Clinical Trials
https://geninvo.com/importance-of-table-listing-and-figures-automation-in-clinical-trials/

GenInvo. (2023). CDISC Standards and Data Transformation in Clinical Trial
https://geninvo.com/cdisc-standards-and-data-transformation-in-clinical-trial/

GenInvo. (2025). Beyond the Buzzwords: How GenInvo's AI/ML Tools Are Transforming Life Sciences
https://geninvo.com/ai-ml-tools-life-sciences-geninvo/

GenInvo. (2025). How Meaningful Synthetic Data Generation Tools Are Transforming AI Development
https://geninvo.com/how-meaningful-synthetic-data-generation-tools-are-transforming-ai-development/

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Inderjeet Arora
GenInvo
1408 East Empire Street 61701
Bloomington (Illinois), United States
+1 309-807-5006

inderjeet.a222@geninvo.com