Paper TT10

From Static Archives to Analytical Intelligence: Transforming Statistical Analysis Value Through Standardization

Ilan Carmeli, Beaconcure, Boston, MA, USA
Achinoam Ravet Perel, Beaconcure, Boston, MA, USA

ABSTRACT

Current models driving the generation, delivery, and reuse of statistical analyses are ill-equipped to maximize the full strategic value of the insights generated in tables, listings, and figures (TLFs). Static displays lock value in the highly variable, non-standard structures of PDFs, RTFs, HTMLs, and other proprietary output formats, when a more structured storage approach would yield far greater returns on investment. Statistical output rendered into standardized, machine-readable formats allows for applications to more fully exploit the value generated in the statistical analysis phase of clinical trials. This paper explores practical implementations of AI-driven table analysis, standardized metadata frameworks, and automated generation capabilities that demonstrate measurable improvements in efficiency, traceability, and analytical reusability. Each of these applications promises reduction in study startup time, greater efficiency through automation and standardization, and resource optimization through Write Once, Read Many methodologies.

INTRODUCTION

The drug development industry stands at an inflection point. For decades, 'standard of care' models driving the generation, validation, delivery, and archiving of statistical analyses have become antiquated and increasingly ill-equipped to unlock the full potential value inherent in tables, listings, and figures (TLFs). Single-use static displays that freeze valuable analytical insights within highly variable, non-standard file structures represent missed opportunities for maximizing return on the significant investment made in conducting the analyses. This challenge extends beyond output formatting concerns. When study artifacts and statistical output are stored in fixed and disparate formats, it limits an organization's ability to drive automation, enhance traceability, and extract value in multiple ways. By contrast, results data that are structured (assisted by human expertise) in a consistent, meaningful way facilitate AI-enabled tools in statistical analysis functions such as validation, meta-analysis, traceability, and even TLF generation. Flexibility flows from this structure.

Better outputs begin with better inputs. The journey to high-quality standardized outputs begins with equally high-quality standardized foundational documents. Traditional approaches to protocol development, statistical analysis plan (SAP) creation, and TLF specifications often result in documents that vary significantly in structure, nomenclature, and level of detail. This variability creates downstream challenges when attempting to link study objectives, analytical approaches, and output specifications.

Building more standardized precursor study artifacts begins with consistent vocabulary across interrelated file types, standardized templates for analysis specifications, and machine-readable document structures that facilitate automated pattern discovery, analysis, and presentation. When protocols clearly define analysis populations using consistent terminology, SAPs specify analytical methods with standardized language, and TLF specifications utilize uniform formatting approaches, the resulting foundational bedrock becomes far more useful to most AI tools. The more standardized and specific the inputs, the higher quality the output.

Currently, analysis results are aggregated as a combination of static Word, Excel, HTML, RTF or PDF formatted table packages. Display structures for commonly presented summaries have an identifiable form but vary by sponsor. Similarly, in many cases, multiple iterations of the same statistical analyses are generated for functional units on the study team that each use the statistics for slightly different purposes. Each single output requires production, validation, and review.

Figure 1: For each ADaM dataset and for each team's purpose, statistics must be generated, validated, and reviewed multiple times

Beaconcure's Verify platform was developed to offer users a powerful toolkit that parses, traces, and transforms TLFs and other study artifacts into a multi-use, linked, and standardized results database. From this structured database, it becomes possible to manage study progress to completion, conduct validation, extract specific results data for bespoke reporting by other groups, and simplify future use and reuse cases such as meta-analysis and TLF generation.

FROM DOCUMENTS TO DATA: OUTPUT DIGITIZATION AND METADATA EXTRACTION

The volume of data collected in clinical trials is a natural target for taking advantage of AI-enabled computing power. In harnessing this technology, high volume, repetitive tasks can be done quickly, efficiently, and are infinitely repeatable to a high degree of accuracy. Machine learning algorithms become better over time, and are not limited by the parameterization that current macro-enabled validation employs. As a table set grows through the duration of the study, the breadth of the display types also grows. With this, an AI model is able to ‘understand’ more of the connections that link outputs.

When tables are uploaded to Verify, the platform first parses, or digitizes, the tables to ‘harvest’ relevant metadata from the headers and footers, such as populations, datasets, treatment information, datetime stamps, etc. This metadata is attached to the cells in the body of the table, and is compared to referents such as a table of contents or a mock shell, or is used to search for like entities in the digitized representation of other tables in the set. Figure 2 illustrates how the structure and content of each display contain contextual information that facilitates understanding.

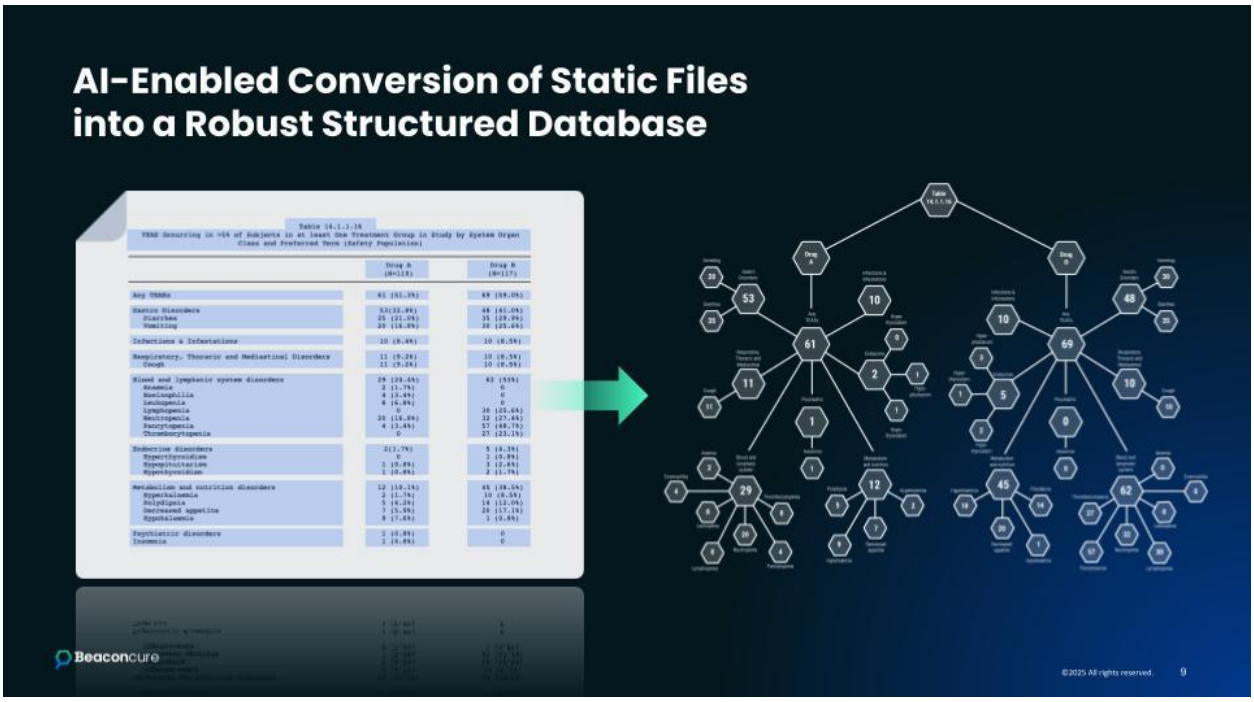

Figure 2: A table is parsed by Verify for inclusion in the AI-enabled digital database.

A combination of techniques including Named Entity Recognition (NER), Retrieval Augmented Generation (RAG), and rule-based methods, is used during parsing processes to extract metadata from TLF sets, specifications, and tables of contents, in order to create a robust metadata layer that can be used for both current and future study development. Programmers as human-in-the-loop (HITL) engage in fine-tuning metadata linkages in the Metadata Curator interface, reviewing discrepancies as they arise and refining outputs based on their expertise. This human-directed process ensures that all automated results align with analysis plans, CDISC standards, and good programming practice. By extracting and storing all related metadata in a digitized database, it can be used in different ways, and can also be exported to a digitized format.

PRACTICAL IMPLEMENTATIONS AND OUTCOMES

AI-DRIVEN TABLE ANALYSIS AND VALIDATION

Verify's parsing technology transforms the static output paradigm from traditional table-by-table validation to broad-spectrum systematic validation strategies. Rather than treating each TLF in a set as a unique entity requiring individual validation, Verify employs a validation method where common analytical patterns (format checks, within-table checks, across-table checks, across-study checks, et al) are applied globally to the TLF set, then released for reviewers to confirm. This approach significantly reduces manual validation time while maintaining high analytical quality benchmarks.

This automation reduces the time and effort required for manual review and cross-table visual checking, and enables consistent handling of similar table structures across different studies and sponsor organizations.

AI techniques such as natural language processing (NLP) are paired with HITL domain expertise and contextual understanding so that automated processes withstand the scrutiny of scientific rigor. AI, backed by HITL feedback, forms the foundation of this approach. Figure 3 below shows how Verify automated generation of a table for the purpose of validation, using extracted metadata and the corresponding ADaM dataset, with the AI model having the capability to choose the relevant ADaM variables and values. In doing so, traceability is provided through the presentation of the entities and ADaM variables used to generate the table. Reviewers or programmers can review the metadata alignment for accuracy, and either confirm Verify's AI-generated table or make entity alignment updates and regenerate the table accordingly for revalidation. Traceability is maintained within the Metadata Curator interface, and changes are recorded in the automated Activity Log.

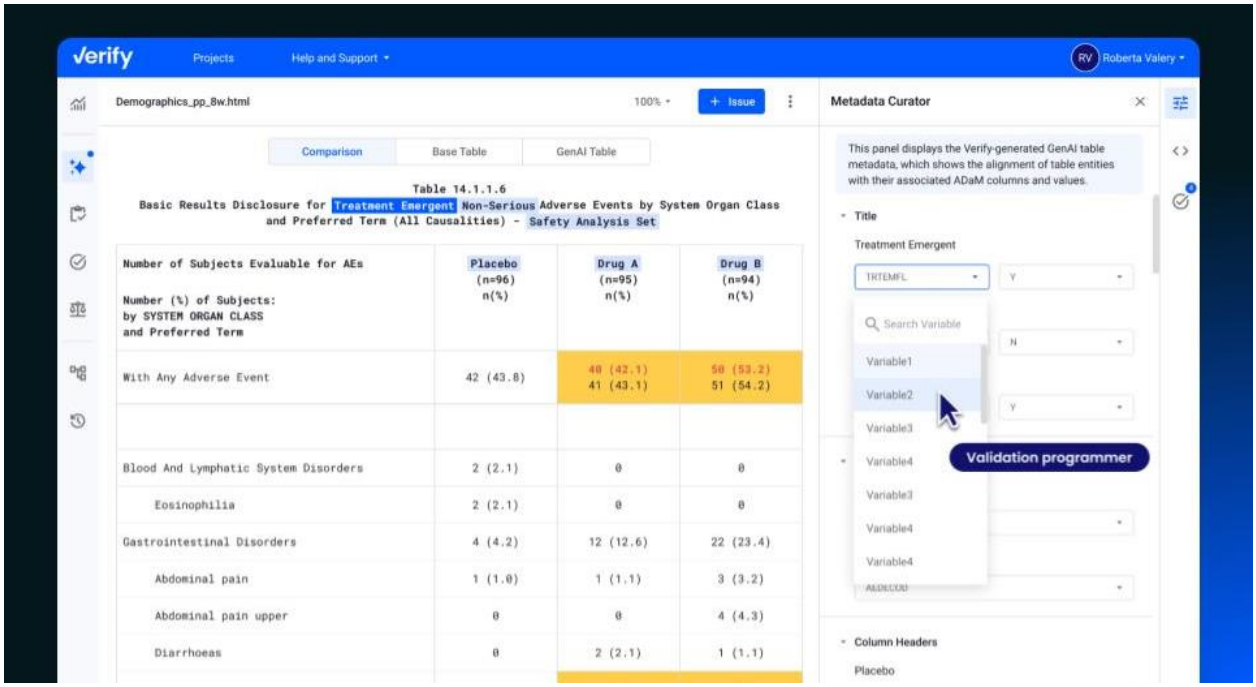

Figure 3: Verify parses outputs, identifies key metadata elements of a table, and represents outputs in standard format for review and study-specific modifications.

In a regulated industry, changing a time-tested paradigm requires developing trust in a new system equal to that of the existing process. An automated Activity Log that generates immutable records of all actions and decisions enables critical traceability and audit readiness. This capability addresses one of the most challenging aspects of traditional statistical programming: maintaining comprehensive documentation of analytical decisions and modifications throughout the development process.

STANDARDIZED METADATA FRAMEWORKS

Like SDTM and ADaM, CDISC's newly released Analysis Results Standard (ARS) provides users with a comprehensive framework that captures everything from metadata to statistical methods, to cell level results.

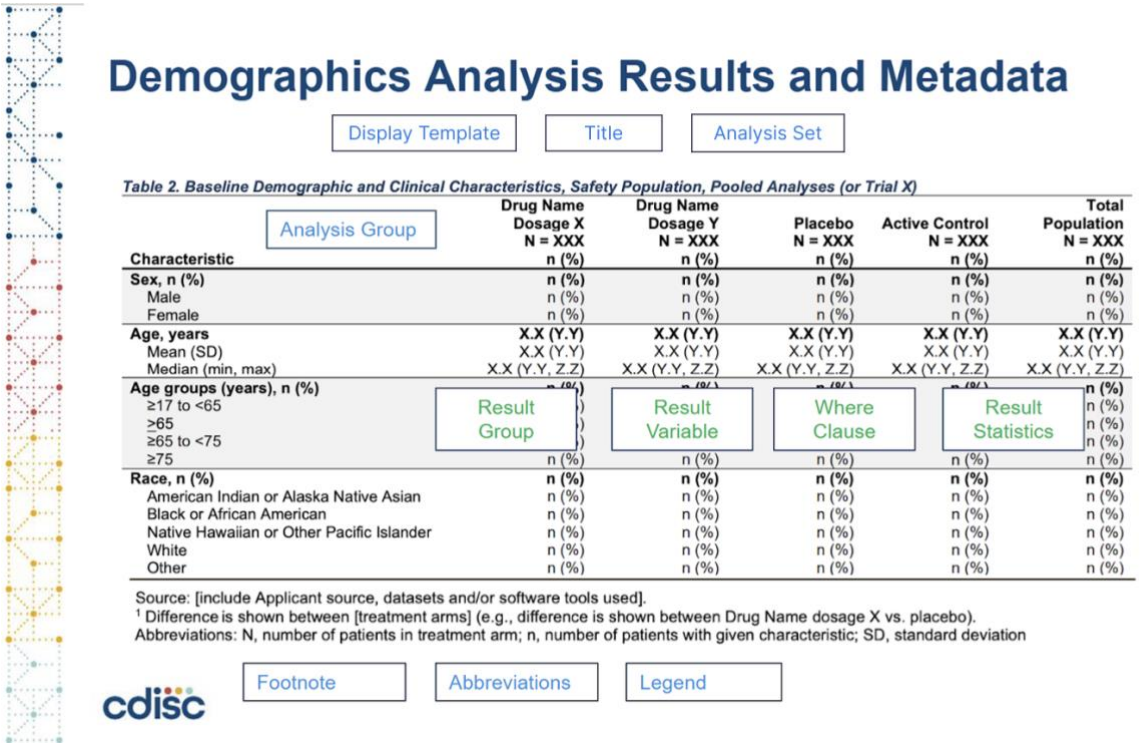

*Figure 4: CDISC ARS provides **technical specifications** for analysis results metadata to facilitate and automate TLF planning and production.[1]*

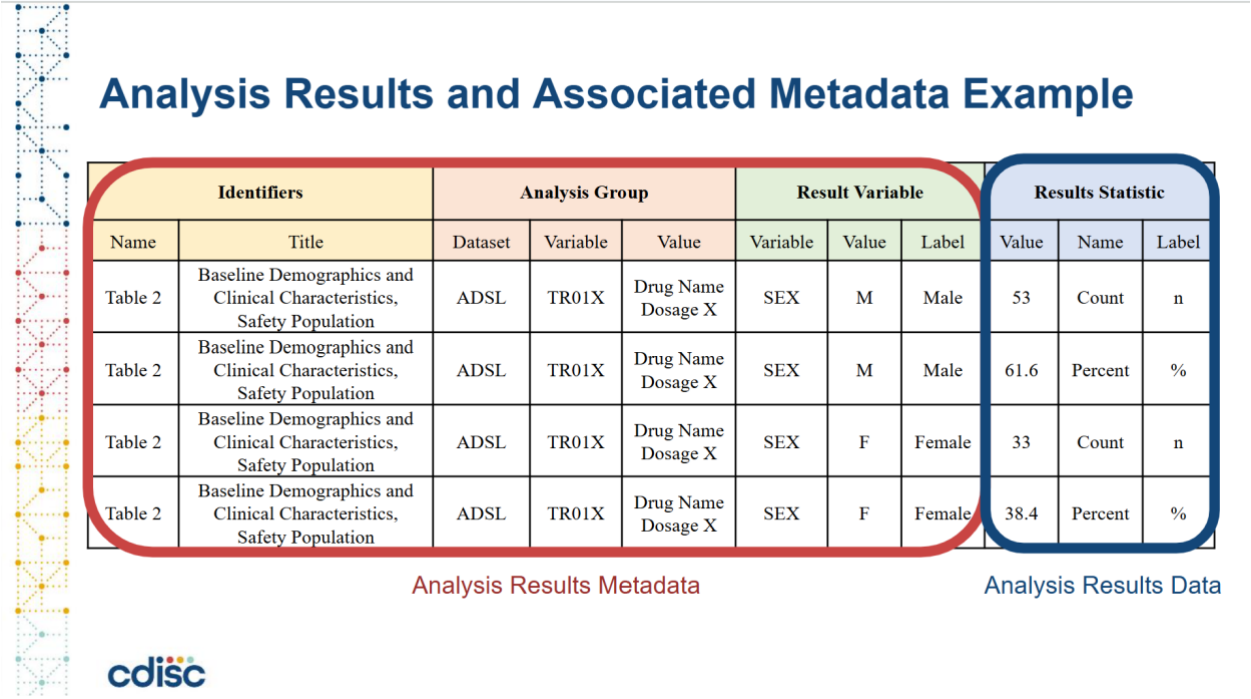

*Figure 5: CDISC ARS provides a **standard structure** to represent analysis results and qualifying metadata to support traceability, reproducibility, reusability and quality.[1]*

In a more flexible output framework, statistical analyses would be generated and validated once and stored in a structured results database, for each of the functional business units to extract as needed for their use. Results in the body of the table are stored with associated metadata such as population flags, data subsetting information, analysis algorithms, titles, and footnotes, to be used and reused for iterations of TLFs for the current deliverable, or for future work such as meta-analysis or TLF generation.

The ideal scenario is that a set of analyses is written once, validated once, then made available for multiple use cases – in programming parlance, the “WORM” principle: Write Once, Read Many. A single iteration of production and validation of statistical analysis, instead of three. Fewer programs to create, validate, and review translates to shorter development cycles, less complex deliverables, and greater confidence in the conclusions being drawn.

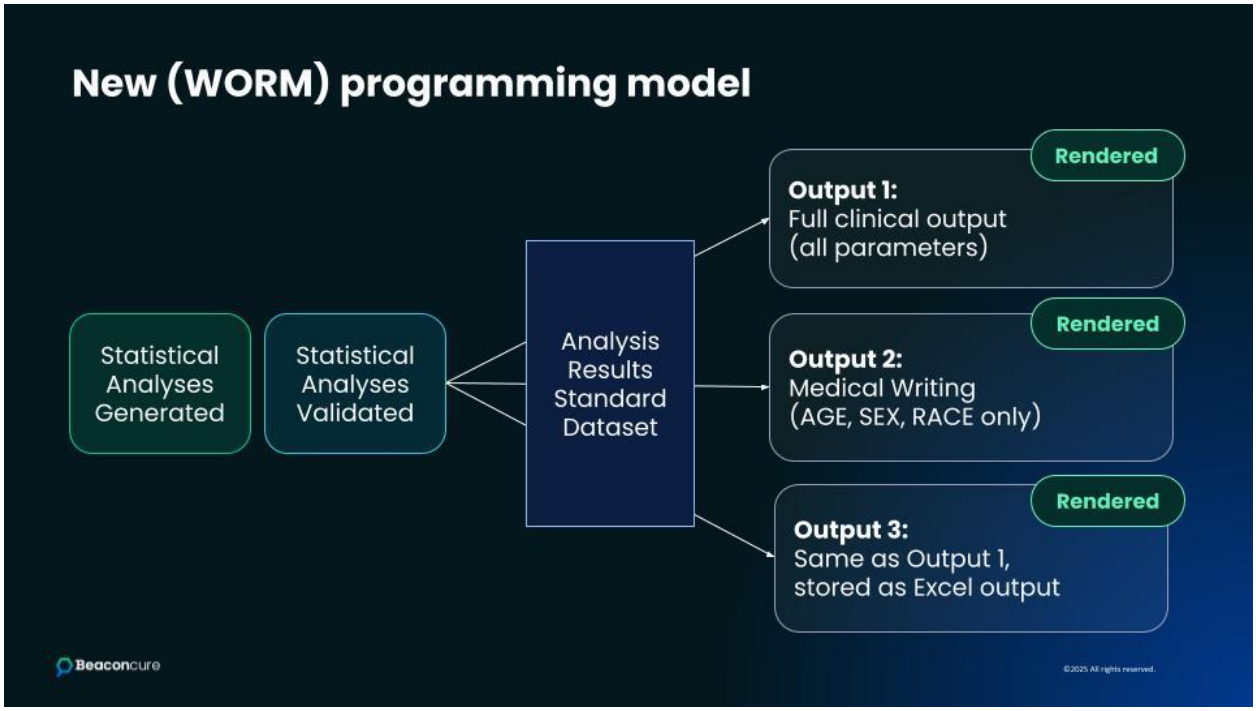

Figure 6: An ideal programming model follows the Write Once, Read Many (WORM) principle: a set of analyses is written once, validated once, then made available for multiple use cases.

ARS facilitates direct linkages back to precursor components of the clinical documentation (Statistical Analysis Plans, SDTM and ADaM dataset specifications) and creates an auditable trail that links analytical results to their foundational specifications and data sources. When Verify parses tables, the AI algorithm links entities from table titles and headers to the corresponding ADaM variables and values. This information is stored in a structured database containing a semantic understanding of the outputs. It then becomes simpler to render this resultant database into a standardized format for reuse. When ARS is paired with Verify's parsing and metadata repository, Verify has the potential to support the process of using the WORM principle to create standardized Analysis Results Standard Datasets.

AUTOMATED OUTPUT GENERATION

By leveraging standardized ADaM datasets, AI-driven software like Verify can move beyond static reporting toward analytical intelligence. Using extracted metadata, the platform generates statistical tables by dynamically identifying the correct variables, parameters, and values within ADaM. This can accelerate table creation while enabling alignment with CDISC standards and study specifications. Critically, the AI further validates outputs generated by production programmers by cross-checking variable selection and value mappings, detecting inconsistencies that may compromise data integrity. Through this combination of automation, standardization, and validation, organizations can transform statistical outputs from static archives into living analytical assets that deliver continuous value across studies and therapeutic areas. As Verify 'sees' more and more output (and by extension, a reviewer confirms its decision-making), Verify's accuracy increases.

Ideally, structured results and metadata also create another opportunity: the use of existing (previously parsed) data as an aid in validation and in the prospective mapping and/or generation of statistical renderings. Once a table is parsed and mapped, table entities are reviewable by team members as SMEs, and can even be updated when necessary. The audit-ready log also captures changes in mappings, discussions between team members, and discussion points to be noted in regulatory filings.

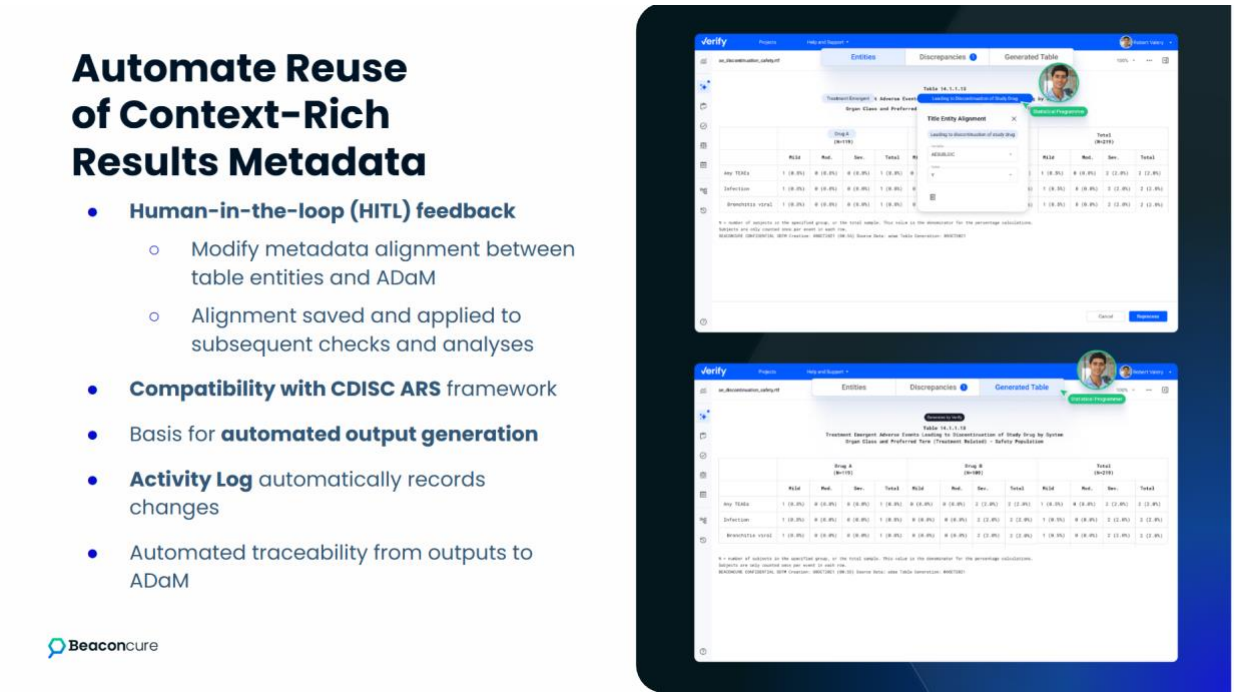

Figure 7: Verify transforms statistical analysis value through multiple standardization paths, enabling automated validation, metadata alignment traceability and CDISC ARS compatibility.

Leveraging the parsing and metadata extraction capabilities of the Verify platform represents a fundamental shift away from programming toward the goal of a static output, to programming with the goal of multi-use statistical outputs in dataset form that facilitates multiple output vectors.

The capability to rapidly adapt to new data and display formats unlocks generalizability: as the platform encounters new table structures, it can quickly incorporate appropriate checking based on its accumulated knowledge base. Results metadata can then be managed and reused in future deliverables and cross-study analyses, creating a growing repository of validated statistical results.

CONCLUSION

Real world implementations of analysis results metadata (ARM) suggest measurable benefits across the various functions involved in clinical development.

The transformation of outputs from static archives to flexible and multiple-use analytical intelligence represents a transformational shift in clinical statistical analysis and reporting. Through AI-driven table recognition, creation of standardized metadata frameworks, and systematic reuse capabilities, organizations can unlock substantial additional value in their statistical work products, while at the same time accelerating development timelines and reducing costs.

The technical frameworks supporting this transformation (document parsing, entity extraction, relationship mapping) are emerging through initiatives like CDISC ARS and platforms like Verify, with organizations beginning to implement solutions that demonstrate tangible benefits and intuitive tools for reviewing and mapping ARM. The benefits extend beyond incremental efficiency gains to include enhancements in flexibility in how analysis results are consumed, quality improvements through standardization, and added strategic value through reuse.

As the pharmaceutical industry faces increasing pressure to accelerate development while still maintaining the highest standards of quality and compliance, systematic implementation and use of analysis results metadata represents not just an opportunity but an imperative for an organization's sustained competitive advantage. The journey is complex and challenging, but tangible implementations demonstrate that the potential rewards justify the investment.

REFERENCES

1. Bhavin Busa, Bess LeRoy, and Richard Marshall., PharmaSUG US, MM-327: All You Need to Know about the New CDISC Analysis Results Standards!, 2023.

ACKNOWLEDGMENTS

The authors would like to thank Steve Ross for his indispensable contributions in the development and production of this paper.

CONTACT INFORMATION (HEADER 1)

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

Author Names: Ilan Carmeli and Achinoam Ravet Perel
Company: Beaconcure
Address: 867 Boylston Street, 5th Floor #1491, Boston, MA
Emails: Ilan@beaconcure.com; Achinoam@beaconcure.com
Website: www.beaconcure.com
LinkedIn: http://www.linkedin.com/company/beaconcure