Paper ET03

OPS Builder: Navigating the Cosmos of Study Deliverables with Stellar Automation

Suhas Kirani Ravindra, GSK, Bengaluru, India

Disclaimer: The opinions expressed in this paper are solely those of the author and not necessarily those of GSK.

ABSTRACT

The Output & Programming Specifications (OPS) Builder, a groundbreaking Automated Clinical Data Transformation enabler product, transforms how study teams plan, execute, and manage deliverables. By functioning as an orchestrator, it mainly drives the end-to-beginning approach, OPS Builder automates the curation of essential information for study analysis and reporting processes, thus streamlining and accelerating pipeline delivery. Our long-term strategy is to provide a comprehensive digital application that simplifies the selection and definition of analysis deliverables, including TFL (Tables, Figures, and Listings) digital mock-shells, ADaM Specifications, and derivation registrations.

Key benefits of OPS Builder include enhanced accessibility of OPS content for downstream systems, facilitating metadata-driven automation that significantly reduces human intervention. Transitioning from 100% to 50% human input in development and maintenance, OPS Builder leverages machine learning to prescribe TFL-ToC based on protocol, SAP, standards, and therapy area knowledge. This approach not only accelerates study deliverables but also increases the adoption of standards, providing analysis recommendations powered by historical data and logical connections to SAP or protocol.

By empowering Clinical Programmers and Statisticians of varied experience levels to contribute more effectively which is designed to think end-to-beginning approach, OPS Builder elevates study deliverable management, ensuring rapid and precise execution aligned with regulatory guidelines and stakeholder inputs.

This paper focuses on what GSK has achieved till now in this journey of automating clinical data transformation.

INTRODUCTION

The biopharmaceutical industry is under constant pressure to accelerate the delivery of new medicines to patients. A significant bottleneck in this process lies in the manual transformation of clinical data into submission-ready analyses. These traditional workflows are not only resource-intensive but also prone to errors, limiting the potential for rapid, high-quality outputs. In an increasingly competitive environment, automation has become a critical theme for staying ahead.

At GSK, we are transforming our approach by creating a more efficient and interconnected programming ecosystem. This ecosystem links various data standards, enables automated validation, and accelerates the journey from data collection to submission. A cornerstone of this transformation is the OPS Builder.

THE VISION: A DIGITAL APPLICATION FOR STUDY ANALYSIS AND REPORTING DELIVERABLES

Our vision is to provide a comprehensive digital application that simplifies the selection and definition of analysis deliverables for clinical studies. This includes the creation of digital mock-shells for Tables, Figures, and Listings (TFLs), defining Analysis Data Model (ADaM) specifications, and registering derivations and definitions. The ultimate goal is to move from a state of 100% human intervention in development and maintenance to 50% where machines prescribe the TFLs and associated metadata based on a variety of inputs.

KEY CONCEPTS OF AUTOMATION

To achieve this vision, our strategy is built upon four key pillars:

- **Standardization**: The availability of core end-to-end analysis standards is fundamental.
- **Linked Metadata**: Utilizing linked metadata is crucial to fill the gaps between SDTM, ADaM, and Analysis Results.
- **Modularization**: We are creating modularized reporting tools that leverage open-source technologies.
- **Automated Testing**: Automated processes are in place to identify and mitigate data variability.

OPS BUILDER: THE ORCHESTRATOR OF AUTOMATION

The OPS Builder is a revolutionary product that functions as an orchestrator for automation. It meticulously curates all necessary information for study analysis and reporting. This allows for a shift from manual, time-consuming processes to a streamlined, automated workflow.

KEY BENEFITS:

- **Automation:** OPS content is made easily accessible in a digital format for downstream systems, enabling metadata-driven automation.
- **Acceleration:** The OPS Builder facilitates a hyper-acceleration of pipeline delivery, optimizing both the speed and reliability of study execution.

By leveraging the OPS Builder, study teams can achieve unprecedented levels of efficiency and adaptability. The development time for OPS has potential reduction from approximately two weeks of manual, repetitive work to about three days of parallel, GxP compliant processing.

ARCHITECTURE AND WORKFLOW

The OPS Builder is a central component in a larger automated clinical data analysis and reporting framework. It integrates with various modules, including:

- ADaM Spec Designer and Planner: For creating and managing ADaM specifications.
- TFL Planner and Mock-shell Editor: For planning and designing TFLs.
- OPS Co-Pilot: An AI powered component for recommending analyses.

This integrated system utilizes fully machine-readable, linked metadata to enable a "write once, render multiple times" approach.

CHALLENGES IN THE PATH TO AUTOMATION

Transitioning from a deeply entrenched manual process to a metadata-driven automated ecosystem presents a series of significant challenges. These hurdles span technology, process, and people, and addressing them is critical to the success of the OPS Builder and the wider automation initiative.

1. **Deconstructing Traditional Outputs into Machine-Readable Metadata**

   A primary challenge lies in translating traditional, document-based artifacts into fully machine-readable metadata. Historically, a mock-shell for a table or listing existed as a static representation in a document, with critical details like titles, footnotes, and variable derivations embedded in a format intended for human eyes. The process relied on programmers and statisticians to act as a bridge, manually "connecting the dots" between the protocol, the SAP, and the final output.

   The key challenge was to deconstruct this visual mock-shell into discrete, structured metadata elements, such as Analysis Results Metadata (ARM) and Analysis Output Metadata (AOM). This required a fundamental shift from viewing a mock-shell as a picture to seeing it as a collection of queryable, linkable metadata components that can be versioned, reused, and understood by a machine.

2. **Ensuring a High-Quality User Experience**

   For any new system to be adopted, it must be intuitive and efficient for the end-users. A significant challenge noted during the development of the OPS Builder was centered on the user experience. Statisticians and programmers are accustomed to document-centric workflows. Introducing a suite of new digital tools like the TFL Mock-shell Editor and the ADaM Spec Designer requires not only technical proficiency but also a change in mindset. The system must be designed to guide the user through the new process seamlessly, reducing the cognitive load and demonstrating clear value over the manual methods it replaces.

3. **Achieving Reliability in AI-Driven Recommendations**

   The integration of Artificial Intelligence and Machine Learning (AI/ML) to recommend analyses introduces its own set of complex challenges. While initial experiments with TFL recommendations yielded promising results, a key finding was that the accuracy of these recommendations varied significantly across different clinical domains.

   This inconsistency poses a major hurdle, as trust is paramount for adoption. An unreliable recommendation engine would fail to provide the intended efficiency gains and could even introduce errors. Consequently, the implementation of this feature was paused pending further refinement. The ongoing challenge is to develop robust, metadata-driven automated workflows that can be used to retrain and optimize the model, ensuring it becomes a reliable and trusted "Co-Pilot" for study teams.

4. **Overcoming the Inertia of Existing Processes**

   The traditional approach, while inefficient, is familiar and well-established. The move from a manual, sequential process to a GxP-compliant, parallel workflow requires significant organizational change management. It involves redefining roles, establishing new governance models for standards, and demonstrating tangible benefits—such as reducing development time from weeks to days—to overcome the natural resistance to change and build momentum for the new ecosystem.

CURRENT STATUS AND FUTURE DIRECTION

We are currently in the process of transforming our ecosystem. While we are not yet at the end of our journey, we have made significant strides. The OPS Builder is a testament to our ambition to challenge the status quo and leverage technology to accelerate the delivery of new medicines. We are actively monitoring initiatives like digital Protocol and digital SAP to inform our next steps and continue to refine our automated workflows.

A CALL TO ACTION: RECOMMENDATIONS FOR A MODERNIZED INDUSTRY

The journey to create the OPS Builder has provided us with critical insights that extend beyond the walls of GSK. The challenges we face and the solutions we are building are not unique; they represent a necessary evolution for the entire biopharmaceutical industry. To collectively accelerate the delivery of drugs to the patients we believe in,

1. **Embrace a "Metadata-First" Mindset**

   For decades, our industry has operated on document-centric workflows, where protocols and Statistical Analysis Plans (SAPs) are created as prose in static documents. This is the root cause of many downstream inefficiencies. We call upon the industry to shift from a document-first to a **metadata-first** mindset.

Recommendation: Let us move towards authoring key study documents like protocols and SAPs as structured, machine-readable data from their inception. Initiatives like electronic Protocol (eProtocol) and eSAP are steps in the right direction, but this philosophy must be adopted universally. Instead of creating a digital picture of a table (a mock-shell), we should be defining its components—titles, footnotes, data sources, and derivations—as discrete metadata elements like ARM and AOM. This is the foundational layer upon which true automation can be built.

2. **Adopt a Pragmatic and Foundational Approach to AI**

The promise of Artificial Intelligence (AI) and Large Language Models (LLMs) is immense, but their application must be grounded in reality. Our experience with TFL recommendations showed that AI is only as good as the data it is trained on. Applying AI to interpret unstructured, inconsistent documents will inevitably lead to variable accuracy and hallucinations, eroding trust.

Recommendation: The industry should prioritize building a robust, linked metadata foundation *before* attempting to deploy AI at scale for analysis planning. AI should be the consumer of clean, structured, and standardized metadata, not the interpreter of messy documents to enable higher accuracy that is acceptable. Let's use AI as a "Co-Pilot" that leverages a trustworthy foundation to make reliable recommendations, rather than asking it to perform magic on unstructured text. This foundational approach is the surest way to reduce hallucination and build the trust necessary for widespread adoption.

3. **Invest in Integrated Digital Ecosystems, Not Siloed Tools**

The full power of automation is unlocked when metadata flows seamlessly from one step of the process to the next. A standalone mock-shell tool or a separate specification editor provides only incremental benefits. The real transformation comes from an integrated ecosystem where the output of one tool becomes the direct, machine-readable input for the next.

Recommendation: We encourage companies to think in terms of a connected **digital ecosystem** for clinical trial reporting. This involves creating (or adopting) tools that are natively designed to communicate with each other through linked metadata. The goal is to create a digital thread that connects the protocol to the final analysis result, enabling an "execute" button that can run the entire workflow, from ADaM creation to final TFL generation, based on a single source of truth.

By collectively embracing these principles, we can move beyond simply accelerating old processes and instead transform *how* we work, paving the way for a future of faster, more efficient, and higher-quality clinical data analysis and reporting for the benefit of all.

CONCLUSION

The OPS Builder represents a paradigm shift in how we approach clinical study analysis and reporting. By embracing automation, standardization, and the power of linked metadata, we are creating a more efficient, accurate, and accelerated pathway from data to submission. This not only eases the manual burden on our programmers and statisticians but also positions us to stay ahead in a competitive industry.

CONTACT INFORMATION

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

Author Name: Suhas Kirani Ravindra
Company: GSK
Email: suhas.k.r@gsk.com