PhUSE US Connect 2019

Paper DV06

Delivery of Data Visualizations within GlaxoSmithKline Clinical Development and Partners[1]

Jason W. Szewczyk, GlaxoSmithKline, Collegeville, PA

ABSTRACT

The successful Enterprise Level Delivery (ELD) of Data Visualizations is a complex endeavor, and for Clinical Development, requires the successful collaboration of: Centralized Data Visualization Operations, Clinical Programming, Data Management, Statistics, Safety, Medical Monitoring, and GSK's External Partners.

Recently, GSK's Data Visualization Operations (DVO) has been evolving a Strategic Initiative to broadly provide Data Visualization Tools across Clinical Development. DVO has developed a scalable, ELD model to support Safety, Risk Based Monitoring, and Clinical Teams. Using a single tool, Standardized, Program-Level Templates support both internal/external users, and provide the ability to seamlessly transition between summarized and detailed, subject-level anonymized data to also integrate multifunctional data groups.

Data Visualization provides rapid answers to key clinical questions with increased: clarity, efficiency, and cost/time savings. This ELD approach allows study-level innovations to escalate across the Clinical landscape, via Standardized Data Visualization Templates, to evolve organizational best practices for data interrogation.

INTRODUCTION

The implementation of Data Visualization (DV) in large organizations is a challenging endeavor. At this time, the benefits DV provides is recognized by both individuals and organizations, however the challenge is how to best deliver DV tools at an enterprise level in a scalable and reproducible way. The Visualization goal for most enterprise-level organizations is to have DV Tools broadly available covering key business practices, and team members using those tools routinely in support of day to day business activities. Currently, many organizations have not achieved this high level of integration and fall into two main categories. The first are companies that have invested and made DV Tools widely available, however for a variety of reasons (e.g. culture and/or training) the organization has not been able to achieve widespread use of these tools nor have they been successfully embedded in daily activities. For the second group, stakeholders are currently using some rudimentary tools and are actively seeking opportunities for further adoption, however organizational restrictions (e.g. funding, infrastructure or insufficient expertise) prevent widespread delivery and full adoption.

At GSK, we are on the journey to achieve the highest level of DV Tool integration. In the last 3 years, in GSK Clinical Development, there has been significant progress in the broad delivery DV Tools across all therapeutic areas, strong adoption by stakeholders, and increased demand/embedding of DV Tool utilization during study conduct. Herein, we describe the development of a scalable Enterprise Level Delivery (ELD) Model for DV Tools, methods to drive integration, and key lessons learned from this implementation.

DEVELOPMENT OF A SCALABLE ENTERPRISE LEVEL DELIVERY MODEL

In 2015, GSK made the strategic decision to centralize all DV Operations into a single organization to develop and deliver Clinical Data Visualizations (DV) and Consultancy services across R&D Clinical Development. This formed a global team of highly trained Data Visualization experts based in the both the US (Upper Providence) and UK (Stockley Park) with off-shore support from TCS India to meet all GSK business and stakeholder needs.

The first step for the development of any delivery model capable of supporting the complex needs of a functionally diverse department, like GSK Clinical Development, was to review previous DV efforts. Legacy approaches for Tool delivery were focused at the study level. These legacy approaches led to high initial adoption and singular, elegant DV Tool solutions, however issues with long term support, scalability, and reuse led to unacceptable levels of adoption and reuse.

The lessons learned from these efforts provided guidance on the shape, scope, and mandate of the Centralized Data Visualization Group. For a small, centralized group to successfully deliver DV Tools across Clinical Development, standardization of all aspects of DV Operations would be critical to ensure both initial delivery and long-term support/customizations. These best practices and lessons were codified into four key organizational principles and used to drive the development of the ELD Model for Data Visualization. Specifically, for successful implementation of the ELD Model it was necessary to: standardize the DV Programs supported, align these Programs to essential business practices, develop a Standard Template for each Program, standardize the data types supported, and develop a scalable process to support

delivery and customization of DV Tools (Table 1).

Table 1. Four Key Principles for a Successful DV Enterprise Level Delivery Model

1. Define DV Programs Supported and Alignment to Essential Business Operations
2. Standardize Collection of Data Visualizations (Standard Templates) for each Program
3. Standardize Data Types Supported
4. Develop and Standardize Scalable & Reproducible Process

1) Standardize DV Programs Supported

The first step to developing an ELD Model for DV was to clearly define the scope and number of functional DV Programs to be supported and ensure these Programs are aligned with essential business operations. Program alignment to specific Clinical activities ensures that programming effort is clearly directed to support critical operations and allows the key stakeholder groups to be clearly identified. Specifically, identification of the key Departmental Stakeholders helps in requirement gathering and ensures the mode of operation for the DV Tools within each Program is fit for purpose for each Department's needs.

For GSK Clinical Development, there are 3 DV Programs currently supporting key business activities, specifically: In-Stream Clinical Monitoring & Data Review (CSTV Template), Clinical Safety Review (SRT Template), and Risk Based Monitoring (RBM Template/Tool). While each program utilizes the same underlying datasets, the way the data are aggregated, and the outputs of the Visualizations are quite different in order to support these diverse functional Programs (Figure 1). For In-Stream Clinical Monitoring, a typical output is the swim lane plot shown below, CSTV: Treatment Duration (Figure 1, CSTV.) In this case, each individual patient or "swimmer" is represented as a horizontal line on the Y-axis and important patient specific events/milestones are arrayed chronologically across the X-axis. Similarly, for Safety Review, Visualizations are designed to assess the overall safety profile for all subjects and provide rapid identification of potential at risk patients for further investigation. For example, the SRT: Hy's Law plot in Figure 1 assesses the overall safety profile for all patients against two risk criteria. Potential, at risk patients are easily identified when meeting one of the two criteria, by falling into either the upper left or lower right quadrant. Patients at high risk, who satisfy both risk both risk criteria, are sorted into the upper right quadrant, and readily identified (shown as two orange dots). For RBM: Visits Entered Late, these outputs are not assessing individual patients, but rather Sites performance. In this case, each site is evaluated against the average performance of all sites within the trial. Sites are classified either red, amber, or green (RAG) by a set of proprietary statistical algorithms with the specific weightings and thresholds for each study set by the study team. Figure 1 provides an example sorting of Sites by RAG status. In this case, it is expected for Sites to converge over time, and Sites requiring further monitoring are identified as outliers as shown by the orange dot.

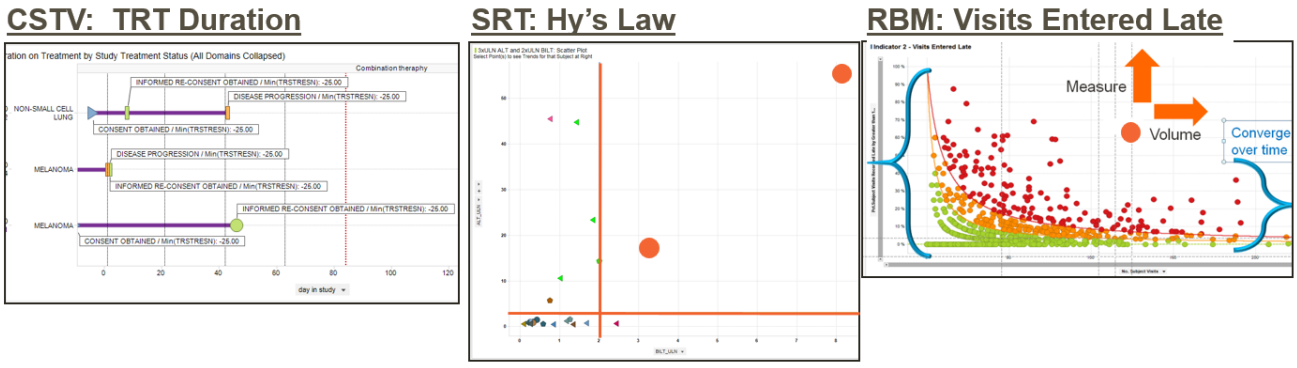

Figure 1.* Example Data Visualization outputs from the 3 main DV Programs, i.e. In-Stream Clinical Monitoring (CSTV Template), Safety Review (SRT Template) and Risk Based Monitoring (RBM Template Tool).

* The graphs rendered herein are not referenced to any specific GSK study or asset, contain no personally identifiable information, and are

shown for demonstrative purposes only.

2) Develop Standard Templates

The next step is developing a series of Standard Templates for each Program. Keeping the number of Standard Templates low is critical for delivery, since as the number of Templates increase so does the operational overhead. Specifically, regular updates and enhancements are an important part of the maintenance of the Standard Templates, thus minimizing the number of Templates allows rapid up-versioning and minimizes any potential impact to DV Tool delivery or operations.

To successfully achieve the smallest number of Standard Templates required to support each DV Program, Template development focuses on two key design criteria. The first is that all Standard Template Outputs are designed to transcend all Therapeutic Areas and Clinical Trial Phases. This requires a high level of collaboration across all stakeholder Departments and the centralized Data Visualization Group; i.e. Statistics, Clinical Programming, Data Management, Safety, and Clinical. Collation of all Stakeholder requirements ensures that the Templates provide the desired content and that the Analytic Outputs are of high utility. The second is that each Standard Template follows the 80:20 rule. This means that the content of each Standard Template on Setup should provide 80% of the Visualizations required for the life-time of that trial. The other 20% of Visualizations required are delivered via the Customization Request Process.

For each Template, the required content is captured and delivered across no more 25-30 pages. Requiring each template to adhere to a page range provides several key benefits, namely: managing QC overhead, assisting in facilitating user navigation, and minimizing the amount of training required for users operate the DV Tools.

3) Standardize Data Types Supported

Standardization of supported data types is an essential step to support the ELD model, as standardizing the number of data inputs allows a large number of studies to be supported using only a minimum number of standard templates. Currently, support is focused on 2 data types: System Independent (SI, GSK Proprietary IDSL) and SDTM (CDISC Standard). This allows programmers to utilize a single, standard process for data updates helping to ensure both rapid delivery and high quality. Importantly, data type standardization offers a number of advantages: investment in upfront programming and coding, rapid and efficient Setup timelines, high study to study reproducibility for speed and quality, and assurance of output validation. Notably, data standardization allows senior Data Visualization Analysts the opportunity to address most study-to-study variations/issues at the Template level, allowing junior Analysts to be able to successfully deliver Setups with minimal troubleshooting. Additionally, Setup timelines are minimized, since reoccurring study issues are now addressed at the Template level. Currently, only 6 Standard Templates are needed to support all Programs across Clinical Development, namely: 2 CSTV Templates for In-Stream Clinical Monitoring (SI/SDTM), 3 SRT Templates for Safety Review (SI Blinded/Unblinded, SDTM), and one for Risk Based Monitoring (RBM Tool).

4) Develop Scalable & Reproducible Process

To support ELD of Data Visualizations, a single, scalable process was developed to govern the entire DV Portfolio (Figure 2). This process is in two parts: Part A is focused on Setup of the Standard Template, while Part B is focused on Workflow Customization. The Setup and Customization Process steps are serial by design to ensure the fastest possible delivery of the existing DV Tools to the study teams. Rapid deployment of the Standard Tools releases some of the initial timeline pressure and allows teams space to collaborate and consider what customizations/enhancement would be the most impactful. Importantly, as the Figure 2 shows, GSK maintains full Oversight throughout the Setup and Customization Process.

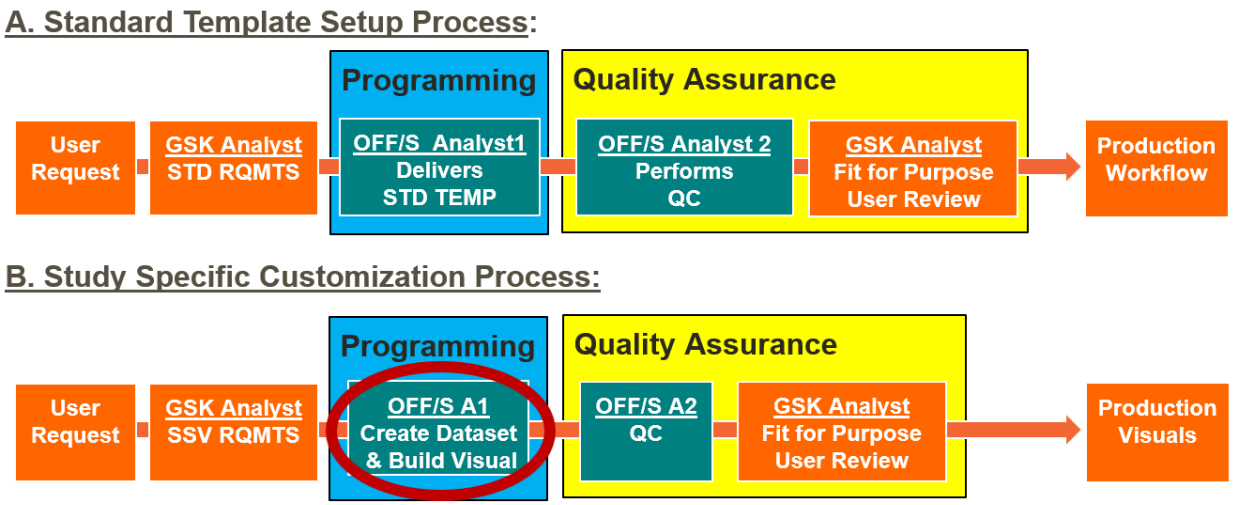

Figure 2. A. Schematic depicting the Setup Process for Standard Templates. B. Schematic depicting the study specific customization process following delivery of the Standard Template.

To support this process, the centralized Data Visualization Operations Team utilizes a flexible resourcing model composed of a matrixed On/Off-Shore Team (ON/S, OFF/S). This provides the best use of resources allowing routine programming work to be moved Off-Shore, while providing capacity for On-Shore Staff to devote additional resources to the more challenging Customizations/Updates. Importantly, the Off-Shore staff is cross-trained in other areas allowing demand-based expansion or contraction of the team, while ensuring that GSK On-Shore resourcing is independent of pipeline changes.

The Setup Process (Figure 2A) begins with a User request that is prioritized and assigned. The assigned Data Visualization Analyst reviews the request against the requirements of the requested Standard Template and ensures the necessary criteria are met for programming to begin. Programming is done Off-Shore, followed by a two-step QC Review by On and Off-Shore Staff. Once the Workflow has passed internal review, it is released to study team to assess if the deliverable is Fit-For-Purpose (FFP). Once FFP is granted, the Workflow is then moved to Production status.

The Customization Process (Figure 2B) is very similar to Setup process with 2 notable differences. The first is that the requirements for the customizations/enhancements requested are developed in a highly collaborative effort with the study team. The Data Visualization Analyst holds a requirement's meeting and gathers input from the key stakeholders (i.e. Statistics, Clinical Programming, Data Management, Safety, and Clinical). Secondly, the Analyst then translates this guidance into Programmable Requirements and works closely with the programming team Off-Shore to build the necessary data architecture and visualizations. A key to success has been the engagement and collaboration of the DV Analysts with Statistics and Clinical Programming (S&P). S&P has key study knowledge to help develop innovative Analytic Visualizations and align In-Stream DV Outputs with end of study reporting to provide the most timely and accurate information available.

All information for the Setup and Customization of a Standard Template is captured in a single document, i.e. Study Specific Workbook (SSW). The SSW is a comprehensive record detailing all changes made to the Standard Template as it has been configured for a study, detailing: Issues, Bugs, Coding Changes, Data Architecture, Enhancements, and New Visualizations. The SSW forms a key organizational process asset forming part of change control. Importantly, the SSW is a key input to the DV Improvement Process for the Up-versioning and Enhancement of the Standard Templates. The Improvement Process uses information in the SSWs to identify key study level innovations from each Therapy Area, and then incorporates these enhancements into the Standard Template Updates for delivery across the Portfolio.

CONCLUSION

Centralization of Data Visualization Operations at GSK provides rapid access to DV Tools with efficient management of Tool Setup and Customization. These interactive Data Visualizations are powerful engines for data interrogation. Successful organizational embedding of Data Visualization requires scalable plans for adoption & responsible Governance/Quality Plans ensuring risk/benefit balance maintained. Centralized Data Analytics has delivered key benefits of knowledge sharing and reuse; where elegant visualizations, developed at Program-Level, are integrated centrally to drive innovation at the Template-Level to augment in-stream analysis.

REFERENCES

1. The graphs rendered herein are not referenced to any specific GSK study or asset, contain no personally identifiable information, and are shown for demonstrative purposes only.

ACKNOWLEDGMENTS

The authors would like to acknowledge colleagues Jonathan Keach and Mark Andrews who have contributed their knowledge to our work.

CONTACT INFORMATION

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

Jason W. Szewczyk, PhD
GSK
1250 S. Collegeville Road
Collegeville, Pennsylvania, 19426-0989
610.917.4570
jason.w.szewczyk@gsk.com