Paper DV07

Translating Data into Tableau Insights: Bridging the Gap Between Programmers and Medical Physicians

Agnieszka Bel, SGS Health Science, Mechelen, Belgium
Roman Radelicki, SGS Health Science, Mechelen, Belgium
Annelies Van Zeveren, SGS Health Science, Mechelen, Belgium

ABSTRACT

Tableau dashboards provide valuable insights into clinical trial data. Our extensive library of standard dashboards covers common scenarios across different studies. However, end users often require custom data visualizations tailored to their needs and trial's nuances. The understanding of these requirements and effective communication between diverse study teams, especially medical physicians and programmers, are essential to meet these needs.

This presentation will outline our methodology for creating and validating custom Tableau dashboards and addressing potential challenges. The core feature in the process is setup of the demo environment, accessible by the end user. There, we can perform rapid prototyping, and the dashboards can be viewed and validated internally, and by the end user. This allows us to quickly understand if we are going in the right direction. Result? The final dashboards align with user expectations, offer fast results, and eliminate the need for reprogramming, ultimately benefiting all parties.

INTRODUCTION

Data visualizations are a powerful tool during the conduct of clinical trials for many stakeholders. Tableau enables the creation of dynamic, and interactive dashboards, providing insights in the chosen data in a secure environment. The clinical data domains (such as Adverse Events, Concomitant Medications, Vital Signs, Laboratory Test Results, ...) and scenarios (Adverse Events vs Concomitant Medications, Adverse Events vs study drug exposure, …) commonly present over different studies can be covered by a wide variety of standard dashboards, catering to different user's needs. However, sometimes the need arises for tailor-made visualizations. That's why SGS has developed a process of creation and validation of custom dashboards, ensuring that the final product is as desired and available in a timely manner.

WHY TABLEAU?

Nowadays, there are many data visualization tools on the market. Tableau, Microsoft PowerBi, Tibco Spotfire, Qlik, … to name a few. Also, more programming languages are providing ways to visualize data, like R Shiny or Python for instance. Each tool or language has its own strengths and considerations.

At SGS, we chose Tableau as the data visualization tool of clinical data for sponsors and medics for a couple of reasons. Tableau is a widely used platform known for its user-friendly interface and powerful visualization capabilities. It allows creation of interactive and dynamic dashboards without requiring extensive programming knowledge.

Some advantages of using off-the-shelf data visualization software like Tableau for dashboard creation in clinical trial research:

EASE OF USE AND PROTOTYPING

The drag-and-drop interface makes it accessible to a wide range of users, including non-programmers. With the quick prototyping and iteration of dashboards, we can rapidly create and modify visualizations. This is especially useful when we are working with end users like medical physicians to develop custom dashboards. We can quickly show some mock dashboards based on dummy data to get everyone on the same track as quickly as possible.

INTERACTIVITY

Tableau offers a variety of interactive features, such as filters, tooltips, and drill-down capabilities. These features enhance data exploration and facilitate on-the-fly data analysis.

DATA CONNECTION

The software supports the connection to various data sources including databases, spreadsheets cloud-based platforms and SAS datasets (sas7bdat). This is essential for integrating diverse clinical trial data.

While one could argue that programming language like R can offer dashboards with more complexity because users can leverage the full capabilities of R programming to pre-process the data, perform statistical analyses, many off-the-shelf data visualization tools also support and easily integrate Python and R code nowadays. It generally requires a stronger programming background compared to Tableau and therefore has a steeper learning curve. The choice ultimately depends on the specific scope, needs and skill sets of the clinical trial research team.

TABLEAU DASHBOARDS – SGS APPROACH

To ensure the quality of our dashboards, SGS follows the Software Development Life Cycle (SDLC) process when developing them. Meaning that for each set of dashboards we carefully analyze the request, plan, design and develop the dashboards; test, validate and deploy them in production. The lifecycle of such dashboards is all documented in the SDLC system.

The process of creation and validation of the dashboards is defined by a SOP. The dashboards exist in three different environments for development, testing and production. The creation of the dashboards happens in the development stage, based on the created test data, and validation is performed after moving to the testing stage.

SECURITY

With Tableau we can guarantee strictly separated environments per client. In Tableau-speak, the term “site” is used as a collection of users, groups, and content (workbooks, data, sources) that’s walled off from any other groups and content. Another way to say this is that Tableau Server supports multi-tenancy by allowing administrators to create sites on the server for multiple sets of users and content.

All server content is published, accessed, and managed on a per-site basis. Each site has its own URL and its own set of users. Each site’s content (projects, workbooks, and data sources) is completely segregated from content on other sites.

SGS STANDARD DASHBOARDS

Over time, SGS has developed an extensive library of standard dashboards built on dummy data (it is a ‘live’ library, continuously being improved and growing, depending on the findings and feedback we receive from various end users). They cover the domains and scenarios commonly present over different studies and contain variety of visualizations, suitable for different user’s needs (e.g., data cleaning, medical safety review / safety meetings, medical monitoring, patient narratives, …). Some of the dashboards are dedicated to one specific domain, but the library also contains cross-domain dashboards, allowing to view information from different datasets at the same time, and analyze potential relationships. This includes patient profiles, providing a complete overview of patient’s core safety data: adverse events (AE), concomitant medications (CM), medical history (MH), laboratory data (LB), vital signs (VS), electrocardiography results (EG) and more.

Our preferred and most efficient way to develop standard dashboards is to base them on SDTM data. Since it’s a standardized format, these dashboards can be used for different end users and studies. Another advantage is, that dashboards that have been once validated, can be re-used for other projects without the need for re-validation, as long as the same data structure is used, and no dashboard updates were necessary (the source SDTM data is validated by a different Quality Control (QC) process). Hence, they can be deployed in production shortly after the SDTM production datasets are available. The refresh of the data is incorporated in SGS DM data conversion process. That means that the dashboards are automatically refreshed with the new data at the frequency agreed upon with a study team.

Nevertheless, various data sources other than SDTM can be used for Tableau visualizations (or combination of the data sources): raw eCRF data, Excel, SAS… files (for example, transfers from an external vendor). Also, depending on the source and availability of the validated test data, dashboards can be available shortly after the data from the first patient is available (Figure 1).

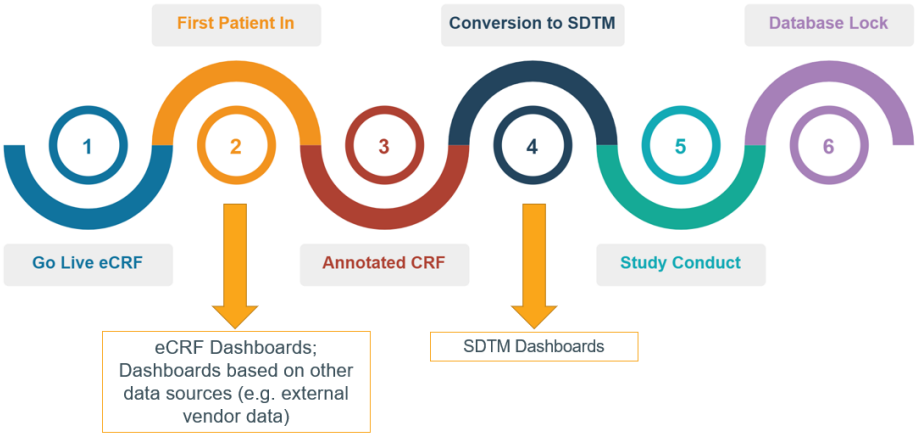

Figure 1: Availability of dashboards during setup of a trial

TABLEAU DASHBOARDS IN THE PROJECT

When dashboards are requested for a clinical study, the first step is setting up a meeting with relevant team members to introduce Tableau, demonstrate the standard dashboards and explain the software functionalities.

The main parties involved are: end users/medical physicians, Tableau programmers (the creators of the dashboards) and the Tableau coordinator (the liaison between all involved teams). Depending on the study type and initial project assumptions, during the demonstration the focus is mainly put on the visualizations that are potentially of the biggest use for the study. The team is also made aware of the possibility to create the custom dashboards. Because of the number of the standard dashboards present in the SGS library, one meeting is not enough to explain in detail the characteristics and purpose of each individual dashboard. After the meeting, the library is shared with the chosen team members (who are granted temporary access) within a dedicated secure Tableau site created specifically for the project, for agreed period. During this time, they can browse through all the standard dashboards and choose, which will be used for the study. This also allows them to get an idea, what kind of dashboards and visualizations can be created. At the same time, a tutorial present in Tableau itself is shared, with explanation of various dashboards functionalities and how to use them. It helps the users to get acquainted with working with the dashboards (and user interactivity, like sorting, filtering, downloading the data, …) and allows to fully utilize the advantages of all visualizations.

While standard dashboards, as mentioned, cover a wide range of the common data present across different trials, often there is a need for a specific dashboard, tailored to the needs of a specific team / study. That's where the custom dashboards come in play.

CUSTOM TABLEAU DASHBOARDS

Standard dashboards are a simple “take what you want” solution. When talking about the custom Tableau dashboards, there is a wide range of scenarios, depending on end user’s needs and trial nuances. They can be based on existing standard dashboards that get modified, or completely new dashboards can be made from scratch. A big advantage is, once a custom dashboard is created, it can be standardized and re-used for other similar studies. If the structure of the data source is the same, for the next study the set-up (including the validation process) of such dashboard is much easier and faster, similar to the standard dashboards.

The main challenge when creating the custom dashboard is making sure that all parties are on the same track. On one side, programmers responsible for creation of the dashboards, who operate with “dry” data and formulas. On the other side, team members, like medical physicians, who look at the data from a safety point of view and are not necessarily aware what happens behind the scenes of the dashboards, programming-wise. That includes combining the requests like amount of data to include / number of domains / visualizations / potential pre-processing / desired functionalities and interactivity, with the technical possibilities and limitations of the software. That’s where the Tableau coordinator plays the big role, working as a “bridge” between different teams.

The whole process from request, until the final product, consist of few phases, described below: understanding the needs – drawing the roadmap; creation of the specifications and creation and validation of the dashboards.

UNDERSTANDING THE NEEDS – DRAWING THE ROADMAP

The starting point in the creation of the custom dashboard is gathering information about the needs of the end user, such as how the data should be displayed (e.g., type of graph, ...), if there are study timelines constraints, and for which tasks or purpose the dashboards will be used in the study. Depending on the experience of the team and the study scope, the received information can differ a lot. Sometimes their requirements might be clearly specified (with example graphs or elaborated descriptive specifications), the team might have some idea based on a standard dashboard, or they might just provide very basic descriptive information. It is also important to have an idea when the dashboard should be available to ensure that expectations are aligned. The same dashboard, showing exactly the same information can be created from various data sources, but the data source might be available at different stages of the study setup (as depicted in Figure 1). Also, the format of the data source and amount of the information to be included, impacts the complexity of the dashboard and the time needed for development and validation.

After we have received the information from the team; we discuss the approach with the Data Visualization team. Having worked on number of various dashboards, the team analyses where potential challenges and issues can occur. If a deeper insight into trial details is needed (e.g., SDTM annotations), Tableau coordinator reaches out to other SGS teams involved in the project (Data Management, Medical Safety, Electronic Data Collection…) to gather their expertise. If needed, additional discussion with the end user can be scheduled.

CREATION OF THE SPECIFICATIONS

Before the technical part of the SOP-defined process for the custom dashboard creation begins, SGS creates specifications, to ensure that programmers and end users are on the same track. The initial approach was to keep the specifications very descriptive and without technical details, to ensure that they are understandable for non-programmers. The specifications were covering all the necessary information (data source, types of graphs…) and functionalities (filters, different display options…) to be included in the dashboard. Afterwards they were sent to the end user for review and approval. After approval, the creation and validation steps were started.

With the first custom dashboards created by SGS, it quickly turned out that although the specifications were approved by the end user, they were too abstract to provide full understanding on how the dashboard will look like and work. Once the dashboards were programmed, validated internally, and sent for study team validation, we received a lot of requests for updates. This resulted in additional programming and validation effort and extended the study timelines to get the dashboards approved and released in production.

BRIDGING THE GAP – MOCK VISUALIZATIONS

This issue was addressed by the introduction of mock visualizations in the specifications in addition to the description. They were created by using available test datasets of the trial (for example, dummy test data used for QC of the SDTM data conversion), or by creating new test datasets (done manually or by using a test data generator).

This solution provides better insight for the end user on how the visualization will look like, rather than just descriptive specifications. Also, the advantage for the programmers is, that this extra step can be used as a sort of “test phase”, allowing rapid prototyping in an agile way and limiting the (final) programming effort. Based on the mock programming outcome, it’s possible that certain requests are not feasible/efficient and changes in the dashboards are required, or the end user concludes that the dashboard doesn’t fulfill their expectations or needs. In this case the updates can be discussed earlier, and new mock visualization can be programmed at this stage, rather than discussing this at a later stage when the full dashboard has been programmed and validated internally, which was the case before.

Data source: ECG Data csv. file+ SUPPDM (SDTM datasets)

For all graphs, the average value of all repetitions per visit/timepoint and subject will be taken.

Line plots – with filter allowing to change between the views:

1. View per subject (y-axis – test result + lines indicating the lower and upper limit of the normal range; x-axis – visit/timepoint name)
2. Mean profile of selected subjects (y-axis – test result + lines indicating the lower and upper limit of the normal range; x-axis – visit/timepoint name)
3. View per subject + mean profile

- Filters: cohort, subject name, test; possibility to display multiple tests
- Display options: absolute value or change from the baseline

Example view:

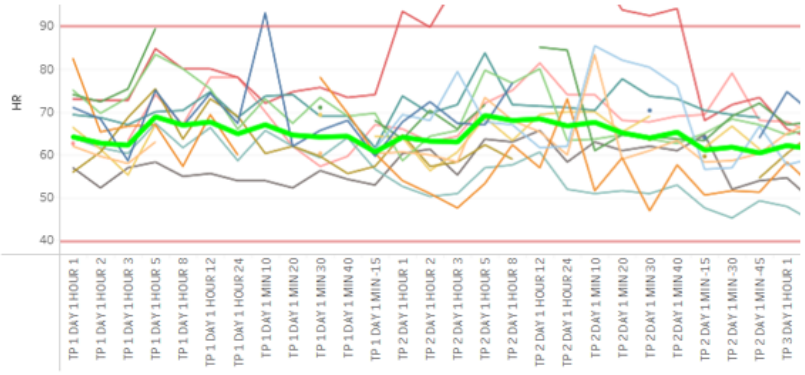

Figure 2: Example custom dashboard specifications with mock visualization

CREATION AND VALIDATION OF THE DASHBOARDS

As mentioned before, the creation of the dashboards starts in the development environment, based on available/created SDTM datasets / eCRF extract / other data sources (as applicable for the dashboard). It is important that structure-wise, the test data source is stable and like the future production datasets and content-wise – that it resembles realistic study safety data scenarios to allow end users to evaluate and validate the dashboard correctly. This prevents issues when the actual study data is uploaded.

The next step is the **internal validation** in the test environment. On one side, the dashboard is verified against the specifications, to ensure that all agreed upon visualizations and functionalities have been implemented. On the other side, the content of the dashboard is checked with the test data: is all data uploaded in the correct places, is there nothing missing, in case of data manipulations done in pre-processing (like taking the average value of the repetitions in the example specifications in Figure 2), are the calculations done correctly, ….
Additionally, the practical working aspect of the dashboard is verified – do filters work correctly, are the visualizations readable, are the legends clear, etc.

Still, the most important thing is that the dashboard is suitable for the end user. That's why we foresee sufficient time for the **external validation** (taking into account dashboard complexity). It ensures that the team members can verify if the dashboard fulfills all their needs, get acquainted with it and get a real user-friendly feeling during validation. The dashboard for the end user's review is released from the test environment into a special demo folder, located within their site.
The dashboards are designed to be intuitive and simple in use. At the same time, at this stage some team members might not yet be experts in using Tableau and require some extra guidance and time to get acquainted with all aspects of the system and dashboards. To support them, they receive a detailed description of the working principle of the custom dashboards in addition to the tutorial. If desired, a training session can be scheduled.

From that step onwards, in case of any validation comments, the process of dashboard programming → internal validation in test environment → external validation in demo environment continues, until the end user confirms that they approve the dashboard, at which point it can be released in the production environment.
When the first release of live study data occurs, the dashboard is checked once again by SGS team, to make sure that there are no unexpected issues. Afterwards, the dashboard is ready for use in the study.

Also at this stage, we have improved our process compared to the first custom dashboards created. Initially, we were ensuring that enough test data is available to validate the dashboards and that all information is covered. From the

programming point of view, there were no issues. However, since this data wasn't always resembling the actual study scenarios, the medical physicians were not getting the real-life experience of working with the dashboards. Only when the dashboards were released in production, we got the requests for updates to improve working with the visualizations.

BRIDGING THE GAP – TEST DATA

SGS addressed the issue above by making the test data reflecting the real-life situation as much as possible. This way the dashboard meant for validation closely resembles the actual dashboard that will be used in the study (for example, the number of visits and assessments foreseen in the clinical study protocol, the right chronology of the events, using realistic values…) and gives the end user possibility to verify if it suits them.

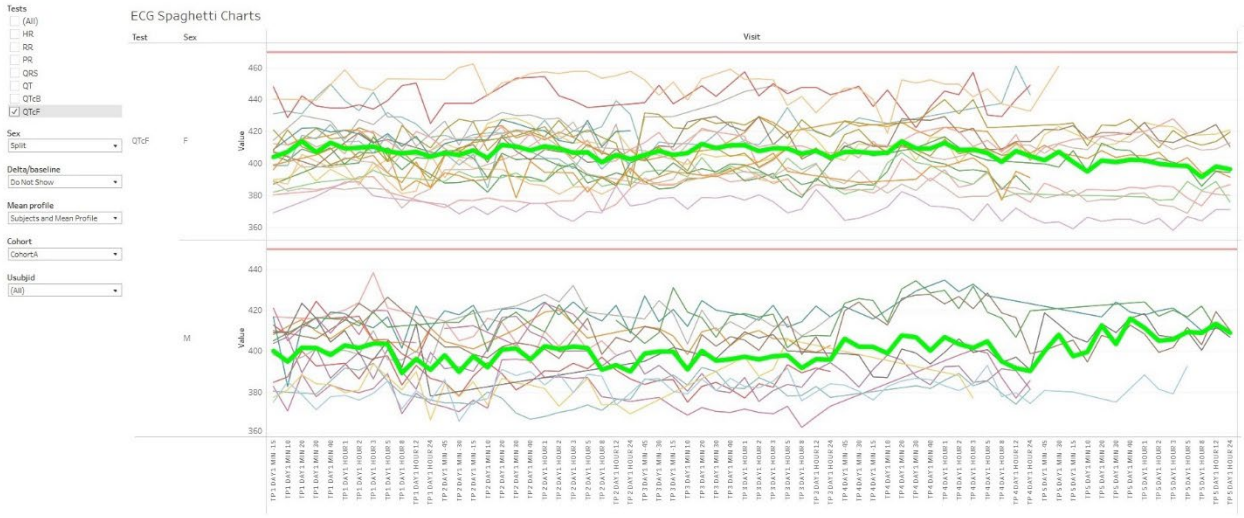

Figure 3: Example custom dashboard – ECG results, based on a test .csv file

Figure 4: Example custom dashboard – oncology swimmer plot, based on test SDTM datasets

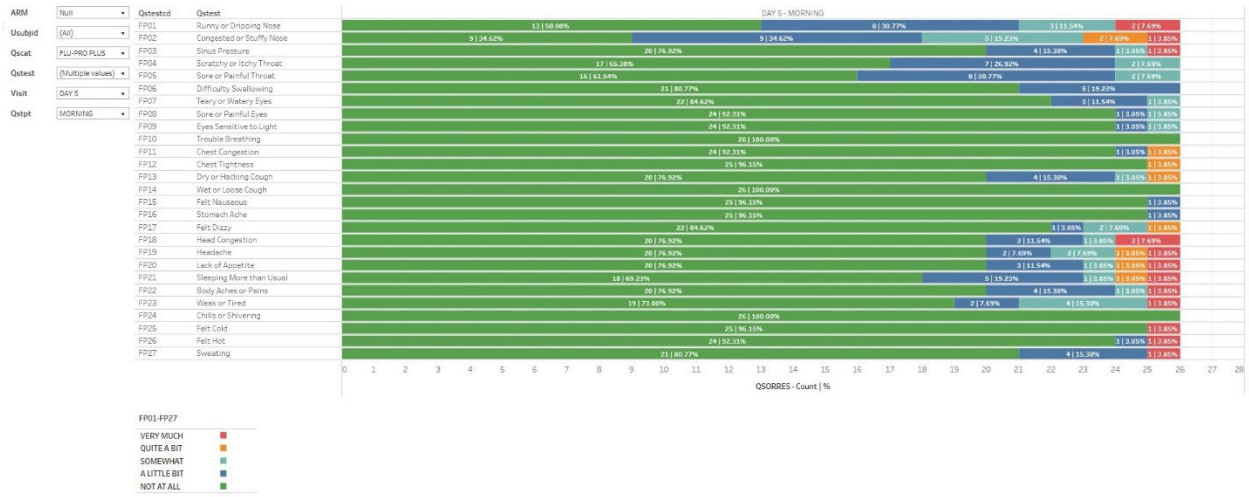

Figure 5: Example custom dashboard – FLU-PRO questionnaire results, based on test SDTM datasets

CONCLUSION

The versatility of Tableau makes it a very suitable tool for creation of the custom data visualizations, covering different needs of the end users. The most important hurdle during the whole process of creating custom dashboards, is the communication between different teams and end users involved in the project. This is addressed and facilitated by the Tableau coordinator who ensures that there is a pro-active approach, a clear communication, and an understanding of each other's needs. On one side, the programmers need to be aware of what they need to create. On the other side, the end users are informed about the possibilities and limitations of the software (for example, if the underlying SDTM datasets are not clean, the output in Tableau might seem "wrong", even though the dashboard works correctly). The Tableau coordinator needs to understand both sides and act as a bridge between both teams to ensure they are all on the same track.

Over time, we have adapted our process, based on the experience with the first custom dashboards. There were two main improvements implemented by the SGS team to bridge the gap between the programmers and the medical physicians.

1. We introduced mock visualizations at the stage of development of the specifications for the custom dashboard. On one side, this allowed end user to have immediate idea on how the final visualization will look like. On the other side, it limits the final programming and validation efforts of the SGS team, in case of updates.

2. Basing the dashboards for validation on test data closely resembling the real study data, enables the team members to get the feeling of working with the real dashboards and clearly realizing if they suit their needs.

Thanks to all these elements, we ensure that the road from the request until the production dashboard is as smooth as possible, and the end product is delivered in a short turnaround time with the highest quality.

CONTACT INFORMATION

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

Agnieszka Bel
Roman Radelicki
SGS Health Science
Generaal De Wittelaan 19A bus 5
BE - 2800 Mechelen
Work Phone:
EUROPE: +32 15 27 32 45
AMERICAS: + 1 877 677 2667
Email: clinicalresearch@sgs.com
Web: www.sgs.com/cro