

Paper 4270-2020

Using SAS® Visual Analytics Reports for Operational Reporting in Clinical Trials

Jun Wen, Jack Shostak, Duke Clinical Research Institute, Durham, NC

ABSTRACT

Our presentation shows how to use SAS® Visual Analytics to implement data visualization for operational reports in clinical trials.

In clinical trials, operational reports are often used to analyze and monitor site characteristics and performance metrics as well as patients' enrollment, safety, and follow up. Data visualization can track the site and enrollment information in a dynamic and interactive way, and provide significant time savings with no decrease in accuracy compared to traditional static reporting. At the Duke Clinic Research Institute, we have successfully implemented data visualization for operational reports using SAS Visual Analytics. In our operational reports, we use SAS Visual Analytics Geo Maps to monitor site enrollment distribution. In our Geo Maps, we add the display rules to identify at risk sites, and the time series to show the site activation/enrollment by time. We add Hierarchy (patients' status) on pie charts so we can track patients' status from screening to end of study by drilling down into the pie charts. We also add drop down lists to subset enrollment/site reports, so the reports can be used to compare sites by milestones, or enrollment by desired frequency (monthly/weekly). We use interactions to link aggregated reports (line chart, bar chart, pie chart et al.) and detailed listing tables, so for any interesting summary information, we can easily identify the individual patient/site level information.

INTRODUCTION

A clinical trial is a prospective study comparing the effect and value of intervention(s) against a control in human beings^[1]. In clinical trials, operational reports are often used to analyze and monitor site characteristics and performance metrics as well as patients' enrollment, safety, and follow up.

Compared to traditional static reporting, data visualization provides a dynamic and interactive way to present operational reports. Munsaka et al^[2] used visual analytics approaches to enhance evaluation of drug safety. Shneiderman et al^[3] explored using interactive visualization to improve healthcare. Tong et al^[4] and Duke et al^[5] found that visualization for trials could provide significant time savings and achieve similar accuracy as traditional static reporting. Goldman et al^[6] explored using data visualization for survival and other time event analysis. At the Duke Clinic Research Institute, we have successfully implemented data visualization for operational reports using SAS Visual Analytics.

In this paper, we will demonstrate reports for 1) enrollment monitor; 2) site monitor; 3) Consort diagram; 5) site report card; 6) Sankey diagram and Correlation matrix.

ENROLLMENT MONITOR

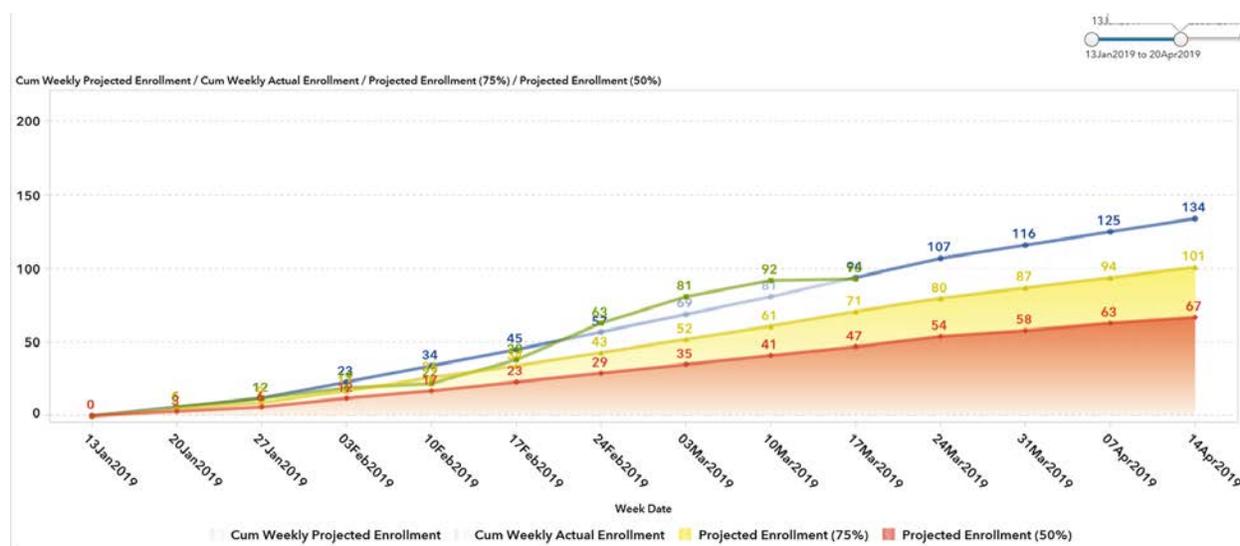
Meeting the projected enrollment is vital for a clinical trial's success, and a site's performance is normally tied to their ability to enroll subjects. The study team closely monitor enrollment and sites using tables, listings, and figures. The enrollment monitor often includes comparing projected and actual enrollment, actual enrollment by site by desired frequency, and patients' follow up status.

PROJECTED ENROLLMENT AND CUMULATIVE ACTUAL ENROLLMENT

Projected enrollment is defined based on a sample calculation, and the analysis will not have enough statistical power if we cannot reach the projected enrollment. Based on the study design, enrollment period, and monitoring frequency, the projected enrollment is often presented by month or by week. For governmental trials, 75% and 50% of projected enrollment are two critical stages. If the actual enrollment is lower than 75% of the projection, often times a modified enrollment plan will be required. If the actual enrollment is lower than 50% of the projection, the trial might be terminated.

Since the enrollment is so vital for a trial, the project leadership closely monitors subject enrollment. Figure 1 shows the comparison between the actual enrollment and projected enrollment. The blue line represents the projected enrollment by week through the whole enrollment period and the green line represents the actual cumulative weekly enrollment. The yellow line and red line and associated shading represent the 50% and 75% projected enrollment levels. In the upper right corner, a time slider shows a preset period, normally from the first day of enrollment to the last day of the study. Using the time slider, you can zoom into the line chart for a chosen time range for a closer look, which can be very useful when you have a lot weeks.

Figure 1: Weekly Cumulative Projected Enrollment vs Weekly Cumulative Actual Enrollment



Note: The blue line represents the projected enrollment, and the green line represents the actual enrollment.

WEEKLY ENROLLMENT BY COHORT BY SITE

To encourage increased site enrollment, enrollment is normally monitored by site level, and often broken down by additional factors. Figure 2 shows an example of actual enrollment by site. The figure includes a bar chart for enrollment by site by treatment cohort, an interactive listing for patient level information, and a drop down listing for each week.

The drop down list contains all weeks for the enrollment period, and the date is the Sunday of each week. The number associated with each week to the right of the date represents the number of subjects enrolled for each week, but you can also choose percentage if you prefer. The percentage will represent the percent of the site's enrollment over the study enrollment.

The bar chart represents the enrollment by each site. The number on the top of the bar is the total enrollment, and if you hover over each section (blue for treatment cohort, and green for placebo cohort), you can see the number for the selected cohort as you can see in Figure 2b.

The list table shows patients enrolled. To make the table convey more information, we use "Display Rules" adding color for patients' treatment and follow-up status. The rows highlighted in green represent patients in placebo cohort, and rows without highlights represent patients in treatment cohort. The cell highlight in red represents a patient that died.

The bar chart and the list table is interactive with the drop down list. When "Clear Filter" is selected, the bar chart and list table will show cumulative enrollment. If a specific week is chosen (e.g. 20Jan2019), the bar chart and the list table will be updated to include only enrollment at that week (Figure 2a). In addition to the drop down list, the bar chart is also interactive with the list table. Once a specific site is chosen in the bar chart, the list table will be updated to include only enrollment at that week and for only that site (Figure 2b).

Figure 2: Weekly Enrollment by Cohort by Site

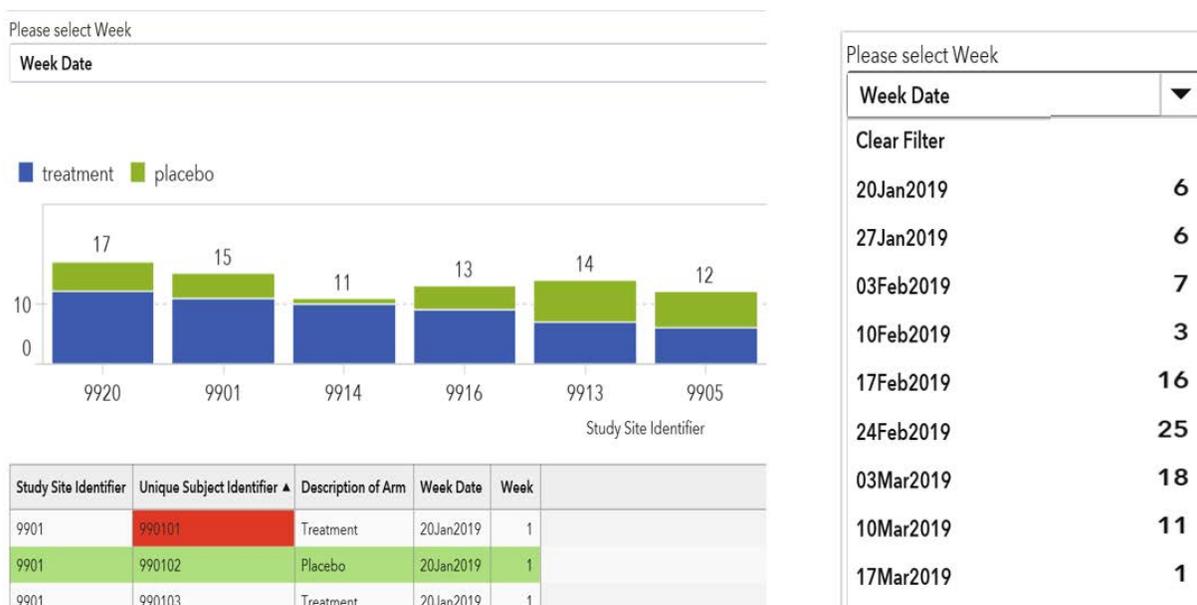


Figure 2a: Weekly Enrollment by Cohort by Site for Selected Week

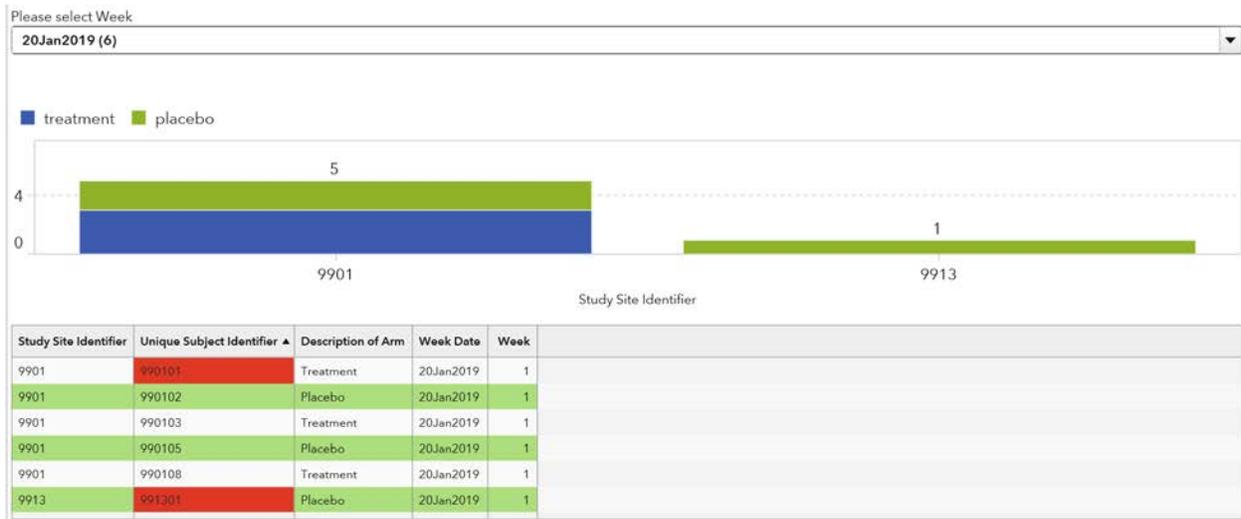
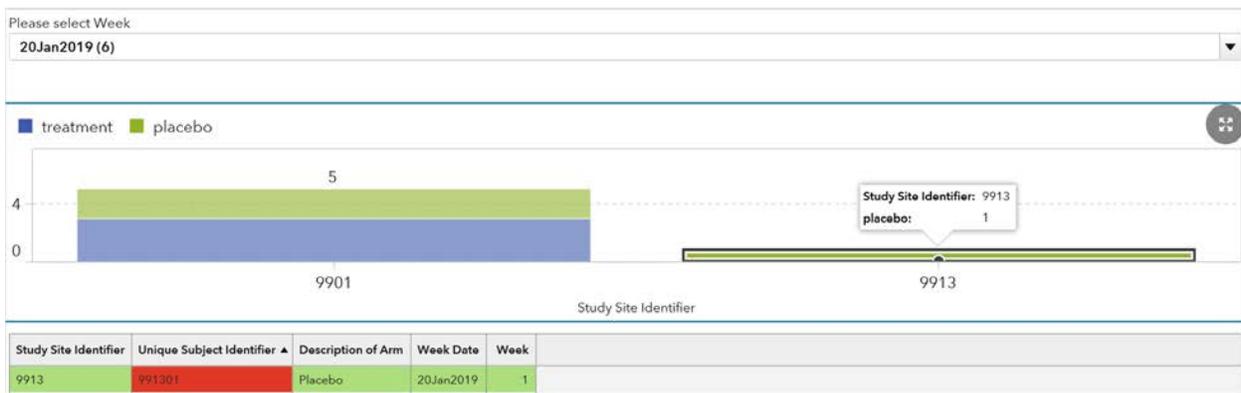


Figure 2b: Weekly Enrollment by Cohort for Selected Week and Site



WEEKLY ENROLLMENT BY SITE BY WEEK

A report that shows enrollment trend by site is also very helpful encourage site enrollment. Figure 3 shows the weekly enrollment by site by week. The figure includes a drop down list, a line chart, and a list table for patients' detail information.

The drop down list includes all sites, and the number associated on the right with each site represents the total enrollment for the site.

The line chart displays the enrollment by week, and the list table below it contains all enrolled patients. Just as we see with Figure 2, the drop down list controls the information presented for the line chart and the list table. The list table is also interactive with the line chart. Select any point at the line chart, the list table will be updated to include enrollment at that week for the chosen site, and the cell highlight in red represents a patient that died (Figure 3a).

Figure 3: Weekly Enrollment by Site by Week

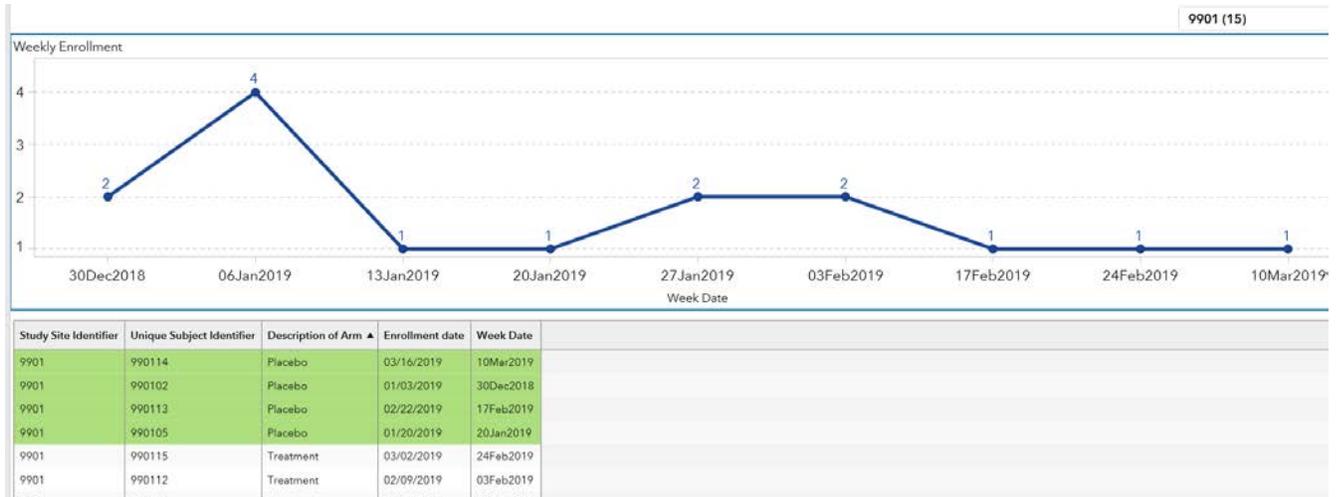


Figure 3a: Weekly Enrollment by Site for Chosen Week



SITE MONITOR

Sites recruit, retain, and collect all patients' clinical information. Having enough actively enrolling sites, and having the sites geographically located per design are important for a trial. In this paper, we want to show a site monitor report for site distribution, site milestones, and site performance associated with enrollment.

SITE DISTRIBUTION

To make sure the study enrolls a variety of patients, sites' locations are carefully chosen for the study. Figure 4 demonstrates a geo map for site distribution, and uses the bubble size and color to represent the enrollment and Key Risk Index (KRI). KRIs are the threshold set up by project leadership to monitor at risk factors, such as enrollment, safety. In Figure 4, the location is pinpointed using each site's zip code, the bubble size is controlled by the size of the site's enrollment, and the color is defined using the Display Rule function based on the enrollment KRIs. Red represents enrollment <5, yellow represents enrollment 5-20, and

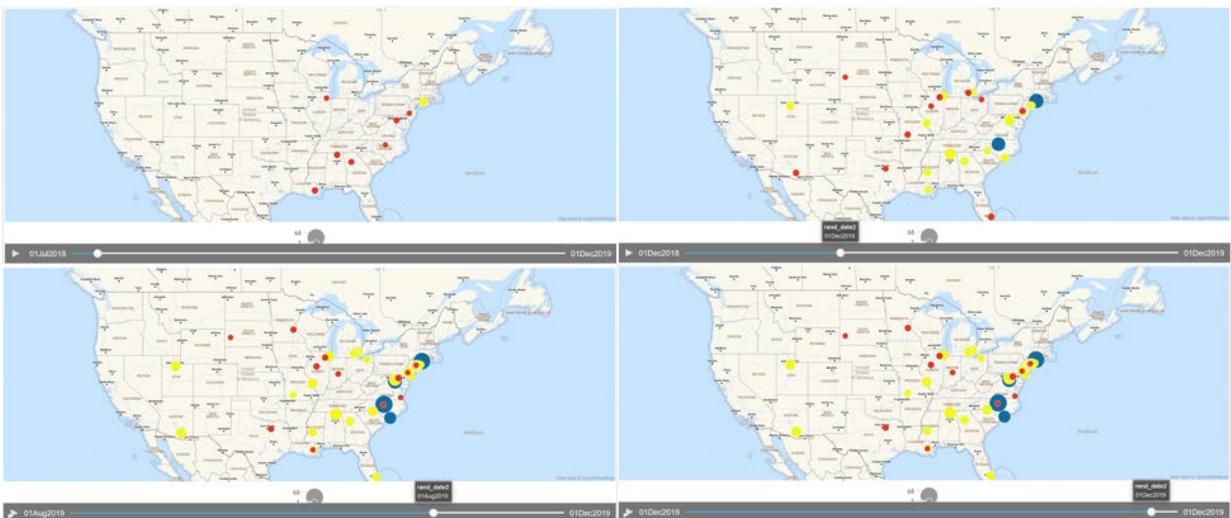
blue represents enrollment >20. In this way, it's easier to have an overall assessment for the sites.

Figure 4a demonstrates the site distribution over time using the animation function in VA. Using the time slide, you can see the site distribution changing during the enrollment period.

Figure 4: Geo Map for Site Distribution



Figure 4a: Geo Map for Site Distribution – Animation

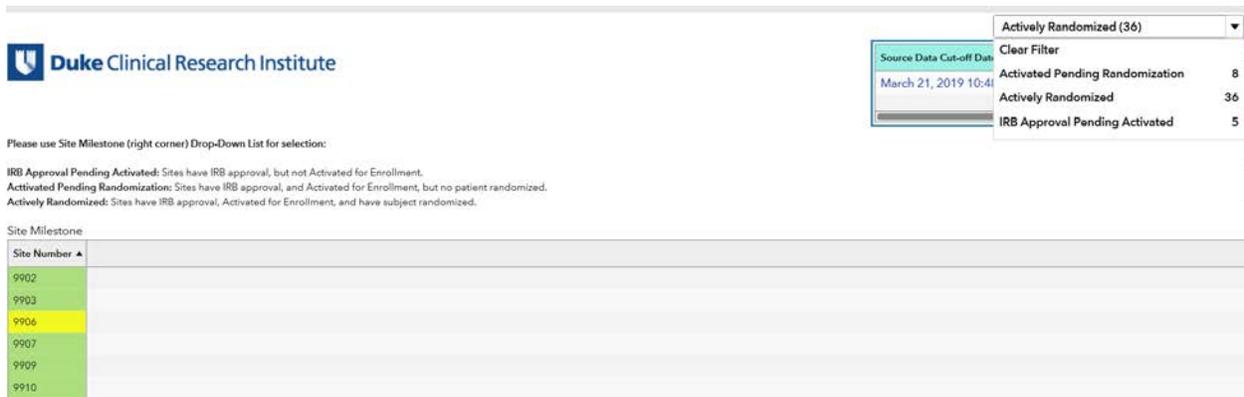


SITE MILESTONE

Often times there are site milestones that are key in determining a site's enrollment performance. Figure 5 shows a dynamic table for site milestone monitor using VA Controls. The Drop-Down List on the upper right corner shows the milestones for this study: IRB Approval Pending Activated, Activated Pending Randomization, and Actively Randomized.

The number displayed to the right of each milestone shows the number of sites, and the Drop-Down List controls the list table below. If you choose "Clear Filter", the table will show all sites, and if you choose any specific milestone, the list table will be updated to only include the sites having that milestone. You can also use Display Rule to add color on the list table for additional information. For example, in Figure 5, sites that have been activated 90 days but having had no randomization will be shown in red, 30-90 days in yellow, and less than 30 days in green. The List Table currently only included Site Number for demonstration, but you can have site PI, coordinates, date for each milestone et al.

Figure 5: Dynamic List Table for site milestone monitor



SITE PERFORMANCE

Figure 6 shows a bar chart and an interactive list table for site enrolment monitor. For the bar chart, each bar represents how many sites fall within the specific range of subjects randomized. For detail, the bar can be used to drill down. For example, if you are interested in a site list of sites with zero enrollments, just click the bar with 0, and the below list table will be updated to the 8 sites with 0 enrollments (Figure 6a).

Figure 6: Interactive Bar chart and List Table for site performance

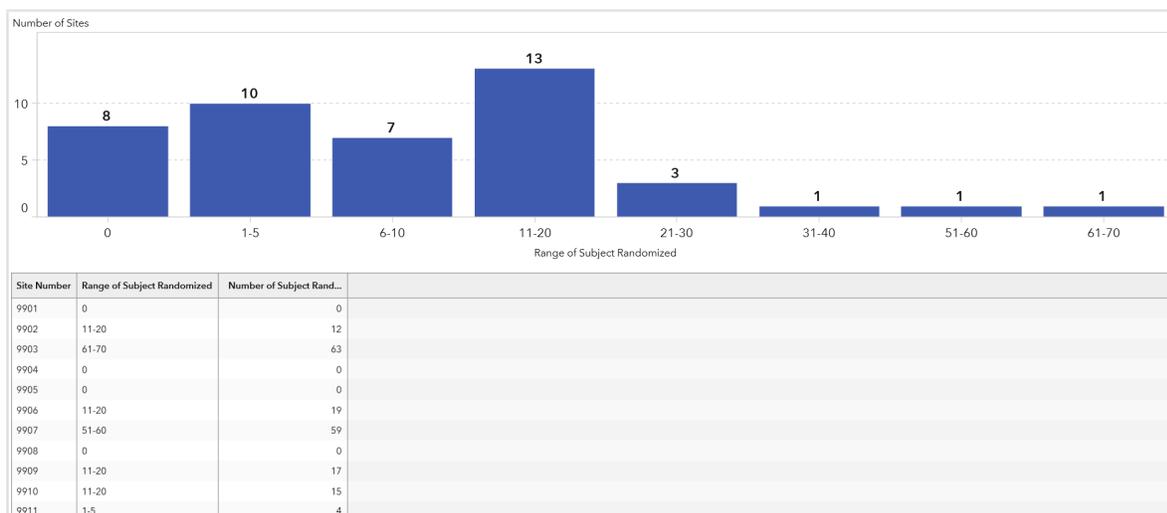
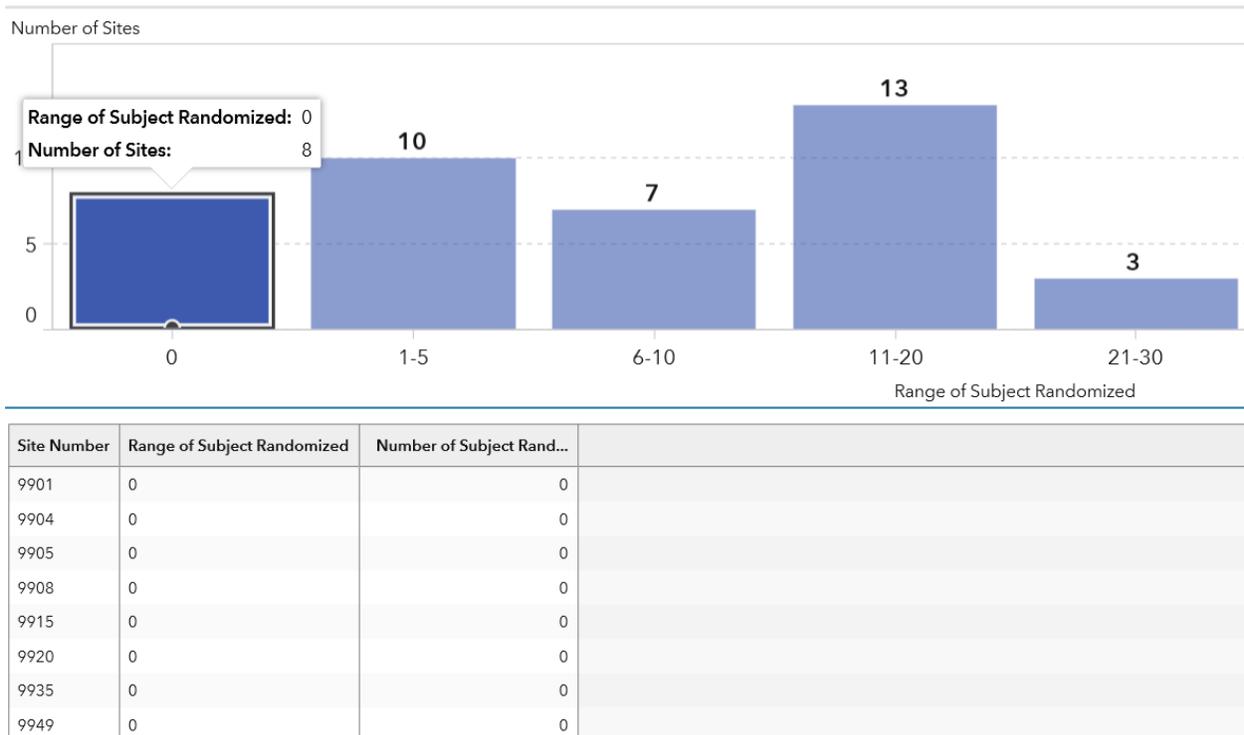


Figure 6a: Interactive Bar chart and List Table for site performance at specific range



CONSORT DIAGRAM AND PATIENT DISPOSITION

The following section shows some consort diagram examples and patient disposition displays. The idea here is to show how SAS Visual Analytics can be used to show patient progress through the course of a clinical study.

CONSORT DIAGRAM

The Consort diagram^[7] is a flow diagram of a patient's progress through the phases of a trial. A traditional consort diagram is not something that is easily done with Visual Analytics version 7.4. However, using list table and image objects, one can construct a traditional consort diagram as seen in Figure 7. The list tables are the summary boxes and the images are lines that were created in Microsoft Word and imported into SAS VA. The data that supports Figure 7 is simple one record per subject data with population flags as you might find in the CDISC ADaM ADSL dataset. Figure 7a shows the ability to drill down into each list table population presented and see a list of associated variables of interest. There you can see the completed, lost to follow up, and withdrawn status of the 229 active treatment arm subjects when clicking in that box.

Figure 7: Consort diagram using List Table and Image

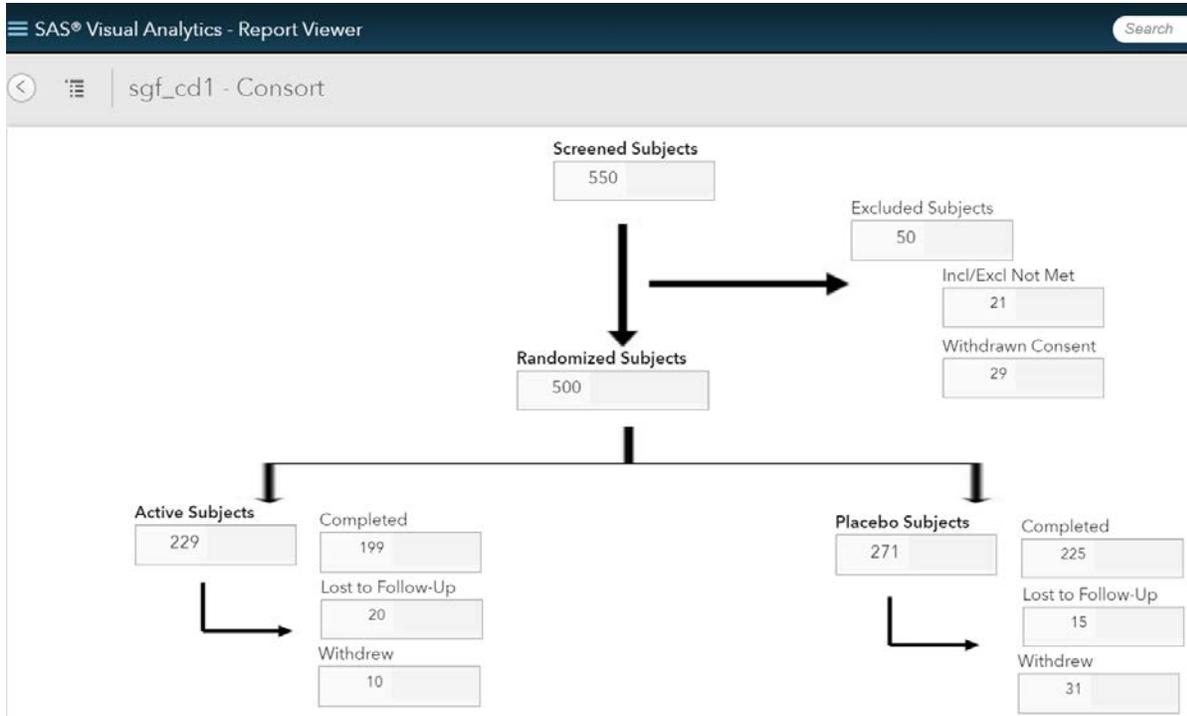


Figure 7a: Consort diagram using List Table and Image Drill Down

SAS® Visual Analytics - Report Viewer

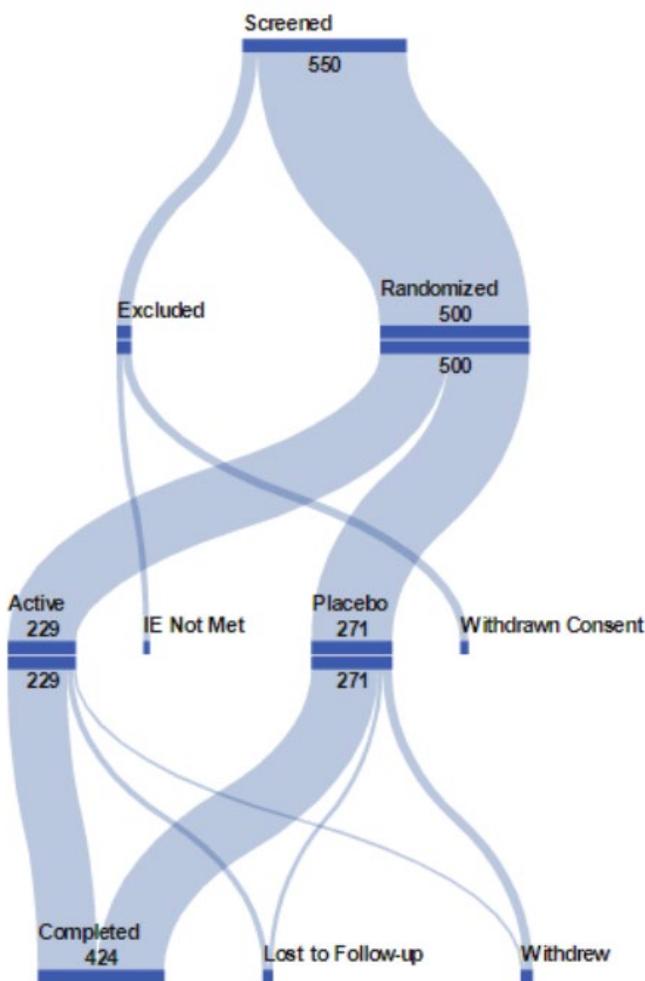
sgf_cd1 - Active Subjects

usubjid	Planned treatment	Completed flag	Lost to follow up flag	Withdrawn flag
185	Active	1	.	.
186	Active	1	.	.
187	Active	1	.	.
188	Active	1	.	.
189	Active	1	.	.
190	Active	1	.	.
191	Active	1	.	.
192	Active	1	.	.
193	Active	1	.	.
194	Active	1	.	.
195	Active	1	.	.
196	Active	1	.	.
197	Active	1	.	.
198	Active	1	.	.
199	Active	1	.	.
200	Active	.	1	.
201	Active	.	1	.

Another way to present this Consort diagram in a similar way is via a Sankey Diagram's path analysis display. In Visual Analytics version 7.4, the Sankey Diagram can be found within the Data Explorer tool. The same data in Figure 7 is presented in Figure 8 via a Sankey Diagram. Do note that some summarization of the ADSL dataset driving Figure 7 must be done before the Sankey diagram can be done. Also note that although the Sankey diagram is more eye pleasing, you do not have control over how SAS Visual Analytics draws the paths. The display is mostly automated.

Figure 8: Alternate Consort Diagram Via Sankey Diagram

Sankey Diagram of event



PATIENT DISPOSITION

Patient disposition refers to a patient's progression throughout a clinical trial at various time points. Such a display of patient dispositions is very similar to a Consort diagram.

Figure 9 shows a drill down pie charts for tracking patients' disposition by treatment. The pie chart shows the patients current disposition at any interested time point. On the right, an interactive list table shows patients detail information. Patients in placebo are highlighted

in green for the row, and discontinued patients are highlighted in red for a selected field. If you are interested in a specific disposition (e.g., Completed), you just click that pie slice, and the list table will be updated to include only completed patients (Figure 9a). You can also use the Hierarchy function to organize multiple layers of information. The number of layers and the information for each layer can be determined by study relevance. In our example, we are interested in treatment counts for each disposition. As shown in Figure 9b, for any selected disposition (e.g., Completed) and using drill down, the pie chart and the list table are updated to show treatment for completed patients.

Figure 9: Patient disposition

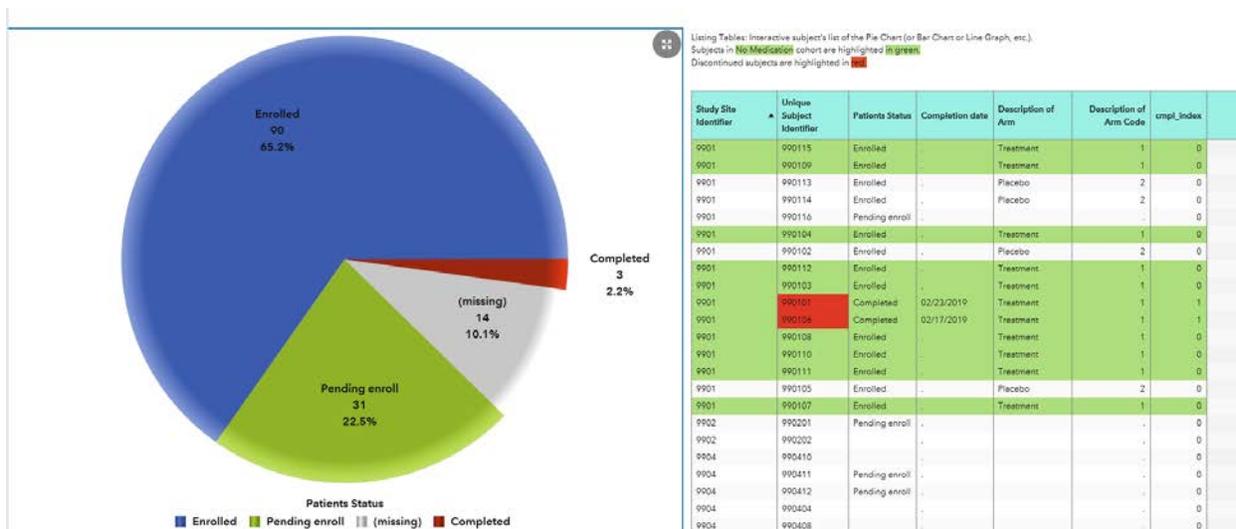


Figure 9a: Patient disposition for selected disposition

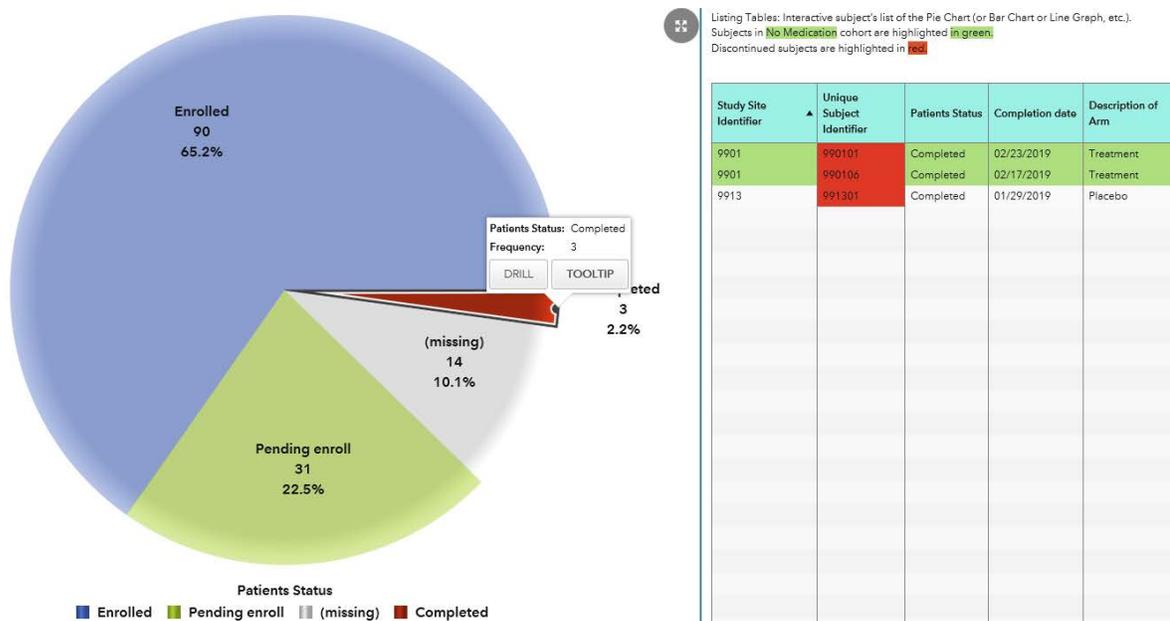
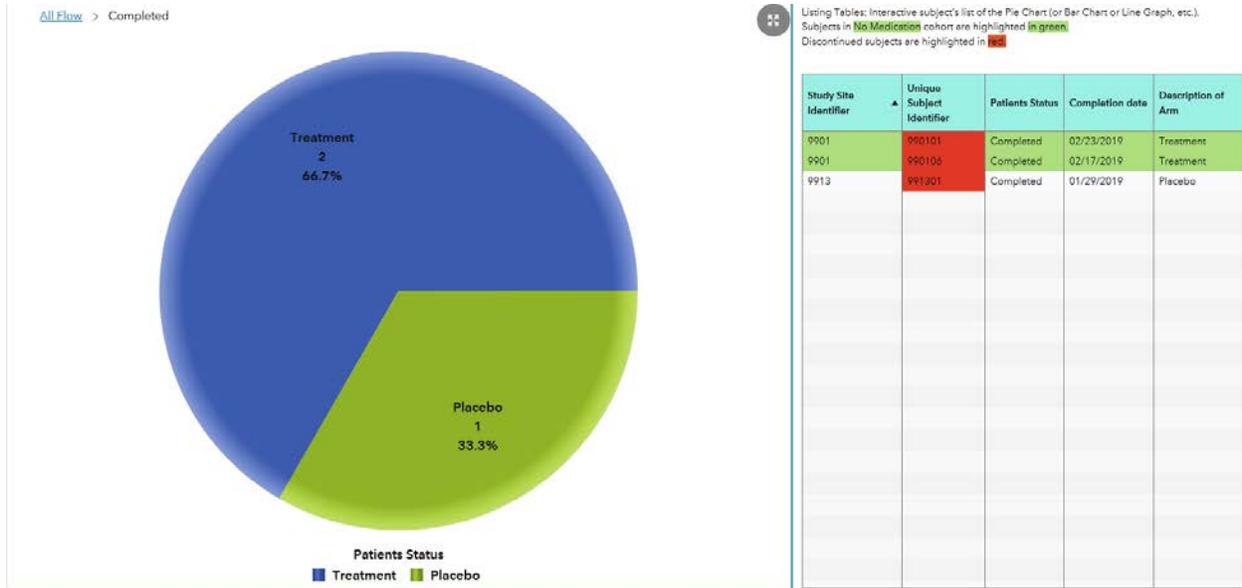


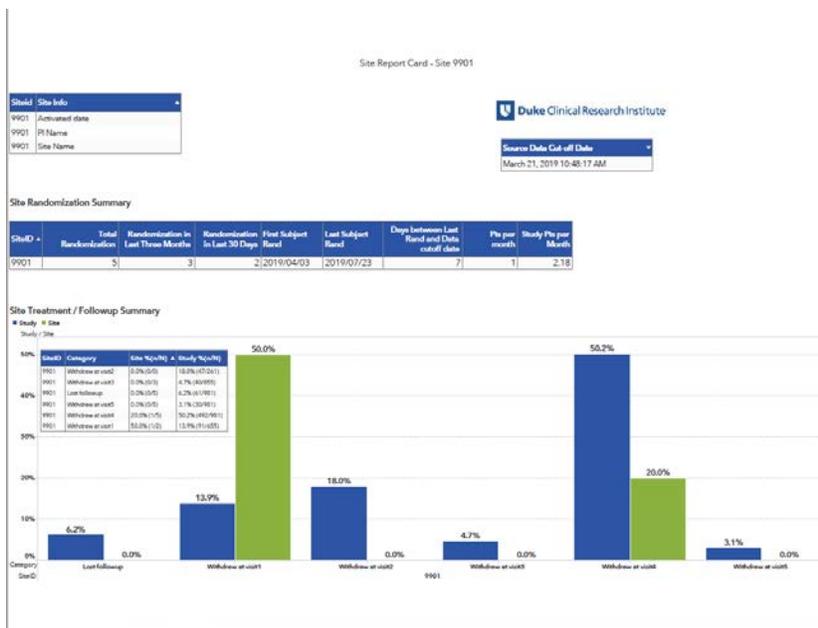
Figure 9b: Patient disposition by treatment for selected disposition



SITE REPORT CARD

The Site Report Card is a widely used overall summary method for monitoring site performance, and it often includes enrollment, patient follow up, and some study specific key risk indicators. As Figure 10 shows, the report starts with the site information (site name, Primary Investigator (PI) name, and site activation date), study logo, and the source data date (the date the underlying data has been harvested). The list table include the site enrollment performance. The site report card is not only providing a summary for monitors, but also for the site to check its performance and contribution to the study. A bar chart is included to show the comparison between the site and the study for selected performance measurements. The site report card is created using the Precise Layout in VA as this allows us to put the summary information exactly where we want it.

Figure 10: Site Report Card



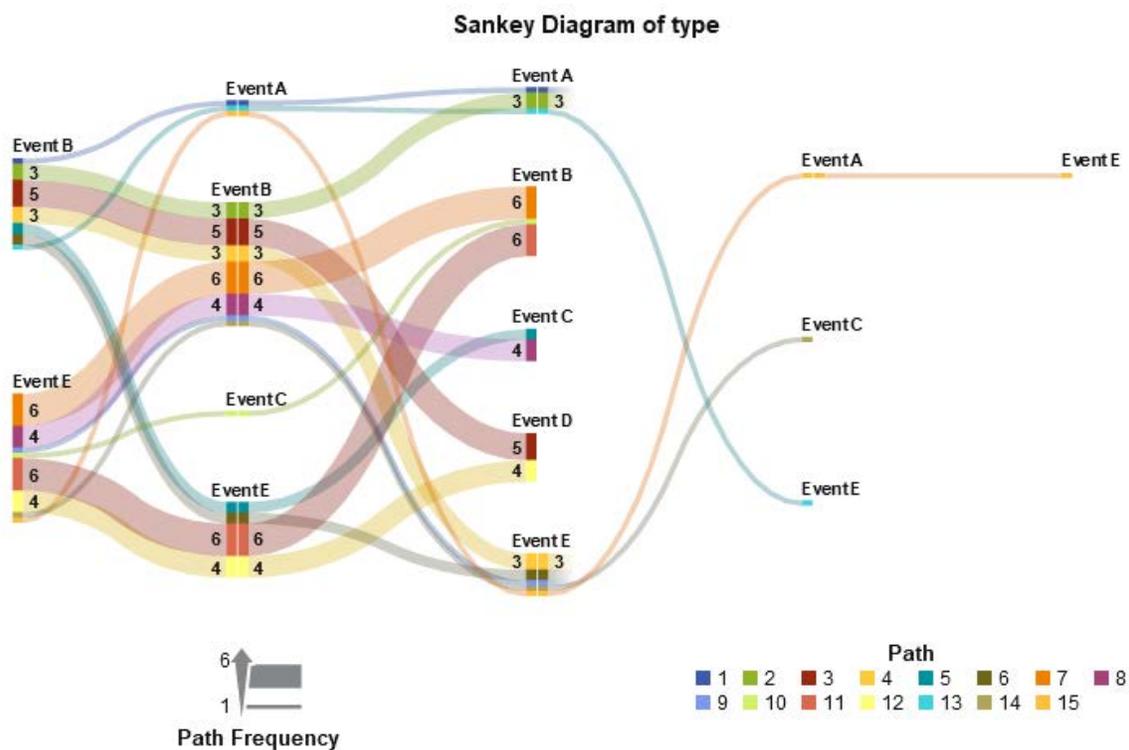
ADDITIONAL DATA EXPLORER DISPLAYS

Data Explorer in SAS VA also provides some data analysis function. We want to share some examples that were very helpful in our studies.

SANKEY DIAGRAM FOR EVENT SUMMARY

A Sankey diagram^[8] enables you to perform path analytics. Path analytics displays flows of data from one event (value) to another as a series of paths. Figure 11, 11a, and 11b shows how to use Sankey diagrams to identify the path for crossover events. In this example, a patient can have 4 possible events, and any changing event from the starting event is a crossover. Figure 11 shows an overall view of all crossovers at all visits, and researchers can easily see that the majority of crossovers starting from Event E ended at Event B. Figure 11a shows crossovers starting at a specific event, and Figure 11b shows crossovers ending at a specific event.

Figure 11: Sankey Diagram for Crossovers



An artificial sequence order was generated for 3 paths that contain simultaneous events.

Filters: None

Ranks: None

CORRELATION MATRIX

A correlation matrix [9] displays the degree of correlation between two intersections of measures as a matrix of rectangular cells. Each cell in the matrix represents the intersection of two measures, and the color of the cell indicates the degree of correlation between those two measures. Figure 12 shows the correlation matrix for 5 vital measures and 1 event variable. The color shows the degree of correlation between those two measures. Figure 12a shows a detailed correlation between two selected variables. Using correlation matrix, it's convenient to study the linear relationship among variables.

Figure 12: Correlation matrix for interested variables

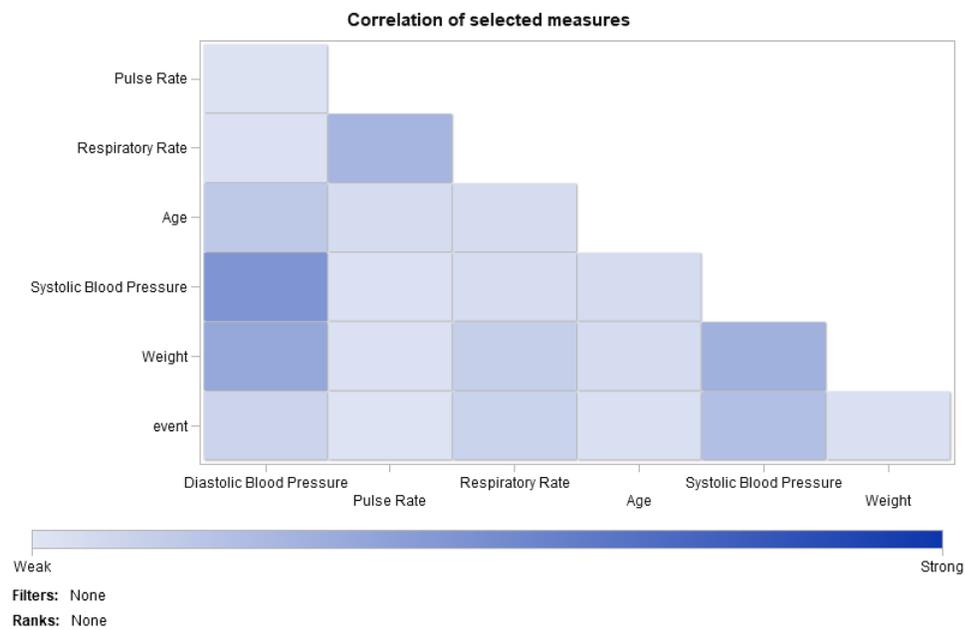
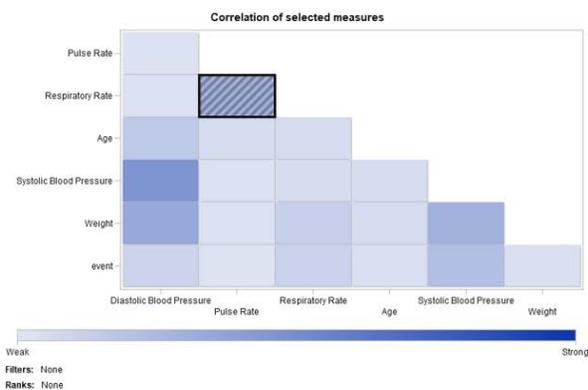
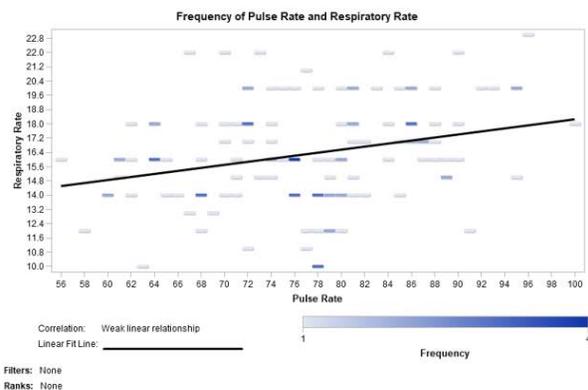


Figure 12a: Correlation of selected measures

Correlation of selected measures



Frequency of selected measures



CONCLUSION

At the DCRI, to better serve our customers as the data coordinating center for many industrial and governmental trials, we wanted to provide more dynamic and user-friendly reports. Implementing SAS VA for operational reports slowed us to do this, and we received a lot of positive feedback from all stakeholders.

It is easy to distribute SAS VA reports. SAS VA reports are web-based, and the reports can be shared as a web link. As long as you keep the same report name, the link will remain the same. For the study team, once we set up the schedule to refresh the reports, we don't need to distribute static reports by emails or web portals anymore. All team members can access the updated report just using the bookmarked website.

It is easy to identify the information of interest due to dynamic nature of SAS VA. SAS VA provides many functions so that information can flow according to logical thinking. We can set up a set of display objects (e.g., graphic, list tables) in a report based on how a series of questions will be answered, and the information can be presented by just a few mouse clicks away. The "Controls" and "Filter" features can subset the reports in any desired sub group; the "Interactions" can link the summary graphs and detailed list tables; the "Hierarchy" can build the drill-down graphs.

It is easy to make eye-catching reports by changing the VA report's cosmetics. SAS VA provides your choice of Font/Color, and various themes for the graphs under "Styles". You can make the reports look modern without complicated programming. The "Display Rules" function also allows for color customization of data points which adds to emphasis of specific data.

It is easy to arrange the objects in a report. Using Precision Layout for a report, you can pin point the objects effortlessly in a report. This is very useful for designing site report card and consort diagram, because both reports require a specific layout for the objects.

So far, we have found that VA reports have been successful in creating dynamic and visually appealing operational reports. In the future, we hope to extend using VA reports in other aspects of clinical reporting such as with data safety and monitoring boards. We also want to make the VA reports more timely by automating the flow of information exchange from source data to the final VA reports.

REFERENCES

1. Friedman L, DeMets D, Furberg C. 2010. Fundamentals of Clinical Trials. New York: Springer-Verlag New York
2. Munsaka M, Zhou K, Kracht K. 2017. "Enhancing Visual Analytics Approaches in Safety Monitoring". JSM – Biopharmaceutical Section: 826-839
3. Shneiderman B, Plaisant C, Hesse B. 2013. "Improving Healthcare with Interactive Visualization". COMPUTER: 58-66
4. Tong M, Hsu W, Taira RK. 2016. "Evaluating a Novel Summary Visualization for Clinical Trial Reports: A Usability Study". AMIA Annual Symposium Proceeding, 2016 – ncbi.nlm.nih.gov
5. Duke J, Li X, Grannis S. 2010. "Data visualization speeds review of potential adverse drug events inpatients on multiple medication". Journal of Biomedical Informatics: 326-331
6. Goldman A. 2012. "Event charts: Visualizing Survival and other Timed-Events Data". The American Statistician: 13-18

7. EQUATOR network. "CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials". <https://www.equator-network.org/reporting-guidelines/consort/>.
8. Chawla V, Luppi R. 2015. "Sankey Diagrams in SAS Visual Analytics". <https://support.sas.com/resources/papers/proceedings15/SAS1808-2015.pdf>
9. SAS Visual Analytics 7.4: User's Guide. <http://support.sas.com/documentation/cdl/en/vaug/69957/HTML/default/viewer.htm#p0gjyd81hbwwrtn1f00ld85hg3rt.htm>

CONTACT INFORMATION

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

Jun Wen
Senior Analyst Programmer
Duke Clinical Research Institute
(919)-668-4074
Jun.Wen@duke.edu

Jack Shostak
Associate Director
Duke Clinical Research Institute
(919)-668-8832
Jack.Shostak@duke.edu