

Statistical Graphics and Reporting in Drug Development

Michael O'Connell, Insightful Corp., Durham, USA

Sven Knudsen, Insightful Corp., Basingstoke, UK

ABSTRACT

With current trends of diminishing drug pipelines, safety concerns and lengthy drug development processes, there is a critical need for efficiencies in drug development. The use of concise, compelling and standardized graphical analyses and reports can help create such efficiencies, while focusing attention on sound, scientific interpretation of clinical data.

In this paper we present some primary elements of statistical graphics for the analysis of clinical data, and a graphics taxonomy whereby the graphics are categorized via data and metadata typing. We then apply the graphics types in context of the major clinical data sources through the analysis of efficacy endpoints, adverse events, laboratory measurements and vital signs. In each of these applications we show how the primary elements may be used to develop compelling graphical analyses and reports for clinical drug development data.

The statistical graphics are based on S-PLUS® Trellis™ and S-PLUS Graphlets™, two of the graphics systems available in the S-PLUS software. S-PLUS Graphlets enable drill-down, brushing, and display of meta-data and are well-suited to exploratory analysis and clinical data review. The report graphics are produced using the S-PLUS Clinical Pack for SAS® Users, which enables the integration of S-PLUS Trellis (and S-PLUS Graphlets) graphics into the SAS environment for production of clinical study reports and/or incorporation of graphics into SAS ODS output.

INTRODUCTION

Clinical trials are complex, representing many years of research, and comprising data from the many dimensions associated with a clinical study e.g. efficacy endpoint measurements, dosing, adverse events, lab measurements, medical history and concomitant medications. Statistical graphics play a crucial role in the interpretation of such complex clinical drug development data by enabling accurate and rapid interpretation of the multifaceted data.

In this paper we present some primary elements of statistical graphics for the analysis of clinical data, and a graphics taxonomy whereby the graphics are categorized via data and metadata typing. We then apply the graphics types through the analysis of safety data including adverse events, laboratory measurements and vital signs. In each of these applications we show how the primary elements may be used to develop compelling graphical analyses and reports for clinical drug development data.

GRAPHICS FOR STATISTICAL ANALYSIS, DATA REVIEW AND STUDY REPORTING

In clinical trial research three key areas are (i) exploratory data analysis (EDA) of clinical data, (ii) informal review of results by clinicians and management, and (iii) formal clinical study reports, scientific papers and presentations. In exploratory analysis and data review, the goals are to quickly extract, display and review key safety and efficacy information in the data. In these settings there are minimal formatting restrictions and it is possible to use interactive plots to help facilitate interpretation. In clinical study reports the graphics must adhere to regulatory (e.g. FDA) reporting standards and company style guides. For example, strict standards regarding headers, footers, fonts, symbols, lines, axes, legends, annotations and the graphic file itself are typically enforced.

STATISTICAL GRAPHICS PRINCIPLES

Excellence in statistical graphics consists of complex ideas communicated with clarity, precision and efficiency (Tufte, 1983). The goal is to communicate the salient information in the data and the channel is reader perception of data values and data patterns. It turns out that both accuracy-of-data-presentation and speed-of-interpretation are required for accurate and retained information extraction.

Pattern perception from a graph involves three key tasks: (a) detection – recognition of geometry, (b) assembly – grouping of detected symbols; and (c) estimation – assessment of relative magnitudes. Some estimation tasks are easier than others for the human brain. The following tasks are listed in order of ease

PhUSE 2006

- + Position along common scale [most accurate]
- + Position along identical non-aligned scales
- + Length
- + Angle and slope [pie charts are hard to interpret!]
- + Area, Volume
- + Color: hue (R, G, B), saturation (pale, deep) - hue can give good discrimination but poor ordering – can't process many hues [most difficult]

Tufte (1983, and in his widely attended seminars) espouses key graphics principles for accurate and retained information extraction:

- + Show all the data
 - + Many numbers in small space – multivariate data and metadata
- + Induce the viewer to think about the substance rather than the graphic design – maximize the data-to-ink ratio
- + Avoid distorting what the data are saying
- + Make large data sets coherent
- + Encourage the eye to compare different pieces of data - leverage investment by showing multiple plots of same type
- + Reveal the data at several levels of detail, from a broad overview to the fine structure
- + Serve a clear purpose: description, exploration, tabulation
- + Be closely integrated with the statistical and verbal descriptions of a data set
- + Use gray scale and color sparingly

Applying these graphics principles to clinical analysis results in the following tips for comparative graphical analysis and presentation:

- + Rotate data when curves are steep [Bland-Altman: $(Y-X) \vee (Y+X)/2$, rather than $Y \vee X$]
- + Set a sound aspect ratio [e.g. banking to 45 degrees]
- + Include (feint) grid lines and reference lines as appropriate
- + Sort categories e.g. on dot plots, Pareto plots (see below)
- + Use group and multi-panel displays (with same scale)
- + Minimize use of legends by annotating the graph
- + Don't use more than 3 hues
- + Use contrast hues (blue-orange/yellow); avoid color-blind hues (red-green)
- + Use black and gray for 2 group comparisons (treatment and control)
- + Don't use too many dashed line types – they are difficult to distinguish

S-PLUS TRELLIS GRAPHICS

S-PLUS Trellis graphics (Cleveland, 1993, Becker et al., 1996) embody the graphics principles outlined above. While the name "Trellis" refers to the layout of graphs in panels resembling a garden trellis, the depth and power of Trellis graphics lies in their simple model-based specification and their internal object-oriented nature.

Trellis graphs are objects with attributes and methods. In S-PLUS, Trellis graphs are specified like a statistical model; for example just like a linear model is specified as $\text{lm}(y \sim x \mid z)$, a Trellis scatter plot is specified as $\text{xyplot}(y \sim x \mid z)$. In both specifications z is a conditioning variable; and can represent combinations of conditioning variables, each with multiple levels. As such, Trellis graphs are particularly helpful in representing multivariate data. Trellis graphics leverage a reader's investment in interpreting one graph, by showing multiple graphs of same type. Trellis graphics provide a powerful mechanism for understanding and interpreting relationships between and within subgroups and subjects in the graphical analysis of safety and efficacy endpoints. For single-panel and multi-panel graphs alike, Trellis graphics provide a solid framework for enforcing consistent styles between and within graph types; with the graph type collection readily managed as a library of Trellis panel functions.

STATISTICAL GRAPHICS – SOME PRIMARY ELEMENTS

Clearly a graph has an x-variable a y-variable, axis labels, axis tick marks and labels, a title/header and a footer. Graphs also often include reference lines, legends and free text annotations. A well-formed statistical graphic typically includes some of the following primary elements:

- + Grouping Variables - symbols and colors to show groups in the data e.g. Treatment v Placebo
- + Trellis Plots and Conditioning Variables - panels for subjects or sub-groups in the data e.g. adverse event preferred terms or body systems
- + Clinical Metadata – in addition to the primary endpoint being plotted (y-variable), data on supporting

PhUSE 2006

variables can provide a richer picture of the data e.g. the use of ticks and regions to communicate adverse events on an efficacy endpoint graphic.

- + Graphics Drilldown – in exploratory analysis (as opposed to clinical study reporting) it is convenient to browse and review with drilldown from one graph to another e.g. in the sequence of Population Lab to Subject Lab to AE to demographics

These primary elements are used in the examples provided below.

TAXONOMY OF STATISTICAL GRAPHICS

It is convenient to organize statistical graphics in terms of the data and metadata they support. We organize statistical graphics in the following categories:

- + Summary statistics (univariate categorical variables)
- + Distributions (univariate continuous variables)
- + Relations: Continuous v Continuous
- + Relations: Continuous v Categorical
- + Relations: Categorical v Continuous
- + Relations: Time and Survival

Examples of statistical graphics for display of safety and efficacy data in these categories include:

- + Summary statistics
 - + Dot plot
 - + Barchart [data-to-ink ratio can be low]
 - + Pie chart [hard to interpret angles]
- + Distributions
 - + Box plot
 - + Empirical Cumulative Distribution Function (eCDF) plot
 - + Density plot
 - + Histogram
 - + QQ plot
- + Relations: continuous v continuous
 - + Line plot (means, s.e.'s, c.i.'s, raw data)
 - + Risk plot
 - + Scatter plot
 - + Scatter plot matrix
- + Relations: continuous v categorical
 - + Box plot
 - + Line plot
 - + Surface, contour and image plots
- + Relations: categorical v continuous
 - + Dot plot
 - + Forest plot
- + Relations: time and survival
 - + Kaplan Meier plot
 - + Event chart
 - + Shift plot
 - + Delta plot

Some of these graphs are described in Heiberger and Holland (2004).

GRAPHICS FOR PRESENTATION OF SAFETY DATA

There has been much recent work on development of statistical methodology for clinical trial design, power, and analysis methods for efficacy endpoints. In contrast, safety data are collected as concomitant information, analyzed as an afterthought, and reported as simple tables and listings. Analysis of adverse event rates, for example, is typically done without much attention to basic statistical issues such as multiple comparisons and without much common sense regarding the adverse events' practical significance. Adverse event data are often presented in tables with basic p-values and then much effort is spent explaining why adverse events with significant p-values are not important. In short, the current approach to safety data analysis in clinical trials is rudimentary at best.

PhUSE 2006

The consumers of safety data analysis and clinical study reports need to rapidly and accurately interpret the safety content and have short available time windows to do so. DSMBs typically have one-day meetings, and paging through tables is inefficient. Regulatory agencies such as the FDA want transparency of analysis and consistency of results; they are sensitive to any gaming of results through opaque modeling. Finally, internal stakeholders such as sponsors, CROs and partners want fast and accurate interpretation of the safety data.

Statistical graphics, combined with simple summary statistics such as relative risk or attributable risk, are well suited to the analysis of safety data. Statistical graphics can present safety data clearly and unambiguously. In exploratory data analysis using Trellis graphics, there is little chance of missing a signal because all of the data can be displayed efficiently. In clinical reporting using Trellis graphics, a well-formed statistical graph can communicate the safety message in the data clearly and quickly. Such graphics can anchor the clinical study report from the safety perspective.

In what follows, we present some graphical analyses of adverse events, labs (liver) and vitals (QT prolongation) using S-PLUS Trellis graphics from the S-PLUS Clinical Reporting Solution™.

ADVERSE EVENTS

The primary goals of adverse event analysis include identifying which adverse events may be elevated in treatment vs. placebo, and the rapidity of their onset in treatment vs. placebo. These are inferential questions relating to treatment effects and patterns; as such, this is a population level analysis.

We illustrate some graphical analyses using data from a vaccine trial in babies from Mehotra and Heyse (2001). Figure 1 shows a simple *grouped dot plot* comparing adverse event rates in treatment and control; created using the Trellis graphics library in S-PLUS and the S-PLUS Clinical Reporting Solution.

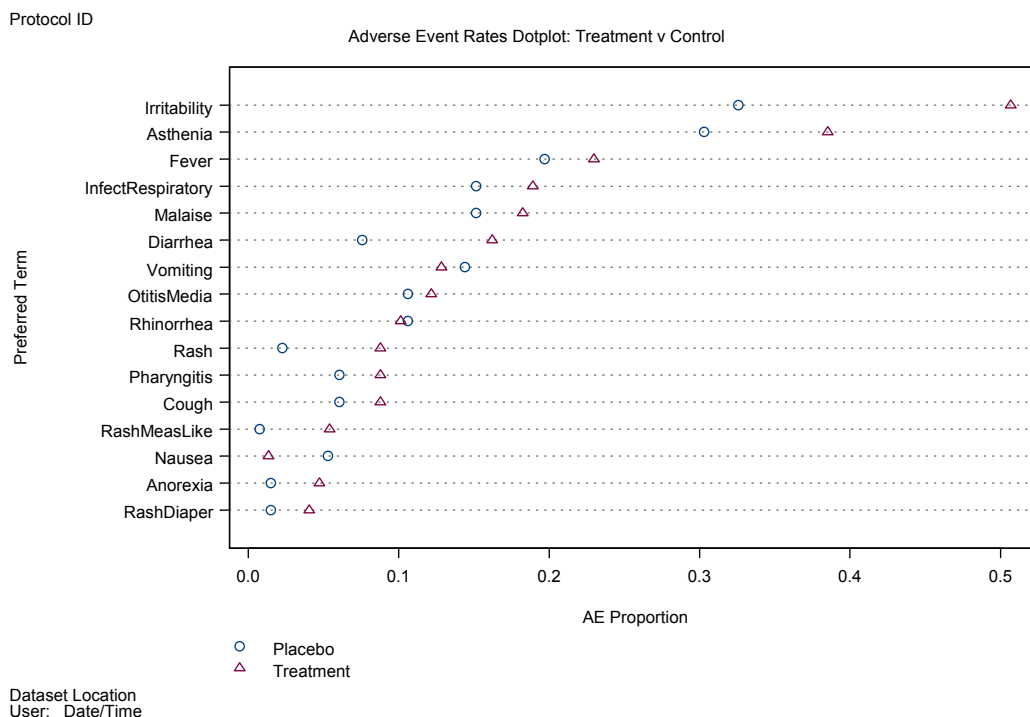


Figure 1: Dot plot of adverse rates in vaccine trial (Mehotra and Hayes, 2001); grouped on Treatment (blue circle = Placebo, red triangle = Treatment).

Figure 2 shows a *grouped risk plot* for the Mehotra and Heyse data; created using the Trellis graphics library in S-PLUS. This graph was first introduced in the analysis of adverse event data by O'Connell (2006) and shows the relative risk or ratio of the adverse event rates on the x-axis and the p-value comparing treatment and control on the y-axis. The additional information on the p-value of the treatment effect is important since it incorporates the number of observed events and confidence in the treatment effect. As shown in the figure, Anorexia has a 3.5X risk but only 9 observed events. Diarrhea on the other hand has a 2X risk, but with 34 observed events is a stronger potential signal than Anorexia. Variations on this plot include use of the risk difference rather than ratio on the x-axis; this has the nice interpretation of attributable risk i.e. the proportion of risk attributable to the treatment. This can be used to provide a simple estimate of the number at risk in an at-large population.

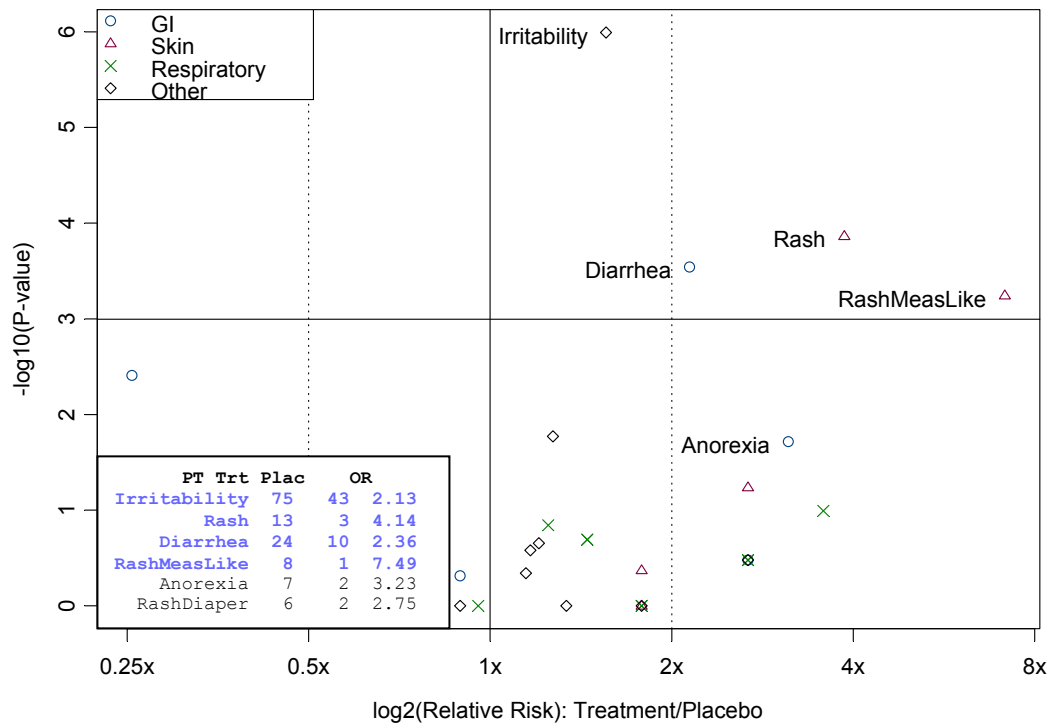


Figure 2: Risk plot of adverse event rates in a vaccine trial (Mehotra and Hayes, 2001); grouped on system organ class (SOC).

LABORATORY MEASUREMENTS

The primary goals in the analysis of lab data such as liver panels, include (a) identifying which subjects have elevated (liver) labs, (b) identifying subjects with elevation on multiple labs, and (c) incorporating adverse event information for subjects with elevated lab values. While patterns of effects on lab values across time and treatment are of interest, subject level analysis is important.

Figure 3 shows a *grouped scatter plot matrix* for a panel of three liver lab measurements; created using the Trellis library in S-PLUS and displayed as an interactive S+Graphlet using the java.graph device in S-PLUS. The plot is grouped on treatment with blue and red symbols representing two doses of drug. Variations on this plot include the use of upper limit of normal (/ULN) rather than raw scale. Such a scale has the advantage of easy comparability and associated reference to phenomena such as the Hi effect.

Figure 4 shows the same grouped scatter plot matrix for the panel of three liver lab measurements. In this plot a subject with elevated labs is highlighted. This particular subject has elevated results on all liver labs. The highlighting is achieved by clicking on a point, and the metadata in the top right hand corner of the graph is obtained through mouseover.

Additional clicking on this graph leads to a patient-level graph of the same lab measurement over time which also shows any adverse events. Such a sequence is typically concluded with a drill-down to patient-level demographics and medical history.

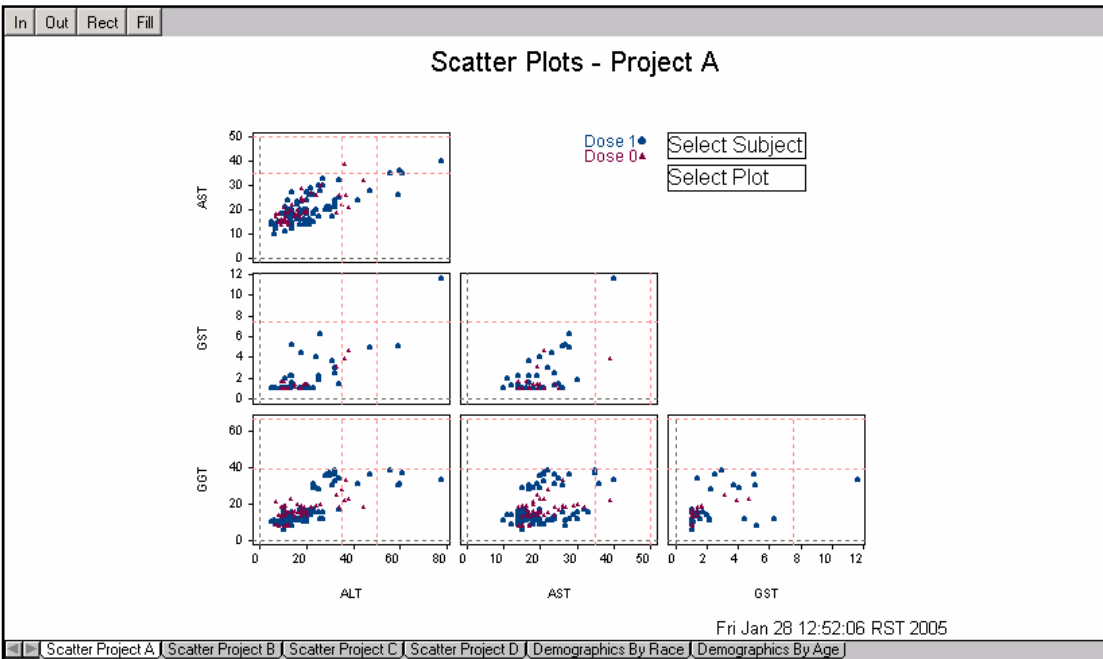


Figure 3: Scatter plot matrix for liver panel; grouped on Treatment (blue and red points are different doses). Reference lines are included at clinically meaningful values. An alternate form of this graph uses values normalized to their upper limit of normal (ULN).

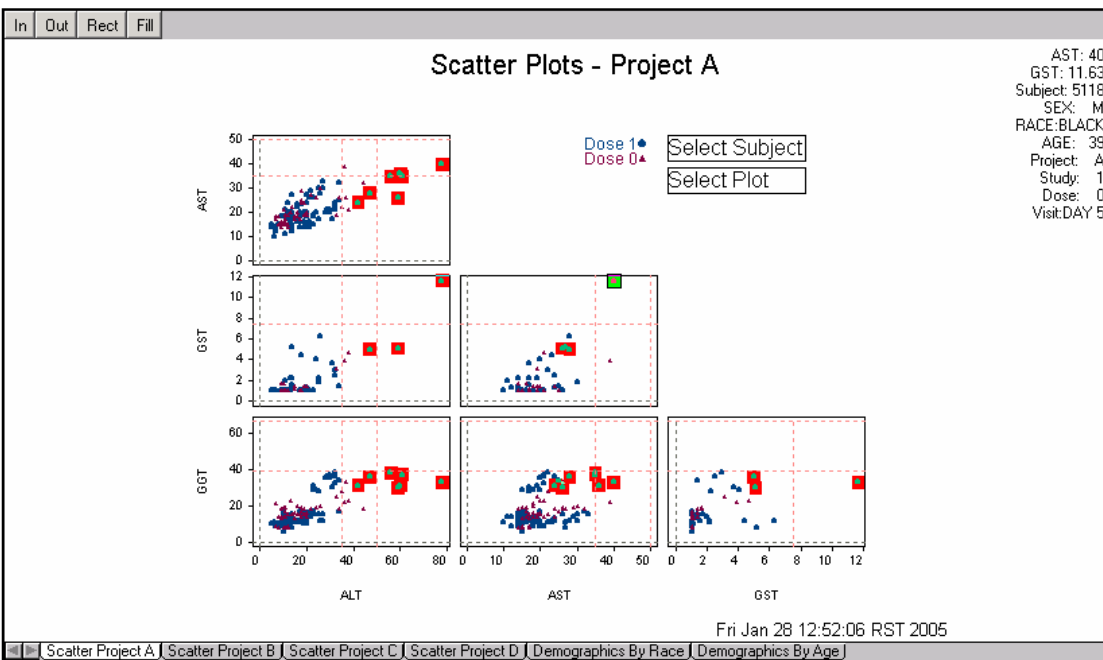


Figure 4: The same scatter plot matrix for liver panels in Figure 3; with an individual subject's values highlighted. Note the metadata about the highlighted subject in the top right hand corner of the graph.

VITAL SIGN MEASUREMENTS

As outlined in the recent FDA guidance (US Dept. Health and Human Services, 2004) on clinical evaluation of QT/QTc prolongation, some goals in the analysis of QT intervals include assessment of (a) elevation of corrected QT (QTc) for the drug treatment, with 450, 480, 500 ms providing reference values of interest, (b) elevation of QTc v BaseQTc, for the drug treatment, of 30 or 60 ms, and (c) elevation of QTc(Drug) v QTc(Placebo) of 5 or 10 ms. Analysis of vital sign data involves both subject level and population level analysis, referred to as categorical and central tendency analysis in the guidance.

Figure 5 shows a line plot of mean QTc, grouped on treatment. Guidance values of 450, 480 and 500 ms are shown as dotted reference lines. Note the confidence intervals around the mean are staggered to prevent overlap and allow assessment of both the point and interval estimates for the two treatments in the study.

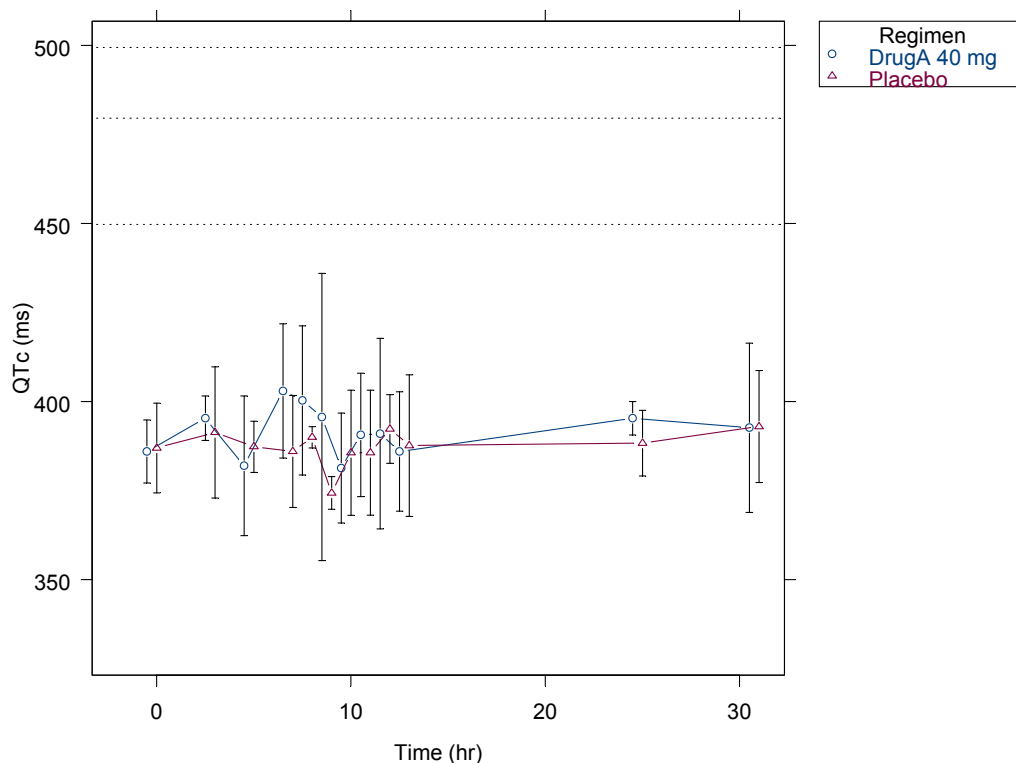


Figure 5: Grouped line plot of QTc; grouped on treatment. Confidence intervals are staggered so as not to overlap. Reference lines are added at 450, 480 and 500 ms in accordance with the FDA guidance.

An alternative to the graph in Figure 5 is a grouped boxplot, showing quantiles of the empirical distributions rather than confidence intervals. One advantage to such a plot is that subject outliers are included and can be assessed and compared to the reference lines. Additional drill down on any outliers could link to metadata regarding AEs, labs and/or demographic information.

Figure 6 shows a scatter plot of QTc delta, grouped on treatment. Guidance values of 30 and 60 ms are shown as dotted reference lines. Points represent clinical trial subjects. Note that there are two values in the first panel (drug treatment) with potentially prolonged QTc. Drill down and metadata on the elevated points on this graph could provide additional relevant subject level information e.g. on labs and/or AEs.

An alternative to the graph in Figure 6 is an empirical CDF plot of the delta QTc, grouped on treatment. One issue to consider in all QTc and lab value plots is whether to include all observations or just the maximum (or maximum delta).

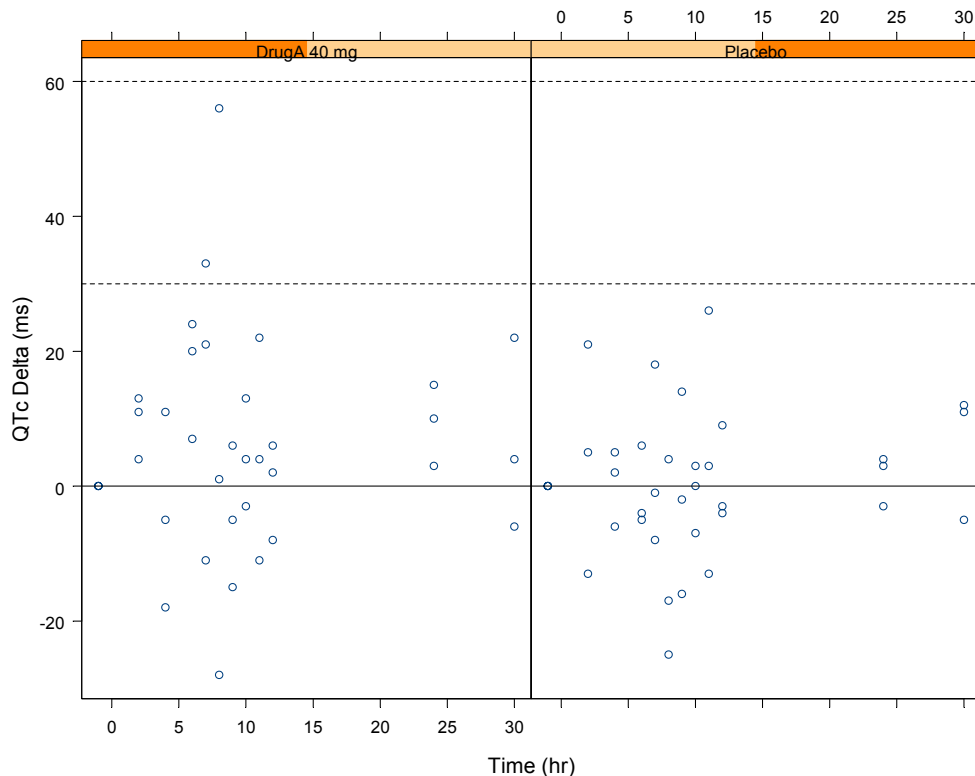


Figure 6: Conditioned scatter plot of QTc Delta; conditioned on treatment. Reference lines are added at 30 and 60 ms in accordance with the FDA guidance.

DEPLOYMENT ENVIRONMENTS FOR STATISTICAL GRAPHICS

We have described the richness and power of S-PLUS Trellis graphics to convey the scientific information in clinical studies. We now turn our attention to deployment of S-PLUS Trellis graphics. Insightful offers three S-PLUS client environments: a point-click Windows GUI, an Eclipse IDE and a Batch utility; and two server environments: Insightful Analytic Server and Insightful Reporting Server. For clinical graphics and reporting there are two relevant Insightful applications: the S-PLUS Graphical Workbench Environment (S+GWE) which plugs in to the point-click S-PLUS Windows GUI; and the S-PLUS Clinical Pack for SAS Users (S+ClinPack) which plugs in to the Insightful Analytic Server. Together, S+GWE and S+ClinPack form the S-PLUS Clinical Reporting Solution. This Solution enables graph specification and exploratory data analysis in the Windows desktop environment, and graph publishing for data review and (compliant) clinical study reporting in the UNIX server environment.

The S+ClinPack provides an efficient way to produce validated S-PLUS graphical summaries from within the SAS environment. In practice, a simple macro is submitted in SAS, and the S-PLUS graph requested is automatically generated based on SAS data. The plot is then either inserted into an ODS report or written to a file on the file system. The S-PLUS (verbose) log may be either inserted into the SAS log or managed separately. This framework provides a flexible and rigorous approach for advanced graphical reporting that satisfies SOP requirements in both development and production environments.

S-PLUS GRAPHICS FROM WITHIN THE SAS ENVIRONMENT – THE S+CLINPACK

An example of SAS ODS with embedded S-PLUS graphics is displayed in Figure 7. The graph is based on simulated Albumin (G/DL) measurements for two treatment groups, Placebo and A mg, at days 0, 10 and 20 post treatment. There are 20 patients in the study, 10 in each treatment arm. Using the XYPLOTTRELLIS macro from the S+ClinPack, the user can quickly produce the scatter plot of Albumin measurements by time for each patient group; the code is given below.

Similar to this conditioning on treatment group, one can condition on types of measurement. In this way we can obtain displays of, say, Albumin, Alkaline Phosphate, ALT, AST, Blood urea nitrogen by lab center – all in a single graph. The conditioning is also very useful in model diagnostics since we can condition residual plots on treatment groups, or demographic groups, and quickly obtain insight into variability patterns.

The S+ClinPack works seamlessly with the SAS ODS HTML, PDF, PS, WMF and RTF graphic formats and includes macros for all of the Trellis functions in S-PLUS including scatter plots, dot plots, box plots, histograms, bar and pie charts, level plots and surface plots. These graphics are suited to data presentation and model diagnostics applications.

```

/* Scatter plot of Albumin by time conditioned on treatment group */
FILENAME report &reportFile;
ODS HTML BODY=report;
%XYPLOTTRELLIS(dataLibrary = mylib,                /* SAS libname      */
               dataSet = randata,                  /* SAS dataset name */
               x = time,                           /* Explanatory var   */
               y = response,                       /* Response var      */
               condition = trt,                    /* Conditioning var  */
               graphTitle = 'Albumin by Treatment', /* Plot main title   */
               xAxisLabel = 'Time Post Treatment (days)', /* X-axis label     */
               yAxisLabel = 'Albumin (G/DL)',        /* Y-axis label      */
               outputFilePrefix = simpleScatter);    /* Output file name  */

ODS HTML CLOSE;

```

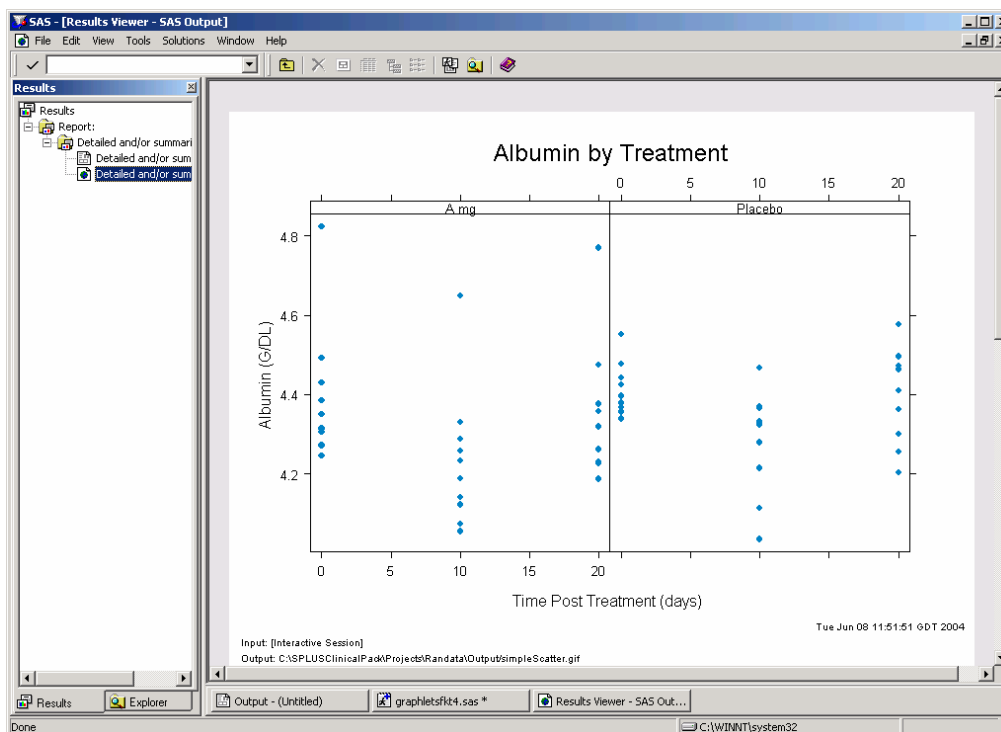


Figure 7: SAS Results Viewer showing plot of Albumin by treatment group. The graph has been stamped with time, session mode, and output destination. The SAS Results left-pane is updated with a link to the graph for easy access in the SAS session.

CUSTOM TRELLIS GRAPHICS FROM WITHIN SAS

The S+ClinPack includes all of the Trellis graphics, as illustrated above. S+ClinPack can also be used to run custom S-PLUS graphics through the SAS environment. For example, one may include fitted curves, annotations, and custom plot types within the Trellis panels. Panel functions can be included directly in the SAS Macros that implement the Trellis graphics in the S+ClinPack. For example in the XYPLOTTRELLIS macro, creating a new plot takes only a few lines of SAS code. The S+ClinPack collection of macros makes it easy to handle data, create a graph, write a graph file to the file system in a variety of formats, and/or include a graph in a SAS ODS report.

The following SAS macro shows an example of using a panel function to customize a plot. This example produces a scatter plot with a regression line of orthodontic distance versus age for each subject. The panel function and the call to panel.lmline() is responsible for creating the regression line on the plot. In this example you can see that it takes very little programming to add advanced details to plots within the SAS environment.

```

/* Create a scatterplot with regression line */
/* Of orthodontic distance vs. age, conditioned by subject */
/* Making use of PANEL function */
%XYPLOTTRELLIS (dataLibrary = mylib,                /* SAS libname */
               dataSet = pr,                        /* SAS dataset */
               x = age,                             /* Explanatory var */
               y = y,                               /* Response var */
               condition = person,                  /* Conditioning var */
               graphTitle = Orthodontic Distance - plot for each person, /* Plot title */
               xAxisLabel = Age [year],             /* X-axis label */
               yAxisLabel = Distance [mm],          /* Y-axis label */
               panel = function(x,y,...) {;          /* Panel function */
                   panel.xyplot(x,y);
                   panel.lmline(x,y) ;;,
               /* (Add line) */
               outputFilePrefix = &reportFile);      /* Output file name */

```

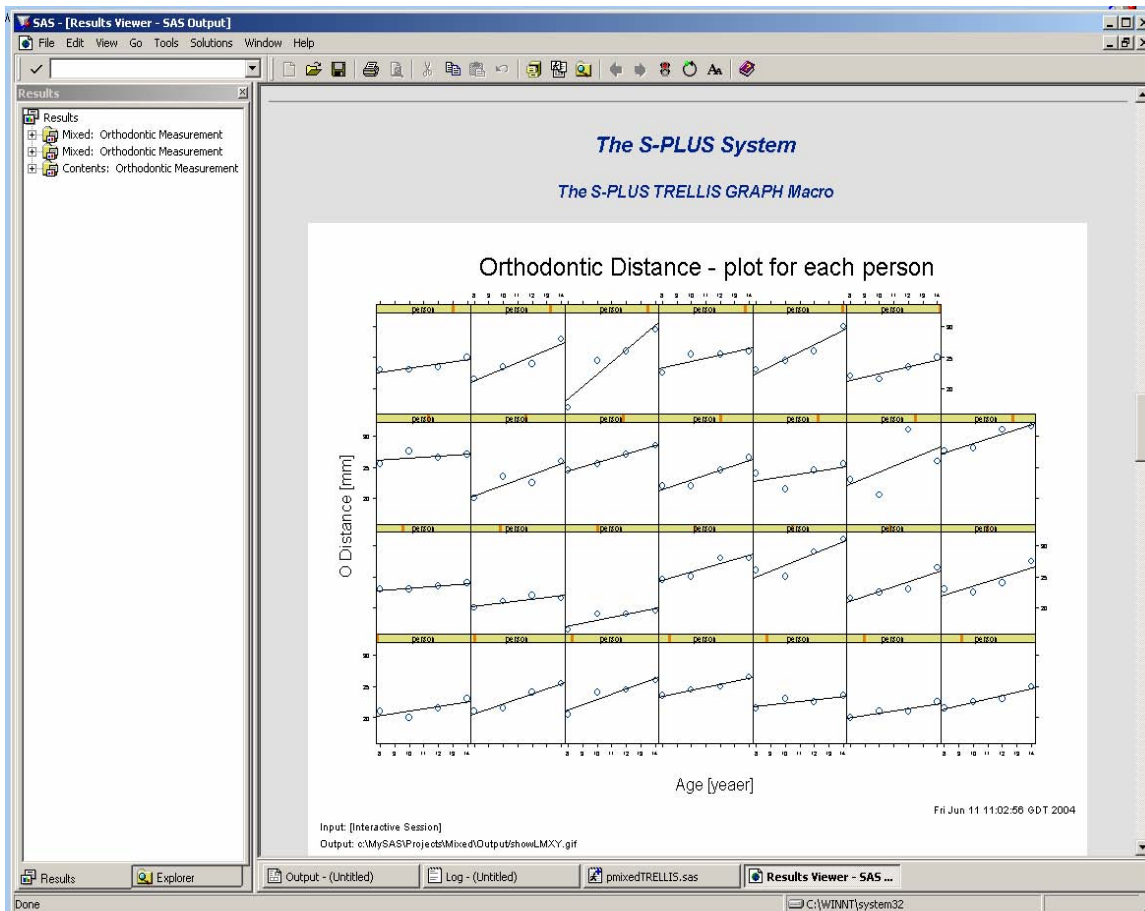


Figure 8: SAS Results Viewer showing plot of Orthodontic Distance by Subject with fitted lines on each panel. The graph has been stamped with time, session mode, and output destination. The SAS Results left-pane is updated with a link to the graph for easy access in the interactive session.

INTERACTIVE S-PLUS GRAPHLETS FROM WITHIN SAS

S+Graphlets provide interactive viewing and drilldown for data browsing and review. S+Graphlets use exactly the same base graph code as regular S-PLUS (Trellis) graphs, but the graphs are presented as lightweight java applets that are created by simply changing the graphics output device (to `java.graph()`). Additional S-PLUS commands are used to form the drill-down and menus e.g. `java.identify()` and `java.menu.action()`.

PhUSE 2006

S+Graphlets may be (a) interactively viewed in a browser, (b) moused-over to show metadata, (c) clicked for drill-down to additional graphs or data listings e.g. patient demographics, medical history, and (d) simply customized to add menus and drill-down from menus of hot-links for graphs, tables and/or metadata. S+Graphlets do not require the end-user to have access to an S-PLUS license or environment for viewing or interacting with the graphs. S+Graphlets provide a comprehensive graphical view into the clinical trial data and provide immense value for clinicians and other stakeholders in the review of clinical data. The following example SAS macro creates a Graphlet to be deployed within the SAS ODS system.

```
%BoxCompareSPLUSGRAPHLET(dataLibrary = mylib,  
                           dataSet = randata,  
                           taskName = lftglet,  
                           value = response,  
                           visit = time,  
                           group = trt,  
                           protocol = Blinded,  
                           population = Post Treatment,  
                           userID = IFUL,  
                           reportFile = &reportFile);
```

The resulting graph is displayed in Figure 9. The box plot shows the distribution of Albumin over time by treatment group. It also includes a total over time. The points on the plot represent the maximum value of Albumin for a particular time point. Figure 9 shows a menu item that has been embedded in the plot. The menu item is connected to one of the outlier points and is invisible until the user right clicks on the point.

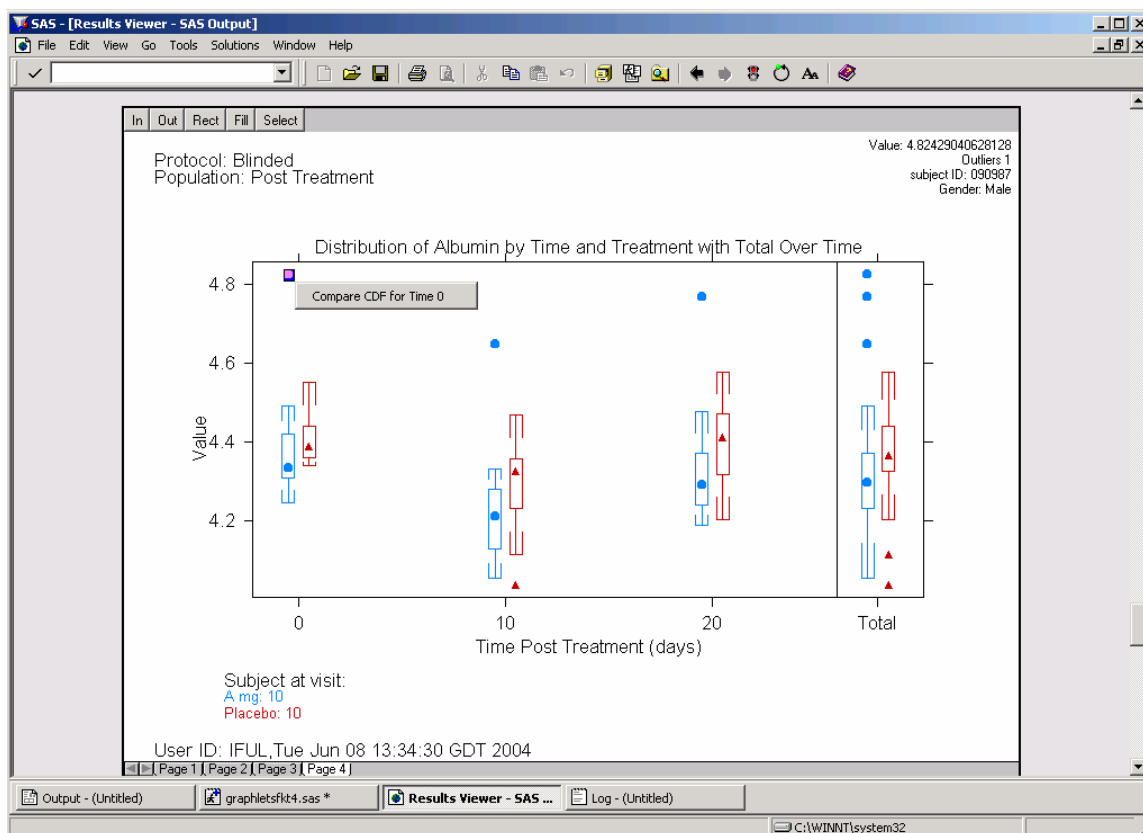


Figure 9: SAS Results Viewer showing an S+Graphlet containing a box plot of the distribution of Albumin by time and treatment with total over time. You can see that the point in the upper left hand corner is selected and allows you to view a CDF plot for that time point.

PhUSE 2006

When the user selects the menu item shown in Figure 9 it displays a new graph. This new graph, shown in Figure 10, displays a cumulative distribution of VALUE for time zero. By attaching menu items to these plots we can effectively create drill down graphics that will allow the end-user (e.g. clinician) to connect associated graphs in very flexible ways.

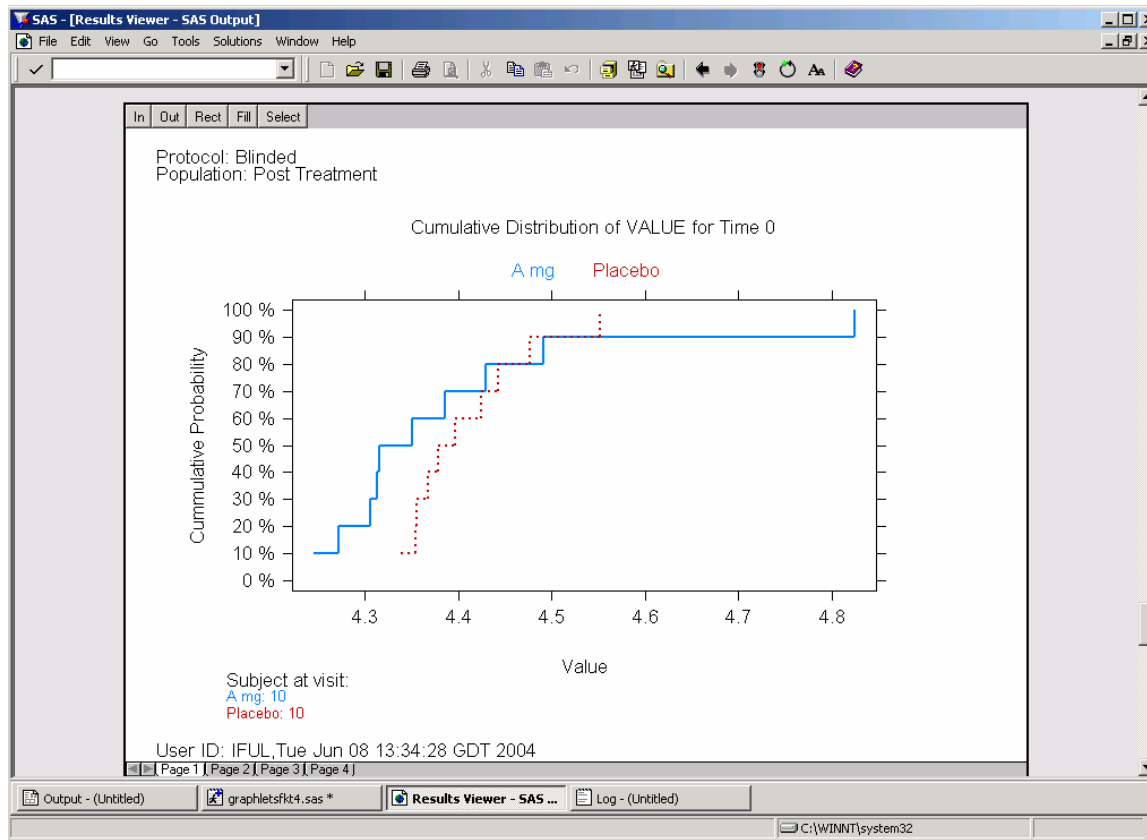


Figure 10: SAS Result Viewers showing an S+Graphlet containing a CDF plot of VALUE for time 0. This plot was viewed by selecting the menu item displayed in Figure 9.

CONCLUSION

Statistical graphics are core functionality to the efficient analysis and interpretation of clinical / safety data. Statistical graphics (a) enable rapid and accurate comparative analysis of treatment effects, (b) enable exploratory analysis and understanding of safety and efficacy data, (c) enable data browsing and clinical data review, (d) can anchor the safety and efficacy message in clinical study reports, and (d) provide clear and effective presentation of study results in scientific and marketing applications.

Consistency in the deployment of statistical graphics principles and standards is key to the value created by statistical graphics. Considerable efficiencies are created throughout an organization when key statistical graphs are specified consistently for clinical study reports, data browsing and review, and scientific presentation. S-PLUS Trellis graphics provide such consistency both within and between graph types.

SAS and S-PLUS bring unique strengths to clinical data analysis. SAS offers robust data management capabilities and well-accepted analysis methods. S-PLUS is known for its publication quality graphics and modern statistical methods. For these reasons, many clinical data analysts use both SAS and S-PLUS in their daily work. As a compliment to this existing environment, Insightful has developed the S-PLUS Clinical Reporting Solution, including the Graphics Workbench Environment on Windows (S+GWE) for graph specification and exploration and the S+Clinpack on UNIX for creation of graphics for clinical study reports. This S-PLUS Clinical Reporting Solution provides an easy way for SAS users to access the powerful graphical methods in S-PLUS without having to learn the S-PLUS language.

With the S+Clinpack, the S-PLUS graphics are either inserted into an ODS report or written to a file e.g. a clinical

PhUSE 2006

study report template. The S-PLUS (verbose) log may be either inserted into the SAS log or managed separately. Using the S+Clinpack, a SAS programmer can easily integrate Trellis graphics, customize these graphics, and create drill-down S-PLUS Graphlets with meta-data, all from within SAS. S-PLUS analyses are easily validated against SAS by writing relevant data values to the log and output files, like many SAS analyses are validated.

In the not-too-distant future, we see a world where statistical graphics are widely used to anchor the message from clinical studies; where grouping, conditioning, safety/efficacy metadata and drill-down are commonly used techniques; where statistical graphics plug in to standardized clinical data sources (e.g. CDISC) seamlessly; where statistical graphics are suggested based on data types using the clinical graphics taxonomy described herein; and where a statistical graphics palette is standardized across industry and the FDA. We invite you, the reader of this article, to be part of this world.

REFERENCES

- Becker, R.A., Cleveland, W.S., and Shyu, M.J. (1996): The visual design and control of Trellis displays. *JCGS*, 5 (2): 123--155
- Cleveland, W. (1993): *Visualizing Data*. Hobart Press, Summit, NJ
- Heiberger, R. and Holland, B. (2004) *Statistical Analysis and Data Display: an Intermediate Course with Examples in S-PLUS, R, and SAS*. Springer-Verlag, NY. <http://springeronline.com/0-387-40270-5>
- Mehotra, D.V. and Heyse, J.F. (2001). Multiplicity considerations in clinical safety analysis. *Statistical Methods in Medical Research* 13, 227-238.
- O'Connell, M. (2006). Statistical graphics for design and analysis of clinical development studies. Insightful webcast. http://www.insightful.com/news_events/webcasts/2006/07clinical
- Tufte, E.R. (1983). *The Visual Display of Information*, Graphics Press.
- U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research Guidance (2004). The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic potential for Non-Antiarrhythmic Drugs

ACKNOWLEDGEMENTS

We would like to acknowledge the important contributions made by Insightful colleagues in developing the S-PLUS Clinical Reporting Solution and its components the S-PLUS Graphical Workbench Environment (S+GWE) and the S-PLUS Clinical Pack for SAS users (S+ClinPack). In particular Peter McKinnis, Robert Collins, Patrick Aboyoun, Bob Treder, Chris Thornton and Chris Wiggins have all made important contributions to these Insightful solutions.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. The contact author can be reached as follows:

Michael O'Connell
Insightful
2505 Meridian Pkwy Suite 175
Durham NC 27713
919-5725545x3006
moconnell@insightful.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.

S-PLUS and all other Insightful Corp. product or service names are registered trademarks or trademarks of Insightful Corporation in the USA and other countries.

® Indicates USA registration.

Other brand and product names are trademarks of their respective companies.