



Waterfall Plot with Heatmap

Sandhya Yalla, Statistical Analyst II,
Chennai, 19Jul2014



The views and opinions expressed in the following slides are those of the individual presenter and should not be attributed to any organization with which the presenter is employed or affiliated

Agenda



- Background
- Example
- Purpose
- Interpretation
- Procedures
- Challenges & Conclusion

Background : Waterfall Plot



- Waterfall graphs is a data visualization technique, normally used for understanding how an initial value is affected by a series of intermediate positive or negative values.
- The first formal description of waterfall model article is by Winston w.royce in 1970 and then Waterfall charts were popularized by the strategic consulting firm McKinsey & Company .
- In oncology clinical trial data analysis, this can be used for performance or response analysis, especially for understanding or explaining the overall tumor response of a parameter which can vary with the effect of multiple factors such as change from baseline.

Background : Heat Map



- A heat map is nothing more than a graphical representation of data where the individual values are represented as a matrix of colors.
- Software designer Cormac Kinney originally coined and trademarked the term "heatmap" in 1991 to describe a 2D grayscale display of financial data.
- A simple heatmap provides an immediate visual summary of information. More elaborate heat maps allow the viewer to understand complex data sets.
- Heatmap use cell size and color to display complex information in an **intelligible way**.



- The example below illustrates how the values in the table are displayed as color gradients in the heat map cells.

Test	Time 1	Time 2	Time 3
Test A	2.02	3.21	5.57
Test B	2.92	4.37	6.02
Test C	2.64	5.02	7.19
Test D	2.37	3.48	8.21
Test E	2.21	3.12	5.38
Test F	2.43	8.84	6.47

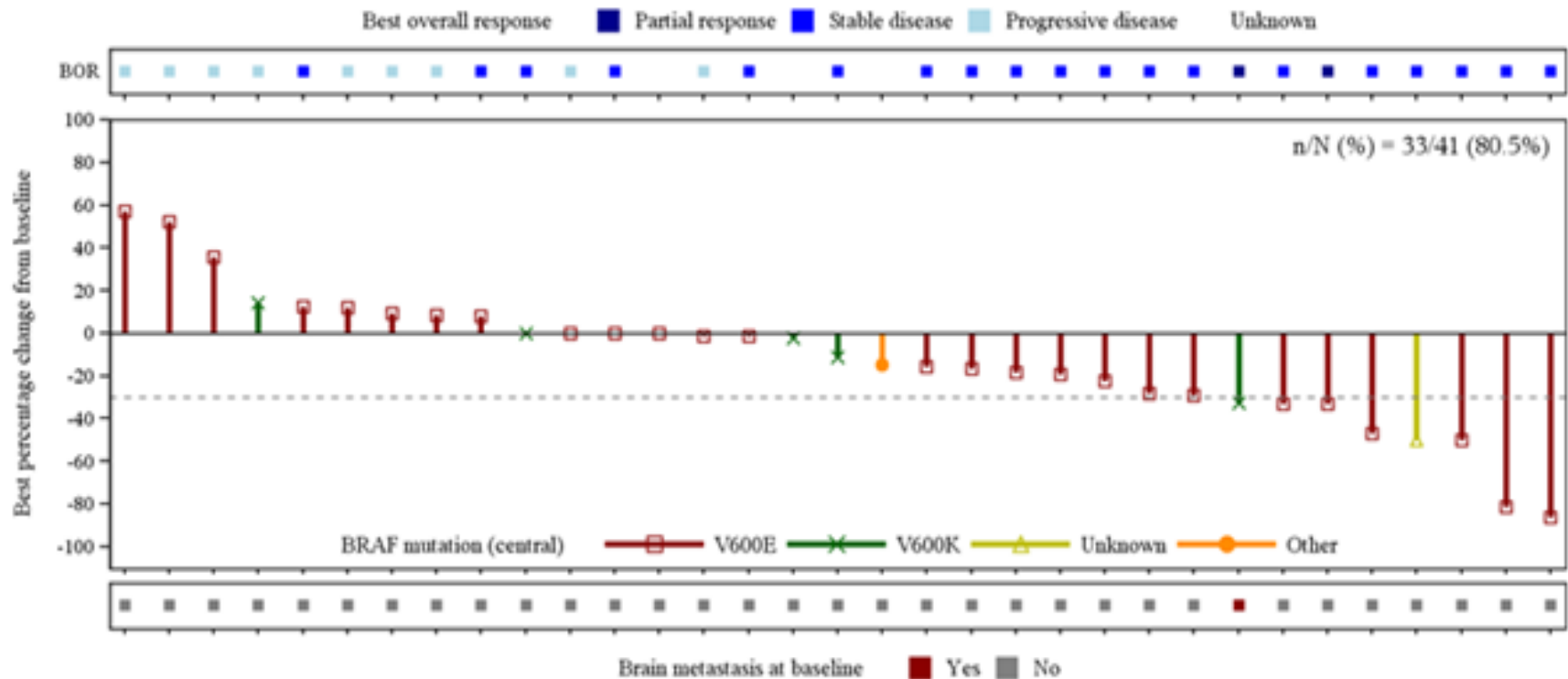


Waterfall Plot with Heatmap



- “Waterfall plot with heat map” are especially effective at demonstrating the effects of the **study drug** for a large population.
- There are ways in which a Waterfall can be combined with any other type of graph (such as Heatmap) to understand some correlations between different kinds of data.
- Here we see an example of Waterfall plot with Heatmap to understand if there are any relationship between Mutations and the Best overall response or brain metastasis.

Example: Co-relation of Mutation and overall response



Source: Dummy data representing a Waterfall with Heatmap.

Sample code:



```
/*Waterfall;*/
...|
    ScatterPlot X=$f_wh_xvar Y=_yvar2 /
                primary = true
                Group    = watervar
                GROUPORDER = ASCENDING
                NAME      = "SCATTER"
                markerattrs = (size=$waterfall_marker_size)
                ;

    NeedlePlot X=$f_wh_xvar Y=$f_wh_yvar /
              group=watervar
              GROUPORDER = ASCENDING
              NAME="NEEDLE"
              lineattrs = (thickness=$waterfall_line_thickness)
              ;

...
/*Heatmap*/
...
/* Increment heatmap counter to define appropriate y heatmap variable */;
%let __heatmap_counter = %eval($__heatmap_counter + 1);

    ScatterPlot X=$f_wh_xvar Y=$$f_wh_yheatvar$__heatmap_counter /
                primary      = true
                Group        = heatmapvar$_j
                GROUPORDER   = ASCENDING
                markerattrs  = (size=$$heatmap$_j._marker_size)
                LegendLabel  = " "
                NAME          = "S$_j";
    ;
```

Source: Dummy data representing a Waterfall with Heatmap.

Purpose



- At first, we take detailed measurements of all target lesions and assess patient disease status at a specified visit. Later, we consider an overall response assessment at the end of the trial.
- Waterfall plots show the distribution of quantitative variable(such as change from baseline).
- Usually a criteria showing how a patient responded to a treatment is used on the y axis, and the variable must be ordered.
- On the x axis you have 1 line /needle by patient showing the patient response.
- The heatmap is used to details on the population displayed. It adds another layer of information.
- It visualizes relations between baseline characteristics(mutations, brain metastasis) and variable of interest (change from baseline).



- This graph is widely used in oncology. Usually best percentage change from baseline in sum of tumors diameters(target lesions) is displayed.
- This allows to quickly see clinical activity as all values below zero which show patients for which there is a tumor shrinkage.
- -30% is the threshold above which a patient is considered as responder for a target lesion.
- Heat map show additional information such as best overall response(on top), or presence/absence of brain metastasis (which can be a prognostic factor).

Procedures



Procedures type	SAS/Graph - GPLOT	Statistical Graphics (SG) – SGPLOT	Graphics Template Language (GTL) – PROC TEMPLATE with SGRENDER
Supporting Version	sas 9.1	sas 9.2	sas 9.3
Preferable based on the requirement	Built-in flexibility of SAS that allows the creation of custom graphs (requires annotation). Complex to create custom graphs	If we run into limitations using SG procedures more than twice, it is time for us to learn GTL.	We can produce impressive stylish graphs with different data by running PROC SGRENDER over and over. GTL will empower us to control the elements associated with the graphic display.

Challenges



- To learn the Procedure
- Adjusting the size of heatmap based on data
- Setting the right color gradient



- To communicate relationships between data values that would be much harder to understand may not be obvious from the viewpoint of tabular reports if presented numerically in a spreadsheet or tables.
- The advantage of displaying data as “Waterfall with Heat map” provides the ability to convey a large amount of information in a relatively small space for speedy visual identification of any existing patterns or trends.



- SUGI Papers
- SAS Helps
- Google search

Questions... ?

