

## Swimmer Plot by Graphic Template Language (GTL)

Baiming Wang, Pharmaceutical Products Development (PPD), Wilmington NC

### ABSTRACT

Time-to-Event analysis is one of the key analyses for oncology studies. Besides the Kaplan-Meier survival step curve to summarize overall drug performance, investigators and sponsors are also very interested in individual subject response situations. Many oncology clinical trials request Swimmer plots to show individual response trends on a continuing basis. Graphic Template Language (GTL) is a powerful approach for all kinds of visual clinical data presentations. Swimmer plots can presents all subject time-to-event data such as Complete Response Start, Partial Response Start, Continued Response after Study Drug Exposure, Disease Progression, Death, tumor types, etc. All these data can be shown on a Swimmer bar with different symbols/color combinations. Different colors can represent different tumor types or tumor stages. Subject IDs can be shown on each bar for easy review. Each event has data labeled on each marker to show exact days for that event. Every component of the plot will have different data modules easily programmed, step-by-step, adding each part to the plot.

### INTRODUCTION

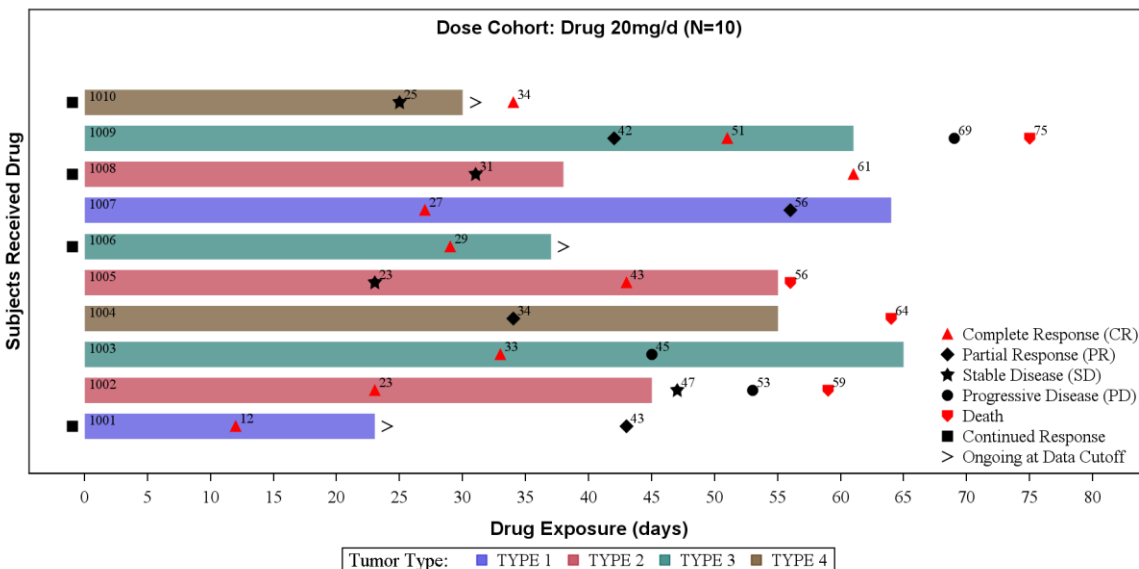
Swimmer Lane plots can visually tell each subject's response story with exact data labels, such as, which time point has Complete Response, or which point has Disease Progression (PD), or Death. Different colored bars distinguish different types of tumor or stage. It also provides the pattern of individual responses that help sponsors make evidence based decisions. This paper demonstrates different modules on the plot, step by step, making it easy to be adopted. Any programmer who wants to adopt this approach will need some basic SAS 9.2 version Graphic Template Language (GTL) knowledge. A SUG paper (PharmaSUG 2014 - Paper DG07) has demonstrated the steps by PROC SGPLOT and annotated datasets to create a similar style of Swimmer plot figures. By comparing the two totally different approaches, it will show that GTL is much simpler, straightforward, with additional useful features that SGPLOT does not have, such as the data label on event markers, subject IDs on the bar, etc.

The paper starts with a brief description and example of a Swimmer plot, and then proceeds to the step-by-step process of generating the plot with the GTL plot statement. This paper demonstrates a program using version SAS 9.2 with very basic plot statements, which benefits those sponsors not having their SAS version updated yet. Versions 9.3 and 9.4 allow for even further simplified statements due to plot statement enhancements.

### DESCRIPTION AND EXAMPLE OF SWIMMER PLOT

A Swimmer lane-like bar can be used to visually simulate subject time-to-event data in an oncology study in one glance. The example plot below shows the following six pieces of information: length of treatment duration, when each response episode(s) started (time to response), when the response episode(s) ended or resulted in an ongoing response. All CR, PR, SD, PD, Death, those related subject first dose date can be marked on the Swimmer Lane Bar. Reviewers can quickly check any interested subject's response situation and also compare the different treatment group response trend.

The example plot below shows the following each pieces of information: length of treatment duration, when each response episode(s) started (time to response), when the response episode(s) ended or resulted in an ongoing response,



Each bar represents one subject in study; Bar length represents study drug exposure time (days).  
 Bar color represents tumor type at baseline; Marker data label represents event start since first dose date.

Figure 1. Time-to-Event Swimmer Plot

## STEP-BY-STEP PROGRESSION OF SWIMMER PLOT WITH CODE AND PLOT EXAMPLES

SAS Graphic Template provided multiple selection to draw a Swimmer Lane Bar, Here we use VECTORPLOT for the bar and SCATTERPLOT combination for the bar legend. In the below sections, there are details about the input dataset structure followed by a description of how to break the construction of the plot into four phases: BASIC BAR, CUSTOMIZED PLOT TO PUT SUBJECT ID ON THE BAR, PLOT TIME-TO-EVENT MARKS WITH DATA LABEL, PLOT MARK LEGEND. Explanations of the statements used in each step are presented as well.

### INPUT DATA STRUCTURE

The following one record per subject data structure is used. 10 subjects are used to show different response situations.

- SUBJID: Subject identifier
- XTIME: Subject study drug exposure length (days)
- ONGOING: Flag to indicate still ongoing on study drug after cutoff date
- CRDY: Subject start complete response (CR) episode since first dose date (days)
- PRDY: Subject start partial response (PR) episode since first dose date (days)
- SDDY: Subject first stable disease (SD) episode since first dose date (days)
- PDDY: Subject start disease progress (PD) or death date since first dose date (days)
- DTHDY: Subject death days since first dose date
- CONRSPFL: Flag to indicate subject continued response after date cut off
- TUMTYPN/TUMTYP: Tumor type number or actual tumor type description

Here is the step to produce the dataset:

```
DATA SW;
INPUT SUBJID $ XTIME ONGOING CRDY PRDY SDDY PDDY DTHDY CONRSPFL $ TUMTYPN TUMTYP $;
datalines;
1001 23 1 12 43 . . . Y 1 TYPE1
1002 45 0 23 . 47 53 59 N 2 TYPE2
1003 65 0 33 . . 45 . N 3 TYPE3
1004 55 0 . 34 . . 64 N 4 TYPE4
1005 55 0 43 . 23 . 56 N 2 TYPE2
1006 37 1 29 . . . . Y 3 TYPE3
1007 64 0 27 56 . . . N 1 TYPE1
1008 38 1 61 . 31 . . N 2 TYPE2
1009 61 0 51 42 . 69 75 N 3 TYPE3
```

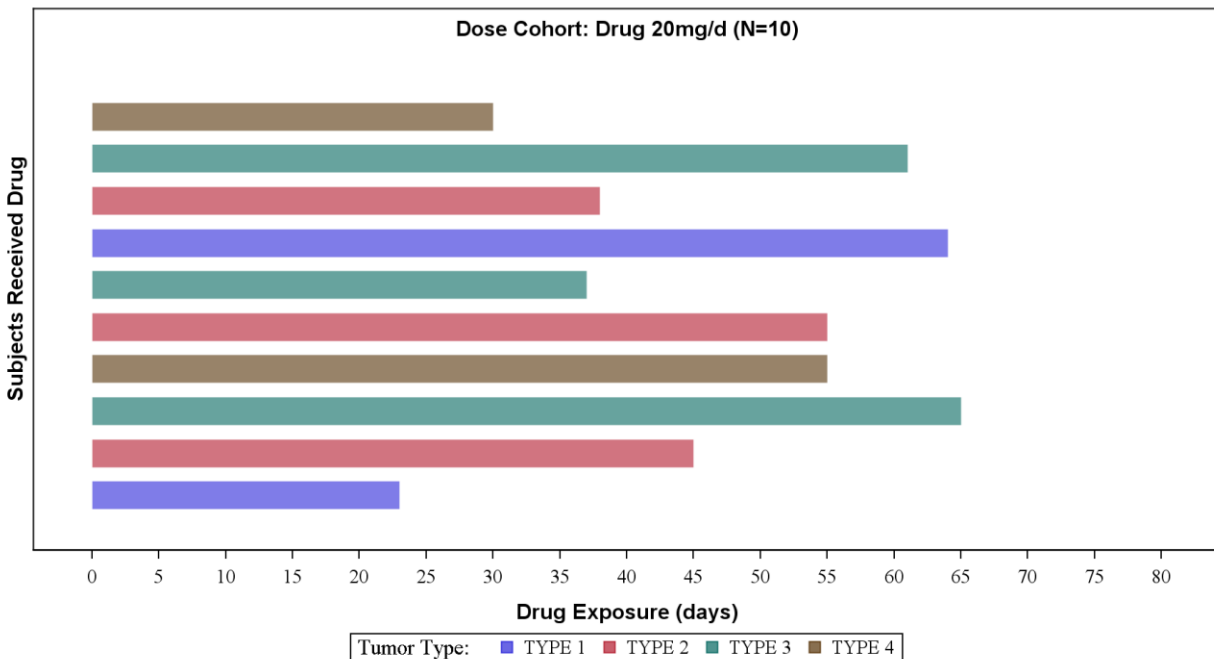
Swimmer Plot by Template Graphic Language (GTL), continued

```
1010 30 1 34 . 25 . . Y 4 TYPE4
;
RUN;
```

	SUBJID	XTIME	ONGOING	CRDY	PRDY	SDDY	PDDY	DTHDY	CONRSPFL	TUMTYPN	TUMTYP
1	1001	23	1	12	43	.	.	.	Y	1	TYPE1
2	1002	45	0	23	.	47	53	59	N	2	TYPE2
3	1003	65	0	33	.	.	45	.	N	3	TYPE3
4	1004	55	0	.	34	.	.	64	N	4	TYPE4
5	1005	55	0	43	.	23	.	56	N	2	TYPE2
6	1006	37	1	29	.	.	.	.	Y	3	TYPE3
7	1007	64	0	27	56	.	.	.	N	1	TYPE1
8	1008	38	1	61	.	31	.	.	N	2	TYPE2
9	1009	61	0	51	42	.	69	75	N	3	TYPE3
10	1010	30	1	34	.	25	.	.	Y	4	TYPE4

Display 1. Input Data for Swimmer Plot

**BASIC PLOT**



**Figure 2. Time-to-Event (Basic Bar Plot by VECTORPLOT)**

This basic plot demonstrates the three most important analyses regarding the Time-to-Event data:

1. Total study drug exposure time (represented by bar length)
2. Number of subjects who received study drug (represented by number of bars and N=xx text)
3. Tumor Type (or Disease stage) at baseline (represented by the color of the bars)

The basic bar plot is generated by VECTORPLOT and SCATTERPLOT:

```
VECTORPLOT x=XTIME Y=YVAR xorigin=0 yorigin=yvar/group=TUMTYPN
DATATRANSPARENCY=0.4 ARROWHEADS=FALSE LINEATTRS=(PATTERN=1 THICKNESS=20);
```

```
SCATTERPLOT x=eval(XTIME+999) Y=YVAR/group=TUMTYPN Name="TYP" DATATRANSPARENCY=0.3  
markerattrs=(symbol=squarefilled size=9);
```

```
DISCRETELEGEND "TYP"/Title="Tumor Type:" Border=1 across=4 location=outside;
```

### Statement 1: VECTORPLOT

VECTORPLOT creates a horizontal bar based on a drug exposure time XTIME for each subject. Both VECTORPLOT and BARCHART can be used to create this bar. VECTORPLOT uses a THICKNESS option to control the bar width. It sets a consistent bar width regardless of how many bars are on that page, BARCHART uses BARWIDTH option to control bar width. This bar width has a ratio of max width between 0 and 1, so the width depends on how many bars are on that page. If there are multiple pages by treatment group, it is very likely each group will have a different number of subjects. Using BARCHART, it is not possible to get a consistent bar width for all pages. This is the main reason VECTORPLOT is used to draw the bar.

- XTIME (required): specifies the total drug exposure that marking the end of the bar
- YVAR (required): Record number to determine the bar order on plot.
- TUMTYPN (optional): is the group variable to assign different color for each tumor type.
- ARROWHEAD=FALSE show no arrow on the bar.
- THICKNESS option determines the bar width

### Statement 2: SCATTERPLOT

This statement is optional for the purpose of setting the legend symbol as square instead of long bar.

Note, `x=eval(XTIME+999)` draws the SQUAREFILLED mark out of the plot x-axis range since it uses a large number such as 999. Here, we just want borrow its legend information.

### Statement 3: DISCRETELEGEND

This statement will draw the bar legend from the SCATTERPLOT step, it looks like a square mark that color matches the bars.

## CUSTOMIZED PLOT WITH SUBJID ON THE BAR

Once the basic plot, i.e., the respond bar, is constructed, further key information of the subject identification needs to be built in. This will help the sponsor to easily check one particular subject's response situation. Otherwise the bar seems useless for telling a specific individual response story.

The code to draw the SUBJID on the bar is:

```
Scatterplot x=eval(XTIME-XTIME) y=eval(YVAR)/Group=TUMTYPN DataLabel=SUBJID  
DataLabelattrs=(Color=black) markerattrs=(size=0) ;
```

- `x=eval(XTIME-XTIME)`: put the SUBJID at beginning of bar,
- `DataLabel=SUBJID`: use the data contents to label the mark.
- `markerattrs=(size=0)`: set mark size to zero, instead of using its datalabel function, we don't want draw any mark on the plot.

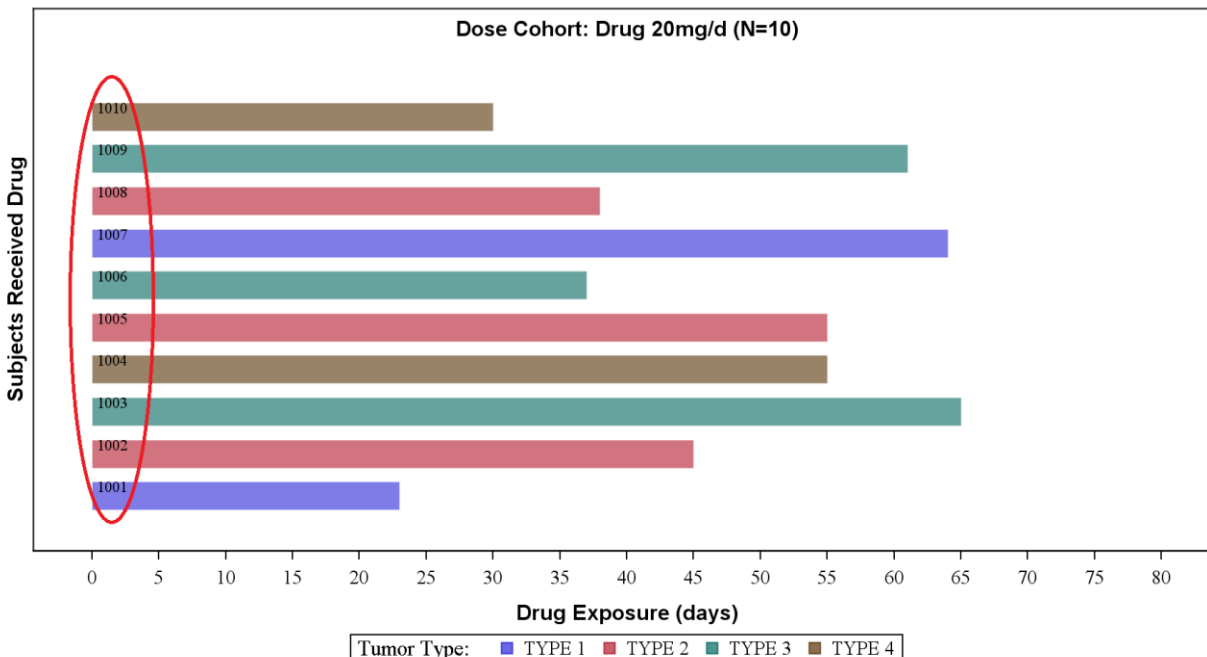


Figure 3. Time-to-Event (Customized Plot with SUBJID on the Bar)

### CUSTOMIZED PLOT WITH ALL TIME-TO-EVENT MARKS

These statements draw five different marks and legend on plot:

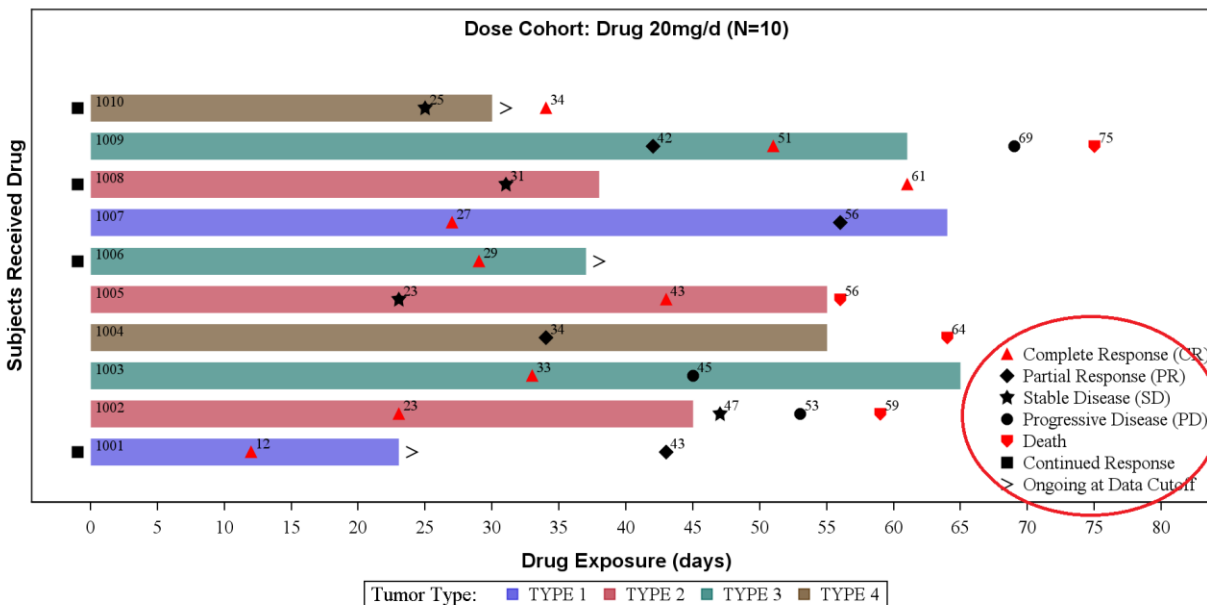
```

SCATTERPLOT X=CRDY Y=YVAR/datalabel=CRDY name="MK1" LEGENDLABEL="Complete response
  start" MARKERATTRS=(SYMBOL=trianglefilled SIZE=9 COLOR=red);
SCATTERPLOT X=PRDY Y=YVAR/datalabel=PRDY name="MK2" LEGENDLABEL="Partial
  response start" MARKERATTRS=(SYMBOL=diamondfilled SIZE=9 COLOR=black);
SCATTERPLOT X=PDDY Y=YVAR/datalabel=PDDY name="MK3" LEGENDLABEL="Disease
  Progress/Death" MARKERATTRS=(SYMBOL=circlefilled SIZE=9 COLOR=black);
SCATTERPLOT X=XONGO Y=YVAR/name="MK4" LEGENDLABEL="Continued response"
  MARKERATTRS=(SYMBOL=squarefilled SIZE=9 COLOR=black);
SCATTERPLOT X=XCONRSP Y=YVAR/name="MK5" LEGENDLABEL="Ongoing at data cutoff"
  MARKERATTRS=(SYMBOL=GreaterThan SIZE=9 COLOR=black);
    
```

```

DISCRETELEGEND "MK1" "MK2" "MK3" "MK4" "MK5" /Border=0 across=1 location=inside
HALIGN=RIGHT VALIGN=BOTTOM;
    
```

- x=CRDY, PRDY, PDDY, XONGO, XCONRSP define the mark x-axis position
- datalabel option shows actual data on the mark
- Note Flag variable XONGO and XCONRSP show no datalabel
- Symbol option defines the shape of marker, Size option defines the marker size
- SCATTERPLOT statement is used to generate legend,
- LEGENGLABEL option is used to define each legend title
- SYMBOL option defines the shape of marker both in plot and legend
- COLOR option define the color of symbol both in plot and legend
- SIZE option defines the size in the legend; it can be any reasonable size.
- DISCRETELEGEND statement combines all 5 markers legend together and put lower right corner.



Each bar represents one subject in study; Bar length represents study drug exposure time (days).  
Bar color represents tumor type at baseline; Marker data label represents event start since first dose date.

Figure 4. Time-to-Event (Customized Plot with Event Marker and Legend)

## DISCUSSION HOW TO HANDLE MISSING DATA

This paper demonstrates the perfect data situation in which all four expected tumor types exist. It is possible using the GROUP option to assign groups to a color automatically. Also, the legend can show all types/colors. But in a real data situation, it is likely that one or more categories of data will be missing. For example, it is expected to have four tumor types, but an actual dataset may only have three types recorded. In this situation we still want to show four colors in the legend, even when there are only three different colors of bars. There is a different approach to handle such data and will result in complicated data handling processes. This topic is beyond the scope of this paper and it will be discussed as another direction of data processing. We should be aware of such a situation when we adopt this approach to the real data.

## BRIEF DISCUSSION REGARDING THE USE OF HIGHER VERSION SAS 9.3 or 9.4

The example demonstrated in this paper uses version 9.2 basic plot statement to produce Swimmer Lane plots. Version 9.3 and version 9.4 have many updates to version 9.2. We will only concentrate to those options that relate to this plot. With the higher version, there are more choices to create this plot, following are the options that version 9.2 does not support:

- 1) LEGENDITEM statement: with this statement, individual legend items can be created without being related to the input data; the missing category data situation can be easily handled, e.g., the data plotted has only used three colors for three tumor types, LEGENDITEM can be used to add a fourth color for the missing tumor type in the legend so that the complicated data processing steps can be avoided.
- 2) In version 9.4, HIGHLOWPLOT can be used to draw the bar with many sophisticated functions and statements; an INTERVALBARWIDTH option can be used to define the fixed bar width; it can produce a bar-like strip with a consistent appearance. For a continued response mark, a line end cap can be used for better analysis interpretation; here we are only interested in drawing the bar.
- 3) MERGEDLEGEND in version 9.4 makes it possible to combine the bar and marker legends together.

## COMPLETE CODE FOR CUSTOMIZED SWIMMER PLOT BY GTL

```
DATA SW;
INPUT SUBJID $ XTIME ONGOING CRDY PRDY SDDY PDDY DTHDY CONRSPFL $ TUMTYPN TUMTYP $;
datalines;
1001 23 1 12 43 . . . Y 1 TYPE1
1002 45 0 23 . 47 53 59 N 2 TYPE2
```

Swimmer Plot by Template Graphic Language (GTL), continued

```

1003 65 0 33 . . 45 . N 3 TYPE3
1004 55 0 . 34 . . 64 N 4 TYPE4
1005 55 0 43 . 23 . 56 N 2 TYPE2
1006 37 1 29 . . . Y 3 TYPE3
1007 64 0 27 56 . . . N 1 TYPE1
1008 38 1 61 . 31 . . N 2 TYPE2
1009 61 0 51 42 . 69 75 N 3 TYPE3
1010 30 1 34 . 25 . . Y 4 TYPE4
;
RUN;

Proc FORMAT;
  value TUMTYP_
    1 = "TYPE 1"
    2 = "TYPE 2"
    3 = "TYPE 3"
    4 = "TYPE 4" ;
RUN;

DATA Final;
  SET SW;
  by SUBJID;
  YVAR = _n_;
  if ONGOING = 1 then do;
    XONGO = -1;
  end;
  if CONRSPFL = "Y" then
    XCONRSP = XTIME + 1 ;
  else
    XCONRSP = . ;
  FORMAT TUMTYPN TUMTYP_ . ;
RUN;

Proc SORT data=Final; by SUBJID YVAR XTIME ; run;

*ODS output, set graph orientation, size, RTF file name;
OPTIONS nodate nonumber orientation=landscape papersize='LETTER' ;
ods graphics on / height=5.2 in width=9.5 in ;
ods listing close;
ods rtf file = "SWIMMERPLOT_VECTORPLOT.rtf" NOGTITLE NOGFOOTNOTE style=tlout;

* Output graphic template;
Proc TEMPLATE;
  define STATGRAPH GTL_GRAPH;
    BEGINGRAPH /BORDER=False;
    ENTRYFOOTNOTE halign=left textattrs=(Family="Arial" color=black
size=11pt) "Each bar represents one subject in study; Bar length represents study drug
exposure time (days)." ;
    ENTRYFOOTNOTE halign=left textattrs=(Family="Arial" color=black
size=11pt) "Bar color represents tumor type at baseline; Marker data label represents
event start since first dose date." ;

    LAYOUT OVERLAY /
      YAXISOPTS= (DISPLAY=(LABEL) OFFSETMIN=0.1 OFFSETMAX=0.2
LABEL="Subjects Received Drug"
      LABELATTRS=(FAMILY="Arial" WEIGHT=bold SIZE=11pt)
      LINEAROPTS=(VIEWMIN=1 viewmax=10 TICKVALUESEQUENCE=(start=1
end=10 INCREMENT=1)))
      XAXISOPTS=(OFFSETMIN=0.05 OFFSETMAX=0.05 LABEL="Drug Exposure
(days)"
      LABELATTRS=(FAMILY="Arial" weight=bold size=11pt)
TICKVALUEATTRS=(COLOR=BLACK)
      LINEAROPTS=(VIEWMIN=0 VIEWMAX=80 TICKVALUESEQUENCE=(START=0

```

## Swimmer Plot by Template Graphic Language (GTL), continued

```
END=80 INCREMENT=5));
    ENTRY HALIGN=center TEXTATTRS=(FAMILY="Arial" COLOR=black SIZE=11pt
WEIGHT=bold) "Dose Cohort: Drug 20mg/d (N=10)"/valign=top ;

    * draw swimmer lane bars;
    /*
    barchart x=YVAR y=XTIME /Group=TUMTYPN Name="TYP" BARWIDTH=0.6 BARLABEL=0
stat=sum orient=horizontal;
    */

    VECTORPLOT X=XTIME Y=YVAR XORIGIN=0 YORIGIN=yvar/GROUP=TUMTYPN
    DATATRANSPARENCY=0.4 ARROWHEADS=FALSE LINEATTRS=(PATTERN=1 THICKNESS=20);

    SCATTERPLOT X=eval(XTIME+999) Y=YVAR/GROUP=TUMTYPN Name="TYP"
DATATRANSPARENCY=0.3 MARKERATTRS=(SYMBOL=squarefilled SIZE=9) ;

    * draw SUBJID on the bar;
    SCATTERPLOT X=eval(XTIME-XTIME) Y=eval(YVAR-0.05)/GROUP=TUMTYPN
DATALABEL=SUBJID DATALABELATTRS=(COLOR=black) MARKERATTRS=(size=0) ;

    * draw marks and assign each legend label;
    SCATTERPLOT X=CRDY Y=YVAR/datalabel=CRDY name="MK1" LEGENDLABEL="Complete
Response (CR)" MARKERATTRS=(SYMBOL=TriangleFilled SIZE=9 COLOR=red);
    SCATTERPLOT X=PRDY Y=YVAR/datalabel=PRDY name="MK2" LEGENDLABEL="Partial
Response (PR)" MARKERATTRS=(SYMBOL=DiamondFilled SIZE=9 COLOR=black);

    SCATTERPLOT X=SDDY Y=YVAR/datalabel=SDDY name="MK3" LEGENDLABEL="Stable
Disease (SD)" MARKERATTRS=(SYMBOL=StarFilled SIZE=9 COLOR=black);

    SCATTERPLOT X=PDDY Y=YVAR/datalabel=PDDY name="MK4" LEGENDLABEL="Progressive
Disease (PD)" MARKERATTRS=(SYMBOL=CircleFilled SIZE=9 COLOR=black);

    SCATTERPLOT X=DTHDY Y=YVAR/datalabel=DTHDY name="MK5" LEGENDLABEL="Death"
MARKERATTRS=(SYMBOL=HomeDownFilled SIZE=9 COLOR=red);

    SCATTERPLOT X=XONGO Y=YVAR/name="MK6" LEGENDLABEL="Continued Response"
MARKERATTRS=(SYMBOL=SquareFilled SIZE=9 COLOR=black);
    SCATTERPLOT X=XCONRSP Y=YVAR/name="MK7" LEGENDLABEL="Ongoing at Data Cutoff"
MARKERATTRS=(SYMBOL=GreaterThan SIZE=9 COLOR=black);

    * Tumor type legend;
    DISCRETELEGEND "TYP" /TITLE="Tumor Type:" BORDER=1 ACROSS=4
LOCATION=outside;
    * Mark legend;
    DISCRETELEGEND "MK1" "MK2" "MK3" "MK4" "MK5" "MK6" "MK7"/BORDER=0 ACROSS=1
LOCATION=inside HALIGN=RIGHT VALIGN=BOTTOM;
    ENDLAYOUT;
ENDGRAPH;
end;
run;

Proc SGRENDER DATA=Final TEMPLATE=GTL_GRAPH;
```

## CONCLUSION

This paper has shown that the complex time-to-event Swimmer Plot can be produced by GTL overlay different components together. Each component can be independently created and costumed. Data Label is very useful information that provides exactly event time point. It also demonstrated how to produce the legend symbol without actually draw the marker on graph. The GTL provided by SAS has the powerful flexibility to put custom text or legends on plots in any desired location without the involvement of complicated, annotated datasets. Through the technique of overlaying plots on the top of each other, oncology swimmer plot can be easily created with several



simple SAS GTL statements.

## REFERENCES

SAS® 9.2 Graph Template Language Reference, Second edition, SAS Institute  
PharmaSUG 2014 - Paper DG07  
Swimmer Plot: Tell a Graphical Story of Your Time to Response Data Using PROC SGPLOT  
Stacey D. Phillips, Inventiv Health Clinical, Princeton, NJ

## ACKNOWLEDGMENTS

Thanks Jen Nash for her great support during this paper development. Also thanks Amy Caison for her former PharmaSUG paper examples. Ken Borowiak's encouragement and the guideline about the PharmaSUG conference procedures are very helpful for this paper at its final stage.

## CONTACT INFORMATION

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

Name: Baiming Wang  
Enterprise: Pharmaceutical Products Development (PPD)  
Address: 929 North Front Street  
City, State ZIP: Wilmington, NC 28401  
Work Phone: (910) 558-3190  
E-mail: baiming.wang@ppdi.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. ® indicates USA registration.

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