Paper 357-2013

Introducing and Producing Thunderstorm or Rain-drop Scatter Plots Using the SAS/GRAPH® Annotate Facility

Charlie C. Liu, Allergan, Inc.

ABSTRACT

A new type of plot, the thunderstorm or rain-drop scatter plot is introduced. Such a plot allows for viewing data with 2 or more values on the y-axis corresponding to one value on the x-axis for each of several subjects in a population. The resulting plot looks like rain-drops, with each rain-drop representing data for a single subject. When data for many subjects are plotted, it resembles a thunderstorm, hence the name. A thunderstorm/rain-drop scatter plot is a useful tool for data visualization and outlier detection. It has applications in many areas including clinical research, agriculture, and finance, etc. Using examples from clinical research, this paper shows how to create thunderstorm/rain-drop scatter plots using the SAS/GRAPH® annotate facility, a convenient tool to connect the 2 or more values on the y-axis of the same subject to form rain-drops. The dataset structure is described and the SAS programs producing the sample plots are provided and discussed.

INTRODUCTION

Data visualization using attractive and meaningful figures to display the whole data points to allow us to see the overall data patterns, study the inherent relationships and detect outliers has been an interesting yet challenging job for statisticians (Cleveland 1985; Tufte 1983, 1987). Many different types of plots have been introduced and produced, including those especially named for their shapes, like the bag, bar, bike, box, bubble, pie, pyramid, spaghetti, rainbow, violin, and waterfall plots, etc. (Allison 2012; Hyndman and Shang 2010). This paper introduces and illustrates how to produce a new type of plot, the thunderstorm or rain-drop scatter plot. Such a plot allows for viewing data with 2 or more values on the y-axis corresponding to one value on the x-axis for each of several subjects in a population. The resulting plot looks like rain-drops, with each rain-drop representing data for a single subject. When data for many subjects are plotted, it resembles a thunderstorm, hence the name. This is the first time the concept of thunderstorm/rain-drop scatter plots is introduced and real figures are produced using the SAS/GRAPH® annotate facility, a convenient tool to connect the 2 or more values on the y-axis with one value on the x-axis of the same subject to form rain-drops.

A thunderstorm/rain-drop scatter plot is a useful tool for data visualization and outlier detection. It has applications in many areas including clinical research, agriculture, and finance, etc. For example, it can be used to display efficacy measurement before and after treatment for patients in clinical trials, yield production before and after fertilizer application for crop varieties in agriculture, and buying power before and after a certain promotion measure for customers in a financial company, etc.

To illustrate the application and production of thunderstorm/rain-drop scatter plots, two examples are presented in the paper based on clinical research in the glaucoma therapeutical area, one for two types of intraocular pressure (IOP) values on the y-axis by corresponding central corneal thickness (CCT) on the x-axis for each eye of 100 patients, and the other for mean diurnal IOP at baseline and week 12 on the y-axis by subjects grouped within investigator sites on the x-axis. The thunderstorm/rain-drop scatter plots are produced using the simulated dataset and the clinical meaningfulness cannot be interpreted by the sample figures shown. The data structure and SAS annotated dataset producing the sample plots are discussed. The actual SAS programs producing the sample figures are provided and discussed.

EXAMPLE 1: TWO IOPS ON THE Y-AXIS BY CCT ON THE X-AXIS

The Goldmann applanation tonometor (GAT) is the gold standard for measuring IOP, and has been the most widely used IOP measurement technology (Goldmann 1957). However, GAT IOP has been known to be affected by eye physical ocular properties, especially the CCT (Ehlers et al. 1975; Whitacre and Stein 1993). Recently, the Pascal dynamic contour tonometer (DCT, Swiss Microtechnology AG, Port, Switzerland) has been developed to provide automated IOP readings that are less affected by CCT compared with those measured by GAT (Boehm, et al. 2008). A thunderstorm/rain-drop plot is a good tool to study the agreement/difference between GAT and DCT IOPs and see whether they are affected by the corresponding CCT values.

We simulated a dataset including 100 patients, with IOP values by GAT and DCT and the corresponding CCT value for each eye (see data structure in Table 1). The thunderstorm scatter plot displaying the two IOPs (GAT and DCT) on the y-axis by its corresponding CCT value on the x-axis at the eye level of all 100 patients (200 eye data points) is shown in Figure 1. The figure allows us to view the two IOP values (GAT and DCT) of the same eye and directly see it's agreement or difference by CCT values. If you feel the thunderstorm plot is too crowded with too many data points (rain-drops), we can separate it into several figures by sub-setting the data by the patients (e.g. with each figure displaying data from 20 eyes in 10 patients) or by the CCT values (500 to 510, 520 to 540 microns etc.). The new figures then might not be qualified as the "thunderstorm" anymore, but as "rain-drop" plots (Figure 2 displays data from 20 eyes of 10 patients). As the figures are produced from the simulated data, please refrain from interpreting the agreement/relationship between GAT and DCT IOP by CCT values.

Figure 1. GAT and DCT IOP by CCT Value: All Data from 200 Eyes in 100 Patients

(A Thunderstorm Scatter Plot)

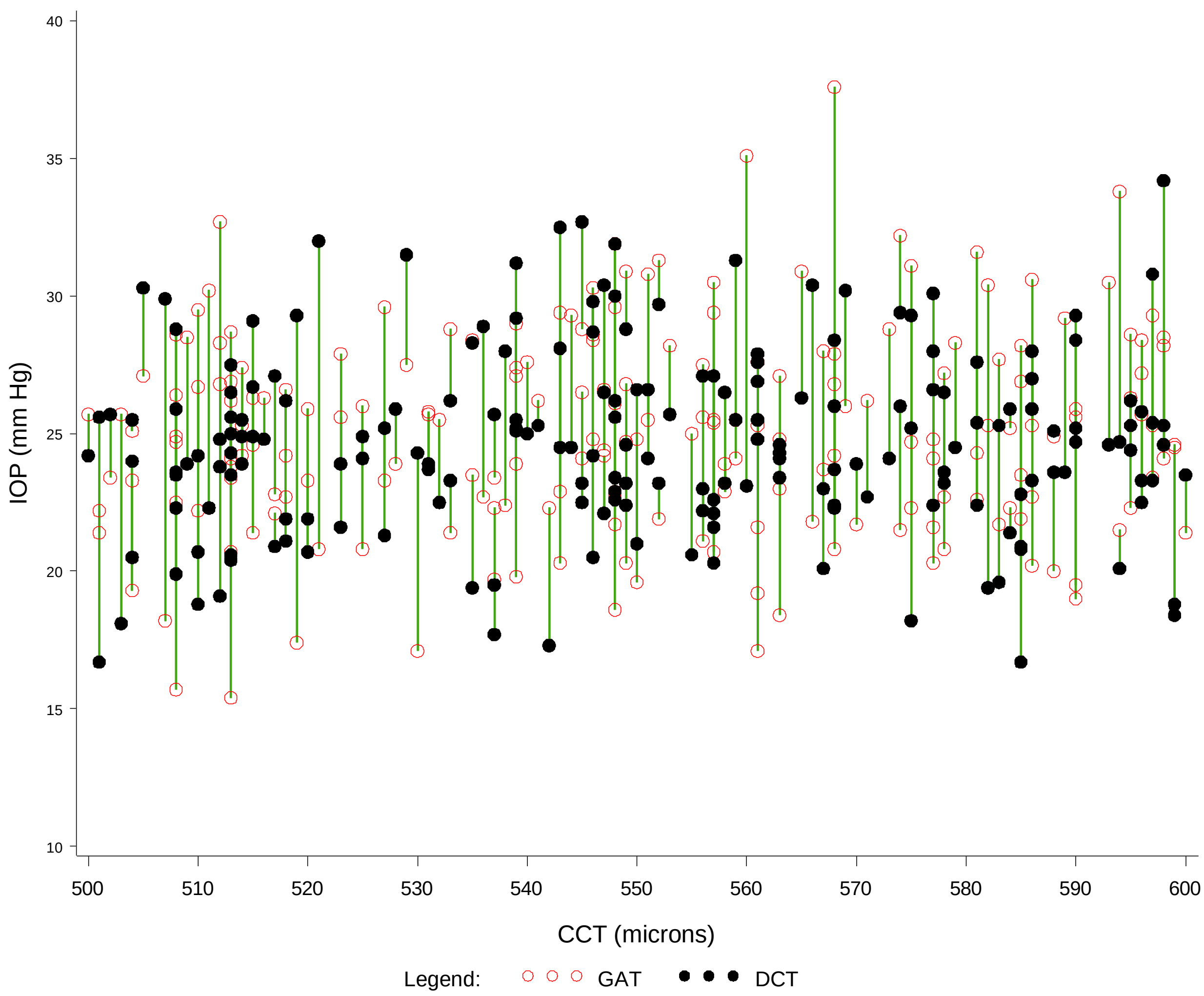
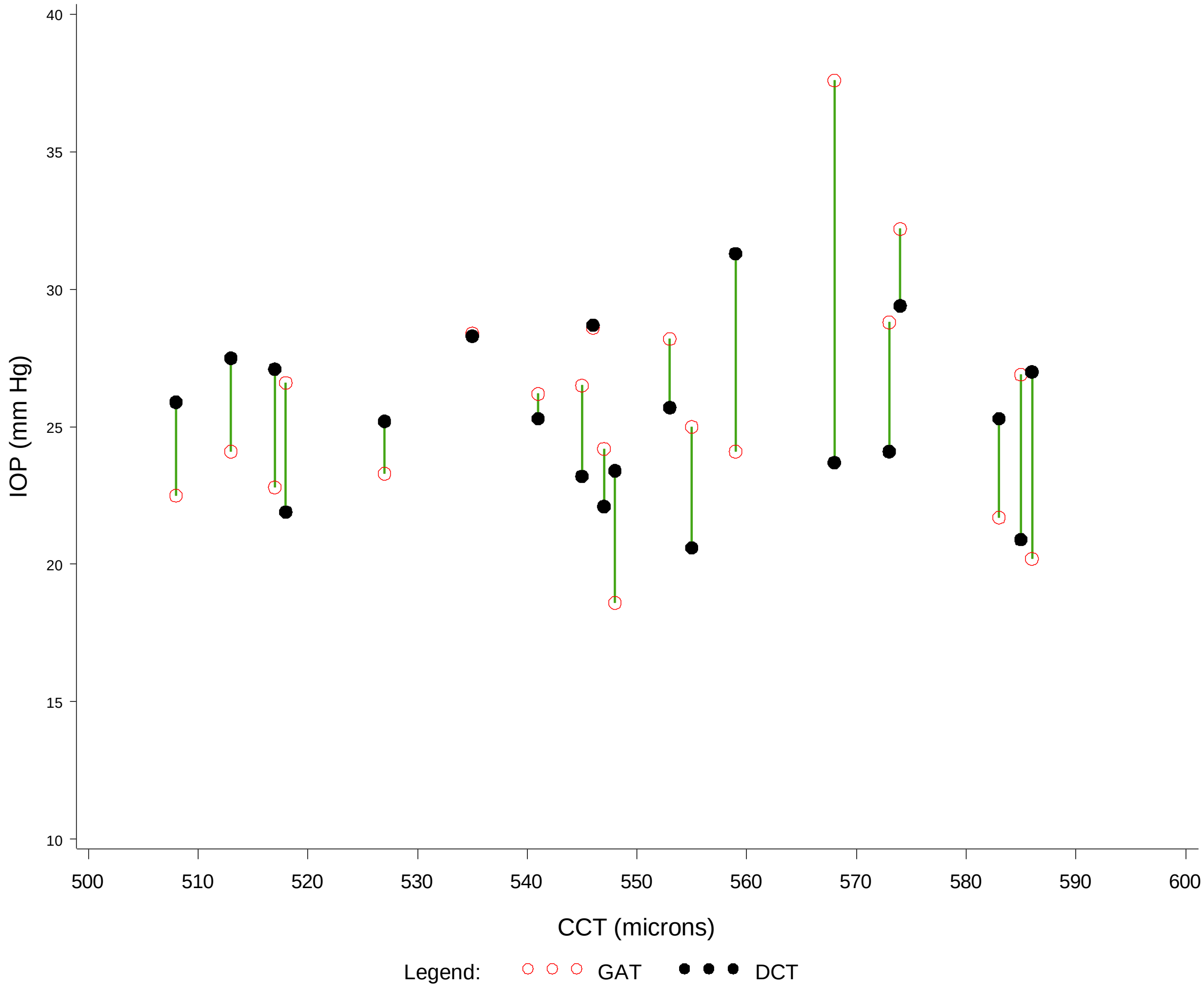


Figure 2. GAT and DCT IOP by CCT Value: Data from 20 Eyes in 10 Patients

(A Rain-drop Scatter Plot)

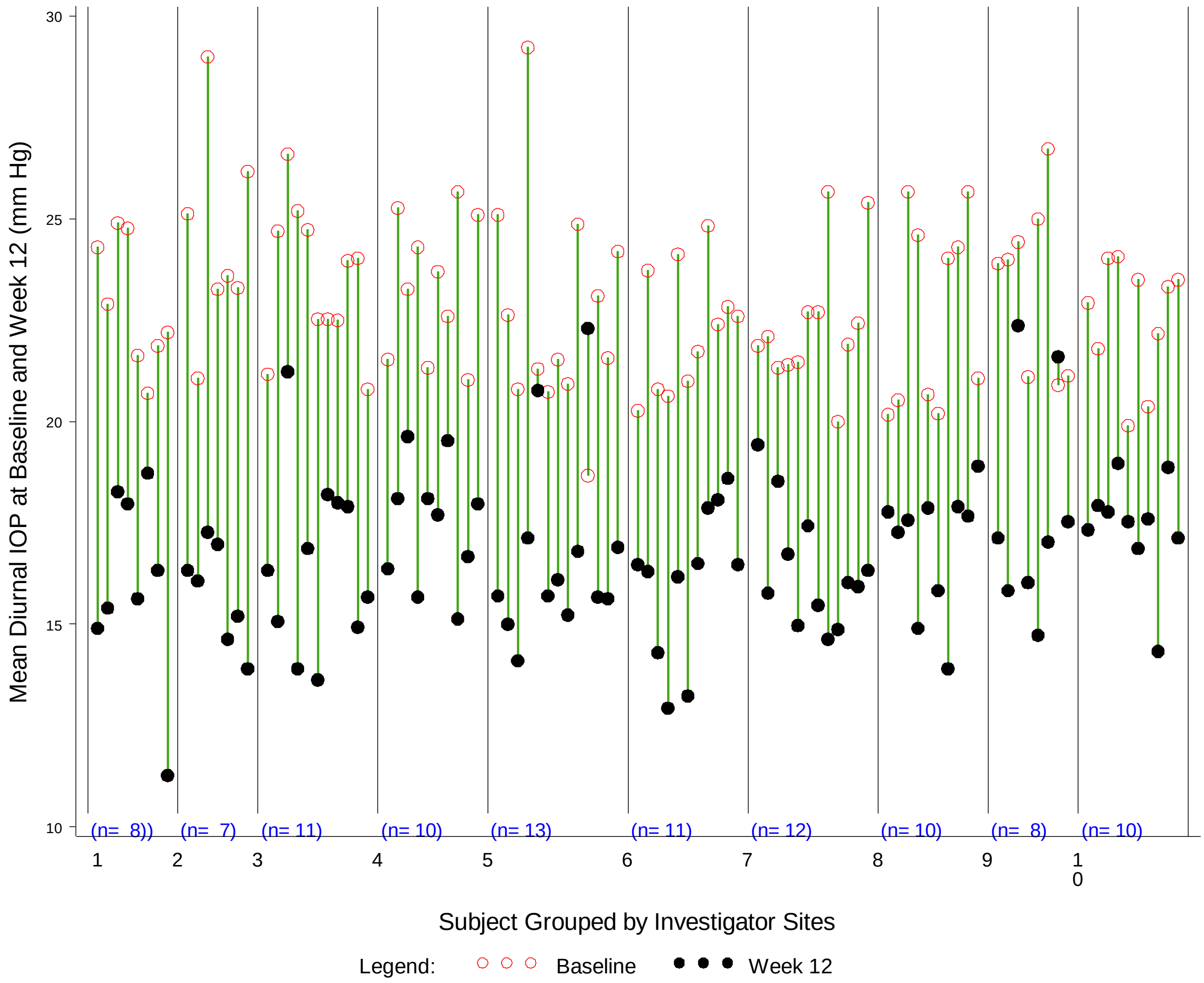


EXAMPLE 2: BASELINE AND WEEK 12 MEAN DIURNAL IOP ON THE Y-AXIS BY SUBJECTS GROUPED BY INVESTIGATOR SITES ON THE X-AXIS

In one of the phase III clinical trials at Allergan, subjects' IOP values are measured at hours 0, 2, and 8 at baseline (day 0), and at weeks 4, 8 and 12 post-baseline. Mean diurnal IOP, the mean IOP values of hours 0, 2 and 8 at week 12 in the study eye is the primary analysis variable. The thunderstorm scatter plot in Figure 3 (based on simulated data) provides a good visualization of the mean diurnal IOP values at baseline and week 12 for subjects within each investigator site. The plot allows us to see the overall data pattern for all subjects at both visits within each investigator site and helps the clinical team to issue data queries. For example, there are two subjects (one at investigator site 5 and the other at site 9) whose IOP values at week 12 are higher than at baseline, which are considered as outliers as the subjects' IOP values are expected to decrease after treatment following baseline.

Figure 3. Study Eye Mean Diurnal IOP at Baseline and Week 12

(A Thunderstorm Scatter Plot)



DATA STRUCTURE AND SAS ANNOTATED DATASET

Part of the simulated dataset with 100 subjects' IOP and CCT values at the eye level used to produce the thunderstorm/rain-drop scatter plots in Example 1 is shown in Table 1. Each eye of a subject (OD and OS) has IOP values by GAT and DCT and the corresponding CCT value. Each subject has 4 records of values in the dataset (each eye with 2 different IOP and one corresponding CCT values). There are 100 subjects with a total of 400 records in the whole simulated dataset.

Table 1. Part of the Simulated Dataset to Produce the Thunderstorm Scatter Plot for GAT/DCT IOP by CCT

SUBJID	EYE	IOP_TYPE	IOP	CCT
1001	OD	GAT	26.9	585
1001	OD	DCT	20.9	585
1001	OS	GAT	24.1	559
1001	OS	DCT	31.3	559

SAS/GRAPH® annotate facility is used to connect the 2 IOPs of the same eye together by a straight line. The record for one subject in the annotated dataset is displayed in Table 2. For the right eye (OD), SAS at first 'move' the point to the position at (585, 26.9) (CCT and GAT IOP values in the X and Y axis respectively) then 'draw' a line to the point (585, 20.9) (CCT and DCT IOP values in the X and Y axis respectively) in vivid green color (VILG). The line type and size is controlled by the value of variables "line" and "size". For a more detailed introduction of SAS annotate facility, please refer to Carpenter's book (Carpenter, 1999).

Table 2. Part of the Annotated Dataset to Produce the Thunderstorm Scatter Plot for GAT/DCT IOP by CCT

SUBJID	EYE	CCT	IOP_GAT	IOP_DCT	RECID	function	xsys	ysys	x	y	color	size	line
1001	OD	585	26.9	20.9	1	move	2	2	585	26.9			
1001	OD	585	26.9	20.9	1	draw	2	2	585	20.9	VILG	2	1
1001	OS	559	24.1	31.3	2	move	2	2	559	24.1			
1001	OS	559	24.1	31.3	2	draw	2	2	559	31.3	VILG	2	1

NOTES TO SAS PROGRAMS

In order to for SAS to produce high-quality figures, it's important to choose the right GOPTIONS, especially the device drivers and type of graphs. In the sample programs, SASEMF is used as the device driver to produce figures in EMF (Enhanced Window Metafile) format, one of the vector formats with high quality. Figures in vector formats (e.g. EMF, PS, EPS, etc.) can be re-sized without quality loss, and figures produced in EMF can be inserted into word documents for the ease of editing. The other figure format category SAS/GRAPH® can produce is bitmap (e.g. GIF, JPEG, PNG, etc.), which are composed of 'bits' or pixels and the quality is reduced when the figure is re-sized. Another important setting in GOPTIONS is the font for titles/footnotes and text in figures, "Times-Roman", one of the hardware fonts, is used here for high-quality figures. This is done by setting up a macro variable for title and footnote justifications and fonts and then call the macro variables in the title and footnote statements (below).

```
** Define macro variables for title and footnote format: justification and fonts;
%LET FMTTITLE =H=2.5 JUSTIFY=CENTER FONT="&FONTNAME";
%LET FMTFOOT  =H=2   JUSTIFY=CENTER FONT="&FONTNAME";

** Set-up titles/footnotes with the preset justification and fonts;
title1 &FMTTITLE "Figure 1. GAT and DCT IOP by CCT Value";
title2 &FMTTITLE "All Data from 100 Patients (200 Eyes)";
title3 &FMTTITLE "Thunderstorm Scatter Plot";
footnote1 &FMTFOOT "&pgmloc.";
```

In both examples, a SAS dataset is simulated by using DO loops and SAS function RANNOR () to generate random numbers that are normally distributed with pre-set sample means and standard deviations. The fixed seed number is used so that each independent run will produce the same datasets and the same figures. It's also a good practice to use different seed at each DO loop, this is achieved by incorporating the loop numbers into the seed by using (&seed. + i + j) as the seed at each loop.

SAS annotate facility is a very useful tool to draw a line between 2 points in a figure, it is done by preparing an annotate dataset with instructions on "Move" and "Draw" lines between 2 points and using the "anno=" option in proc gplot. Please note that the "RecID" variable in the annotated dataset of Example 1 is based on the eye level and each subject has 2 records This is important in sub-setting the dataset by subjects in producing Figure 2, the rain-drop scatter plots with 10 subjects but 20 records in eyes.

In the second example, SAS function "call symput ()" is used to save the number of subjects at each site to macro variables to be displayed in the figure. Reference line positions are also saved in macro variables and used to group subjects by investigator sites by using "href= " option in the plot statement. The investigator site number is shown in the figure by using proper format. A SAS macro is used to produce thunderstorm scatter plots for baseline and each of the post-baseline visits (weeks 4, 8 and 12) although only one figure is shown for baseline and week 12.

CONCLUSION

This is the first time that the concept of a thunderstorm/rain-drop scatter plot is introduced and real figures are produced. A thunderstorm/rain-drop scatter plot is a useful data visualization tool to see overall data pattern and detect outliers for data with 2 or more values on the y-axis corresponding to one value on the x-axis. It can be used in many areas including clinical research, agriculture, and finance, etc.

SAS PROGRAMS PRODUCING THE THUNDERSTORM/RAIN-DROP SCATTER PLOTS

```
options mprint symbolgen validvarname=v7;

** You might need to change the locations to run the SAS program in your computer;
%let outloc = C:\Advanced Graphs in SAS\Figures;
%let pgmloc = C:\Advanced Graphs in SAS\SASPGMS\Thunderstorm Scatter Plots.sas
&sysdate &systime SAS V&sysver;

%let seed = 1699;

* Define device driver, file name and font name for figures;
%LET DRIVER=sasemf;
%LET EXT=emf;
%LET FONTNAME=Times-Roman;

** Define macro variables for title and footnote format: justification and fonts;
%LET FMTTITLE  =H=2.5 JUSTIFY=CENTER FONT="&FONTNAME";
%LET FMTFOOT =H=2   JUSTIFY=CENTER FONT="&FONTNAME";

* Define the options needed for the graphs.;
%LET GRPHOPTS=%STR(
goptions
  reset   = (title)           /* RESETS titles only */
  GUNIT=PCT
  rotate  = landscape
  gsfmode = replace
  gsfname = GSASFILE          /* OUTPUT FILE REFERENCE */
  device  = &DRIVER
  lfactor = 1
  hsize   = 8 in
  horigin = 0 in
  vsize   = 6.5 in
  vorigin = 0 in
  ftext   = "&FONTNAME"
  htext   = 10pt
  ftitle  = "&FONTNAME"
  htitle  = 10pt
 );

******************************************************************;
* Produce Sample Thunderstorm Scatter Plots Shown in Example 1     ;
******************************************************************;
proc format;
  value eyedf
       1 = 'OD'
       2 = 'OS';
  value iopdf
       1 = 'GAT'
```

```
      2 = 'DCT';
run;

** Simulate 100 subject's IOP data in both eyes;
data iop;
  do i=1 to 100;    ** 100 subjects;
    do j = 1 to 2; ** 2 eyes;
       do k = 1 to 2; ** 2 IOP type;
         subjid = 1000 + i;
         eye = j;
         iop_type = k;
         iop = round((RANNOR(&seed. + i + j + k)* 3.5 + 25), .1);
         output;
       end;
    end;
  end;
  format eye eyedf. iop_type iopdf.;
  drop i j k;
run;

** Simulate 100 subjects' CCT data ;
data cct;
  do i=1 to 100;    ** 100 subjects;
    do j = 1 to 2; ** 2 eyes;
      subjid = 1000 + i;
         eye = j;
      CCT = round((RANNOR(&seed. + i + j)* 50 + 550), 1);
        ** CCT between 500 to 600 microns;
         if 300 <= cct < 400 then cct = cct + 200;
         if 400 <= cct < 500 then cct = cct + 100;
         if 600 < cct <= 700 then cct = cct - 100;
         if 700 < cct <= 800 then cct = cct - 200;
         output;
    end;
  end;
  format eye eyedf. ;
  drop i j;
run;

proc sort data=cct;
  by subjid eye;
run;

data iopcct;
  merge iop cct;
  by subjid eye;
run;

proc transpose data=iopcct out= iopcct2 prefix=iop_;
  by subjid eye cct;
  var iop;
  id iop_type;
run;

data iopcct2;
  set iopcct2;
  iop_diff = iop_dct - iop_gat;
  recid = _N_;
run;

** Annotate dataset used to to connect the 2 IOPs of the same eye by a line;
data annoa;
  set iopcct2;
```

```
  by subjid;
   function='move'; xsys='2'; ysys='2'; x=cct; y=iop_gat; output;
   function='draw'; xsys='2'; ysys='2'; color='VILG'; x=cct; y=iop_dct; size=2;
line=1;
   output;
run;


&GRPHOPTS;
title1 &FMTTITLE "Figure 1. GAT and DCT IOP by CCT Value";
title2 &FMTTITLE "All Data from 100 Patients (200 Eyes)";
title3 &FMTTITLE "Thunderstorm Scatter Plot";
footnote1 &FMTFOOT "&pgmloc.";

SYMBOL1 H=2.5 C=RED CO=RED INTERPOL=none W=2 L=1 VALUE=circle;
SYMBOL2 H=2.5 C=BLACK CO=BLACK INTERPOL=none W=2 L=2 VALUE=dot;

AXIS1  MINOR=NONE OFFSET=(1,1) order=(10 to 40 by 5) LABEL=(FONT="&FONTNAME" ANGLE=90
       HEIGHT=2.5 font="&FONTNAME" "IOP (mm Hg)" ) VALUE=(H=1.5) ;
AXIS2  MINOR=NONE OFFSET=(1,1) order=(500 to 600 by 10) LABEL=(FONT="&FONTNAME"
       HEIGHT=2.5 font="&FONTNAME" "CCT (microns)") VALUE=(H=2);

FILENAME GSASFILE "&OUTLOC./Figure 1..&EXT";
proc gplot data= iopcct;
  plot       iop * cct = iop_type  / HAXIS=AXIS2 VAXIS=AXIS1 noframe anno=annoa;
  label iop_type = 'Legend:';
run;

title1 &FMTTITLE "Figure 2. GAT and DCT IOP by CCT Value";
title2 &FMTTITLE "-. Data from 10 Patients (20 Eyes)";
title3 &FMTTITLE "Rain-drop Scatter Plot";
footnote1 &FMTFOOT "&pgmloc.";

FILENAME GSASFILE "&OUTLOC./Thunderstrom..&EXT";
proc gplot data= iopcct;
  where 1001 <= subjid <= 1010;
  plot       iop * cct = iop_type  / HAXIS=AXIS2 VAXIS=AXIS1 noframe anno=annoa (where
= ( 1 <= recid <= 20));
  label iop_type = 'Legend:';
run;

*******************************************************************;
* Produce Sample Thunderstorm Scatter Plots Shown in Example 2     ;
*******************************************************************;
%let sitenum = 10;

proc format;
  value visdf
    1 = 'Baseline'
    2 = 'Week 4'
    3 = 'Week 8'
    4 = 'Week 12';

  value hrdf
      1 = 'Hour 0'
      2 = 'Hour 2'
      3 = 'Hour 8';
run;

** Generate the required number of subjects;
data site_subj;
  do i = 1 to &sitenum.;
    do j = 1 to 15;
```

```
      siteid = i;
       subjid = i * 1000 + j;
       shuffle = ranuni (&seed. + i + j);
      output;
    end;
  end;
  drop i j;
run;

proc sort data=site_subj;
  by shuffle;
run;

** Drop 50 subjects so each site might have different subject numbers;
data site_subj2;
  set site_subj;
  if _N_ <= 50 then delete;
run;

proc sort data=site_subj2;
  by siteid subjid;
run;

** Distribution of subjects by site;
proc freq data=site_subj2 noprint;
  table siteid/out=subj_freq (drop=PERCENT);
run;

** Save the subject number at each site to macro variables;
data _NULL_;
  set subj_freq;
  if siteid = 1 then call symput ('n_site1', put(count, 3.));
  if siteid = 2 then call symput ('n_site2', put(count, 3.));
  if siteid = 3 then call symput ('n_site3', put(count, 3.));
  if siteid = 4 then call symput ('n_site4', put(count, 3.));
  if siteid = 5 then call symput ('n_site5', put(count, 3.));
  if siteid = 6 then call symput ('n_site6', put(count, 3.));
  if siteid = 7 then call symput ('n_site7', put(count, 3.));
  if siteid = 8 then call symput ('n_site8', put(count, 3.));
  if siteid = 9 then call symput ('n_site9', put(count, 3.));
  if siteid = 10 then call symput ('n_site10', put(count, 3.));
run;

** Set up the study eye IOP values per subject based on visits/timepoints: ;
** assuming 6 to 8 mmHg reduction in IOP at post-bsl visits;
%let SD = 3.5;
data iop;
    set site_subj2;
    do i = 1 to 4; ** 4 visits;
       do j = 1 to 3; ** 3 timepoints/visit;
         visit = i;
         hour = j;
         if i = 1  then do;  ** Baseline;
              if j =1 then iop = round((RANNOR(&seed. + i + j)* &SD. + 25), .1);
              if j =2 then iop = round((RANNOR(&seed. + i + j)* &SD. + 23), .1);
              if j =3 then iop = round((RANNOR(&seed. + i + j)* &SD. + 22), .1);
         end;
         else if i > 1 then do; ** Post-baseline: with 6 to 8 mmHg reduction
              if j =1 then iop = round((RANNOR(&seed. + i + j)* &SD. + 17.5), .1);
              if j =2 then iop = round((RANNOR(&seed. + i + j)* &SD. + 16.5), .1);
              if j =3 then iop = round((RANNOR(&seed. + i + j)* &SD. + 16.2), .1);
         end;
          output;
```

```
      end;
  end;
  drop i j shuffle;
  format visit visdf. hour hrdf. ;
run;

proc sort data=iop;
  by siteid subjid visit hour;
run;

** Mean diurnal IOP at each visit: mean of hours 0, 2 and 8 IOP;
proc transpose data=iop out=iop_t ;
  by siteid subjid visit;
  var iop;
  id hour;
run;

data iop_MnDiur;
  set iop_t;
  iop_MnDiur = round (mean (Hour_0, Hour_2, Hour_8), .01);
  drop _NAME_ ;
run;

** Set the subject index number to be used in X-axis;
data subjid;
  set site_subj2;
  subj_id = _N_;
  if siteid = 1 then subj_index = _N_;
  if siteid = 2 then subj_index = _N_ + 1; ** leave space for reference line;
  if siteid = 3 then subj_index = _N_ + 2;
  if siteid = 4 then subj_index = _N_ + 3;
  if siteid = 5 then subj_index = _N_ + 4;
  if siteid = 6 then subj_index = _N_ + 5;
  if siteid = 7 then subj_index = _N_ + 6;
  if siteid = 8 then subj_index = _N_ + 7;
  if siteid = 9 then subj_index = _N_ + 8;
  if siteid = 10 then subj_index = _N_ + 9;
  keep siteid subjid subj_index;
run;

** Set macro variables to hold the reference line locations in x-axis;
data _NULL_;
  refn1 = &n_site1. + 1;
  refn2 = &n_site1. + &n_site2. + 2;
  refn3 = &n_site1. + &n_site2.  + &n_site3. + 3;
  refn4 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + 4;
  refn5 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + &n_site5. +5;
  refn6 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + &n_site5. + &n_site6. + 6;
  refn7 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + &n_site5. + &n_site6. +
       &n_site7. + 7;
  refn8 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + &n_site5. + &n_site6. +
       &n_site7. + &n_site8. + 8;
  refn9 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + &n_site5. + &n_site6. +
       &n_site7. + &n_site8. + &n_site9. + 9;
  refn10 = &n_site1. + &n_site2.  + &n_site3. + &n_site4. + &n_site5. + &n_site6. +
       &n_site7. + &n_site8. + &n_site9. + &n_site10. + 10;

  call symput ("Ref1", refn1);
  call symput ("Ref2", refn2);
  call symput ("Ref3", refn3);
  call symput ("Ref4", refn4);
  call symput ("Ref5", refn5);
  call symput ("Ref6", refn6);
```

```
  call symput ("Ref7", refn7);
  call symput ("Ref8", refn8);
  call symput ("Ref9", refn9);
  call symput ("Ref10", refn10);
run;

data mndiur_fig;
  merge iop_MnDiur subjid;
  by siteid subjid;
run;

proc transpose data=MnDiur_fig out=iop_MnDiurT prefix=vst_;
  by siteid subjid subj_index;
  var iop_MnDiur;
run;

data  mndiur_anno;
  set iop_MnDiurT;
  drop _NAME_ ;
  rename vst_1 = BSL vst_2 = WK4 vst_3 = WK8 vst_4 = WK12;
run;

proc format;
  value siteid
    1 = '1'
    &ref1. = '2'
    &ref2. = '3'
    &ref3. = '4'
    &ref4. = '5'
    &ref5. = '6'
    &ref6. = '7'
    &ref7. = '8'
    &ref8. = '9'
    &ref9. = '10'
    OTHER = ' ';
run;

&GRPHOPTS;
%macro Thunderstorm_Plot (vstnum1=, vstnum2=, var1=, var2=, vardes1=, vardes2=,
figname=);

** Annotate dataset to connect the 2 IOPs of the same subject by a straight line;
data annob;
  set mndiur_anno;
  by subj_index;
   function='move'; xsys='2'; ysys='2'; x=subj_index; y=&var1.; output;
   function='draw'; xsys='2'; ysys='2'; color='VILG'; x=subj_index; y=&var2.; size=2;
line=1;
   output;
run;

SYMBOL1 H=2.5 C=RED CO=RED INTERPOL=none W=2 L=1 VALUE=circle;
SYMBOL2 H=2.5 C=BLACK CO=BLACK INTERPOL=none W=2 L=2 VALUE=dot;

AXIS1  MINOR=NONE OFFSET=(1,1) order=(10 to 30 by 5) LABEL=(FONT="&FONTNAME" ANGLE=90
       HEIGHT=2.5 font="&FONTNAME" "Mean Diurnal IOP at &vardes1. and &vardes2.
       (mm Hg)" ) VALUE=(H=1.5) ;
AXIS2  MAJOR=NONE MINOR=NONE OFFSET=(1,1) order=(0 to 110 by 1)
       LABEL=(FONT="&FONTNAME"  HEIGHT=2.5 font="&FONTNAME" "Subject Grouped by
       Investigator Sites") VALUE=(H=2) reflabel=(position=bottom c= blue
       font="&FONTNAME" h=2 j=r "(n=&n_site1.))" "(n=&n_site2.)" "(n=&n_site3.)"
       "(n=&n_site4.)" "(n=&n_site5.)" "(n=&n_site6.)" "(n=&n_site7.)" "(n=&n_site8.)"
       "(n=&n_site9.)" "(n=&n_site10.)");
```

```
FILENAME GSASFILE "&OUTLOC./&figname..&EXT";
proc gplot data= mndiur_fig;
  where visit in (&vstnum1., &vstnum2.);
  plot iop_MnDiur * subj_index = visit  / HAXIS=AXIS2 VAXIS=AXIS1 noframe anno=annob
       href=0 &ref1. &ref2. &ref3. &ref4. &ref5. &ref6. &ref7. &ref8. &ref9. &ref10.;
  label visit = 'Legend:'; format subj_index siteid.;
run;
quit;

%mend Thunderstorm_Plot;

** Mean diurnal IOP at Baseline and Week 12;
title1 &FMTTITLE "Figure 3. Mean Diurnal IOP at Baseline and Week 12";
title2 &FMTTITLE "Thunderstorm Scatter Plot";
footnote1 &FMTFOOT "&pgmloc.";
%Thunderstorm_Plot (vstnum1=1, vstnum2=4, var1=Bsl, var2=Wk12, vardes1=Baseline,
vardes2=Week 12, figname=Figure 3);

/* The 2 figures below are not displayed in the manuscript   */
** Mean diurnal IOP at Baseline and Week 8;
title1 &FMTTITLE "Figure 4. Mean Diurnal IOP at Baseline and Week 8";
title2 &FMTTITLE "Thunderstorm Scatter Plot";
footnote1 &FMTFOOT "&pgmloc.";
%Thunderstorm_Plot (vstnum1=1, vstnum2=3, var1=Bsl, var2=Wk8, vardes1=Baseline,
vardes2=Week 8, figname=Figure 4);

** Mean diurnal IOP at Baseline and Week 4;
title1 &FMTTITLE "Figure 5. Mean Diurnal IOP at Baseline and Week 4";
title2 &FMTTITLE "Thunderstorm Scatter Plot";
footnote1 &FMTFOOT "&pgmloc.";
%Thunderstorm_Plot (vstnum1=1, vstnum2=2, var1=Bsl, var2=Wk4, vardes1=Baseline,
vardes2=Week 4, figname=Figure 5);
```

REFERENCES

- Allison, Robert. 2012. *SAS/GRAPH®: Beyond the Basics*. Cary, NC: SAS Institute, Inc.
- Boehm AG, Weber A, Pillunat LE, et al. 2008. "Dynamic contour tonometry in comparison to intracameral IOP measurements," *Invest Ophthalmol Vis Sci.* 2008;49:2472–7.
- Carpenter, A. 1999. *Annotate: Simply the Basics*. Cary, NC; SAS Institute.
- Cleveland, William. 1985. *The Elements of Graphing Data. Graphics Press.*
- Ehlers N, Bramsen T, Sperling S. 1975. "Applanation tonometry and central corneal thickness," *Acta Ophthalmol (Copenh)* 53:34e43.
- Goldmann H, Schmidt T. 1957. "Uber applanationstonometrie," *Ophthalmologica.* 134:221e42.
- Hyndman, R.J., and Shang, H.L. 2010. "Rainbow Plots, Bagplots, and Boxplots for Functional Data," *Journal of Computational and Graphical Statistics*. 19(1): 29-45.
- Tufte, Edward. 1983. *The Visual Display of Quantitative Information.* Graphics Press.
- Tufte, Edward. 1997. *Visual Explanations*. Graphics Press.
- Tufte, Edward. 1997. *Beautiful Evidence.* Graphics Press
- Whitacre MM, Stein R. 1993. "Sources of error with use of Goldmann-type tonometers," *Surv Ophthalmol* 38:1e30.

ACKNOWLEDGMENTS

The author would like to thank Paula Bernstein for her excellent editing and valuable comments.

CONTACT INFORMATION

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

Charlie Chunhua Liu, PhD
Biostatistics & Programming

Allergan, Inc.
2525 Dupont Dr.
Irvine, CA 92612
Work Phone: (714) 246.6296
E-mail: Liu_Charlie@Allergan.om