PharmaSUG 2013 - Paper DG07

Communication of Statistical Findings by Tables and Graphs

Howard Liang, inVentiv Health Clinical, Eden Prairie, Minnesota

ABSTRACT

When communicating clinical trial results, tables summarizing data are useful. However, if we have data summarized by graphs in addition to tables, then the presentation will be more effective. Moreover, since the summary statistics do not explain what actually happen to individual subjects, if we add some graphical display of individual data, then we can make our conclusion more convincing. Examples will be used in this paper to show how to convert the summary statistics in a table to graphs and how to create graphs for individual subjects using the SAS®/GRAPH GPLOT procedure.

INTRODUCTION

The proper use of graphical display is important in statistical analysis and interpretation of the results. The examples in this paper will explain how to communicate clinical trial results by a table and graphs when using a simple mixed model repeated measurement (MMRM) analysis on the simulated lab data. The first example converts the statistical results in a table to graphs. The second example creates graphs for individual data.

DATA SIMULATION

Suppose we have a phase 1 clinical trial data with a sample of 20 subjects randomly assigned to placebo and a new drug with a 1:1 randomization ratio. The outcomes are lab measured at four time points. The first measurement is a pre-dose baseline assessment which is followed by three post-dose repeated measurements. The following code produces the simulated lab data:

```
**** INTRODUCE VARIABILITY;
data sample;
  do param="A(mmol/L)" , "B(g/L)" , "C(U/L)" ;
    do subjid=1 to 20;
      if uniform(12345)>0.5 then treatment='pl';
      else treatment='dx';
      do month=0 to 3;
        if param=:"A" then result=10+5*normal(6789);
        else result=5+0.5*normal(6789);
        if month>0 and treatment="dx" then
           if param=:"B" then result=result-0.5*month;
           else if param=:"C" then result=result+0.7*month;
        output;
      end;
    end;
  end;
run;

**** DERIVE CHANGE AND BASELINE VARIABLES;
proc sql;
  create table sample2 as select a.param, a.subjid, a.treatment, a.month,
                                 a.result-b.result as change, b.result as baseline
    from sample(where=(month>0)) as a, sample(where=(month=0)) as b
    where a.param=b.param and a.subjid=b.subjid
    order by param, subjid, month;
quit;

**** FIT A SIMPLE MIXED MODEL REPEATED MEASUREMENT (MMRM) ON THE SIMULATED LAB DATA;
ods output lsmeans=lsm1 diffs=diff1(where=( month=_month and  _treatment='pl'));
proc mixed data=sample2;  by param;
  class subjid month treatment;
  model change=treatment month treatment*month baseline baseline*month/ddfm=kr;
  repeated month/subject=subjid type=un;
  lsmeans treatment*month/cl pdiff;
run;
quit;
```

The data produced by PROC MIXED code above is stored in the work data sets LSM1 and DIFF1, and they can be combined to produce the following typical summary table by PROC REPORT:

Lab Parameter	Month	Placebo LS Mean (95% CI)	New Drug LS Mean (95% CI)	Difference to Placebo LS Mean(95% CI)	p-value
A(mmol/L)	1	-1.52 (-4.33, 1.29)	1.38 (-1.44, 4.19)	2.89(-1.14, 6.93)	0.148
	2	3.55 (0.99, 6.12)	0.54 (-2.02, 3.10)	-3.01(-6.69, 0.67)	0.102
	3	-0.18 (-3.28, 2.91)	0.50 (-2.60, 3.59)	0.68(-3.77, 5.12)	0.751
B(g/L)	1	0.07 (-0.26, 0.39)	-0.55 (-0.91, -0.19)	-0.62(-1.11, -0.13)	0.016
	2	-0.30 (-0.64, 0.05)	-1.09 (-1.47, -0.71)	-0.79(-1.31, -0.27)	0.005
	3	-0.10 (-0.44, 0.25)	-1.53 (-1.91, -1.15)	-1.43(-1.95, -0.91)	<0.001
C(U/L)	1	0.12 (-0.26, 0.51)	0.75 (0.37, 1.14)	0.63(0.07, 1.19)	0.029
	2	0.40 (0.08, 0.72)	1.25 (0.93, 1.57)	0.85(0.38, 1.32)	0.001
	3	0.08 (-0.27, 0.44)	2.34 (1.99, 2.70)	2.26(1.74, 2.77)	<0.001

Table 1. Typical Summary Table

CONVERTING A DATA SUMMARY TABLE TO GRAPHS

```
**** PREPARE DUMMY DATA FOR BASELINE IN THE PLOTS;
data base;
   do param="A(mmol/L)" , "B(g/L)" , "C(U/L)" , "D(fL)";
      do treatment="dx" , "pl";
         month=0;
         estimate=0;
         hilo=0;
         output;
      end;
   end;
run;

**** CREATE OFFSET MONTH FOR NEW DRUG;
**** DEFINE NEW VARIABLES FOR 95% CI AND POSTIONS FOR ESTIMATE AND 95% CI;
data est(drop=lower upper);
    set lsm1(keep=param treatment month estimate lower upper);
        if treatment="dx" then  month=month-0.05;
        pos=1;
        output;
        pos=0;
        hilo=lower;
        output;
        pos=0;
        hilo=upper;
        output;
run;

**** COMBINE AND SORT;
data final;
    set base est;
run;

proc sort data=final;
by param treatment month;
run;

**** ANNOTATE P-VALUE VS PLACEBO;
data pval_vs_pl;
    set diff1(keep=param treatment month probt);
    if treatment="dx"  then  month=month-0.05;
run;
```

```
data position;
    set  est;
    if pos=1;
    keep estimate treatment month param;
run;

proc sort data=position;
   by param treatment month;
run;

proc sort data=pval_vs_pl;
   by param treatment month;
run;

data anno;
    length function color $8 text $26 sign $25;
    retain xsys '2' ysys '2' hsys '1' when 'A';
    merge position pval_vs_pl;
         by param treatment month;
            if  .<probt<=0.001  then sign="***";
            if 0.001<probt<=0.01  then sign="**";
            if  0.01<probt<=0.05  then sign="*";
            function='Label';
            color='black';
            size=2;
            text=" "||compress(sign);
            position='D';
            x=month;
            y=estimate;
            if treatment = "dx";
run;

**** GET A PARAMETER DATA SET FOR CALL EXECUTE TO LOOP THROUGH;
**** GET MINIMUM P-VALUE TO FLAG FOOTNOTE IF P<0.05;
proc sql;
       create table param as select unique param
              from sample
              group by param
              order by param;

       create table pv as select unique param, min(probt) as minp
              from anno
              group by param
              order by param;
quit;

*** CREATE PLOT 1;
ods listing close;
options nocenter;
options orientation=landscape nocenter pageno=1;

ods rtf file="&outpath./plot1.rtf";

goptions reset=goptions device = png target=png
         xmax=12in ymax=10in ftext = simplex
         htext = 1.5 vsize=6.4 in hsize=8 in
         gsflen = 80 gprotocol = sasgpasc
         htitle=1.4 gsflen=150;

goptions reset=symbol;
symbol1  h=2 i=j font=marker   v='P'    c=red     w=1  h=2 l=1;
symbol2  h=1.5 i=j font=marker   v='U'    c=blue   w=1  h=1.5 l=1;
symbol3  i=hilot v=none   c=red;
symbol4  i=hilot v=none   c=blue;

legend1 mode=protect across=1 frame fwidth=1 position=(top left inside)
```

```
value=(h=2 j=l) shape=symbol(0.01,2 label=none;

axis1 label=none value=("baseline" "month 1" "month 2" "month 3" ) order=(0 to 3 by 1)
minor=none;

**** FIRST MACRO FOR CALL EXECUTE TO LOOP THROUGH THE VALUES OF
     VARIABLE "PARAM" IN DATA SET "PARAM";

%macro plot(param);

   data anno1;
       set anno(where=(param="&param."));
   run;

   data _null_;
       set pv(where=(param="&param."));
       all symputx("pv",minp,'g');
   run;

   **** DYNAMICALLY GENERATE THE MAXIMUM AND MINIMUM OF Y-AXIS;
   proc sql noprint;
     select ceil(max(hilo)) into: ymax from final(where=(param="&param."));
     select floor(min(hilo)) into: ymin from final(where=(param="&param."));
   quit;

   axis2 label=(a=90 "LS Mean and 95% CI")  order=(&ymin. to &ymax. ) minor=(n=1);
   axis3 label=none major=none  value=none  order=(&ymin. to &ymax. ) minor=none;

   title1 h=2 "Plot of LS Mean Change from Baseline in %trim(&param.)
               by Treatment  over Time";
   footnote1 " ";

   **** CREATE FOOTNOTES BASED ON MINIMUM P-VALUE OF THE PARAMETER;
   %if %sysevalf(&pv.<=0.001) %then %do;
       footnote1 " ";
       footnote2 "  * denote p-value (vs Placebo) <= 0.05";
       footnote3 " ** denote p-value (vs Placebo) <= 0.01";
       footnote4 "*** denote p-value (vs Placebo) <= 0.001";
   %end;
   %else %if %sysevalf(&pv.<=0.01) %then %do;
       footnote1 " ";
       footnote2 " * denote p-value (vs Placebo) <= 0.05";
       footnote3 "** denote p-value (vs Placebo) <= 0.01";
   %end;
   %else %if %sysevalf(&pv.<=0.05) %then %do;
       footnote1 " ";
       footnote2 "* denote p-value (vs Placebo) <= 0.05 ";
   %end;

   proc gplot data=final;
       where param="&param.";
       format estimate best4. treatment $gp. ;
       plot estimate*month=treatment/vref=0 lvref=3 cvref=ltgray
            haxis=axis1 vaxis=axis2 legend=legend1 annotate=anno1;
       plot2 hilo*month=treatment/vaxis=axis3 nolegend;
   run;
   quit;

%mend plot;

**** USE CALL EXECUTE TO CREATE GRAPH BY PARAMETER;
data _null_;
      set param;
      call execute('%nrstr(%plot(%str('||param||')))');
run;
```

```
ods rtf close;
ods listing;
```

The above codes and macro create the summary graphs shown below:

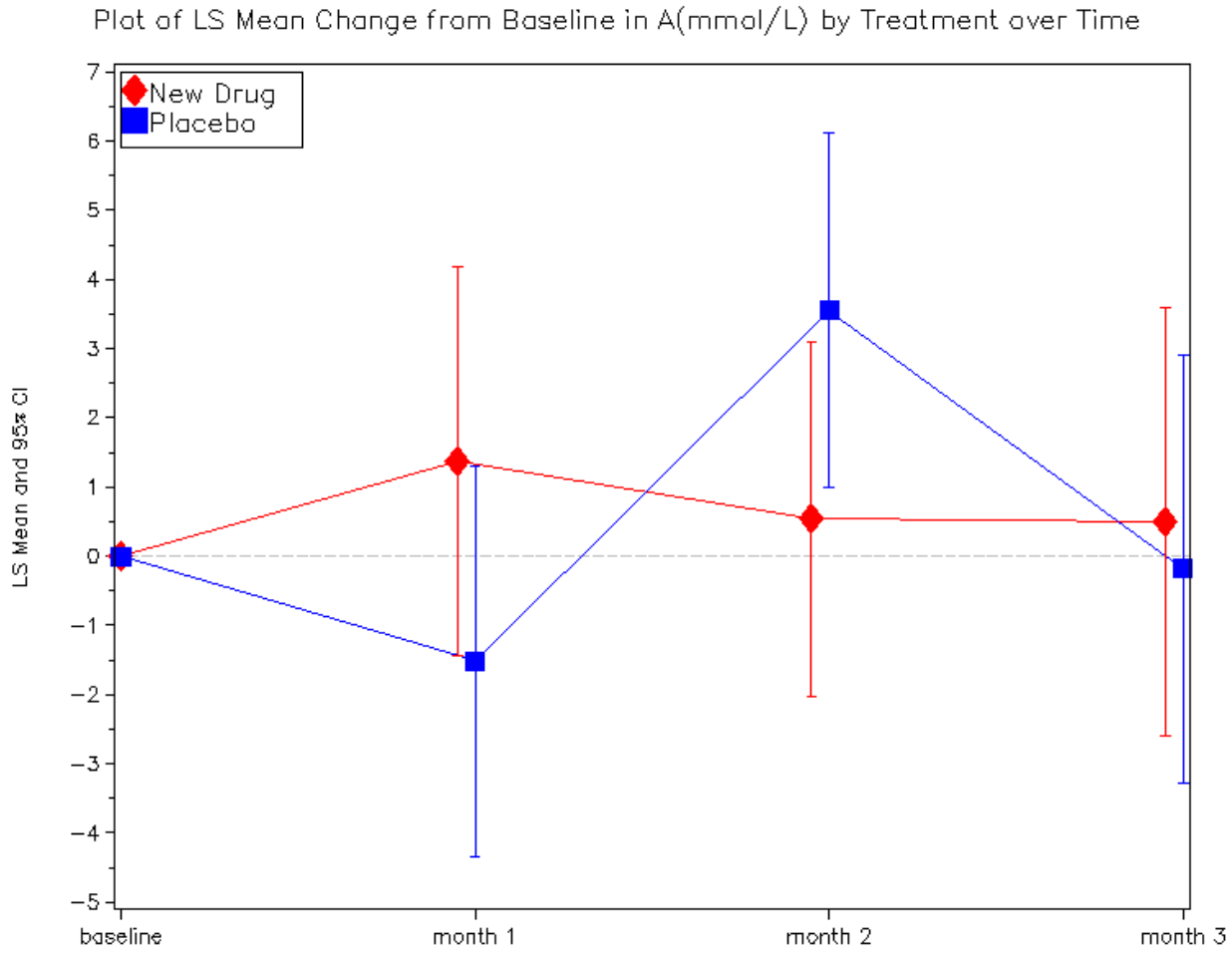

Figure 1. Summary Graph for Parameter A

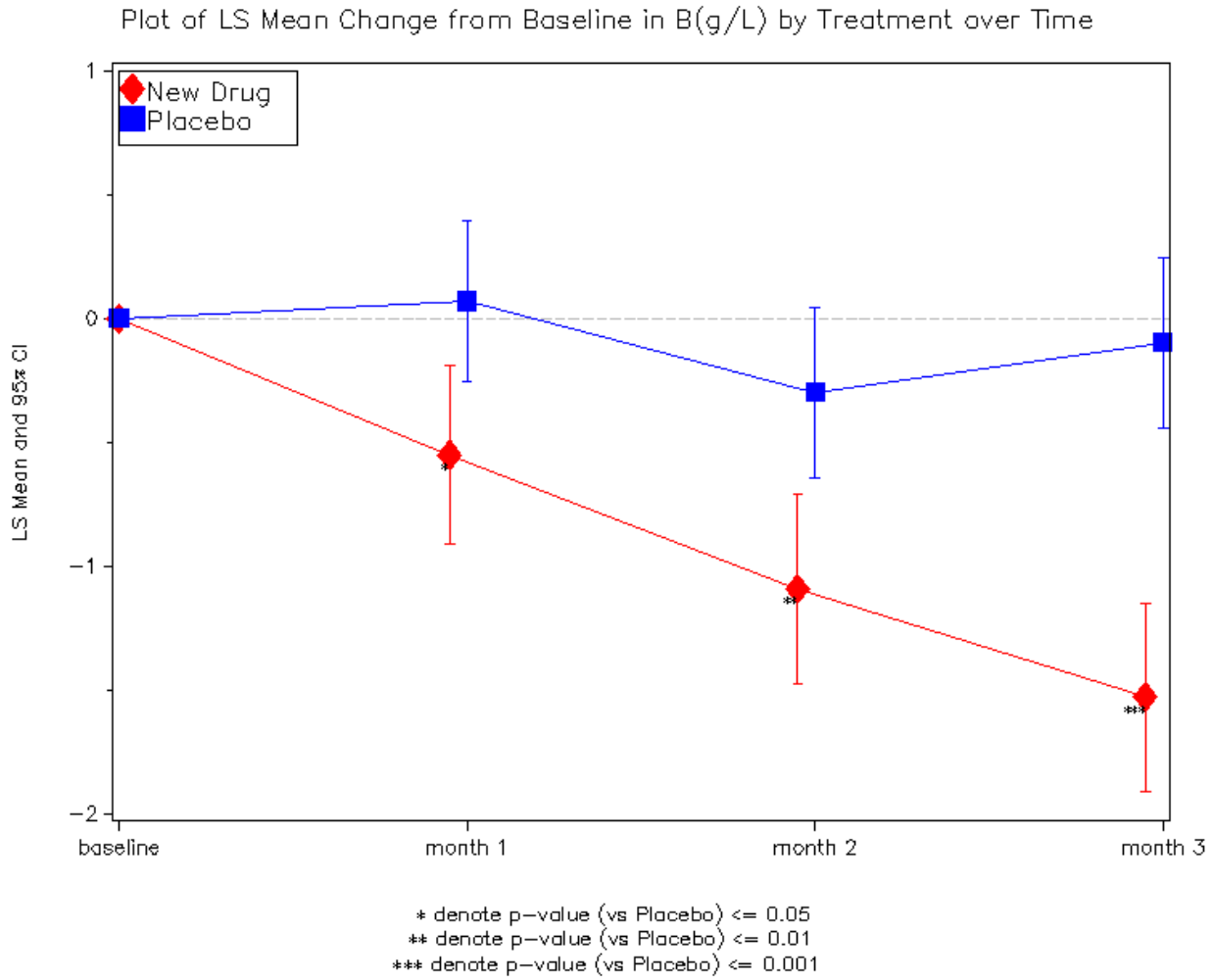

Figure 2. Summary Graph for Parameter B

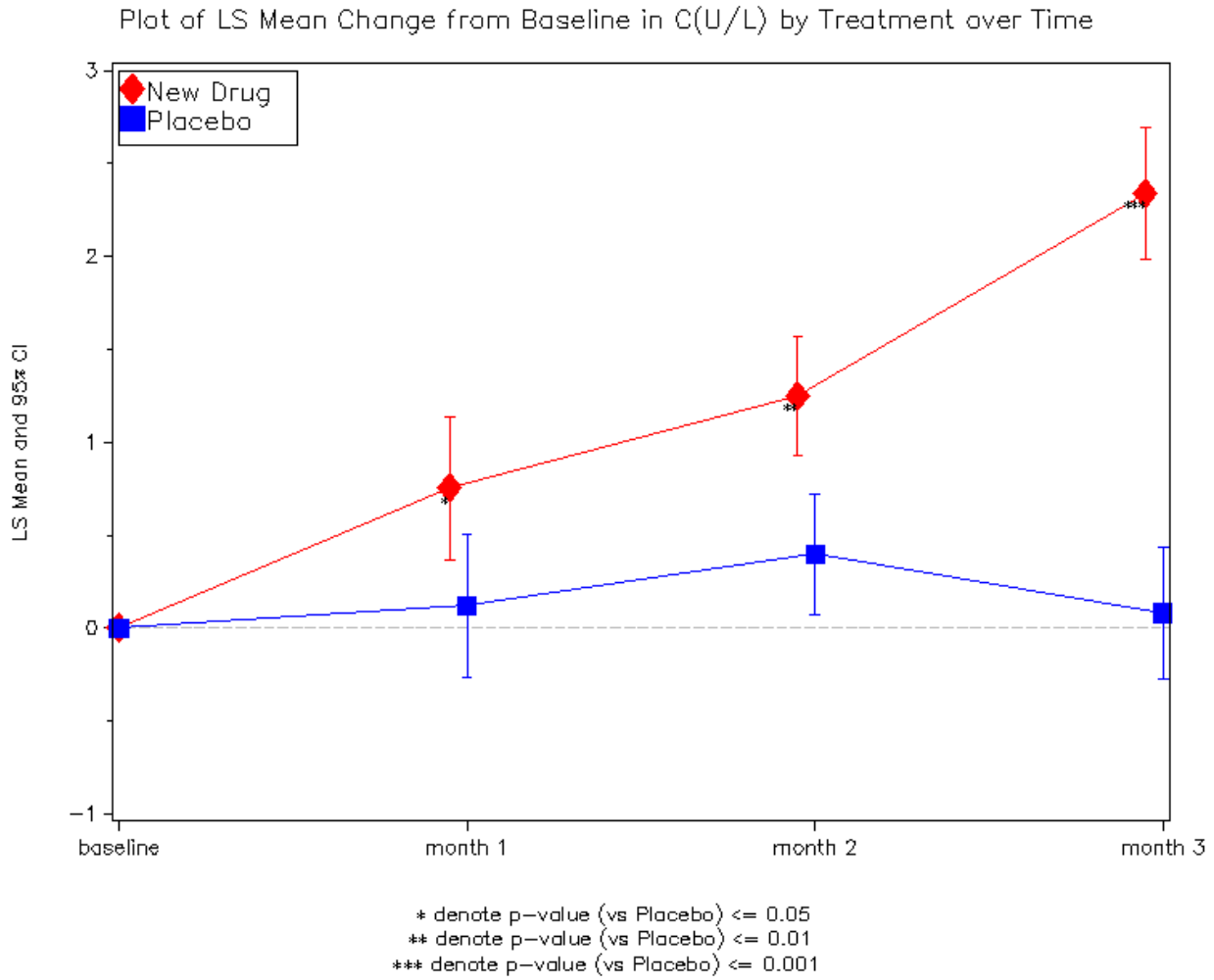

Figure 3. Summary Graph for Parameter C

CREATING GRAPHS FOR INDIVIDUAL DATA

```
**** CREATE OFFSET MONTH IN SEPARATED PANEL FOR NEW DRUG SUBJECTS;
data final2;
    set sample;
    if treatment="dx" then  month=month+5;
run;

proc sort data=final2;
by param treatment month;
run;

*** CREATE PLOT 2;
ods listing close;
options nocenter;
options orientation=landscape nocenter pageno=1;
ods rtf file="&outpath./plot2.rtf";

goptions reset=goptions device = png target=png
         xmax=12in ymax=10in ftext = simplex
         htext = 1.5 vsize=6.4 in hsize=8 in
         gsflen = 80 gprotocol = sasgpasc
         htitle=1.4 gsflen=150;
```

```
axis1 label=(h=2 j=c " |------ Placebo ------|           |------ New Drug ------|")
value=(" " "baseline" "month 1" "month 2" "month 3" " " "baseline"  "month 1" "month
2" "month 3" " ")
order=(-1 to 9) minor=none;

**** SECOND MACRO FOR CALL EXECUTE TO LOOP THROUGH THE VALUES OF VARIABLE "PARAM";
**** IN DATA SET "PARAM";

%macro plot2(param);

   **** DYNAMICALLY GENERATE THE MAXIMUM AND MINIMUM OF Y-AXIS;
   proc sql noprint;
       select ceil(max(result)) into: ymax from final2(where=(param="&param."));
       select floor(min(result)) into: ymin from final2(where=(param="&param."));
   quit;

   options reset=symbol;

   symbol1 h=2 i=j v=dot r=20;

   axis2 label=(a=90 "&param.")  order=(&ymin. to &ymax.) minor=none;

   title1 h=2 "Individual Plot of %trim(&param.) by Treatment over Time";

   footnote1 " ";

   proc gplot data=final2;
       where param="&param.";
       format result best4. treatment $gp. ;
       plot result*month=subjid/href=4 lvref=1 chref=ltgray haxis=axis1 vaxis=axis2
       nolegend;
   run;
   quit;

%mend plot2;

**** USE CALL EXECUTE TO CREATE GRAPH BY PARAMETER;
data _null_;
   set param;
   call execute('%nrstr(%plot2(%str('||param||')))');
run;

ods rtf close;
ods listing;
```

The above codes and macro create the individual graphs shown below:

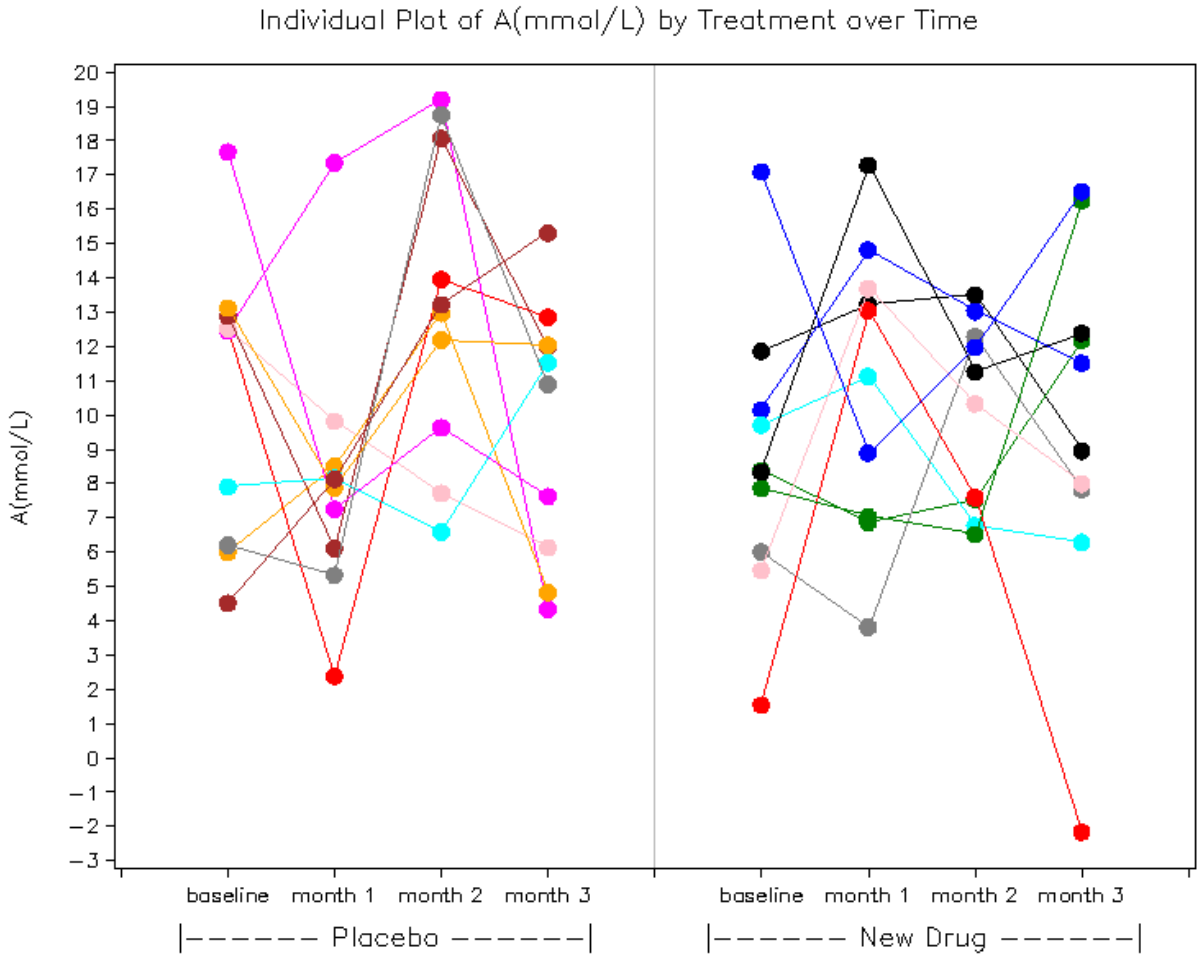

Figure 4. Individual Graph for Parameter A

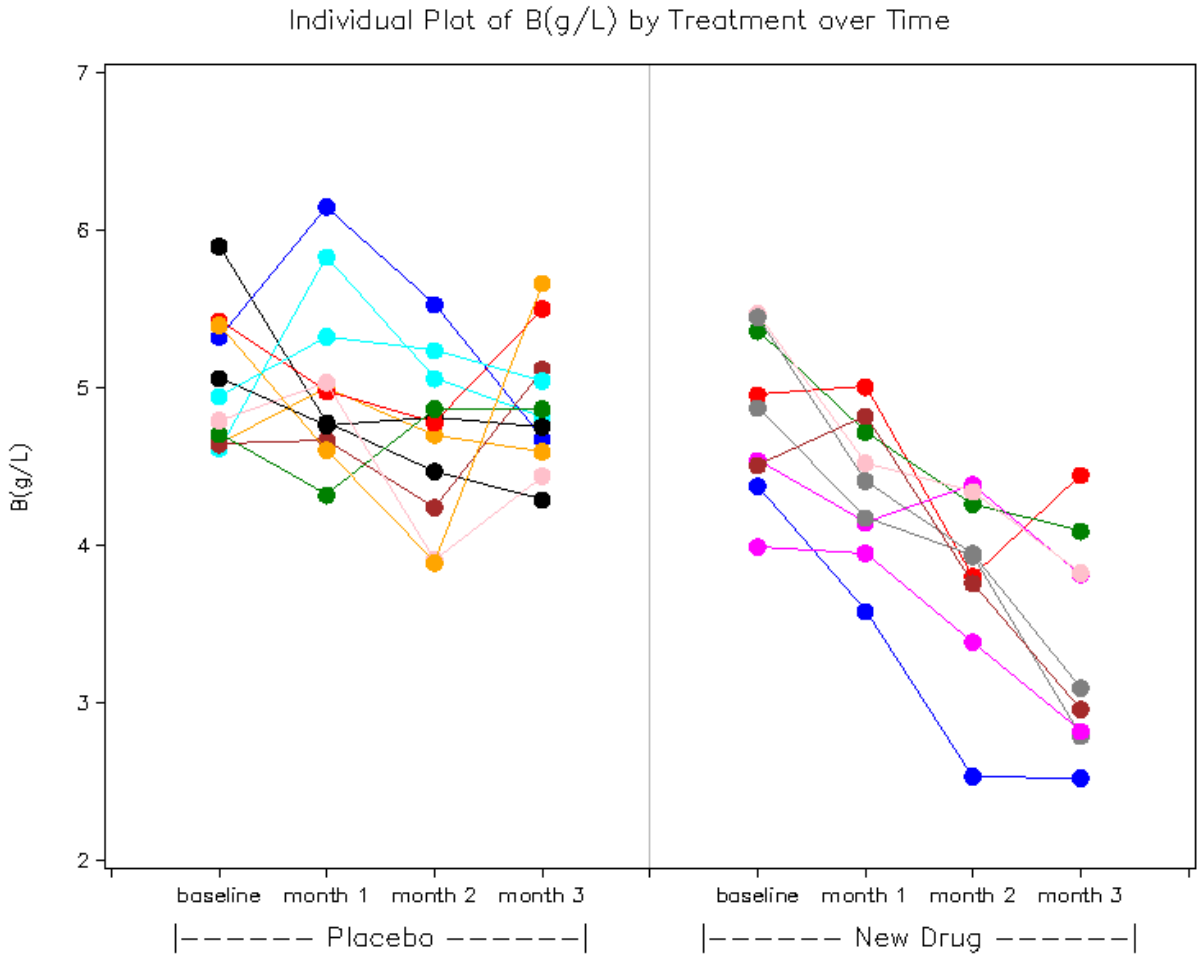

Figure 5. Individual Graph for Parameter B

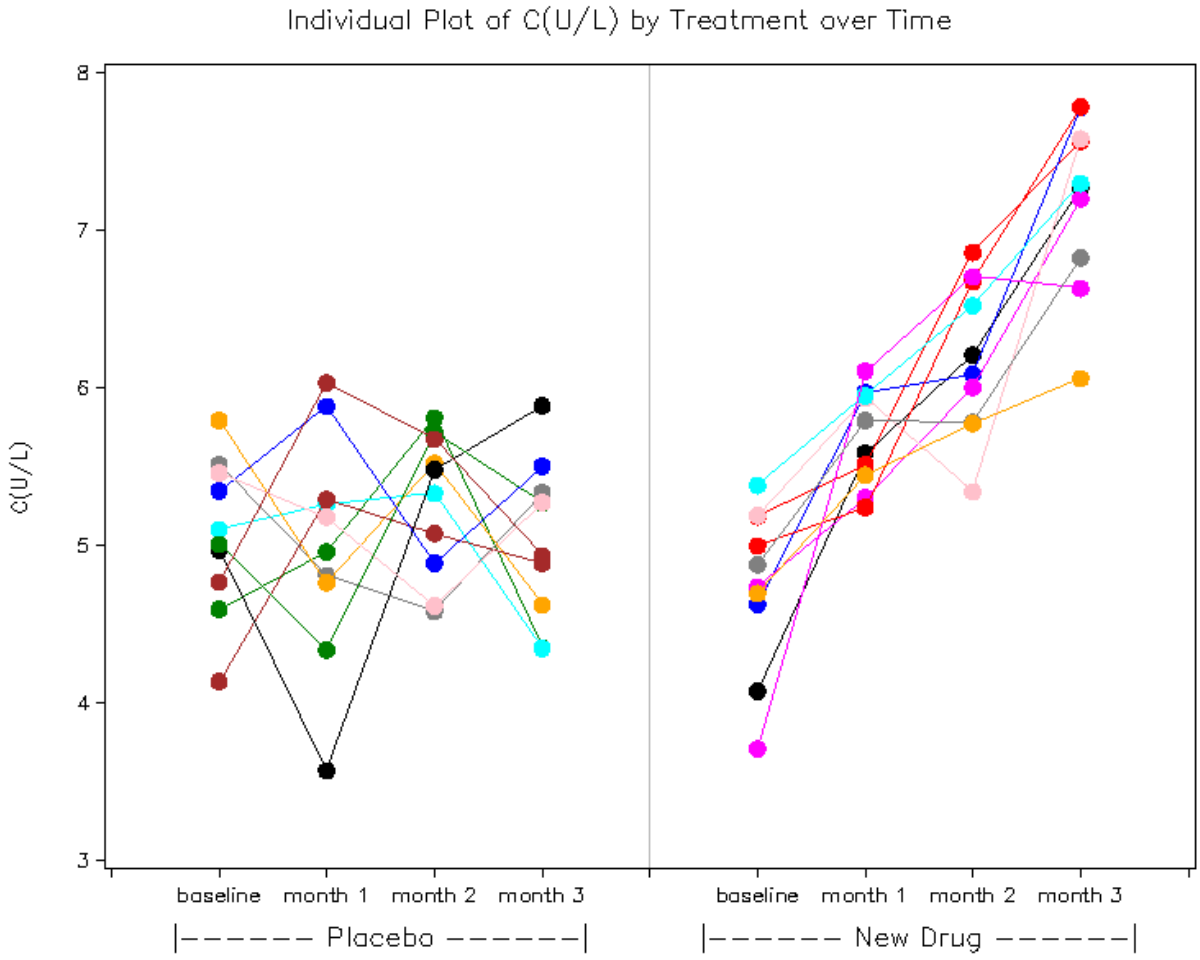

Figure 6. Individual Graph for Parameter C

CONCLUSION

This paper has presented a couple of basic graphs for both statistical findings and individual data that are often encountered in clinical trial analysis and reporting. There exist many more advanced procedures and various tools in SAS/GRAPH which provide help to anyone who wants to use graphical comparisons of treatments to achieve more effective communication of clinical trial results. I hope this paper is useful to statistical analysts and others who perform statistical analysis duties in clinical trials or clinical research.

ACKNOWLEDGMENTS

I greatly appreciate to the contributions of Nancy Brucken who reviewed the paper in draft and made many suggestions for improvement.

CONTACT INFORMATION

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

Name: Howard Liang
Enterprise: inVentiv Health Clinical
Address: 12125 Technology Drive
City, State ZIP: Eden Prairie, MN 55344
Work Phone: (952) 833-7462
E-mail: howard.liang@inventivhealth.com
Web: http://www.inVentivHealthclinical.com