

Graphical Representation of Patient Profile for Efficacy Analyses in Oncology

William Wu, Herodata LLC, San Jose, CA

Xiaoxian Dai, Pharmacyclics, Inc., Sunnyvale CA

Linda Gau, Pharmacyclics, Inc., Sunnyvale CA

ABSTRACT

With the advancement of cancer care and the need to develop new therapies for cancer, oncology clinical trials have become more and more popular. Currently, most of the efficacy analyses in oncology clinical trials are represented by tables and listings. For patients with multiple visits, it is not convenient to read and understand the information presented in the tables and listings. In this paper, we introduce a SAS® macro developed in SAS 9.3 to create patient profile graphs in a multi-page PDF file using the SAS Output Delivery System (ODS) for efficacy analyses. On each page of the PDF file, there are two graphs. In addition, supporting information is added as listings to the graphs using the SAS Annotate Facility. Efficacy data for patients with chronic lymphocytic leukemia (CLL) are used to generate the graphics output using the developed SAS macro. Compared to the tables and listings, the graphs are more intuitive and much easier to understand.

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is the most common leukemia in Western countries (Rozman and Montserrat, 1995). It mainly affects the elderly, with median age of diagnosis at 72 years (Eichhorst *et al.*, 2011). The cause of chronic lymphocytic leukemia is unknown. Currently, many clinical trials are conducted on chronic lymphocytic leukemia. With so much data to be analyzed and reviewed, it is important to present the data in a clear and efficient way so that the information is easier for reviewers to read.

The traditional method of presenting efficacy data in oncology is to use tables and listings. When the data become complicated, tables and listings become a little difficult to be understood, and they are not intuitive. In addition, it is time consuming to present all the useful information for the same patient from different tables and listings. An alternative method is to use graphs to convey the useful information from efficacy data. With the release and enhancements of SAS/GRAPH procedures, generating graphs has become more convenient than before. The combination of SAS/GRAPH and the annotation technique is a powerful tool, which can be used to create a wide variety of high quality graphs. This paper describes a macro developed in SAS 9.3 to create efficacy patient profile graphs in the PDF format using SAS/GRAPH and the annotation technique.

DATA PREPARATION

Before creating a graphic patient profile for the efficacy analyses, it is important to understand the input data sets and related variables used to invoke the macro. Due to confidentiality of the patient data, the data used in this paper mimic the efficacy data for patients with CLL and were randomly generated. The following lists the input data sets and the related variables that might be used to generate the graphics output.

1. Demographics (DM)
 - (a) Subject ID; (b) Safety population flag; (c) Treatment group; (d) First dose date; (e) Last dose date; (f) Other demographic information that needs to be displayed as the subtitle on the top of each page.
2. Hematology (HEM)
 - (a) Subject ID; (b) Lab test name code (the short name for lab tests); (c) Lab test name; (d) Sample collection date; (e) Sample collection time; (f) Visit name; (g) Lab test results.
3. Lymph nodes (LYM)
 - (a) Subject ID; (b) Lesion number; (c) Lesion location; (d) Assessment date; (e) Visit name; (f) Primary diameter; (g) Perpendicular diameter; (h) Lesion status.
4. Investigator's response assessment (RS)

(a) Subject ID; (b) Response assessment date; (c) Investigator's response assessment result.

5. Concomitant medication (CM)

Data for growth factors and blood transfusion are also displayed in the plots. Below is a list of the useful variables.

(a) Subject ID; (b) Medication name; (c) Medication start date; (d) Medication end date.

6. Bone marrow aspirate / biopsy (BP)

(a) Subject ID; (b) Date for the bone marrow aspirate / biopsy.

7. CT scan (CT)

(a) Subject ID; (b) CT scan date.

The input data sets for demographics, hematology and lymph nodes are required for the macro to successfully run, although the macro itself allows some patients to have no valid data. Except for these three data sets, other data sets are optional. If a patient has no valid data, the macro will create a PDF output with 'No Valid Data Available' in the center of the corresponding plot for that patient.

EFFICACY PATIENT PROFILE PLOTTING

As mentioned before, the macro creates the output with multiple pages in the PDF format, with one page per patient. Each page includes two plots, two listings, and the related titles/footnotes which are located outside of the graphic area. Figure 1 is a sample output for the efficacy patient profile graphs. It can be seen from Figure 1 that a lot of useful information can be displayed on the same page for the same patient. The variations in the percent change from baseline over time can be easily understood. Therefore, the graphical representation of the patient profile for efficacy analyses is very intuitive and easy to understand. Below is a sample code used to generate the plot and PDF output.

```
options nodate nonumber orientation=landscape topmargin=1.0in bottommargin=.75in
  leftmargin=.75in rightmargin=.75in;

goptions reset=all device=sasprtc gunit=pct cback=white htext=9pt
  ftext='Times' border;

ods listing close;
ods pdf file="%scan(&outpathf,1,..).pdf" color style=printer bookmarklist=show
  startpage=never pdftoc=2;

%do i=1 %to &_o_;
ods pdf bookmarkgen=yes;
ods proclabel ..;
title1 ..;
title2 ..;
title3 ..;
title4 ..;
footnote1 ..;
footnote2 ..;

data _null_;
  file print;
  put %if &i=1 %then 'A0'x; %else _blankpage_;;
run;

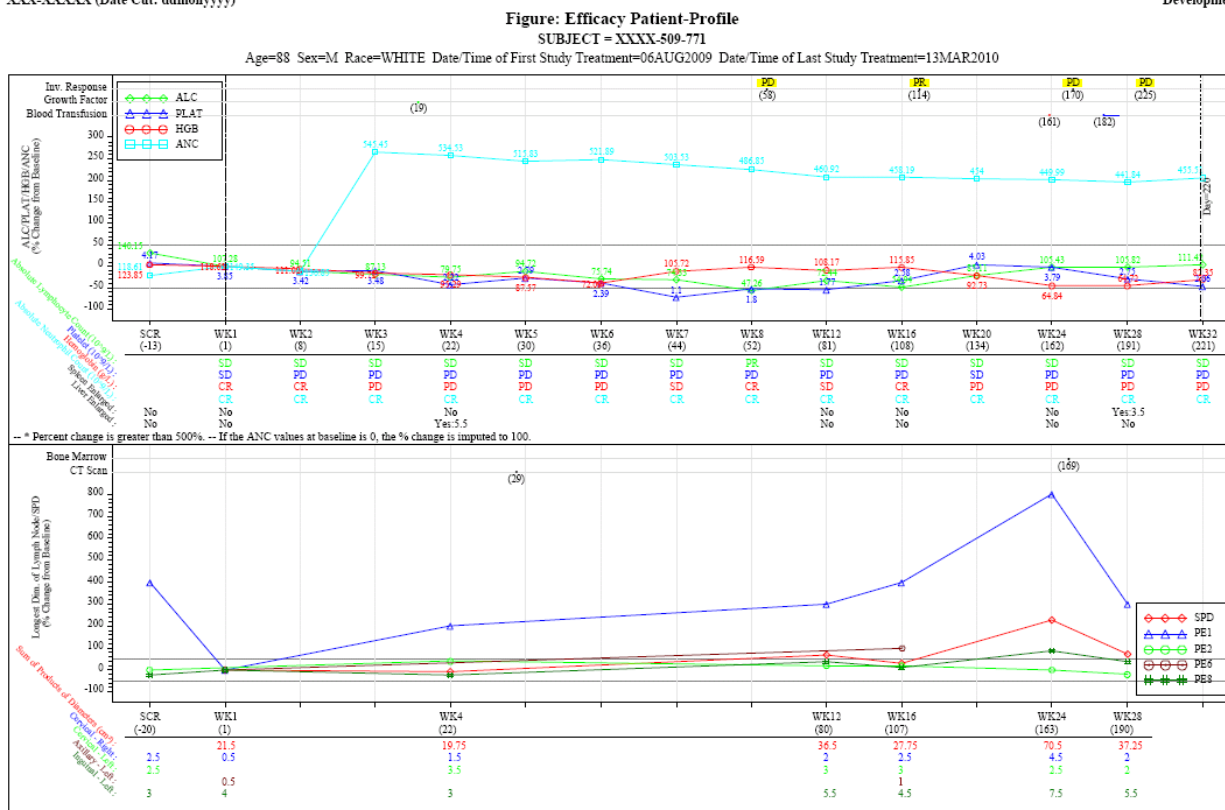
ods pdf bookmarkgen=no;
title1; footnote1;
%do j=1 %to 2;
goptions vorigin=%sysevalf(1.0+3.0*(2-&j))in horigin=.75in vsize=3.0in hsize=9.5in;
```

Graphical Representation of Patient Profile for Efficacy Analyses in Oncology, continued

```
proc gplot data=T&j._&i ANNO=A&j._&i;
  plot _PCHC*_VSTNO=_TESTN_ /vaxis=axis2 haxis=axis1 frame
  ANNO=_B&j._&i %scan(no|legend=,1+(&o_1=Y),|)legend;
run;
quit;
%end;
%end;
ods pdf close;
ods listing;
```

XXX-XXXX (Date Cut: ddmonyyyy)

Development



Baseline is the last non-missing value prior to first dose date. Last available measurement of each visit was displayed on the graph.

<Server Name>: <path/program name.sas>

wwu Created on: 17SEP2013:15:59:54 Page 56 of 98

Figure 1. Efficacy patient profile graphs for patients with CLL. The percent change from baseline for hematology or lymph nodes / sum of the product of the diameters (SPD) is plotted against the visit name. The information for investigator’s response assessment, growth factors, blood transfusion, bone marrow aspirate / biopsy and CT scan are also shown at the top of the corresponding plots. The response categories for hematology, the enlargement of liver / spleen, and the primary diameter for lymph nodes are shown as listings under the corresponding plots. Different colors and symbols are used to represent different lab tests / lesion numbers.

As shown above, the Gplot procedure is used to generate the plots. The annotation data set `_A_&j._&i` is used to annotate the dots on the graphs for ‘Inv. Response’, ‘Growth Factor’, ‘Blood Transfusion’, ‘Bone Marrow’ and ‘CT scan’ (Figure 1). The annotation data set `_B_&j._&i` is used to annotate the listings located under the plots. If the macro variable `&o_1` resolves to Y, legend will be shown on the plot. Otherwise, no legend will be shown.

In terms of the creation of the PDF output, ‘device=sasprt’ directs the graphic output to the default printer. In order to display multiple plots on the same page, ‘startpage=never’ is used to tell SAS not to insert page breaks, even before the graphic procedure. ‘%do j=1 %to 2;’ is used to generate two plots on the same page, with `&j=1` for the hematology data on the top part of the page and `&j=2` for lymph nodes data at the bottom part of the page. ‘%do i=1 %to &o_1;’ is an iterative %DO statement used to generate the PDF output from page 1 until page `&o_1`, where `&o_1`

resolves to the last page number of the PDF output. Character 'A0'x is a 'No-break-space' and is used to display titles/footnotes on the first page. '_blankpage_' moves the printer to the first line of a new page, despite that 'startpage=never' is used.

ADDING ANNOTATION TO A LISTING

As shown in Figure 1, there is a listing under each plot. The listing is added through the annotation technique. Display 1 is an example of the observations in the annotation data set.

O	SUBJECT	GRP_	TESTN_	TESTCD_	STDDY_2	VSTNO_	VSTCD_	LABEL_	PCHG_	YVAL_1	YVAL_2
771	11 XXXX-184-013	1	1	ALC	-7	0	SCR	Absolute Lymphocyte Count (10 ⁹ /L)	22.829	136.5	
772	11 XXXX-184-013	1	1	ALC	1	1	WK1	Absolute Lymphocyte Count (10 ⁹ /L)	0	111.13	SD
773	11 XXXX-184-013	1	1	ALC	8	2	WK2	Absolute Lymphocyte Count (10 ⁹ /L)	13.651	126.3	SD
774	11 XXXX-184-013	1	1	ALC	15	3	WK3	Absolute Lymphocyte Count (10 ⁹ /L)	14.182	126.89	SD
775	11 XXXX-184-013	1	1	ALC	22	4	WK4	Absolute Lymphocyte Count (10 ⁹ /L)	11.167	123.54	SD
776	11 XXXX-184-013	1	1	ALC	29	5	WK5	Absolute Lymphocyte Count (10 ⁹ /L)	-10.42	99.55	SD
777	11 XXXX-184-013	1	1	ALC	36	6	WK6	Absolute Lymphocyte Count (10 ⁹ /L)	30.451	144.97	SD
778	11 XXXX-184-013	1	1	ALC	43	7	WK7	Absolute Lymphocyte Count (10 ⁹ /L)	36.174	151.33	SD
779	11 XXXX-184-013	1	1	ALC	51	8	WK8	Absolute Lymphocyte Count (10 ⁹ /L)	66.679	185.23	PD
780	11 XXXX-184-013	1	1	ALC	81	9	WK12	Absolute Lymphocyte Count (10 ⁹ /L)	80.383	200.46	PD
781	11 XXXX-184-013	1	1	ALC	107	10	WK16	Absolute Lymphocyte Count (10 ⁹ /L)	92.36	213.77	PD
782	11 XXXX-184-013	1	1	ALC	135	11	WK20	Absolute Lymphocyte Count (10 ⁹ /L)	97.525	219.51	PD
783	11 XXXX-184-013	1	1	ALC	165	12	WK24	Absolute Lymphocyte Count (10 ⁹ /L)	76.559	196.21	PD
784	11 XXXX-184-013	1	1	ALC	191	13	WK28	Absolute Lymphocyte Count (10 ⁹ /L)	69.414	188.27	PD
785	11 XXXX-184-013	1	1	ALC	220	14	WK32	Absolute Lymphocyte Count (10 ⁹ /L)	94.727	216.4	PD
786	11 XXXX-184-013	1	2	PLAT	-7	0	SCR	Platelet (10 ⁹ /L)	222.15	5.09	
787	11 XXXX-184-013	1	2	PLAT	1	1	WK1	Platelet (10 ⁹ /L)	0	1.58	SD
788	11 XXXX-184-013	1	2	PLAT	8	2	WK2	Platelet (10 ⁹ /L)	-20.25	1.26	PD
789	11 XXXX-184-013	1	2	PLAT	15	3	WK3	Platelet (10 ⁹ /L)	-4.43	1.51	PD

Display 1. Sample data in the annotation data set used to create the listings below the plots. There is one-to-one correspondence between variables _O_ and SUBJECT, _TESTN_ and _TESTCD_, _VSTNO_ and _VSTCD_. _STDDY_2 is the study day for the corresponding visit. _GRP_ equals 1 for hematology and 2 for lymph nodes. _YVAL_1 is the lab test results and _PCHG_ is percent change from baseline. _VSTCD_ is the visit number for the X-axis while _YVAL_2 is the response categories for lab parameters displayed in the listing.

The following is a sample code which creates the listings for the response categories for hematology lab tests or lymph nodes / SPD under the plots.

```
%annomac(nomsg);
data _A_4;
  set _A_3;
  by _O_ SUBJECT _TESTN_ _TESTCD_;
  length TEXT $256;
  retain WHEN 'A' COLOR _HT_ _PY_;
  %dclanno;
  %system(3,3,3);
  if first.SUBJECT then do;
    _RN_ = 0;
    _HT_ = input(symget(cat('_R',_GRP_,'_',_O_)),best5.);
    %line(9.5,_HT_-9.2,99,_HT_-9.2,black,1,.25);
  end;
  if first._TESTCD_ then do;
    _RN_ + 1;
    COLOR = scan("&CLR",_TESTN_);
    _PY_ = _HT_-8.8-3.3*_RN_;
    %label(8.7,_PY_,catt(_TESTCD_,' :'),*, -45,0,2.5,,<);
  end;
  POSITION = scan('+ >',_GRP_, ' ');
  _PX_ = 11.5+85.5*_VSTNO_/ _MV_- .434*( _GRP_-1)*
    (length(scan(cats(_YVAL_1),1, '.'))-5);
  %label(_PX_,_PY_,ifc(_GRP_=2,catt(_YVAL_1),_YVAL_2),*,0,0,2.9,,*);
run;
```

In the above code, macro '%annomac' compiles annotation macros and makes them available for use. Macro '%dclanno' sets the length and type for all the annotation variables except the TEXT variable. Macro '%system(3,3,3)' defines the annotation reference systems and the XSYS, YSYS, and HSYS variables. Macro '%line' draws a horizontal line between the plot and the listing. The first '%label' is used to display the labels on the left side of the listing, while the second '%label' is used to display corresponding values in the listing. The listing under the plot in Figure 2 is an example of the output from the annotation data set.

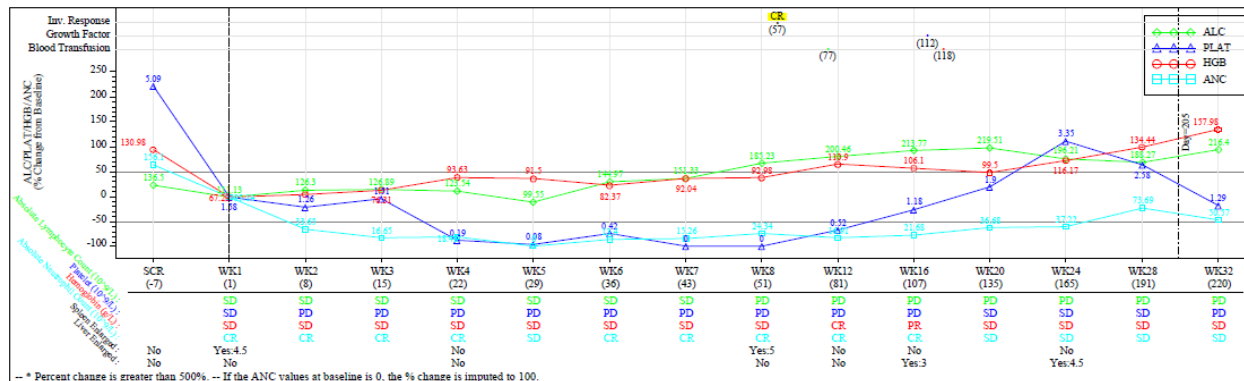


Figure 2. The listing under the plot was generated through the annotation technique.

It can be seen from Figure 2 that the values in the listing are lined up with the visit name on the X-axis. This is mainly achieved through the 'POSITION=' statement. If the variable to be displayed in the listing is a character variable, POSITION='+' is used to center the character values in the listing with the visit name on the X-axis. If the variable to be displayed in the listing is a numeric variable, POSITION='>' is used to line up the numeric values in the listing with the visit name on the X-axis.

MAKING THE GRAPHS DYNAMIC

For a macro developed to create graphs, it is important to make the graphs dynamic. In other words, it should be data driven to allow differences among subjects. In order to develop a macro which creates good-looking graphs, the macro should allow the following statements to be dynamic: (a) The minimum value, maximum value and unit displayed on the X-axis or Y-axis, (b) Location of the legend in the plot.

DYNAMIC X- / Y-AXIS

In order to make the X- / Y-axis dynamic, it is important to have some options in the AXIS statement, such as 'ORDER=', 'MINOR=(n=..)' and 'LENGTH=' for Y-axis. Below is a sample code used to set the AXIS statement dynamically.

```
axis1 offset=(3,3) value=none length=91.5pct minor=none major=(h=1.0pct) label=none
order=(0 to &&_vn_&i) origin=(8.5,&&_R&j_&i)pct;

axis2 offset=(3,%scan(17|13.4,&j,|)) value=(h=6pt)
length=%sysevalf(100-&&_R&j_&i)pct minor=(h=.30pct n=&&_m&j_&i)
label=(angle=90 h=6pt j=c) major=(h=.45pct)
order=(&&_L&j_&i to &&_H&j_&i by &&_u&j_&i);
```

In the above code, several macro variables, such as &&_L&j_&i (minimum value for Y-axis), &&_H&j_&i (maximum value for Y-axis) and &&_u&j_&i (major unit for Y-axis), are set dynamically depending on the individual subject and the individual plot on the same page. The values of these macro variables are calculated before the AXIS statement. Below gives an example to calculate the values for the macro variables which resolve to the minimum value, maximum value, major unit, and number of minor ticks for the axis.

```
proc fcmp outlib=work.funcs.maths;
subroutine hi_low(min, max, scale, UT, MT, LOW, HGH);
outargs UT, MT, LOW, HGH;
_x1_ = (max-min)/scale;
_x2_ = log10(_x1_);
```

```

_x3_ = floor(_x2_);
_x4_ = 10**(_x2_-_x3_);
_x5_ = 1+(_x4_>=2)+3*( _x4_>=5);
MT = 4-(_x5_=2);
UT = _x5_*10**_x3_;
LOW = UT * floor(min/UT);
HGH = UT * ceil(max/UT);
endsub;
run;

options cmlib=work.funcs;

```

In the above code, 'max' represents the maximum value of a variable for a patient while 'min' represents the minimum value for the patient. 'scale' is used to determine the minimum number of major ticks for the axes.

The FCMP procedure will output values for 'UT' (major unit), 'MT' (number of minor ticks for a major unit), 'LOW' (minimum value for the axis), 'HGH' (maximum value for the axis), which can then be converted to macro variables by CALL SYMPUTX statement in a data step. These macro variables can be used in the AXIS statement to determine the minimum value (&&_L&j._&i.), maximum value (&&_H&j._&i.), major unit (&&_u&j._&i.), number of minor ticks (&&_m&j._&i.) for the axis. An example is shown as below.

```

axis2 order=(&&_L&j._&i to &&_H&j._&i by &&_u&j._&i) minor=(h=.30pct
n=&&_m&j._&i)..;

```

DYNAMIC LOCATION FOR THE LEGEND

The location of the legend is dynamic. The legend can be located at one of the four corners of the graphic area depending on the data and the availability of the free space on the plot. The priority from high to low for the location of the legend is lower left corner, lower right corner, upper left corner and upper right corner of the graphic area. If no valid data is available for the patient, there will be no legend displayed. The following code controls the dynamic location of the legend.

```

legend label=none frame mode=share across=1 value=(j=1 h=6pt &&_C&i._&j)
position=(inside %scan(left bottom|right bottom|left top|right
top,&&_G&j._&i,|));

```

The macro variable &&_C&i._&j is the list of the value for variable _TESTCD_ in Display 1. The macro variable &&_G&j._&i (which resolves to 1, 2, 3, 4 with each number representing a position in the plot) is for the location of the legend. The value of &&_G&j._&i depends on the data and can be calculated through several data steps and SAS/SQL.

INVOKING THE MACRO

As mentioned before, the macro requires at least three input data sets (demographics, hematology, lymph nodes) to run successfully. Below is an illustration of the code used to invoke the macro and the explanation of each macro variable. If the input data set lacks one or more of the key variables, DUMMY can be used as the variable name to tell the macro that the variable is missing in the input data set when calling the macro.

```

%efficacy_pp(
  datalist = /* List of input data sets:1. DM 2. HEM 3. LYM 4. RS 5.CM 6.BP 7.CT */
  , where = /* Conditions for subsetting the DM data set */
  , id_var = /* A common variable for patient ID in each data set */
  , core_var = /* List of variables for Demographics data set */
  , hem_var = /* List of variables for Hematology data set */
  , lym_var = /* List of variables for Lymph Node data set */
  , resp_var = /* List of variables for Investigator Response data set */
  , cm_var = /* List of variables for Concomitant Medication data set */
  , bm_var = /* List of variables for Bone Marrow data set */
  , ct_var = /* List of variables for CT scan data set */
  , title = /* Text for Title */
  , footnote = /* Text for Footnote */
  , out_type = /* Type of the output graphs. */
  , outpathf = /* PDF output file name and file path */

```

);

The first part of the macro processes each of the above data sets, including subsetting data sets, merging data sets, defining the baseline, calculating percent change from baseline and creating a variable for response categories for lab parameters based on International Workshop on Chronic Lymphocytic Leukemia (IWCLL) 2008 guidelines and June 2012 e-letter (Hallek et al., 2008 and Hallek et al., 2012). If the input data sets have variables for baseline, percent change from baseline or lab parameter response categories, the macro has the option to use the variables from the input data sets and ignore the part for creating these new variables.

CONCLUSION

In conclusion, this paper describes a macro developed to create efficacy patient profile graphs in the PDF format using SAS/GRAPH and the annotation technique. The ability to display an increased amount of meaningful information, on the same page and for the same patient, enables the reviewers to have a better understanding of that patient's data. Therefore, the graphics output can be used as an alternative to tables and listings for clinical study results interpretation and decision making purposes. Although this paper uses data mimicking the clinical trial data for chronic lymphocytic leukemia, the same technique can be applied to other fields of clinical trials in oncology.

REFERENCES

- Rozman, C. and Montserrat, E. (1995). "Chronic lymphocytic leukemia." *The New England Journal of Medicine*. 333 (16):1052-1057. [Erratum, *The New England Journal of Medicine* (1995). 333 (22):1515.]
- Eichhorst, B., Dreyling, M., Robak, T., Montserrat, E., and Hallet, M. (2011). "Chronic lymphocytic leukemia: ESMO Clinical Practical Guidelines for diagnosis, treatment and follow-up." *Annals of Oncology*. 22 (suppl 6): vi50-vi54
- Hallek, M., Cheson, B.D., Catovsky, D., Caligaris-Cappio, F., Dighiero, G., Döhner, H., Hillmen, P., Keating, M.J., Montserrat, E., Rai, K.R., and Kipps, T.J. (2008). "Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines." *Blood*. 111(12): 5446-5456.
- Hallek, M., Cheson, B.D., Catovsky, D., Caligaris-Cappio, F., Dighiero, G., Döhner, H., Hillmen, P., Keating, M.J., Montserrat, E., Rai, K.R., and Kipps, T.J.(2012). "Response assessment in chronic lymphocytic leukemia treated with novel agents causing an increase of peripheral blood lymphocytes." [e-letter], *Blood*.

ACKNOWLEDGMENTS

The macro was developed while William Wu was an employee of Pharmacyclics, Inc.. The macro is the property of Pharmacyclics, Inc.. The authors would like to thank all the members in the Statistical Programming Department at Pharmacyclics, Inc. for their reviewing of this paper and providing useful comments, especially Raymond Lee, Stephen Chan and Michael Quigley.

CONTACT INFORMATION

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

William Wu, MS
Herodata LLC.
2628 Carlo Scimeca Ct.
San Jose, CA 95132
willywu2001@hotmail.com

Xiaoxian Dai, Ph.D.
Pharmacyclics, Inc.
995 E. Arques Avenue
Sunnyvale, CA 94085
sdai@pcyc.com

Linda Gau, MS
Pharmacyclics, Inc.
995 E. Arques Avenue
Sunnyvale, CA 94085
lgau@pcyc.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.