

## Analyzing and Displaying Study Data in Clinical Trials Made Easy

Denise J. Smith, Mayo Clinic, Cancer Center Statistics, Rochester, MN  
Pamela J. Atherton, Mayo Clinic, Center Statistics, Rochester, MN  
Angelina Tan, Mayo Clinic, Cancer Center Statistics, Rochester, MN  
Paul J. Novotny, Mayo Clinic, Center Statistics, Rochester, MN  
Jeff A. Sloan, Mayo Clinic, Center Statistics, Rochester, MN  
Mayo Clinic QOL Team, Cancer Center Statistics, Rochester, MN

### ABSTRACT

**Background:** Quality of Life (QOL) research is becoming an increasingly integral part of oncology clinical trials. The inclusion of QOL endpoints provides an added uniqueness affecting analytical approaches and presentation of results.

**Objectives:** We use North Central Cancer Treatment Group (NCCTG) oncology clinical trials design to provide examples of analysis techniques and demonstrate macros written to provide unique graphics for easily displaying patient and study outcomes.

**Methods:** SAS<sup>®</sup> macros have been developed for use in SAS v6 and higher to score QOL assessments, aid in the display of longitudinal data and provide summary statistics for study endpoints. We use a hypothetical trial with various endpoints to measure analysis of QOL at a single time point, change in QOL from baseline at a single point, maximum change from baseline, increase (or decrease) in individual QOL scores according to a definition of a clinically significant difference, time to event criteria and area under the curve (AUC).

**Results:** The macros constructed provide versatility so that data can be displayed as per the programmers needs. The macro options allow the selection of either black/white or color graphics, controlling of axes, symbols and formats, printing of statistical results within some graphs, and formatting of graphics to display in mediums other than SAS/GRAPH.

**Conclusions:** The presented macros have potential use in manuscripts, posters, presentations and grants and can be utilized by all programmers of all levels. They provide additional information to the basic study results, as compared to just the p-value alone.

### INTRODUCTION

Quality of Life (QOL) research is becoming an increasingly integral part of oncology clinical trials. The addition of a QOL component to assess patient well-being is typically used in conjunction with the assessment of cancer treatment efficacy. QOL has 5 main domains: Physical, Mental, Spiritual, Intellectual, and Social. Overall QOL or any of the sub-domains can be measured as a study endpoint. To assess these areas, patients are asked specific questions regarding traits which pertain to each domain. Depending on the questionnaire, there may be one question measuring something quite specific or a series of questions for which a computed total score is obtained.

QOL results are unique compared to some clinical results. For example, in the clinical setting, tumor response, survival and blood count endpoints have specific data that needs to be collected to measure the outcome. For QOL, one set of data may be used to measure a variety of outcomes. Different techniques are employed to calculate values for testing QOL hypotheses. Some analysis endpoints include assessing QOL at a certain time point, calculating changes from baseline to determine increases or decreases in QOL, categorizing patients as successes or failures as determined by those increases or decreases in scores, following time-to-event criteria and calculating area under the curve (AUC). The choice of what endpoint to calculate is unique, the analysis is not. The analysis is typically a basic comparison between arms at certain time points and may be conducted by using the Student's t-test or the chi-square test.

The results of any cancer clinical trial are reported in some fashion. The Mayo Clinic Cancer Center Cancer Control and QOL team has devised criteria to promote consistency across studies where the study results are often highly individualistic. The goal of this paper is to demonstrate macros that have proven helpful in the analysis of study endpoints, including macros to score QOL assessments and display both individual results and summary measures. These displays are in both tabular and graphical form, and are demonstrated using a hypothetical study.

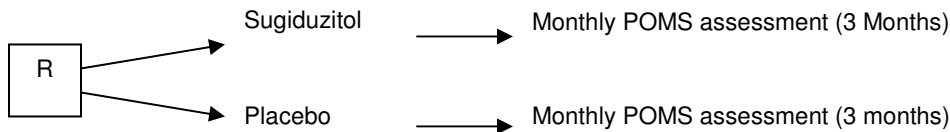
## HYPOTHETICAL STUDY

Assume Sugiduzitol is a new drug which may or may not cure colorectal cancer. To determine Sugiduzitol's effect on the disease, a Phase II protocol is developed in which patients are randomized to receive the treatment drug, Sugiduzitol, or placebo. A previously conducted Phase I study indicates that suspected side-effects of the treatment include fatigue, confusion, gastrointestinal toxicities and loss of fine motor skills. These side-effects are expected to appear within the first month of treatment, resolving themselves after that point with no long-term effects. Therefore, patients are randomized to receive an oral dose of Sugiduzitol or placebo daily for 5 days with toxicity monitoring occurring for 6 months post-treatment. At the end of 6 months, the patient will be assessed as to the curative effect of the treatment. Further, since colorectal cancer patients are likely to show increased levels of carcinoembryonic antigen (CEA), an indicator of tumor distress, CEA levels will be collected weekly.

Patient questionnaires will be completed monthly for 3 months. The QOL questionnaire of choice for this study, the Profile of Mood States (POMS), will be incorporated to quantify the magnitude of these side-effects on patient well-being (1). The Common Toxicity Criteria for Adverse Events (CTCAE v3.0) will be used to measure adverse events.

The POMS questionnaire is an example of a multiple item assessment. This questionnaire asks patients to describe their feelings during the past week. The questions posed require answers ranging from 0 (not at all) to 4 (extremely). In some cases, a response may mean either "extremely good" or "extremely bad." To avoid confusion, it is beneficial to have all scores converted to a meaningful 0-4 scale, where 0 is low (bad QOL) and 4 is high (good QOL).

The POMS contains subscales which address some of the possible side effects of Sugiduzitol and the total score provides a quantity for an overall QOL result. The tension/anxiety (T/A) and depression/dejection (D/D) subscales assess the emotional impact of the study drug on the patient. The confusion/bewilderment (C/B) subscale assesses the impact of the confusion side-effect. The fatigue/inertia (F/I) subscale assesses the impact of the fatigue side-effect. The vigor/activity (V/A) subscale assesses the impact of the loss of motor skills side-effect. The anger/hostility (A/H) subscale assesses any feeling the patient has in relation to having to put up with the treatment or side-effects.



Goals of the study are:

1. To determine the difference in total POMS scores between arms at 1 month post treatment.
2. To determine the difference in the POMS subscale scores at 1 month post treatment (when symptoms should have been experienced).
3. To determine the change from baseline for total POMS and all subscales at 1 month.
4. To determine the maximum change from baseline for the Fatigue/Inertia POMS subscale over entire treatment.
5. To determine the toxicity incidence between treatment arms.
6. To determine whether the Depression/Dejection subscale is measuring the same as the Tension/Anxiety subscale for emotional impact of treatment.
7. To determine the survival profile.
8. To determine the difference between all scores over the course of treatment as measured by the AUC.
9. To determine whether treatment resulted in decreased CEA levels.

## SCORING MACROS

Numerous macros have been compiled which score QOL questionnaires used in a variety of oncology clinical trials. For the sake of brevity, these will not be discussed here but a current listing, along with specific code, is available upon request from the authors.

A scoring macro has been compiled to perform the conversion of the POMS assessment to a meaningful 0-4 scale, compute the 6 subscale scores and compute a patient's total POMS score, which is the sum of the subscales. All seven of these scores are transformed to an interpretable scale of 0-100. Data transformation to the 0-100 scale is often done in QOL studies, allowing for score comparisons within and among assessments which may have a variety of ranges. In addition, this scale allows for ease in the explanation and representation of results.

Figure 1 represents the POMS macro process. There are 65 questions asked on the POMS long version and 30 on the short version. Separate macros are written for each version. Five questions were chosen from the POMS long version for this example (note that all 65 questions are still included for scoring). The macro requires that the input dataset variables have consistent naming conventions, or rather a variable followed by the question number (here, 01-65). In addition, for total or subscale scores to be calculated, 80% of the assessment items must be completed (otherwise the score is set to missing). The results of this macro do not reflect any statistical tests or interpretations. These items must be calculated in a subsequent procedure using the output values from this macro.

**Figure 1. POMS Macro Process**

Original (Raw) POMS Data								
PT_ID	ARM	CYCLE	DATE	POMS01	POMS02	POMS03	POMS04	POMS05
401	B	0	10/15/1998	2	1	0	1	1
401	B	1	11/10/1998	1	1	0	1	1
401	B	2	11/30/1998	0	1	0	1	0
401	B	3	01/20/1999	1	3	1	4	1

↓

POMS Macro Call
%POMS (data = , varname = , idvar = , outdata = );

↓

Transformed POMS Data Example														
Patient Information			Individual POMS Questions					Subscales					Total	
PT_ID	ARM	CYCLE	POMS01	POMS02	POMS03	POMS04	POMS05	TA	AH	CB	FI	DD	VA	POMSTOT
401	B	0	2	3	4	3	3	83.2	76.3	59.2	91.0	88.2	76.5	87.5
401	B	1	1	3	4	3	3	65.9	78.1	97.5	86.5	84.3	84.1	83.2
401	B	2	0	3	4	3	4	75.0	84.1	95.2	64.3	82.1	74.9	86.6
401	B	3	1	1	3	0	3	65.8	91.2	94.4	82.5	76.3	45.7	78.0

## SUMMARY MACROS

A number of macros have been written by the team to create consistent, easy to read summaries of study data. The macros are versatile in the sense that they may be implemented across studies.

### TABLE MACRO

The TABLE macro was designed to display summary statistics for continuous, discrete, ordinal, or survival variables either for all patients in a study or by subgroups. P-values may be included for differences in subgroups. Required parameters include the input dataset name, and the type (continuous vs. discrete) and names of the variables of interest. All parameters in this macro are keyword parameters and must be specified by entering each parameter name followed by an equal sign and the variable corresponding to or the value of the parameter. Formatted variables are printed using their formats. The table may be exported as a .doc file to a directory specified by the user. The user may also specify to have the final data set containing the summary statistics saved as a SAS dataset. Desired statistics may be specified based on variable type, such as obtaining n and % for discrete variables. In addition, one can choose to either print or suppress p-values, and choose which p-value test should be conducted if the default is not appropriate: Kruskal-Wallis for continuous; Chi Square for discrete; Wilcoxon for ordinal; log-rank for survival. Tests available include Fisher's Exact, ANOVA F-test, among others. Missing values may be included, tests used for each p-value may be included as a footnote, or 95% confidence intervals for differences may be printed. Display options include designating titles, changing text sizes, and choosing page orientation. The user may also debug to print values assigned to macro variables and the macro logic.

Figure 2 displays the macro call and the produced tables containing POMS assessment scores by arm at 1 month (Goal 1), changes from baseline of the POMS total and subscale scores at 1 month (Goal 3) (output suppressed to only F/I and A/H

subscales for conciseness), and maximum change from baseline for the F/I POMS subscale over the entire treatment schedule (Goal 4). The input dataset requires one observation per patient for each time point (as in Figure 1).

**Figure 2. TABLE Macro Call and Output Tables**

```

TABLE Macro Call
%TABLE(dsn=_last_, id=, by=, var=, type=, outdoc=, outdat=,list=N, print=N,
number=N, cstats=n mean sd median quartiles range, dstats=n percent, pfoot=n,
surv=, scensor=1, total=Y, pvalues=Y,ptype=, ci=, pop=, incmiss=N, incmiss1=,
comments=, dvar=, dline=, decpct=1, tttitle1=, tttitle2=, tttitle3=, tttitle4=,
foot=, foot2=,foot3=, foot4=, foot5=, date=Y, page=portrait, rules=groups,
frame=hsides, titlesz=10, bodysz=10, footsz=10, titlebld=Y, titlefnt=Times New
Roman, headfnt=Times New Roman, bodyfnt=Times New Roman, ciwd=, pvalwd=100,
levelwd=, datawd=, space=, debug=N);

```



<b>POMS Mean Total and Subscale Scores at Month 1 Comparison Among Arms</b>				
	A (N=30)	B (N=30)	Total (N=60)	p value
<b>POMS Total Score</b>				0.0002
N	30	30	60	
Mean (SD)	84.4 (6.08)	73.8 (11.46)	79.1 (10.53)	
Median	84.7	77.4	81.0	
Q1, Q3	80.6, 87.5	63.8, 83.2	76.5, 86.2	
Range	(73.3-100.0)	(42.7-87.5)	(42.7-100.0)	

<b>POMS Difference in Scores from Baseline to Month 1 Comparison Among Arms</b>				
	A (N=30)	B (N=30)	Total (N=60)	p value
<b>POMS Total Score</b>				0.3346 <sup>1</sup>
N	30	28	58	
Mean (SD)	0.8 (9.70)	4.1 (9.75)	2.4 (9.78)	
Median	2.4	4.5	3.7	
Q1, Q3	-2.6, 7.6	-3.0, 9.7	-2.6, 7.8	
Range	(-28.0-13.3)	(-19.4-26.3)	(-28.0-26.3)	
<b>F/I Subscale</b>				0.7853 <sup>1</sup>
N	30	28	58	
Mean (SD)	-2.0 (23.93)	-3.2 (17.91)	-2.6 (21.06)	
Median	0.0	-3.0	-1.2	
Q1, Q3	-10.8, 14.3	-12.5, 8.9	-10.8, 10.7	
Range	(-53.6-42.9)	(-46.5-28.6)	(-53.6-42.9)	
<b>A/H Subscale</b>				0.0404
N	30	28	58	
Mean (SD)	-0.2 (7.84)	5.9 (11.62)	2.7 (10.23)	
Median	0.0	4.1	2.1	
Q1, Q3	0.0, 4.1	0.0, 11.5	0.0, 6.2	
Range	(-35.4-8.3)	(-14.6-39.6)	(-35.4-39.6)	
<b>F/I Subscale Maximum Change</b>				0.1748 <sup>1</sup>
N	30	28	58	
Mean (SD)	-0.4 (8.72)	4.2 (9.55)	1.8 (9.35)	
Median	0.0	3.3	1.7	
Q1, Q3	-3.4, 5.0	-2.5, 9.2	-3.3, 5.0	
Range	(-23.4-16.7)	(-10.0-31.7)	(-23.4-31.7)	

<sup>1</sup> Kruskal Wallis

The difference in total POMS scores between arms was statistically significantly different at 1 month (84.4 for the Sugiduzitol arm vs. 73.8 for the placebo arm,  $p=0.0002$ ). There was no difference in maximum F/I subscale score changes between arms. The anger/hostility subscale score changes were statistically significant (-0.2 change in scores for Arm A vs. 5.9 change in scores for Arm B,  $p=0.04$ ).

Each of the tables above was produced with a separate macro call. Thus, each table is set out as its own document (if specified using the *outdoc* parameter). A TABLEMERGE macro exists to allow the user to compile the tables of summary measures in a single file. This macro requires the names of the datasets created and saved using the *outdat* parameter and a name for the final document. Other options are available to enhance or modify the look of the final set of tables. Figure 3 gives an example of how one might use the TABLEMERGE macro.

**Figure 3. TABLEMERGE Macro Example**

```
%table(dsn=master, by=arm, var=var1 var2 var3, type=1, outdat=dat1);
                                     ↓
%table(dsn=master, by=arm, var=var4 var5 var6, type=2, outdat=dat2);
                                     ↓
%tablemerge(dsn=dat1 dat2, outdoc=alldat.rtf, number=Y);
```

**PFREQ MACRO**

While an Output Delivery System (ODS) statement may be placed around a FREQ procedure, as of SAS v9.13 the output for a two way cross tabulation is still difficult to read as the key for the cell content is outside of the table. The look is not ideal for summary reports and publications. This macro creates a modification of the FREQ procedure output by producing an extra column containing the cell content (Frequency, Percent, Row Percent, Column Percent) for much easier reading and referencing. The user may choose to include any combination of the aforementioned cell contents and can choose to include any of the statistical tests found in the traditional FREQ procedure such as the Chi-Square Test, Fishers Exact Test, McNemars Test, and the Cochran-Mantel-Haenszel Test. The row and column parameters (*xvar* and *yvar*) may consist of a single variable or a list of variables. The macro requires that the input dataset contain one observation per patient, per x variable (below, toxicity) and y variable (arm). ODS statements are placed around the macro call to save the output as something other than a .lst file. Figure 4 presents the macro call, a table comparing toxicity incidence between arms, and the statistical test table.

**Figure 4. PFREQ Macro Call and Output Tables**

PFREQ Macro Call	
%PFREQ(dsn=, xvar=, yvar=, statopt=1, testopt=1, by=, odstype=rtf, ylevel=10, uselbl=y, debug=N);	

↓  
**Toxicity Incidence by Arm**

		Grade	
		Grade 0/1/2	Grade 3+
Arm			
A	Frequency	351	114
	Row %	75.48	24.52
	Column %	53.67	52.05
B	Frequency	303	105
	Row %	74.26	25.74
	Column %	46.33	47.95
All Arm	Frequency	654	219
	Percent	74.91	25.09

### Chi-Square Statistics

X	Y	P-Value
Arm	Grade	0.6784

Toxicity is graded on a 0-5 scale, where 0 indicates no toxicity and 5 indicates death. According to the chi-square statistic, there is no difference in toxicity incidence (grade 0/1/2 vs. grade 3+) between arms (Goal 5).

## PCORR MACRO

While SAS v9.13 now offers the opportunity to ODS results from the CORR procedure (this is experimental), including correlation graphics, the table produced containing the correlation coefficients, p-values and sample sizes is not clear. The key for the table cells appears as a table title. This macro provides a concise summary of correlation data with an additional column listing the cell content (Correlation, P-Value, N) to aid in interpretation of results. The user still has the choice of the type of coefficient: Pearsons, Spearmans, Heoffding, or Kendall. An ODS statement should be placed around the macro call to save the results as a non-.lst file. Figure 5 provides the macro call and output, which is used to determine whether the Depression/Dejection subscale is measuring the same as the Tension/Anxiety subscale for emotional impact of treatment using a Pearson p-value (Goal 6).

Figure 5. PCORR Macro Call and Output Table

PCORR Macro Call
<code>%PCORR (dsn=, var=, withvar=, type=1, opt=1, uselabel=Y, debug=N, odstyp=rtf) ;</code>



### Relationship between Tension-Anxiety and Depression-Dejection POMS Subscales

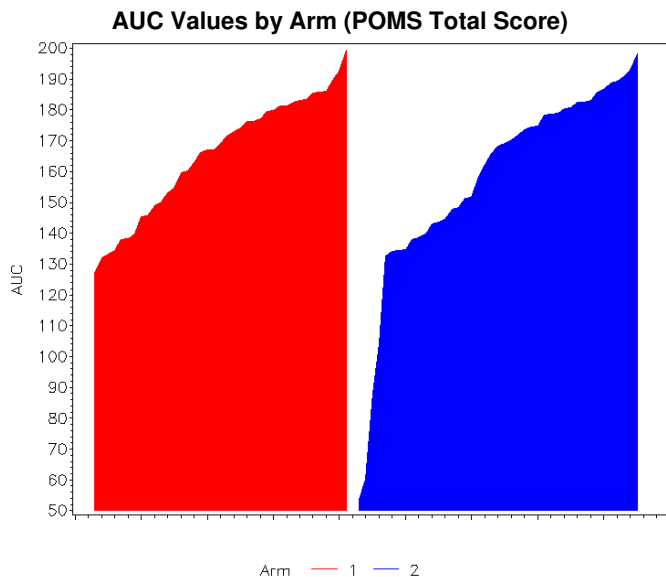
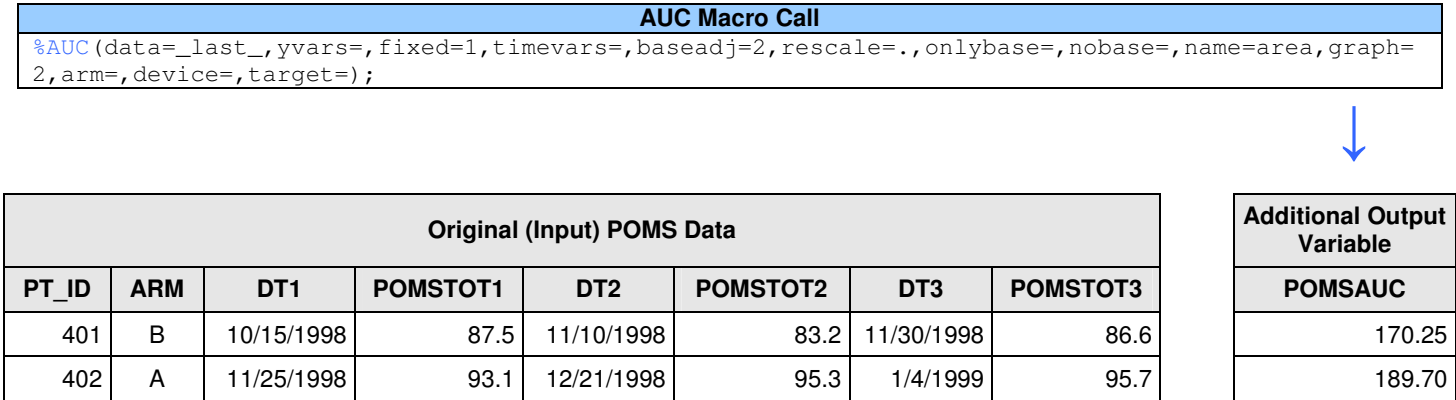
		POMS Depression-Dejection Percentage
POMS Tension-Anxiety Percentage	Correlation	0.81595
	P-Value	<.0001
	N	60

A correlation coefficient less than 0.3 indicates a minimal relationship between assessments, while a correlation between 0.3 and 0.5 indicates a slight correlation. A moderate relationship is a correlation between .5 and .75 while a strong relationship is a correlation greater than .75 (2). The results of the PCORR macro indicate a strong relationship between the T/A and D/D subscales. It can be concluded that for this population, these two items are measuring the same emotional construct.

## AUC MACRO

In longitudinal studies, one of the challenges is summarizing results over time. An AUC summary statistic is calculated as being the area of a 2-dimensional picture constructed by plotting assessment values at individual time points (3). The AUC is interpreted as a measure of a patient's overall on-study QOL. This macro calculates a single AUC value per patient (observation), and allows for specified or fixed time points, baseline adjustments (using the trapezoid method), and rescaling. The input dataset must consist of one observation per patient, with variables indicating the date or time point and the corresponding QOL score. The user has the opportunity to output a graphic of the AUC values for all patients (by arm if applicable), sorted from lowest to highest and shaded to emphasize the area. If there are any missing assessment values, the user should implement imputation methods on their input dataset before running the macro. Figure 6 includes the AUC macro process, including the output dataset and graphic to compare differences in AUC values between arms over the course of treatment (Goal 8).

**Figure 6. AUC Macro Process and Output Graph**

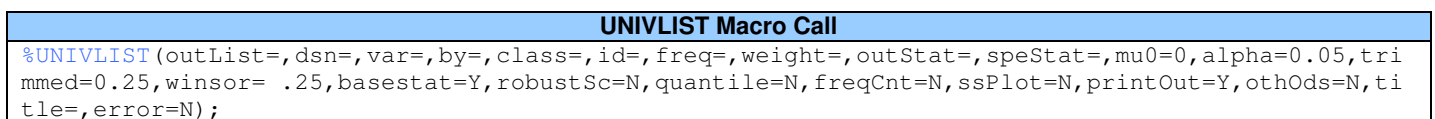


The AUC macro does not produce statistical results. These items must be calculated in a subsequent procedure. The graph indicates a minimal difference in AUC scores between arms.

**UNIVLIST MACRO**

SAS v8 and higher produces proc UNIVARIATE results that take up multiple pages in the SAS output window. This macro condenses the default output so that it takes up 2 landscape pages in its entirety. It provides options to produce subsets of the results, thus eliminating items not needed for analysis and reducing space. The output is divided into 5 separate tables: statistics, robust, percentiles, frequencies and box plots. The macro allows the user to choose one or more tables of interest. Due to the construction of the tables within the macro, the output generated cannot be reproduced as anything other than an ODS .lst file. If the user would like to produce results for a different ODS destination, there are two options: 1) The original tables produced by proc UNIVARIATE may be requested using the *othOds* parameter or 2) the condensed table(s) can be copied from the .lst file and pasted into the proper medium. Summary statistics may be saved to a dataset. Figure 7 provides an example of the macro call and the .lst output as copied into Microsoft Word. The content is a test of the change from baseline for the Sugiduzitol arm of the study to see if there are any changes within the arm.

**Figure 7. UNIVLIST Macro Call and Output**



OUTPUT

The Modified Version of Univariate Procedure

Title : Arm A Poms Diffs  
Statements: VARNAME=dl,

Statistic	Value	Parameter	Estimate	95% CL Assuming Normality		Quantile	Estimate
Nmiss	0	Median	2.4	-2.854	4.394	100% Max	13.3
N	30	Mean	0.77	-2.854	4.394	99%	13.3
Sum Weights	30	Std Deviation	9.704	7.728	13.045	95%	12.9
Sum Observations	23.1	Variance	94.167	59.727	170.177	90%	12.1
Skewness	-1.187					75% Q3	7.6
Kurtosis	1.545	Lowest Value--- Low Obs		Highest Value--- High Obs		50% Median	2.4
Uncorrected SS	2748.63					25% Q1	-2.6
Corrected SS	2730.843	-28	18	8.8	28	10%	-14.65
Coeff Variation	1260.256	-17.2	9	11.2	8	5%	-17.2
Std Error Mean	1.772	-15.1	15	12.1	19	1%	-28
Mode	0	-14.2	13	12.1	21	0% Min	-28
Range	41.3	-10.3	12	12.9	30		
Interquartile Range	10.2	-4.7	3	13.3	7		

Test of Mu0	Stat	----p value----	Count of Mu0	Value	Normality Test	Stat	----p-value----
Student's t	0.435	Pr >  t	0.6671	Num Obs > Mu0	19	Shapiro-Wilk	0.903 Pr < W
Sign	5	Pr >=  M	0.0872	Num Obs ^= Mu0	28	Kolmogorov-Smirnov	0.168 Pr > D
Signed Rank	46.5	Pr >=  S	0.2979	Num Obs < Mu0	9	Cramer-von Mises	0.16 Pr > W-Sq
						Anderson-Darling	0.931 Pr > A-Sq

Stem Leaf	#	Boxplot	Normal Probability Plot
1 12233	5		12.5+ ***** *
0 555889	6	+++++	*****
0 0011223444	10	*-+--*	++++++
-0 431	3	+++++	*****
-0 5	1		+++++
-1 40	2		+++++ * *
-1 75	2		+++++ * *
-2			+++++
-2 8	1	0	-27.5+ *

-----+-----+  
Multiply Stem.Leaf by 10\*\*+1

The t-test for a difference from zero shows there was no statistically significant improvement in the POMS score within the treatment arm from baseline to cycle 1.

### DISPLAY MACROS

The power of SAS/GRAPH to display data is demonstrated by the following display macros. Options allow for either black/white or color displays, printing of statistical results within the graph, and formatting of graphics to display in mediums other than SAS/GRAPH.

### DISPLAYING INDIVIDUAL DATA:

#### BUG PLOT

A novel graph that displays all patient data by treatment arm is the BUG plot, so named because of its resemblance to a squashed bug. The BUG plot macro calculates changes from baseline either as a number or as a percent change from baseline. Displayed are these changes with lines going off in different directions from the center of the graph according to the *by* parameter. The symmetry, or lack thereof, in the resultant plot is an indication of comparability between groups of subjects (and subgroups) over time. The biggest advantage of the bug plot is that it allows the user to see all of the data for every individual on study over time.

The BUG plot macro is unique in that it automatically creates symbols statements according to the dichotomous *by* values and any subgroup (parameter *subgrp*) within this variable. Also automatic is the creation of a legend for any subgroups. The note and symbols statements allow up to 4 levels of the subgroup. Other options are to make color or black and white, choose text font and size, label axes, and save the graph to a file. This macro is for display only. No statistics are calculated. The data set needs to contain a patient identifier, the time point and the values to plot (either as an array or as one observation per patient per time point). Figure 8 depicts changes in POMS total scores from baseline by arm displaying the subgroup of gender.



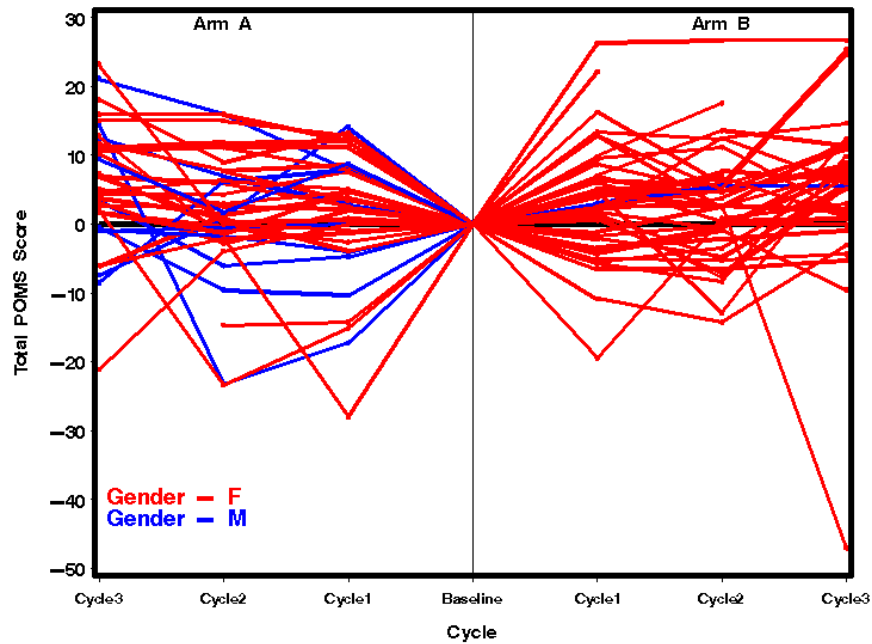
**Figure 8. BUGPLOT Macro Call and Output**

```

BUGPLOT Macro Call
%BUGPLOT (dsn=_last_, class, subgrp, id, time, var, min, max, by, ray, xlabel=%str(Cycle), ylabel=%str(Total
Score), llabel=%str(Arm A), rlabel=%str(Arm B), ylevel, ymin, ymax, yby, awidth=10, httext=2.5, text=swiss,
httit=2.5, type=2, gfile, gtype=xcolor, target, color=y, cline, clist=2, refline, ctick, cback,
ctitle, xcoord=70, ycoord=80, notelab=%srt(Arm), debug=N);
    
```



**POMS (Total Score) Changes from Baseline  
By Gender**



This graph shows there are patients with larger scores on the placebo arm and those with lower scores on the treatment arm. Within the treatment arm, we can see that males on the treatment arm had lower scores than females, but this was not evident in the placebo arm.

**JITTER PLOT**

An alternate to the dot plot is the JITTER plot which is useful if there are multiple observations having the same plotted values so the data points fall on top of one another. A technique is available to “jitter” the data points so they are not superimposed on each other (4). This macro performs the jitter data modification on the continuous time variable so that plotted values are separated, thus making data easier to see and compare visually. The uniqueness of this macro is that it also incorporates the symbols and legend creating techniques implemented in the BUGPLOT macro. Symbols in this macro are dots rather than lines. Parameters are available to produce black and white or colored versions, use up to 4 levels of a ‘by’ variable, and customize titles and axes. To use the JITTER macro, the data will need to have a patient identifier, a continuous time variable, and a continuous measurable variable. Figure 9 displays CEA data collected twice weekly for the first month by arm.

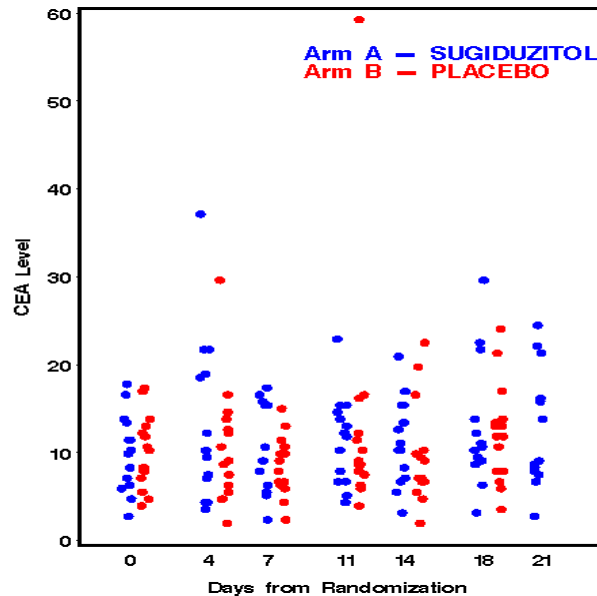
**Figure 9. JITTER Plot Macro Call and Output**

```

JITTER Plot Macro Call
%JITTER (dsn, class, id, time, var, min, max, by, ray, xlabel, ylabel, llabel, rlabel, ylevel, debug, ymin=, ymax
=, yby=, type=);
    
```



CEA Levels by Arm



There are a few outliers but otherwise the CEA values did not decrease as a result of treatment in either arm.

## EVENT CHART

Toxicity or Adverse Events are often monitored in a clinical trial. In this example, GI toxicities were expected to appear. The analysis of the data by assigning patients a binomial endpoint ('yes' if there was grade 3+ toxicity, 'no' if there was not grade 3+ toxicity) has shown that there was no difference between arms. The event chart was developed to provide patient specific information that cannot be discerned from a typical bar chart. An event chart tracks the individual patient experiences over time in order to discover patterns in whole data sets, such as time to toxicity onset or determining more toxic doses of treatment. The EVENT macro allows for the plotting of a variety of different endpoints. Figure 10 includes a toxicity event chart which tracks patient toxicity incidence over time according to the grade of toxicity as defined by the CTCAE v3.0. To use the event chart macro for this purpose, the data set must contain one observation per patient for each event (toxicity), a patient identifier, a toxicity grade and onset (starting date) of each toxicity grade. Further, if the time increment is days from randomization, on study and off study dates are needed.

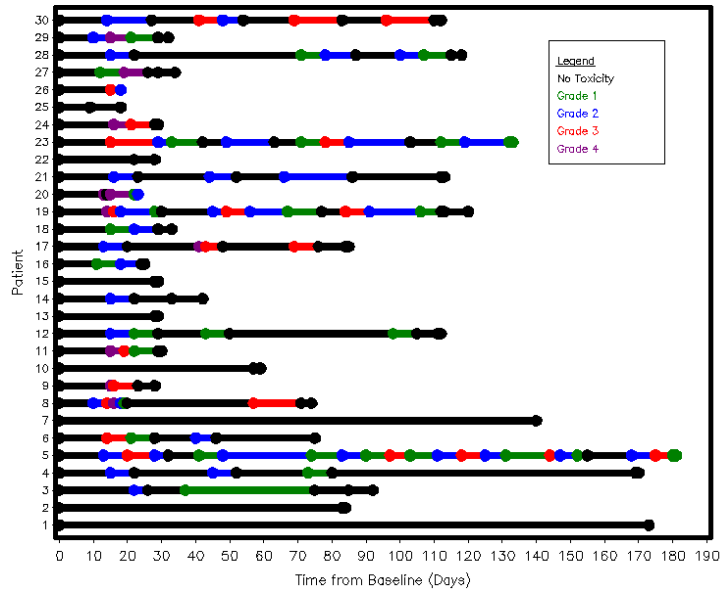
The macro may be modified to follow changes in QOL, which may also be seen in Figure 10. Suppose the study endpoint is not QOL score, but rather whether the patient had a categorical increase in QOL (5, 10 or 15 points). Patient data must be assigned a 'grade' corresponding to an increase in QOL score, much like the toxicity is graded.

Figure 10. EVENT Macro Call and Output Graphs

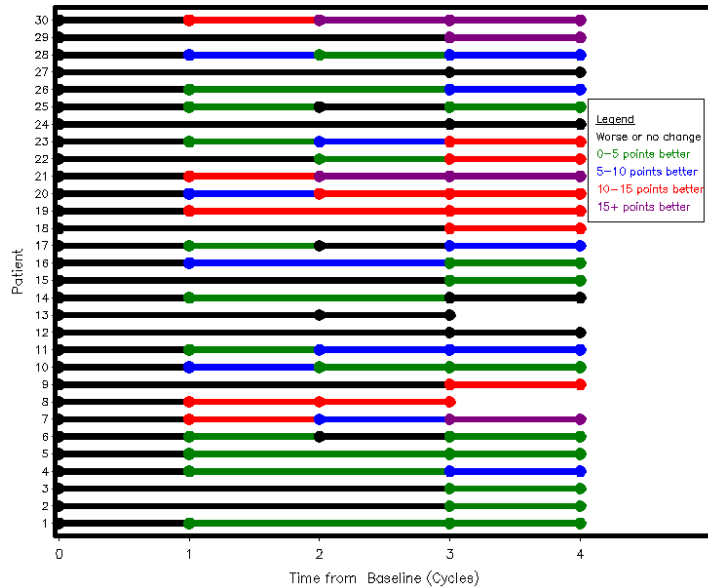
```
EVENT Macro Call
%EVENT (data=_last_,time=time, incre=cycle, onstudy=onstudy, off=off, id=patient, grade=maxgrad, grdfm=
grdfmt, sort=onstudy, linewidth=20, circw=.5, xleg=72, yleg=88, legbox=20, title1=, itle2=, legtitle=Legend, l
egftsz=2, axsftsz=2.5, vref=0, lvref=1, cvref=black, ylabel=Patient, xlabel=Time from
Baseline, order=, colors=black green blue yellow red purple orange);
```



### Individual Patient Toxicity during the Course of treatment Arm A



### POMS Total Score Changes Over Course of Treatment Arm A



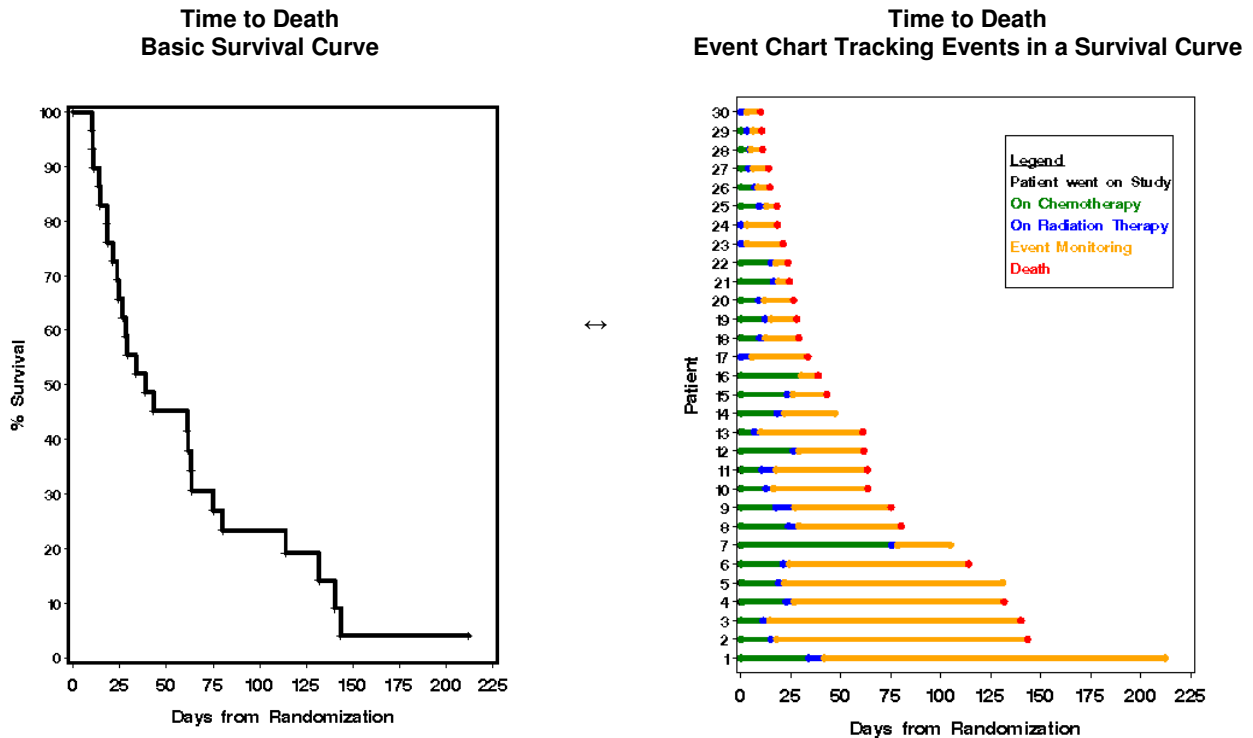
There were only 7 instances of grade 4 toxicity and they were of short duration. Toxicity incidence did not drive the length of time on study as many who stayed on study more than 3 months had some grade 3+ toxicity.

One-third of the patients did not have improving QOL scores after the first cycle of the study. Throughout the entire study, the majority of the patients only exhibited a small (0-5 point) improvement in QOL, but 4 were consistently at scores of 15+ points higher than baseline.

Alternatively, event charts may be used to provide additional information to the traditional survival curve by incorporating study or life events that occurred prior to patient death, rather than just reporting censored status (Goal 7). Suppose treatment on

arm A consisted of chemotherapy followed by radiation therapy, rather than the Sugiduzitol as in the rest of our examples, and that the patients had a survival profile as indicated in Figure 12. The EVENT macro may be used to track study events. In this case, the dataset must contain the variables indicated in the Figure 10 discussion with the events redefined and 'graded', plus the patient follow-up status. Remember, this is not part of the hypothetical study, rather an aside.

**Figure 11. EVENT Macro Used as a Survival Curve**



Three of the 4 patients who are censored (patients 1, 5, 7 and 14), have remained alive for longer than 3 months. One received prolonged doses of chemotherapy. Of the remaining patients, 3 did not receive the chemotherapy (patients 16, 23 and 24) and died within 2 months.

### DISPLAYING SUMMARY DATA:

An alternative to displaying complete data sets is to use plots that show summary statistics, such as those used to calculate the study endpoints. Common plots of this sort are bar charts and plots of means.

### BARChart MACRO

The BARChart macro produces a vertical bar chart of means between arms for a continuous variable. What is unique about this modification to the normal bar chart is that the user may program the macro to perform statistical tests of means between groups (parameter *class*) and print the resulting p-values in the axis statement. All p-values can be printed or only those that are significant. Other options include choosing the type of test statistic, black/white or color, labeling values for the axis statements and labeling the legend. Variables to be plotted on the x-axis must appear as columns in the data. Figure 12 shows means between arms that exist at the 4 week time point for the POMS and its subscales, showing all p-values (Goal 2).

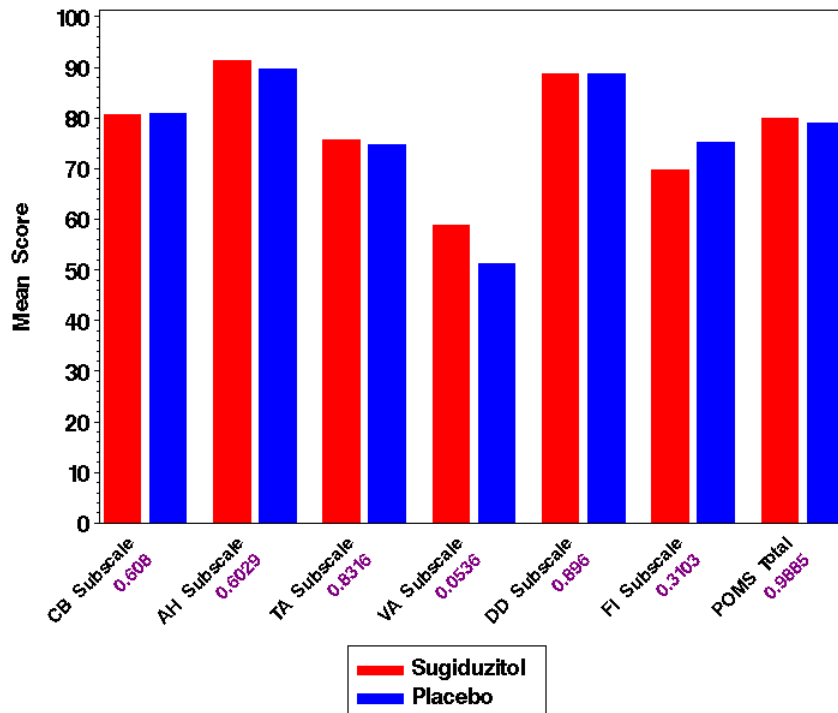
Figure 12. BARCHART Macro Call and Output

```

BARCHART Macro Call
%BARCHART(dsn = , var = ,class=, classlbl=, id = ,color=1,xv1bl=0,xlist=,xangle=45,xtext=2.5,
pctype=,pprint=0,pcolor=gray,phtext=2,awidth=3,ymin=0,ymax=100,yby=10,ylabel=,yangle=90,
htext=3,ftext=swissb,caxis=black,,cback=white,ctext=black,target=,gfile=,gtype=,debug=N);
    
```



Month 1 POMS Subscale and Total Mean Scores  
By Arm



At month 1, there is borderline significance shown between arms for the vigor/activity subscale ( $p=0.0536$ ), where the Sugiduzitol arm shows a higher mean score at this time point.

**PLOTN Macro**

The PLOTN macro produces a line plot of the means between arms, with or without confidence intervals, at each time point (Figure 13). The unique x-axis contains sample sizes for each arm at each time point. The user can plot according to levels of a variable, plot several variables, or plot several variables according to levels of another variable. The input data can be an array of values or multiple observations per patient. The user has the choice of including labels, additional plotting parameters, and color or black and white graphics.

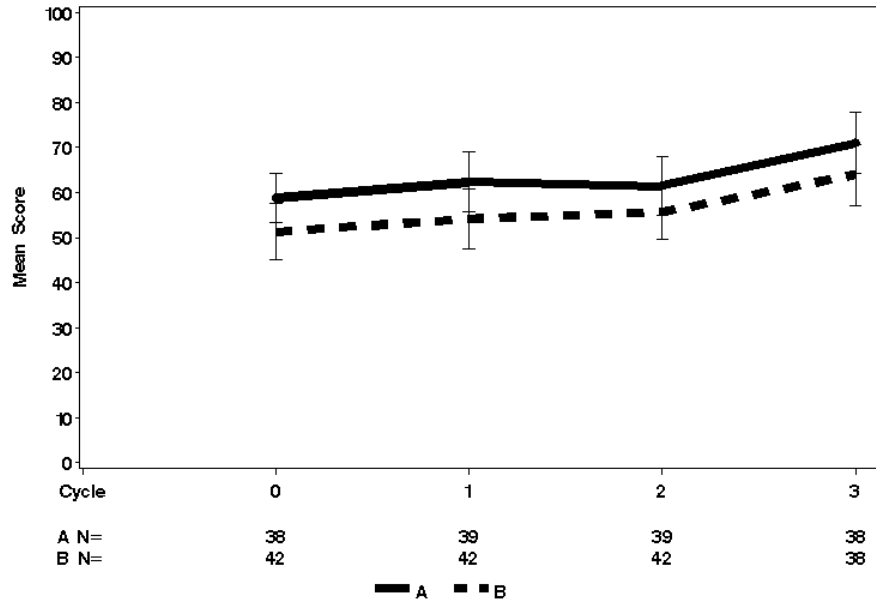
Figure 13. PLOTN Macro Call and Output

```

PLOTN Macro Call
%PLOTN(dsn, arm, time, xmin, xmax, xby, xvalues, score, ymin, ymax, yby, color=0, one=a, stat=, parms=, xlabel=
, ylabel=, ci=0, cimin=, cim�=, star=0, glabel=n, ray=n, rmin=, rmax=, legend=1, symbol=1, printed=n, bold=n
, ctick=, gfile=, gtype=gif, debug=n);
    
```



Mean Vigor/Activity Subscale Scores by Arm



Higher mean vigor/activity subscale scores are seen for the Sugiduzitol arm at each time point. The previous bar chart confirms this result at month 1.

## PULLING IT ALL TOGETHER

The macros presented in the paper, combined with some things we haven't shown you (combined with other SAS procedures) help form the basis of a summary report, which essentially acts as a point in time summary of the frozen data over time. To create this summary report, the tables created are copied and pasted into a final Microsoft Word document, while the graphics (typically saved out as .gif files) are imported into the document by simply using the "insert → picture → from file" menu items.

The tables and graphs are combined with concise explanatory text, much like those included in this paper. The summary report is used to aid in the reporting of study results as presentations, abstracts and manuscripts.

Note: The default for .gif files produced using the device=gif statement in the goptions statement is a file named sasgraph.gif. When producing many graphs in one program, all will be named identically and will write over each other. Thus, only the last graphic created will be saved as sasgraph.gif. Running the following code after each graphic in a program producing multiple graphs has proven quite useful, as the user is able to save each graph out the appropriate study folder.

```
data _null_;  
    call system("mv /path of directory where SAS was opened/sasgraph.gif  
                /path to directory in which to save graphs/NEWGRAPHNAME.gif");  
run;
```

## CONCLUSION

Consistency in analysis styles, and resulting summary reports, aids in the consultation process. The macros presented in this paper, provide the means for multiple statisticians and analysts to incorporate uniformity in reporting styles. The similarity in the summary reports ultimately saves time and money during the analysis phase of a study. The tables and graphs are easy-to-read and interpret requiring minimal explanation to the reader. All macros are formatted for use in manuscripts and other external publications.

Please note that none of the macros shared in this paper are externally validated, but have been internally checked using multiple studies for comparison.

## REFERENCES

1. McNair DM, Lorr M, Droppleman LF. Edits Manual for the Profile of Mood States. Revised 1992. Published by Edits/Educational and Industrial Testing Service. San Diego, CA 92167.
2. Nunally and Bernstein. (1994). *Psychometric Theory*. New York: McGraw Hill. pp. 132 – 133.
3. Sloan JA, Dueck A. Issues for Statisticians in Conducting Analyses and Translating Results for Quality of Life End Points in Clinical Trials. *Journal of Biopharmaceutical Statistics* 2004; 14(1): 73-96.
4. Atherton PJ, Jaspersen B, Nibbe A, Clement-Brown KA, Allmer C, Novotny P, Erlichman C, Sloan JA. What happened to all the patients? Event charts for summarizing individual patient data and displaying clinically significant changes in quality of life data. *Drug Information Journal* 37:11-21, 2003.
5. SAS Institute Inc., SAS System for Statistical Graphics, *First Edition*, Cary, NC: SAS Institute Inc.,1991.
6. SAS Online Doc 9.12 (<http://support.sas.com/onlinedoc/912/docMainpage.jsp>)
7. SAS Customer Support Center (<http://support.sas.com/>)

## ACKNOWLEDGMENTS

Many thanks for the members of the Quality of Life and Cancer Control teams at Mayo, Rochester, MN. Along with the authors, the team includes Mashele Huschka, Cathy Zhao, Jeffrey Huntington, Kelli Burger, Kristy Vierling, and Kara Curry.

## CONTACT INFORMATION

If you would like any macro code, please contact the author. Your comments, suggestions, and questions are welcomed and appreciated!

### **Denise Smith**

Mayo Clinic  
200 1<sup>st</sup> St SW  
Rochester, MN 55905  
Email: [smith.denise1@mayo.edu](mailto:smith.denise1@mayo.edu)