Visualizing Repeated Measures Data in Clinical Trials using SAS/Insight

Jan W. Buzydlowski
American College of Radiology

Abstract

The Radiation Therapy Oncology Group is conducting a Phase II clinical trial of the effect of radiosensitizer plus standard external beam radiation for the treatment of locally advanced prostate cancer. One of the measures of response in the trial is the marker of prostatic-specific antigen (PSA) and its value over time.

This paper discusses the ease of use of SAS/INSIGHT to graphically display the repeated PSA measures of the trial and how the use of a graphic representation as opposed to a previously used report format lead to an important insight which changed the method of doing the analysis.

Introduction

The Radiation Therapy Oncology Group (RTOG) of the American College of Radiology (ACR) conducts and analyzes clinical trials in the area of radiation oncology. The biostatisticians at RTOG are often asked to provide physicians and/or the group at large with information regarding the study, either for feedback during the study's execution for monitoring purposes or in the process of promulgating the results of the trials.

The information may be presented by using SAS/GRAPH to produce survival curves representing events based on the entire study cohort or a subset of interest or by using various techniques and procedures within SAS if other specific information is needed. If a large amount of detailed information is necessary and a report is required, then either SAS programming using PUT statements or PROC REPORT is used.

There are times when either a report is too voluminous or too detailed to be of effective use. In such a situation it is better to either replace or supplement the information with a graphic representation of the data. However, in the past, no easy and flexible technique existed other than SAS/GRAPH to graphically present the data.

A situation presented itself recently when we were asked to report a repeated measure response for each patient in a Phase II clinical trial that involved the use of a radiosensitizer to enhance radiotherapy for the treatment of prostate cancer. The measure of interest was that of prostatic-specific antigen (PSA). One area for analysis was to inspect how the measure changed over time and how the PSA profiles differed among patients.

The first step was to examine the psa trends per patient. Rather than use a previous method of reporting the value in a report format for each of the 36 eligible and analyzable patients, it was felt that a graphical method would improve the examination. It was at this time it was decided that perhaps SAS/INSIGHT could be used to look at the data.

Using SAS/INSIGHT

The first step was to change the data so that the repeated measures were contained within one variable and the dates in another, rather than in separate variables as was currently used. To accomplish this, arrays with a loop and an output statement were used:
* CONVERT HORIZONTAL REPEATED VALUES;
* TO MULTIPLE OBSERVATIONS;
DATA IN2; SET IN1;
DROP DATE1-DATE7 PSA1-PSA7;
ARRAY D DATE1-DATE7;
ARRAY P PSA1-PSA7;
DO I = 1 TO DIM(D);
   PSA = P[I];
   DATE = D[I];
   OUTPUT;
END;
RUN;

The variable DATE represented the date of the follow-up on which the patient was seen, and there were, at most, seven different times. The variable PSA represented the value of the prostatic-specific antigen which was recorded on each follow-up. The variable CN represented the case number which represents a unique key and identifier for each patient.

The next step was to invoke SAS/INSIGHT:

PROC INSIGHT DATA=IN2;

To present the data, a scatter plot of DATE vs. PSA was chosen. The next and most significant step was to group the scatter plot by CN. By simply choosing and clicking on the variable CN as a grouping variable, it easily allowed the graphic representation of the repeated measure over time.

Using SAS/INSIGHT to Explore and Interact

Once it was seen how easily the representation could be done, it was natural to do some interactive exploration. It was felt that it would be better to represent the time variable as the time from randomization (RANDDATE) to the date of the individual follow-ups. By switching to the Program Editor Window and creating a new dataset, it is possible to perform interactive transformations which are not available within SAS/INSIGHT. For example, the following code was executed:

* FIND TIME TO FOLLOW-UP;
DATA IN3; SET IN2;
FU_DAYS = DATE - RANDDATE;

The data was again presented using SAS/INSIGHT on the dataset IN3.

It was then felt that PSA values above 30 units were all similar in their indication of failure; and due to the scaling problem caused by the wide ranges, PSA values above 30 could be categorized into one group:

* CONVERT EXTREME PSA VALUES TO SAME;
* MINIMAL VALUES;
DATA IN4; SET IN3;
IF PSA > 30 THEN NEWPSA = 31; ELSE NEWPSA = PSA;
RUN;

Now, having a proper time and PSA scale and being able to visualize the PSA values over time, it was noticed that some patients had an interesting trend. This was particularly noticeable when a line plot, rather than a scatter plot, was used. The PSA was at first high, gradually came down, began to rise, then suddenly became extremely low. It was suspected that some clinical intervention was being performed. Examining the patient charts revealed that additional hormones were being used when the patient was beginning to have a rising PSA, thus causing the sudden and pronounced decrease in PSA. Having found this, it then became clear that this would need to be factored into any additional analyses.

Using SAS/INSIGHT for Graphical Presentation

Once it was learned that the additional treatment is an important factor in the analysis, and wanting to show this to the physicians, the additional variable ADDRX recorded on each follow-up was added to the dataset using the same methods as shown above. ADDRX is a binary variable indicating whether or not the patient received any additional treatment.

To indicate the additional information, markers were added with the scatter plots. Those NEWPSA values of 31 were selected and given an upward triangle (△) to represent PSA values above 30. A cross (+) was chosen when ADDRX was affirmative.

After using the markers to code the additional information, it was then very easy to re-display the scatter plot of time versus PSA grouped by patient showing the
anomalous trends. To produce a hard copy, the data can be graphed using PROC GPLOT, with each patient represented by a separate graph as the data is already organized to do so. The data can also be easily plotted using PROC INSIGHT directly from the above methods indicated.

Conclusion

The use of SAS/INSIGHT easily allows both interactive exploration and graphical presentation. Individual experimental units can be analyzed by choosing the unit’s unique identifier for use as a grouping variable. Interactive exploration is possible by switching between the SAS/INSIGHT Window and the Program Editor Window using additional work datasets. Additional information can be presented through the use of the markers that are provided.