INTRODUCTION

Clinical trials are scientific experiments carried out in human subjects to test new drugs and treatment regimens. Phase II and Phase III clinical trials are usually designed to establish efficacy and monitor the safety of a new drug or treatment regimen. A major part of the effort to monitor the well-being of the patient consists of the collection of clinical laboratory data. In short term trials for the treatment of acute disease or discomfort, it is not uncommon that clinical laboratory data be collected only twice, once before therapy and once after therapy. Long term trials for treating chronic disease are characterized by multiple dose treatment regimens intended to alleviate the signs and symptoms of disease, induce response, evaluate the duration of response, extend survival time, or extend the disease free interval. Such studies will be monitored by laboratory data collected prior to the beginning of therapy, at one or more times during the course of therapy, and at least once following the conclusion of therapy. The fact that the laboratory data may be collected at irregular intervals in such trials complicates the evaluation of these data. This nonalignment of clinical laboratory data collection times, coupled with the fact that many of the parameters do not exhibit a normal distribution, indicate that the "usual" types of analysis, such as repeated measures analysis, may be inappropriate.

This paper deals with an alternative approach to the evaluation of clinical laboratory data accumulated from clinical trials. This approach is currently in use at Bristol-Myers Pharmaceutical Research and Development and is termed a multi-evaluative approach since the data are analyzed by more than one method. These methods will be illustrated using data collected from three types of clinical studies: an antihypertensive study, an anticancer study, and an antibiotic study. While the first two types of studies represent long term clinical trials, they differ from one another in a number of key aspects. In the antihypertension trial, data were collected at a few regular intervals, whereas in the anticancer trial, an unequal number of sets of laboratory data were collected from patients at irregular intervals. This distinction is not meant to be a criticism of anticancer trials, but rather a reflection of the difficulty inherent in such trials. Another difference between the antihypertensive and the anticancer trial relates to the expectations concerning the clinical laboratory data. In the anticancer trial, toxicities in terms of values beyond the normal range are expected, due to the disease, the treatment regimen, or to both. However, in the antihypertensive trial, abnormal laboratory values might be considered unusual and noteworthy. In contrast to these long term studies, the antibiotic study represents a short term clinical trial in which clinical laboratory data were collected three times for most patients. The multi-evaluative approach that we propose takes all of these differences into consideration in the evaluation of the clinical laboratory data.

THE MULTI-EVALUATIVE APPROACH

For each laboratory parameter, the basic multi-evaluative approach consists of testing for trends over time in the data, comparing pre and post therapy values to the normal range provided by each investigator, and comparing "large" changes in laboratory values from pre to post therapy. Depending on the type of study, these basic approaches can be easily modified to provide the most useful information to the medical monitor concerning the effects of the treatments on laboratory data. This multi-evaluative approach has been successfully applied to randomized clinical trials using a parallel group design and to single treatment group studies that have a retrospective control group. The remainder of this paper describes each of the tests in the basic multi-evaluative approach, shows how these can be modified to test the type of study under consideration, and discusses the statistical interpretation of the results of the analyses. Heuristic justification for the tests that we have chosen is also presented.

I. TREND ANALYSIS

Trends over time are evaluated using the following technique. For each patient's data, fit a linear regression model. The independent variable (abscissa or X-value) is time of observation and the dependent variable (ordinate or Y-value) is the observed laboratory value. Time of observation usually occurs as the date (expressed numerically as a SAS date) that the laboratory rest was done (date of sampling), but this variable could just as easily be the elapsed number of days or hours that the sample was drawn from the patient, using initiation of therapy as time zero. The slope obtained from the linear regression model is the least squares estimate, regardless of the underlying probability distribution associated with the specific laboratory parameter (see Neter and Wasserman, p. 37). This consideration is important in the discussion of testing hypotheses. However, first we shall indicate why we chose the simple linear model

\[ Y_t = B_0 + B_1 X_t + E_t \]

where "t" indexes the laboratory tests for a single patient for the specific laboratory parameter and \( B_0 \) is the unknown intercept, \( B_1 \) the unknown slope and \( E_t \) the error term.

In any regression model over time, one is faced with the question of the appropriate order of the polynomial that may be fit to the data. It is generally accepted that the first order or simple linear model gives reasonable estimates of overall trend. A strong competitor may be the second order model or parabola. For example, such a model would be more appropriate if the effect of the
treatment were to change some parameter over the course of therapy with a return to baseline as therapy tapered off. However, such information is not usually available for most treatments studied. Finally, it is typical in clinical trials that more sets of laboratory data are collected while patients are on therapy than when therapy has ceased. Hence, the tail of an assumed parabola may never be observed. In fact, we have fit parabolas to data from cardiovascular studies and have found that the tests on the second order coefficients have added little to our impression of trends in the data beyond the information obtained from tests on the slope estimate from the simple linear model. Polynomials of degree three or more were excluded because of restrictions on number of tests required and because such functions may well exhibit vagaries of the data rather than overall trends. Hence, based on theoretical considerations and practical experience, we have chosen the slope estimate from a simple linear regression model as an indicator of trend in the data.

Once each patient's data has been reduced to a single indicator of trend, i.e., the slope estimate from the simple linear model, it is possible to test for overall trends both within a treatment group and between treatment groups. The Wilcoxon Signed Rank Test is used to test for trend within each treatment group. Computationally, rank the absolute values of the slopes from smallest to largest and add up the ranks associated with slopes that were positive initially (Hollander and Wolfe, pp. 27-28). For ties, use average ranks. Significance of the test statistic is dependent on both the sign of the slope estimates and on the relative magnitudes (disregarding sign) of the slope estimates. The hypothesis of no treatment effect over time translates to testing the null hypothesis that the median slope is zero. The signed rank test is preferred to the sign test because the latter takes only the sign of the slope into account. On the other hand, while the one-sample t-test takes the actual magnitude of the slope estimates into account, it may be overly sensitive to one or two very large positive or negative slopes. Moreover, normality of the slope estimates is not required for the signed rank test. In practice, the two-sided signed rank test results have correlated well with our visual impression of overall trends in the patients' plotted laboratory values within each treatment group.

Equality of trends between treatment groups translates to testing the hypothesis of equal median slopes in each treatment group. When there are only two treatment groups, the hypothesis is tested using the Mann-Whitney-Wilcoxon Rank Sum Test. See Hollander and Wolfe, pp. 69-69. Since the test statistic is the sum of the ranks of the slope estimates from one sample, where the ranks are obtained from the combined sample, both the sign and relative magnitude of the slope estimates are taken into account. The rank sum test is preferred to the median test because the latter essentially accounts for only the sign of slope estimate relative to the observed median of the combined sample. Moreover, the rank sum test is not as seriously affected as the two-sample t-test by a few very large or very small values and does not require that the slope estimates be normally distributed. In practice, the two-sided rank sum test results have been found to be complementary to those obtained within treatment groups by the signed rank test.

In the event that there are more than two treatment groups, the between group comparison is made using the Kruskal-Wallis Test (Hollander and Wolfe, pp. 115-116). The test statistic is a function of the sum of squared deviations of the average rank within each treatment group from the overall average rank. Again, all slope estimates are combined in a single sample prior to ranking from smallest to largest. As in the two-sample case, the relative magnitudes of the slope estimates are taken into account, but the test is more sensitive than the k-sample median test and less sensitive to outliers than the one-way analysis of variance for normally distributed variables.

II. CLASSIFICATION ANALYSES
   II.A. CHANGES RELATIVE TO NORMAL LIMITS

The following procedure is useful for summarizing the change in a laboratory parameter from a pre-therapy or baseline value to a post-therapy reading. The pre-therapy and post-therapy laboratory values are each categorized according to the investigator's normal range for that parameter as follows:

- **Low** = value less than the lower limit of the normal range
- **Normal** = value within the normal range
- **High** = value above the upper limit of the normal range

The 3 x 3 table formed by the cross classification of the categories of Pre and Post for each treatment group presents an immediate view of the changes in a laboratory parameter over the course of the study.

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Normal</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>High</td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

Cell numbers are indicated in the body of the table.

The three cells on the main diagonal (#1,5,9) indicate no change (N) from Pre to Post relative to the normal range. The three cells below the main diagonal (#4,7,8) represent a decrease (D) in region relative to the normal range, while the three cells above the main diagonal (#2,3,6) represent an increase (I) in region relative to the normal range.

A test of whether or not the changes in region are similar for all k treatment groups can be performed in the following manner. For each group, classify the changes in region as a Decrease (total of cells #4,7,8), No Change (total of cells #1,5,9) and Increase (total of cells #2,3,6). A Chi-square
test can then be used for testing the homogeneity of proportions in the resulting k x 3 contingency table. This test provides additional insight in the comparative trends in laboratory data values between treatment groups.

In some studies, it is not uncommon that most laboratory values for a certain parameter are outside the normal limits. This occurs frequently in anticancer chemotherapy trials. The above procedure may be used with expanded limits to obtain more useful information from the data. For example, the upper expanded cut off might be twice the upper limit of normal and the lower expanded cut off might be one-half the lower limit of normal to keep equal distances on a logarithm scale. For an arithmetic scale, one might consider taking the upper limit plus some function of the normal range (the upper normal minus the lower normal limits), while the lower expanded limit would be the lower normal limit minus the same function of the normal range.

II. B. ANALYSES OF "LARGE" CHANGES

While the analyses of changes relative to normal limits give some information as to trends in the data, there remains a possibility of distorted results. This may occur when large changes in laboratory values are observed within some patients, but each patient's data are all within a single category relative to the normal or expanded normal limits. Another example is when data values just cross these limits, say from the high end of Normal to the low end of High. To compensate for these possibilities, we have adopted a system of identifying "large" changes. At present, we consider a change of greater than 50% of the normal range as a large change. Clinically, this amount has usually been of interest to the medical monitor. Taking Pre and Post values as in the previous section, changes are classified as "large" if greater than 50% of the normal range:

- **Decrease** = "large" decrease from Pre to Post
- **Not Large** = change less than "large" from Pre to Post
- **Increase** = "large" increase from Pre to Post

Treatment groups are then compared using the Chi-square test for homogeneity of proportions in the k x 3 contingency table formed by the Treatment x Large Change in Value cross-classification. The results of this test are not necessarily the same as those obtained in the Chi-square test described in the previous section and provide further insight to the evolution of laboratory data.

III. CRITICAL POINTS ANALYSES

Another type of analysis of changes in laboratory values is performed in studies where the disease and/or treatment may induce toxic values. If laboratory data values are expected to show toxicity, such studies often require laboratory tests every few days during therapy. The analyses described below require at least three sets of laboratory values, one prior to therapy, one post therapy, and at least one during therapy. In the following analyses, "Initial" refers to the most appropriate baseline value prior to or on initiation of therapy and "Final" means the most appropriate final value after therapy. The approach is intended to give some indication of whether or not a regimen may be toxic and to show whether or not recovery is observed.

Having obtained the direction of toxicity for a laboratory parameter from the medical monitor, usually increasing values for serum chemistry and decreasing values for hematology, proceed as follows. Between the initial values (I) and final values (F), find the most toxic value (MT). The treatment group median, minimum, and maximum values are listed for I, MT, and F. For each patient, calculate the change from initial to most toxic as MT-I and change from most toxic to final as F-MT. Positive values denote increases and negative values denote decreases. Within each treatment group, the Wilcoxon signed rank test is used to test that the median of each of the changes is zero. A one-sided test is used. If the direction of toxicity is positive, changes in MT-I are declared significant for large values of the signed rank test and indicate toxicity, while changes in F-MT are declared significant for small values of the signed rank test and indicate significant recovery. The two-sided Mann-Whitney-Wilcoxon rank sum test or Kruskal-Wallis test is used to test for treatment group differences in each change, depending on whether there are two or more than two treatment groups. Significance of the rank sum test on changes MT-I indicates more toxicity in one of the two treatment groups, while the same on changes F-MT indicates more recovery in one of the two treatment groups. The exact level of toxicity may be roughly estimated from the median most toxic value and the median change from initial to most toxic. Similarly, the median change from most toxic to final value may be used to estimate the amount of recovery observed.

IMPLEMENTATION IN SAS

Implementation of the multi-evaluative approach is extremely easy in SAS. The Bristol-Myers Common Clinical System is an IMS-based system. A screen type is defined for each relevant section of the Case Report Form (CRF) for each clinical study. Data are entered and verified by screen, and the data base is updated each evening by a batch program. The SYRACUSE implementation provides a SAS disk data set for each screen type. In addition, there is a screen (Lab Normals) for the normal range of laboratory tests with input fields for investigator, institution, sex, age limits, test name, beginning and end dates for the normal range and the lower and upper limits of normal. The clinical laboratory SAS data set (Lab Data) contains the following fields: unique study/investigator identifier, unique subject, date, test name and value.

The SAS demographic and dosing information data sets can be merged by subject in order to obtain information on treatment group, institution, inclusive dates of therapy, age, sex, height, weight, etc. in a single observation in a SAS data base.
referred to as Study Information. Note that all dates are stored as SAS dates so that numerical comparisons can be made.

The three SAS data sets just described, Study Information, Lab Data and Lab Normals, are all that are needed for the multi-evaluative approach.

In studies with single Pre- and Post therapy laboratory tests, and with no missing data, the Pre and Post values may be easily obtained from Lab Data, as the following sample code illustrates:

```sas
DATA A; SET FILE.LABDATA;
PROC SORT; BY SUBJ TEST DATE;
DATA B; SET A; BY SUBJ TEST;
RETAIN PRE;
IF FIRST.TEST THEN PRE = VALUE;
IF LAST.TEST THEN DO;
   POST = VALUE; OUTPUT;
END;
KEEP SUBJ TEST PRE POST;
```

Finding PRE, POST, INITIAL and FINAL involves much more programming if there are multiple pre or post therapy values. Now, sort and merge Study Information and Data Set B by SUBJ. Sort the resulting data set by SEX and TEST and merge with Lab Normals, keeping only observations with the age limits from Lab Normals that contain the age of SUBJ. It is now a simple matter to categorize PRE and POST according to the normal limits to extended normal limits:

```sas
***NORMAL RANGE:
-1=LOW 0=NORMAL I=HIGH;
CPOST = -1*(POST < LOWER) + 1*(POST > UPPER);
***CHANGE IN REGION:
-1=DECREASE 0=NO CHANGE I=INCREASE;
CHANGE = CPOST-CPRE;
CCHANGE = -1*(CHANGE < 0) + 1*(CHANGE > 0);
***LARGE CHANGES:
-1=DECREASE 0=NOT LARGE I=INCREASE;
DIF = POST-PRE;
LARGE = (UPPER-LOWER)/2;
CLARGE = -1*(DIF < -LARGE) + 1*(DIF > LARGE);
```

One may obtain the 3 x 3 contingency tables for each treatment group (TRT) and the k x 3 table comparing treatments for either change in region (CCHANGE) or large changes (CLARGE):

```sas
PROC SORT; BY TEST TRT;
PROC FREQ; BY TEST TRT;
TABLES CPRE*CPOST;
PROC FREQ; BY TEST;
TABLES TRT*(CCHANGE CLARGE)/CHISQ;
```

Currently, PROC SYSREG is used to obtain the slope estimates from each patient's data after any unwanted data have been stripped from Lab Data. Once the slopes have been obtained, the treatment groups may be compared by using PROC NPAR1WAY to obtain the Wilcoxon rank sum test or Kruskal-Wallis test for each laboratory parameter. The Wilcoxon signed rank test may be easily programmed in PROC MATRIX. The program segment at the end of the paper details the calculation of the signed rank test statistic TPLUS, the normal approximation TNORM, and the P-levels for TNORM. After looping through the macro SGNDKR for each test, merge the test names to and print the data set RESULTS for the group. Reset macro GROUPNO and repeat the process for the next group. The macro SGNDKR is also used in the toxicity analyses to examine the MT-I and F-MT changes within groups, and NPAR1WAY is used for the between group comparisons.

To summarize, the steps needed to apply the multi-evaluative approach are easy to implement in SAS.

**EXAMPLES**

1. **ANTIHYPERTENSION TRIAL**

   In this study, patients were evaluated for two weeks prior to the beginning of the study. There was a four week placebo wash-out, followed by 18 weeks of therapy in this double-blind parallel groups study. Laboratory tests were done at Week 4, 10, 16 and 22.

   Serum cholesterol was chosen for illustrative purposes in this two-arm study. Taking Week 4 as Pre (just prior to initiation of active therapy) and Week 22 as Post (immediately following 18 weeks of therapy), the classification analyses yielded:

   ```plaintext
   DRUG A POST
   L 0 0 0
   N 2 34 0
   H 0 3 3
   DRUG B POST
   L 0 0 0
   N 1 29 3
   H 0 2 2
   ``

   ```plaintext
   CHANGE IN REGION
   DEC UNCH INC
   DRUG A 5 37 0
   DRUG B 3 31 3
   X2 = 3.73 P = .16
   ``

   After fitting a simple linear model to all of the data from each patient, the two-sided Wilcoxon signed rank test and the two-sided rank sum test on slopes yielded:

   ```plaintext
   P-VALUE
   DRUG A .059
   DRUG B .037
   BETWEEN DRUG GROUPS .503
   ```
Using the sign of TNORM, there is some evidence of decreasing cholesterol levels over the course of therapy in each group. However, there is no statistically significant difference between groups, and the classification analyses support the notion that there is very little happening in cholesterol levels in this study. This interpretation appears in the statistical report. The medical monitor uses these results to answer the general question of effect of therapy on cholesterol. Complete patient data listings are provided so that he may address problems in specific patients.

2. ANTICANCER TRIAL

In this solid tumor trial for refractory testicular cancer, Trt A represents a three-drug chemotherapy regimen (N=32). The retrospective control group (N=24), Trt B, received a variety of different regimens, none containing the drug of main interest in Trt A. Both groups of patients had failed first-line chemotherapy, and regimens prior to the study regimens were similar. A course of therapy consisted of administration of the study regimen over 2 - 4 days, and courses were to be repeated every 3 - 4 weeks. Patients were on study for different periods of time, with laboratory data collected at irregular intervals:

Patient 1 had laboratory tests on Week 0, 4, 6, 10, 16, 18 and 22; Patient 2, on Week 0, 2, 5, 10, 12, 15, 18 and 19; Patient 3, on Week 2, 12, 14, and 16; etc. The results of the critical points and trend analyses for hemoglobin are presented below.

HEMOGLOBIN

<table>
<thead>
<tr>
<th>TRT A (N=32)</th>
<th>TRT B (N=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MED</td>
<td>MED</td>
</tr>
<tr>
<td>I</td>
<td>13.3</td>
</tr>
<tr>
<td>MT</td>
<td>8.35</td>
</tr>
<tr>
<td>MT-I</td>
<td>-4.0</td>
</tr>
<tr>
<td>F</td>
<td>11.05</td>
</tr>
<tr>
<td>F-MT</td>
<td>2.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P-VALUES AND DIRECTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT-I</td>
</tr>
<tr>
<td>BETWEEN TRT</td>
</tr>
<tr>
<td>F-MT</td>
</tr>
<tr>
<td>BETWEEN TRT</td>
</tr>
<tr>
<td>TRENDS</td>
</tr>
<tr>
<td>BETWEEN TRT</td>
</tr>
</tbody>
</table>

Significant toxicity was observed in both treatment groups, but so was significant recovery. The treatment groups did not differ significantly in either the amount of toxicity observed or in the amount of recovery. Significant overall trends toward toxicity within each treatment indicated that the recovery observed was less than the toxicity induced. However, the trends could not be distinguished between treatment groups. Hence, similar toxicity occurs in both treatment groups for hemoglobin, but this is generally reversible.

3. ANTIBIOTIC TRIAL

In this two-arm, randomized, parallel groups study, two cephalosporin antibiotics were compared in skin and soft tissue infections caused by susceptible microorganisms. Although therapy could last ten days or more, some patients were on study for only a few days, only long enough to culture the causative microorganism and determine that it was not susceptible to either antibiotic. The liver enzyme, SGPT, results are presented for the classification and trend analyses. Laboratory tests were to be done on Day 0, 4, 10, and at the end of therapy. Most patients had three sets of data.

DRUG A POST

| L | N | H |
|   |   |   |
| 0 | 0 | 0 |

DRUG B POST

| L | N | H |
|   |   |   |
| 0 | 0 | 0 |

CHANGE IN REGION

<table>
<thead>
<tr>
<th>DEC</th>
<th>UNCH</th>
<th>INCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRUG A</td>
<td>3</td>
<td>59</td>
</tr>
<tr>
<td>DRUG B</td>
<td>11</td>
<td>49</td>
</tr>
</tbody>
</table>

\[X^2 = 7.53\] \[P = .02\]

LARGE CHANGE (50% NORMAL RANGE)

<table>
<thead>
<tr>
<th>DECR</th>
<th>NOT LG</th>
<th>INCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRUG A</td>
<td>6</td>
<td>50</td>
</tr>
<tr>
<td>DRUG B</td>
<td>17</td>
<td>41</td>
</tr>
</tbody>
</table>

\[X^2 = 10.10\] \[P = .006\]

TRENDS AND DIRECTION

<table>
<thead>
<tr>
<th>N</th>
<th>DIR</th>
<th>P-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRUG A</td>
<td>101</td>
<td>+</td>
</tr>
<tr>
<td>DRUG B</td>
<td>92</td>
<td>+</td>
</tr>
</tbody>
</table>

\[BETWEEN DRUGS P = .004\]

A significant overall trend toward toxicity was observed for Drug A, while near significance was observed for Drug B. Comparatively, the median slope for Drug A was significantly greater than that of Drug B. There were significantly more increases...
in region for Drug A than Drug B, and there were more increases larger than one-half the normal range for Drug A. Overall, Drug A produced more toxicity in SGPT than did Drug B. Since patients were not followed after therapy ceased, it was not possible to investigate recovery.

**DISCUSSION**

The examples of the previous section highlight the fact that the statistical interpretation is based on all of the analyses carried out in the multi-evaluative approach for a single laboratory parameter. This occurs because there is often no simple answer to the question of what happened for a certain laboratory parameter. While the multi-evaluative approach may be anti-conservative since each test is carried out at the .05 significance level, the approach was instituted to avoid many of the problems inherent with single or repeated paired t-tests within groups and analysis of variance between groups. These approaches often yielded many statistically significant results that were clinically meaningless. This may have resulted from nonnormality of the data, from the presence of a few outliers, or from a combination of these. In contrast, the multi-evaluative approach is nonparametric and is less affected by these events.

While no known analysis accounts for the confounded effects of disease, treatment, time, transfusions and other interventions, the multi-evaluative approach is designed to present a fair evaluation of the overall impact of these on the laboratory data. Moreover, the basic sampling unit for each test in the multi-evaluative approach is the individual patient, just as in the clinical trial. The multi-evaluative approach is adaptable to different clinical trial settings: the trend and classification analyses may be used to look for tendencies when precise measures of effect are not needed; the trend and critical points analyses indicate induced toxicity, investigate recovery, and provide estimates of both. The results of the multi-evaluative approach are usually easily interpreted and correlate well with the medical monitor’s intuition obtained in reviewing data listings and CRFs.

To summarize, the multi-evaluative approach is flexible, easy to use and implement in SAS, and yields interpretable results.

**ACKNOWLEDGEMENTS**

Special thanks to Dr. J.M. White for his help with the parabolic regressions and to Karen R. Yerden for excellent preparation of the manuscript.

**Contact Author:**
William A.T. Archambault Jr., Ph.D.
Assistant Director, Biostatistics
Bristol-Myers Co., Pharmaceutical R & D
P.O. Box 637
Syracuse, New York 13201
(315)326-2694

**REFERENCES**


**PROC SORT DATA=FILE.LABDATA;**
**BY GROUP TEST SUBJ;**
*** OMIT SUBJ WITH ONLY 1 LAB TEST; **
**DATA ONE; SET FILE.LABDATA;**
**BY GROUP TEST SUBJ;**
*** IF FIRST.SUBJ AND LAST.SUBJ THEN DELETE; **
*** USE PRINTTO TO AVOID MATRIX FAGING; **
**PROC PRINTTO UNIT=22 NEW;**
*** CALCULATE SLOPES; **
**PROC SYREG NOPRINT OUTTEST=TREND;**
**BY GROUP TEST SUBJ;**
**MODEL VALU = DAYS;**
*** ACCUMULATE RESULTS-BEGIN W. FAKE; **
**DATA RESULTS;**
*** INPUT TEST & NBS TPLUS TSTAR LOWTAIL UPTAIL; **
**CARDS;**
FAKE 0 0 0 0 0 0

*** DATA SET FOR INTERMEDIATE RESULTS, **
**DATA INTERM;**
**SET RESULTS;**
**DROP TEST;**
*** 3 MACROS DO SIGNED RANK TEST/GROUP; **
*** IDENTIFY GROUP NUMBER; **
**MACRO GROUPNO '1 % **
*** IDENTIFY LAB TEST; **
**MACRO THAME 'HGB' % **
*** USE PROC MATRIX FOR TESTING; **
**MACRO SGNDRK DATA NEW;**
**SET TREND;**
*** IF GROUP=GROUPNO AND TEST=THAME; **
**KEEP DAYS;**
**PROC MATRIX;**
*** FEICH SLOPES DATA=NEW; **
**SIGN = (SLOPES > 0); **
**RANK = RANKABS(SLOPES);**
*** OBTAIN SIGNED RANK STATISTICAL; **
**TPLUS = SUM(SIGN*RANK);**
**M = NROW(SLOPES);**
**TRAR = M*(M+1) / 4; **
**TNORM = (TPLUS-TRAR) / SQRT((2*M+1) / 6); **
**LOWTAIL = PROBNORM(TNORM);**
**UPTAIL = 1 - PROBNORM(TNORM);**
**TWOTAIL = 2*(1-PROBNORM(ABS(TNORM)));**
**MTX = M//TPLUS//TNORM//LOWTAIL//UPTAIL//TWOTAIL;**
**MTX = MTX';**
**COLS = 'NBS' 'TPLUS' 'TSTAR' 'LOWTAIL' 'UPTAIL';**
**OUTPUT MTX OUT=INTERM COLNAME=COLS; **
*** CONCATENATE RESULTS FOR EACH TEST **
**WITHIN A GROUP;**
**DATA RESULTS;**
**SET RESULTS INTERM;*** END MACRO; **
**SGNDRK *** RUNS MACRO FOR HEMOGLOBIN; **
**MACRO THAME 'PLAT' %**
**SGNDRK *** RUNS MACRO FOR PLATELETS;**