A Computer Program for Nonparametric Analysis of Incomplete Repeated Measures from Two Samples

Charles S. Davis
Department of Biostatistics
University of Iowa, Iowa City, IA 52242, U.S.A.

Abstract

RMNP2 is an easy-to-use FORTRAN program for the analysis of repeated measures using the nonparametric two-sample tests of Wei and Lachin [1] and Wei and Johnson [2]. The program compares two groups of subjects or experimental units when measurements are obtained at multiple time points, or under multiple conditions, from each subject. A strength of the methodology is that subjects with missing responses at one or more time points can be included in the analysis, under the assumption that the missing value mechanism is independent of the response. In contrast to other methods that require parametric assumptions concerning the distribution of the outcome variable, RMNP2 is applicable when the response variable is continuous, but not normally-distributed. The program is also useful in the analysis of ordered categorical outcomes when the number of possible responses is too large to permit application of general categorical data methodology. The program can be run on microcomputers, workstations, and mainframe computers. Two examples illustrating the usage and features of RMNP2 are provided.

1. Introduction

A frequently-occurring problem in biomedical research is to evaluate the effectiveness of a new treatment in patients or subjects with a specified disease or condition. One commonly-used design involves selection of eligible subjects, randomization to one of two groups (new therapy, standard treatment), measurement of an appropriate outcome variable following treatment, and statistical comparison of the distribution of the outcome in the two groups. In many such studies, the response variable of interest is measured at multiple fixed time points for each experimental unit. In these situations, it is often desirable to base the comparison of the two treatment groups on the vector of responses from each subject. The statistical analysis is complicated by the dependence among successive observations made on the same experimental unit. In addition, since the investigator cannot usually control completely the circumstances for obtaining measurements, the data may be partially incomplete.

General approaches for the analysis of repeated measurements are now widely available in standard statistical software packages such as BMDP [3] and SAS [4]. Although these methods allow missing data, unbalanced measurement patterns, and multiple discrete and/or continuous covariates, they make the strong assumption that the response variable is normally-distributed. The weighted least squares approach to the analysis of repeated categorical outcomes [5], and subsequent extensions to accommodate missing data [6–8], can also be carried out using standard software [9,10]. The distributional assumptions underlying this approach, however, limit its usefulness to situations in which the sample size is quite large and the number of time points and number of possible values of the categorical response variable are both small. Versatile methods based on extensions of generalized linear models [11] to the repeated measures setting have also been proposed [12–17]. These approaches do not make assumptions concerning the *joint* distribution of the responses across time, but rather assume only a particular form for the *marginal* distribution at each time point, e.g., normal, Poisson, binomial, gamma. Although none of these methods have yet been implemented in commercial packages, software for the GEE approach [12, 13] is available [18–20].

While all of the above methods require assumptions on either the joint or the marginal distributions of the response variable, there are at least three general situations in which nonparametric methods may be useful. First, when the response is continuous, the assumption of multivariate normality may not be reasonable or the underlying distribution may be unknown. In this case, the use of standard parametric procedures is subject to criticism. Second, when the response is an ordered categorical variable with a large number of possible outcomes, the general categorical data methods may be inapplicable due to sample size

limitations. In addition, the restrictive proportional odds assumption of ordinal data methodology may not be justified. Apart from these considerations, there are also situations in which it may be desirable to confirm the results of a parametric analysis using nonparametric methods.

Two-sample nonparametric tests for *complete* multivariate observations have been studied extensively, e.g., [21,22]. This paper describes the FORTRAN program RMNP2 (**R**epeated **M**easures **N**on-**P**arametric **2** samples) which implements two general methods for the analysis of *incomplete* repeated measurements [1, 2]. Although RMNP2 does not require any assumptions concerning the distribution of the response variable, the statistical methodology is limited to comparisons between two groups of experimental units, i.e., multiple covariates can not be accommodated. In addition, the methodology provides hypothesis tests only, and not estimation procedures. Finally, although the missing value patterns in the two groups are allowed to be different, missing responses are assumed to be missing completely at random [23].

Section 2 reviews the Wei-Lachin [1] and Wei-Johnson [2] methodologies and outlines the computational methods, while section 3 describes the RMNP2 program. Section 4 gives two examples illustrating the usage and features of RMNP2. Finally, sections 5 and 6 describe hardware and software specifications and program availability, respectively.

2. Statistical methodology

Wei and Lachin [1] studied a family of asymptotically distribution-free tests for equality of two multivariate distributions. Although their methodology was motivated and developed for multivariate censored failure time data, an important application is to repeated measures with missing observations.

Let n_1 and n_2 denote the number of subjects in groups 1 and 2, respectively, and let $n = n_1+n_2$. Suppose that repeated measurements of an outcome variable Y are scheduled at time points labelled $1, \ldots, t$. Let y_{hij} denote the response at time j from the ith subject in group h, for $h = 1, 2$, $i = 1, \ldots, n_h$, and $j = 1, \ldots, t$. In addition, define indicator variables

$$\delta_{hij} = \begin{cases} 1 & \text{if } y_{hij} \text{ is observed,} \\ 0 & \text{if } y_{hij} \text{ is missing.} \end{cases}$$

Also let $F_h(x_1, \ldots, x_t)$ denote the multivariate cumulative distribution function (cdf) of the repeated observations from group h, for $h = 1, 2$. The Wei-Lachin statistic for testing H_0: $F_1(x_1, \ldots, x_t) = F_2(x_1, \ldots, x_t)$ against the general alternative that $F_1 \neq F_2$ is $X^2_W = W'\widehat{\Sigma}^{-1}_W W$, where $W' = (W_1, \ldots, W_t)$ is a vector of test statistics comparing groups 1 and 2 at each of the t time points, and $\widehat{\Sigma}_W$ is a consistent estimator of $\mathrm{Var}(W)$ given by Theorem 1 of [1]. Apart from a scale factor, the jth component of W equals

$$\sum_{i=1}^{n_1} \sum_{i'=1}^{n_2} \delta_{1ij}\, \delta_{2i'j}\, \phi(y_{1ij}, y_{2i'j}),$$

where

$$\phi(x, y) = \begin{cases} 1 & \text{if } x > y, \\ 0 & \text{if } x = y, \\ -1 & \text{if } x < y. \end{cases} \tag{1}$$

Thus, at each time point j, comparisons between groups 1 and 2 are made for all i, i' for which y_{1ij} and $y_{2i'j}$ are both observed. The asymptotic null distribution of X^2_W is chi-square with t degrees of freedom (χ^2_t).

In many studies, the detection of stochastic ordering of the distributions F_1 and F_2 is of primary interest. For example, the alternative hypothesis H_1 may be that $F_{1j}(x) \leq F_{2j}(x)$ for each pair (F_{1j}, F_{2j}) of marginal cdfs, $j = 1, \ldots, t$. Under this alternative, the observations from group 1 tend to be larger than those from group 2 at each time point. For this situation, Wei and Lachin [1] propose the statistic

$$z_W = \frac{e'T}{\sqrt{e'\widehat{\Sigma}_T e}},$$

where e' is the t-component vector $(1, \ldots, 1)$. The asymptotic null distribution of z_W is normal with mean 0 and variance 1 $[N(0,1)]$. If the alternative hypothesis is $F_1 \leq F_2$ $(F_1 \geq F_2)$, the null hypothesis is rejected when z_W is a large positive (negative) value.

Wei and Lachin [1] derive the preceding methodology based on a commonly used random-censorship model [24] and focus on an omnibus test of equality versus a general alternative. In contrast, Wei and Johnson [2] focus primarily on optimal methods of combining dependent tests and propose a class of two-sample nonparametric tests for incomplete repeated measures based on two-sample U-statistics. Their motivation is that if a researcher wishes to draw an overall conclusion regarding the superiority of one treatment over another (across time), then a univariate one-sided test that combines the results at individual time points is more appropriate than an omnibus two-sided test of $H_0\colon F_1(x_1, \ldots, x_t) = F_2(x_1, \ldots, x_t)$.

The Wei-Johnson [2] test statistic at the jth time point is

$$U_j = \frac{\sqrt{n}}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{i'=1}^{n_2} \delta_{1ij}\, \delta_{2i'j}\, \phi(y_{1ij}, y_{2i'j}),$$

where $\phi(x, y)$ is a "kernel" function. Under mild regularity conditions, the elements of the variance-covariance matrix $\widehat{\Sigma}_U$ of $U = (U_1, \ldots, U_t)'$ can be estimated consistently by

$$\widehat{\sigma}_{jk} = \frac{n}{n_1}\, \widehat{\sigma}_{1jk} + \frac{n}{n_2}\, \widehat{\sigma}_{2jk},$$

where

$$\widehat{\sigma}_{1jk} = \frac{1}{n_1 n_2 (n_2 - 1)} \sum_{i=1}^{n_1} \sum_{l \neq l'=1}^{n_2} \delta_{1ij}\, \delta_{1ik}\, \delta_{2lj}\, \delta_{2l'k}\, \phi(y_{1ij}, y_{2lj})\, \phi(y_{1ik}, y_{2l'k}),$$

$$\widehat{\sigma}_{2jk} = \frac{1}{n_2 n_1 (n_1 - 1)} \sum_{l=1}^{n_2} \sum_{i \neq i'=1}^{n_1} \delta_{1ij}\, \delta_{1i'k}\, \delta_{2lj}\, \delta_{2lk}\, \phi(y_{1ij}, y_{2lj})\, \phi(y_{1i'k}, y_{2lk}).$$

Since the null distribution of U is approximately multivariate normal with zero mean vector and variance-covariance matrix $\widehat{\Sigma}_U$, the hypothesis $H_0\colon F_1 = F_2$ can be tested against a general alternative using the statistic $X^2_U = U' \widehat{\Sigma}_U^{-1} U$, which is asymptotically χ^2_t. A univariate one-sided test that combines the results at individual time points can be based on the linear combination $c'U = \sum_{j=1}^{t} c_j U_j$, where $c' = (c_1, \ldots, c_t)$ is a vector of weights. Under H_0, the statistic

$$z_U = \frac{c'U}{\sqrt{c' \widehat{\Sigma}_U c}}$$

is asymptotically $N(0,1)$

Under the assumption that the test statistics at the individual time points are estimates of a common effect, the optimal weights are given by $c' = e' \widehat{\Sigma}_U^{-1}$, where e' is the t-component vector $(1, \ldots, 1)$ [25,26]. In practice, however, this assumption may not hold. In addition, Bloch and Moses [27] show that the use of simple weights often results in little loss of efficiency. The simplest choice for the vector c is to weight each component equally, i.e. $c' = (1, \ldots, 1)$. Another possibility is to weight by the reciprocals of the variances, i.e., $c' = (1/\widehat{\sigma}_{11}, \ldots, 1/\widehat{\sigma}_{tt})$. Note that if the values of the test statistics differ considerably across time points, the weights $c' = e' \widehat{\Sigma}_U^{-1}$ may give a result which is quite different from that using equal weights or weighting by precision.

Although Wei and Johnson [2] suggest several choices for the kernel function $\phi(x, y)$, we restrict consideration to the function given in equation (1). With this choice, the Wei-Johnson vector of test statistics U and the Wei-Lachin vector of test statistics W are equivalent, apart from a scale factor. The consistent estimators of the variances and covariances of the components of the vector of test statistics, however, are different. The two methods will usually give similar results.

3. Program description and usage

3.1. General description

The FORTRAN-77 program RMNP2 consists of a main program and 14 subprograms. First, four subroutines are called to determine the type of input and output desired (screen-directed or file-directed), to read in the analysis options and data, and to check for errors and inconsistencies in the input parameters. The analysis options are then written to the screen or output file, as desired. The main program then calls three major subroutines to carry out the required computations and to print the results. First, the Wei-Lachin and/or Wei-Johnson vectors of test statistics and their estimated covariance matrices are computed. The omnibus chi-square statistics X^2_W and X^2_U are then calculated and printed, as well as the linear combination statistics z_W and z_U for the three sets of weight vectors described in section 2:

equal	$c' = (1, \ldots, 1)$,
variance reciprocals	$c' = (1/\widehat{\sigma}_{11}, \ldots, 1/\widehat{\sigma}_{tt})$,
optimal	$c' = (1, \ldots, 1)\widehat{\Sigma}^{-1}$.

The remaining seven subprograms handle matrix operations and calculation of normal and chi-square probabilities.

The user can specify that the analysis be carried out using the Wei-Lachin method, the Wei-Johnson method, or both methods. The program output includes the vector of test statistics comparing the two groups at each time point (W, U), the covariance ($\widehat{\Sigma}_W$, $\widehat{\Sigma}_U$) and correlation matrices of the statistics, and the vector of standardized test statistics (statistic/standard error). The raw data may optionally be listed. The results from the two-sided omnibus test of the null hypothesis of no difference between the two groups are printed (X^2_W and/or X^2_U, degrees of freedom, chi-square p-value), as well as the three one-sided tests of stochastic ordering (z_W and/or z_U, $N(0,1)$ p-value).

3.2. Structure of the input data file

The input data file must be a standard text (ASCII) file with no hidden characters or word processing format codes. This file should contain one line per subject (independent experimental unit). The data for each subject must include a group identifier and the values of the outcome variable at each of the multiple time points or measurement conditions. Although each record may contain other variables not to be used in the analysis, such as subject identifier, demographic data, etc., the program is restricted to reading in at most 100 variables per subject. Note that the group identifier and the repeated measurements are not required to be in specific fields.

All data items are read in as real numbers using free format input. Thus, fields must be separated by one or more blanks. Data values less than a user-specified missing data code are interpreted as missing values.

The current version of the program is restricted to at most 125 subjects in each of the two treatment groups. In addition, analysis is restricted to a maximum of eight repeated measurements. Although these limitations are necessary for execution on an IBM-compatible personal computer using the DOS operating system, the source code can be easily changed for use on other systems. Also, reducing the maximum number of repeated measurements permits increasing the maximum number of subjects, and vice-versa.

3.3. Specification of analysis options

The RMNP2 options allow the user to determine input/output modes and to specify the missing value indicator, the number of variables to be read in, the indices of the group identifier and repeated measurements variables, the values identifying the two groups of subjects, and the test statistics to be computed. These analysis options are listed in Table 1. They can be specified in one of three ways, depending on how the program is invoked.

`[INSERT TABLE 1 ABOUT HERE]`

When RMNP2 is invoked, the user is asked:

```
ARE INPUT PARAMETERS INCLUDED AT THE BEGINNING OF YOUR DATA FILE?
(1=YES, 2=NO)
```

The simplest method of specifying the desired options is to respond with a '2' and then enter the options interactively in response to program prompts. When options are specified in this manner, the program checks for validity and consistency with the values of previously-entered options. If errors are detected, the user is prompted to reenter the option.

It is sometimes more convenient to include the analysis options at the beginning of the input data file. In this case, options B, C, D, F, H, and I should be specified on line 1 of the input file, option E should be specified on line 2, and option G on line 3.

On MS-DOS compatible microcomputers and UNIX workstations, it is also possible to include the values of the options in a separate control file. This file must contain one line for each of the 11 options listed in Table 1 and the first line (option A) must contain the value 2. If this file is called `ANALYSIS.CTL`, the command `RMNP2<ANALYSIS.CTL` is then used to invoke the program. If the analysis options are included at the beginning of the data file or in a separate control file, the values of invalid and/or inconsistent options are printed and execution of the program is terminated. Examples of all three types of option specifications are given in the next section.

4. Examples

4.1. Comparison of two treatments for maternal pain relief during labor and delivery

Eighty-three women in labor were randomized to receive an experimental pain medication ($n_1 = 43$) or placebo ($n_2 = 40$). Treatment was initiated when the cervical dilation was 8 cm. At 30 minute intervals, the amount of pain was self-reported by placing a mark on a 100 mm line (0=no pain, 100=very much pain). Data were collected at baseline and at 30, 60, 90, 120, 150, and 180 minutes following treatment.

The response variable (pain severity) is essentially continuous, since the length of the line to the left of the mark was measured to the nearest 0.5 mm. However, the pain severity scores are very non-normal. At early time points, the distribution is skewed to the left and there are many zeros. The marginal distributions tend to be U-shaped at some of the later time points. Thus, standard parametric analysis methods seem inappropriate. In addition, there are numerous missing values at later measurement times.

The data file PAIN.DAT contains the group indicator (1=experimental treatment, 2=placebo), a subject identifier, and pain severity measures at minutes 0, 30, 60, 90, 120, 150, and 180. Since the smallest nonmissing value is 0.0, missing measurements are coded as -1.0. Table 2 displays the data from the first 10 subjects in group 1; the entire data file is listed in Appendix III of [28].

`[INSERT TABLE 2 ABOUT HERE]`

We will compare the two treatment groups using both the Wei-Lachin and Wei-Johnson methods at minutes 30, 60, 90, 120, 150, and 180. The analysis options will be specified interactively and the results will be written to the file PAIN.OUT. The program is invoked by typing RMNP2. The session log and the results of the analysis are displayed in Tables 3 and 4, respectively. In the session log, statements in capital letters denote prompts which appear on the screen; these are followed by user responses (numbers and lower case letters).

`[INSERT TABLES 3 AND 4 ABOUT HERE]`

For both methods, the signs of the test statistics indicate that, at each time point, the pain scores are lower (better) in the treated group than in the placebo group. Although the two methods yield similar conclusions, the Wei-Lachin standardized statistic is larger in magnitude (more significant) than the Wei-Johnson statistic at every time point. The omnibus Wei-Lachin statistic is highly significant, while the omnibus Wei-Johnson statistic is marginally significant ($p = .065$). The one-sided tests based on linear combinations of the individual statistics also indicate that the two treatments are significantly different.

The preceding analysis can also be carried out by including the options in a separate file. For example, given the file PAIN.CTL listed in Table 5, the statement RMNP2<PAIN.CTL produces the same results as shown in Table 4. (Note that the descriptive comments to the right of each line in Table 5 are not part of the file PAIN.CTL.)

`[INSERT TABLE 5 ABOUT HERE]`

4.2. Effect of chenodiol on cholesterol levels in the National Cooperative Gallstone Study

The National Cooperative Gallstone Study studied the safety of the drug chenodiol in the treatment of cholesterol gallstones. This drug dissolves gallstones by altering the metabolic pathway of cholesterol to reduce cholesterol secretion into gallbladder bile. However, it was thought that it also might increase serum cholesterol, a known risk factor for atherosclerotic disease.

In a group of 112 patients with floating gallstones, 64 patients received 750 mg/day of chenodiol and 48 patients received placebo. Serum cholesterol was measured in these patients prior to treatment and at 6, 12, 20, and 24 months of follow-up. Many cholesterol measurements were missing because patient follow-up was terminated, visits were missed, or laboratory specimens were lost or inadequate. The two groups have rather different missing value patterns, mainly because of the termination of follow-up for different reasons. We wish to compare the chenodiol and placebo groups with respect to the increases in cholesterol levels (from baseline) at months 6, 12, 20, and 24.

The analysis options and the data are contained in the file CHOLEST.DAT, which contains 115 records. Table 6 lists the first 13 records from this file (and includes the data from the first 10 subjects in the chenodiol group); the complete data set is listed in Table 2 of [1]. With the exception of the first three records, which contain the analysis options, each record has five variables: group identifier (1=chenodiol, 2=placebo) and the cholesterol increases at months 6, 12, 20, and 24. Since the cholesterol changes may be either positive or negative, missing values are denoted in the data file by the value -999.

[INSERT TABLE 6 ABOUT HERE]

The first record contains the missing value indicator, the total number of variables to be read in, the index of the group identifier, the number of time points at which test statistics are to be calculated, and the indicators for test statistic calculation and raw data listing (options B, C, D, F, H, and I). The options specified indicate that values less than -998 are missing, that five variables will be read in, that the first variable is the group identifier, that four time points will be used in the analysis, that both methods will be used, and that the raw data will not be listed. The second line gives the group identifiers and the third line specifies the indices of the measurement times for which test statistics are to be calculated.

The session log and the results of the analysis are displayed in Tables 7 and 8, respectively. At each time point, the cholesterol increases are greater in the chenodiol group than in the placebo group. For both methods, the p-value from the omnibus statistic is greater than 0.1, while the linear combination with equal weights is significant at the 0.05 level. However, the two other linear combinations give more weight to the last two time points, where the difference between the two groups is diminished. In particular, the optimal set of weights under the assumption that the individual test statistics are estimates of a common effect yields nonsignificant p-values. The substantial differences among the three linear combinations result from the fact that the values of the test statistics differ across time points.

[INSERT TABLES 7 AND 8 ABOUT HERE]

5. Hardware and software specifications

RMNP2 is written in standard FORTRAN-77. It was originally developed for MS-DOS personal computers and compiled using the MICROSOFT (R) FORTRAN Optimizing Compiler Version 5.0. The program has also been compiled and executed on HP APOLLO Series 700 workstations using the HP-UX FORTRAN 77 compiler; no modifications of any kind were required. Intermediate test statistic and Wei-Johnson covariance calculations use integer arithmetic, with the results stored as single precision real numbers; the Wei-Lachin covariance calculations are carried out using double precision arithmetic. The following algorithms from Griffiths and Hill [29] are used:

SYMINV	inversion of a positive semi-definite symmetric matrix;
CHOL	triangular decomposition of a symmetric matrix;
ALNORM	calculation of tail areas of the standard normal distribution.

In addition, algorithms of MacLeod [30] and Shea [31] are used to compute the natural logarithm of the gamma function and the incomplete gamma integral, respectively.

Although general-purpose programs for the nonparametric analysis of incomplete repeated measures have not been made available previously, Makuch, Escobar, and Merrill [32] provide a FORTRAN subroutine for computing the Wei-Lachin omnibus statistic X^2_W and linear combination statistic z_W using the weight vector $c' = (1, \ldots, 1)$. While their interest was in the analysis of multivariate censored failure time data, they provide instructions for adapting their algorithm to the general repeated measures setting. Apart from a sign change, the two algorithms give the same results.

6. References

[1] L.J. Wei and J.M. Lachin, Two-sample asymptotically distribution-free tests for incomplete multivariate observations, Journal of the American Statistical Association, 79 (1984) 653–661.

[2] L.J. Wei and W.E. Johnson, Combining dependent tests with incomplete repeated measurements, Biometrika, 72 (1985) 359–364.

[3] M.D. Schluchter, 5V: Unbalanced repeated measures models with structured covariance matrices, *in:* BMDP Statistical Software Manual: Volume 1, eds. W.J. Dixon, M.B. Brown, L. Engelman, and R.I. Jennrich, pp. 1207–1244 (University of California Press, Berkeley CA, 1990).

[4] SAS Institute Inc., The MIXED procedure, *in:* SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07, pp. 287–366 (SAS Institute Inc., Cary NC, 1992).

[5] G.G. Koch, J.R. Landis, J.L. Freeman, D.H. Freeman, and R.G. Lehnen, A general methodology for the analysis of experiments with repeated measurement of categorical data, Biometrics, 33 (1977) 133–158.

[6] W.M. Stanish, D.B. Gillings, and G.G. Koch, An application of multivariate ratio methods for the analysis of a longitudinal clinical trial with missing data, Biometrics, 34 (1978) 305–317.

[7] R.F. Woolson and W.R. Clarke, Analysis of categorical incomplete longitudinal data, Journal of the Royal Statistical Society, Series A, 147 (1984) 87–99.

[8] J.R. Landis, M.E. Miller, C.S. Davis, and G.G. Koch, Some general methods for the analysis of categorical data in longitudinal studies, Statistics in Medicine, 7 (1988) 109–137.

[9] SAS Institute Inc., The CATMOD procedure, *in:* SAS/STAT Guide for Personal Computers, Version 6 Edition, pp. 189–282 (SAS Institute Inc., Cary NC, 1987).

[10] C.S. Davis, Analysis of incomplete categorical repeated measures, *in:* Proceedings of 17th Annual SAS User's Group International Conference, pp. 1374–1379 (SAS Institute Inc., Cary NC, 1992).

[11] P. McCullagh and J.A. Nelder, Generalized Linear Models, Second Edition (Chapman and Hall, London UK, 1989).

[12] K.Y. Liang and S.L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986) 13–22.

[13] S.L. Zeger and K.Y. Liang, Longitudinal data analysis for discrete and continuous outcomes, Biometrics 42 (1986) 121–130.

[14] L.J. Wei and D.O. Stram, Analyzing repeated measurements with possibly missing observations by modelling marginal distributions, Statistics in Medicine 7 (1988) 139–148.

[15] D.O. Stram, L.J. Wei, and J.H. Ware, Analysis of repeated ordered categorical outcomes with possibly missing observations and time-dependent covariates, Journal of the American Statistical Association 83 (1988) 631–637.

[16] L.H. Moulton and S.L. Zeger, Analyzing repeated measures on generalized linear models via the bootstrap, Biometrics 45 (1989) 381–394.

[17] L.P. Zhao and R.L. Prentice, Correlated binary regression using a quadratic exponential model, Biometrika 77 (1990) 642–648.

[18] M.R. Karim and S.L. Zeger, GEE: a SAS macro for longitudinal data analysis, Technical Report No. 674, Department of Biostatistics, The Johns Hopkins University (1988).

[19] S.R. Lipsitz and D.P. Harrington, Analyzing correlated binary data using SAS, Computers and Biomedical Research 23 (1990) 268–282.

[20] C.S. Davis, A computer program for regression analysis of repeated measures using generalized estimating equations, Computer Methods and Programs in Biomedicine 40 (1993) in press.

[21] M.L. Puri and P.K. Sen, Nonparametric Methods in Multivariate Analysis (John Wiley and Sons, New York NY, 1971).

[22] G.G. Koch, I.A. Amara, M.E. Stokes, and D.B. Gillings, Some views on non-parametric analysis for repeated measurements and selected bibliography, International Statistical Review 48 (1980) 249–265.

[23] D.B. Rubin, Inference and missing data, Biometrika 63 (1976) 81–92.

[24] J.D. Kalbfleisch and R.L. Prentice, The Statistical Analysis of Failure time Data (John Wiley and Sons, New York NY, 1980).

[25] P.C. O'Brien, Procedures for comparing samples with multiple endpoints, Biometrics 40 (1984) 1079–1087.

[26] D. Ashby, S.J. Pocock, and A.G. Shaper, Ordered polytomous regression: an example relating serum biochemistry and haematology to alcohol consumption, Applied Statistics 35 (1986) 289–301.

[27] D.A. Bloch and L.E. Moses, Nonoptimally weighted least squares, The American Statistician 42 (1988) 50–53.

[28] C.S. Davis, Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials, Statistics in Medicine 10 (1991) 1959–1980.

[29] P. Griffiths and I.D. Hill, Applied Statistics Algorithms (Ellis Horwood, Chichester UK, 1985).

[30] A.J. MacLeod, Algorithm AS 245: a robust and reliable algorithm for the logarithm of the gamma function, Applied Statistics 38 (1989) 397–402.

[31] B.L. Shea, Algorithm AS 239: chi-squared and incomplete gamma integral, Applied Statistics 37 (1988) 466–473.

[32] R.W. Makuch, M. Escobar, and S. Merrill III, Algorithm AS 262: a two-sample test for incomplete multivariate data, Applied Statistics 40 (1991) 202–212.

Table 1
Analysis options

Option	Description
A	Method of specifying input options: 1=included at the beginning of the input file, 2=specified interactively or included in a separate file
B	Missing value indicator (data items less than this value are interpreted as missing values)
C	Total number of variables to be read in (at most 100)
D	Index of the group identifier; permissible values are $1, \ldots, C$
E	The two possible values of the group identifier variable D; these define the two groups of subjects
F	Number of time points at which the test statistics are to be calculated; permissible values are $1, \ldots, \min(C-1, 8)$
G	Indices of variables (time points) at which test statistics are to be calculated; the F values specified must be in ascending order and each must be in the range $1, \ldots, C$ and must not equal D
H	Test statistics to be computed: 1=Wei-Lachin only, 2=Wei-Johnson only, 3=Wei-Lachin and Wei-Johnson
I	Code for optional raw data listing: 1=list the data for the time points at which test statistics are to be calculated, 2=do not list the raw data.
J	Name of the input data file (may include a path name under MS-DOS and UNIX)
K	Name of the output data file (may include a path name under MS-DOS and UNIX); enter * if results are to be written to the screen instead of to a file

Table 2
Data from example 4.1: first ten subjects in group 1

```
1  1   0.0   0.0   0.0   0.0  -1.0  -1.0  -1.0
1  2   0.0   0.0   0.0   0.0   2.5   2.3  14.0
1  3  38.0   5.0   1.0   1.0   0.0   5.0  -1.0
1  4   6.0  48.0  85.0   0.0   0.0  -1.0  -1.0
1  5  19.0   5.0  -1.0  -1.0  -1.0  -1.0  -1.0
1  6   7.0   0.0   0.0   0.0  -1.0  -1.0  -1.0
1  7  44.0  42.0  42.0  45.0  -1.0  -1.0  -1.0
1  8   1.0   0.0   0.0   0.0   0.0   6.0  24.0
1  9  24.5  35.0  13.0  -1.0  -1.0  -1.0  -1.0
1 10   1.0  30.5  81.5  67.5  98.5  97.0  -1.0
```

Table 3
Session log from example 4.1

```
LONGITUDINAL DATA ANALYSIS USING NONPARAMETRIC METHODS FOR
INCOMPLETE REPEATED MEASUREMENTS FROM TWO SAMPLES
WEI AND LACHIN (1984, JASA), WEI AND JOHNSON (1985, BIOMETRIKA)

ARE INPUT PARAMETERS INCLUDED AT THE BEGINNING OF YOUR DATA FILE?
(1=YES, 2=NO)
2

ENTER MISSING VALUE INDICATOR
(VALUES LESS THAN THIS VALUE ARE MISSING)
-0.5

ENTER THE TOTAL NUMBER OF VARIABLES TO BE READ IN (MAXIMUM OF 100)
9

ENTER THE INDEX OF THE GROUP IDENTIFIER (MUST BE IN THE RANGE 1- 9)
1

ENTER THE GROUP IDENTIFICATION CODES
1 2

ENTER THE NUMBER OF TIME POINTS AT WHICH THE TEST STATISTICS
ARE TO BE CALCULATED (MAXIMUM OF 8)
6

ENTER THE INDICES OF THE TIME POINTS TO BE USED
4 5 6 7 8 9

INDICATE WHICH TEST STATISTICS ARE TO BE COMPUTED
(1=WEI-LACHIN, 2=WEI-JOHNSON, 3=BOTH)
3

DO YOU WISH TO LIST THE RAW DATA? (1=YES, 2=NO)
2

ENTER THE NAME OF THE INPUT DATA FILE
pain.dat

ENTER THE NAME OF THE OUTPUT DATA FILE
(* IF RESULTS ARE TO BE WRITTEN TO THE SCREEN)
pain.out
```

Table 4
Output from example 4.1

```
LONGITUDINAL DATA ANALYSIS USING NONPARAMETRIC METHODS FOR
INCOMPLETE REPEATED MEASUREMENTS FROM TWO SAMPLES
WEI AND LACHIN (1984, JASA), WEI AND JOHNSON (1985, BIOMETRIKA)

MISSING VALUE INDICATOR:    -.50
TOTAL NUMBER OF VARIABLES:  9
INDEX OF THE GROUP IDENTIFIER:  1
GROUP IDENTIFIERS:  1  2
NUMBER OF TIME POINTS:  6
INDICES OF TIME POINTS:  4  5  6  7  8  9
NUMBER OF SUBJECTS IN GROUP 1:  43
NUMBER OF SUBJECTS IN GROUP 2:  40
```

Table 4 *(continued)*

```
WEI-LACHIN ANALYSIS:

VECTOR OF TEST STATISTICS
-.39409E+00 -.60172E+00 -.75513E+00 -.72868E+00 -.49725E+00 -.29755E+00

ESTIMATED COVARIANCE MATRIX OF VECTOR OF TEST STATISTICS
 .79355E-01  .47911E-01  .28357E-01  .17829E-01  .11357E-01  .57492E-02
 .47911E-01  .58464E-01  .31605E-01  .20765E-01  .15458E-01  .64448E-02
 .28357E-01  .31605E-01  .36803E-01  .19749E-01  .11149E-01  .36276E-02
 .17829E-01  .20765E-01  .19749E-01  .26539E-01  .14794E-01  .53505E-02
 .11357E-01  .15458E-01  .11149E-01  .14794E-01  .13187E-01  .56908E-02
 .57492E-02  .64448E-02  .36276E-02  .53505E-02  .56908E-02  .52491E-02

STANDARDIZED VECTOR OF TEST STATISTICS (ESTIMATE/S.E.)
   -1.39898    -2.48856    -3.93623    -4.47297    -4.33010    -4.10698

CORRELATION MATRIX
    1.00000      .70339      .52472      .38851      .35109      .28169
     .70339     1.00000      .68135      .52716      .55670      .36789
     .52472      .68135     1.00000      .63194      .50609      .26100
     .38851      .52716      .63194     1.00000      .79081      .45333
     .35109      .55670      .50609      .79081     1.00000      .68400
     .28169      .36789      .26100      .45333      .68400     1.00000

LINEAR COMBINATIONS OF TEST STATISTICS

          WEIGHTS                STATISTIC      VARIANCE        Z         P
(1,1,...,1)                     -.3274E+01     .7113E+00    -3.883      .000
RECIPROCALS OF THE VARIANCES    -.1576E+03     .1055E+04    -4.853      .000
(1,1,...,1) X SIGMA INVERSE     -.6293E+02     .2032E+03    -4.415      .000

OMNIBUS CHI-SQUARE TEST STATISTIC =   30.098     DF=6     P=  .000

WEI-JOHNSON ANALYSIS:

VECTOR OF TEST STATISTICS
-.15784E+01 -.24100E+01 -.30245E+01 -.29185E+01 -.19916E+01 -.11918E+01

ESTIMATED COVARIANCE MATRIX OF VECTOR OF TEST STATISTICS
 .13298E+01  .92683E+00  .65567E+00  .41822E+00  .24287E+00  .14334E+00
 .92683E+00  .11200E+01  .77826E+00  .55765E+00  .36251E+00  .21144E+00
 .65567E+00  .77826E+00  .93373E+00  .75114E+00  .49850E+00  .25548E+00
 .41822E+00  .55765E+00  .75114E+00  .77904E+00  .50155E+00  .25277E+00
 .24287E+00  .36251E+00  .49850E+00  .50155E+00  .41888E+00  .22344E+00
 .14334E+00  .21144E+00  .25548E+00  .25277E+00  .22344E+00  .18191E+00

STANDARDIZED VECTOR OF TEST STATISTICS (ESTIMATE/S.E.)
   -1.36876    -2.27724    -3.12995    -3.30661    -3.07717    -2.79427

CORRELATION MATRIX
    1.00000      .75943      .58841      .41089      .32541      .29143
     .75943     1.00000      .76103      .59700      .52925      .46844
     .58841      .76103     1.00000      .88071      .79709      .61989
     .41089      .59700      .88071     1.00000      .87799      .67147
     .32541      .52925      .79709      .87799     1.00000      .80944
     .29143      .46844      .61989      .67147      .80944     1.00000

LINEAR COMBINATIONS OF TEST STATISTICS

          WEIGHTS                STATISTIC      VARIANCE        Z         P
(1,1,...,1)                     -.1311E+02     .1832E+02    -3.064      .001
RECIPROCALS OF THE VARIANCES    -.2163E+02     .4357E+02    -3.277      .001
(1,1,...,1) X SIGMA INVERSE     -.5229E+01     .6163E+01    -2.106      .018

OMNIBUS CHI-SQUARE TEST STATISTIC =   11.864     DF=6     P=  .065
```

Table 5
Analysis options file for example 4.1

2	Analysis options are not part of the data file
-0.5	Missing value indicator
9	Total number of variables to be read in
1	Index of the group identifier
1 2	Values of the group identifier variable
6	Number of time points at which test statistics are to be calculated
4 5 6 7 8 9	Indices of variables (time points) at which test statistics are to be calculated
3	Test statistics to be computed (3=Wei-Lachin and Wei-Johnson)
2	Code for optional raw data listing (2=do not list raw data)
pain.dat	Name of input data file
pain.out	Name of output data file

Table 6
Data from example 4.2 (first 13 lines)

```
-998 5 1 4 3 2
1 2
2 3 4 5
1   68  117   50   96
1    6   24   -9   86
1   47   30   35  107
1   49   22   41  101
1   27   60   37   25
1   31   -9   18   11
1   24  -20  -24  -62
1   13   36   37    5
1   25  -36   10  -32
1  -29    9   30   -8
```

Table 7
Session log from example 4.2

```
LONGITUDINAL DATA ANALYSIS USING NONPARAMETRIC METHODS FOR
INCOMPLETE REPEATED MEASUREMENTS FROM TWO SAMPLES
WEI AND LACHIN (1984, JASA), WEI AND JOHNSON (1985, BIOMETRIKA)

ARE INPUT PARAMETERS INCLUDED AT THE BEGINNING OF YOUR DATA FILE?
(1=YES, 2=NO)
1

ENTER THE NAME OF THE INPUT DATA FILE
cholest.dat

ENTER THE NAME OF THE OUTPUT DATA FILE
(* IF RESULTS ARE TO BE WRITTEN TO THE SCREEN)
cholest.out
```

Table 8
Output from example 4.2

```
MISSING VALUE INDICATOR: -998.00
TOTAL NUMBER OF VARIABLES:  5
INDEX OF THE GROUP IDENTIFIER:  1
GROUP IDENTIFIERS:  1  2
NUMBER OF TIME POINTS:  4
INDICES OF TIME POINTS:  2  3  4  5
NUMBER OF SUBJECTS IN GROUP 1:  64
NUMBER OF SUBJECTS IN GROUP 2:  48

WEI-LACHIN ANALYSIS:

VECTOR OF TEST STATISTICS
 .51211E+00  .53911E+00  .19742E+00  .55682E-01

ESTIMATED COVARIANCE MATRIX OF VECTOR OF TEST STATISTICS
 .65719E-01  .33619E-01  .22034E-01  .15108E-01
 .33619E-01  .49803E-01  .20568E-01  .13778E-01
 .22034E-01  .20568E-01  .28929E-01  .10351E-01
 .15108E-01  .13778E-01  .10351E-01  .18367E-01

STANDARDIZED VECTOR OF TEST STATISTICS (ESTIMATE/S.E.)
    1.99764     2.41572     1.16070      .41087

CORRELATION MATRIX
    1.00000      .58764      .50534      .43484
     .58764     1.00000      .54187      .45556
     .50534      .54187     1.00000      .44905
     .43484      .45556      .44905     1.00000

LINEAR COMBINATIONS OF TEST STATISTICS

          WEIGHTS                STATISTIC      VARIANCE        Z         P
(1,1,...,1)                      .1304E+01     .3937E+00     2.079      .019
RECIPROCALS OF THE VARIANCES     .2847E+02     .2907E+03     1.670      .047
(1,1,...,1) X SIGMA INVERSE      .5879E+01     .6287E+02      .741      .229

OMNIBUS CHI-SQUARE TEST STATISTIC =    7.313     DF=4     P=  .120

WEI-JOHNSON ANALYSIS:

VECTOR OF TEST STATISTICS
 .20911E+01  .22013E+01  .80613E+00  .22737E+00

ESTIMATED COVARIANCE MATRIX OF VECTOR OF TEST STATISTICS
 .10949E+01  .63868E+00  .41113E+00  .23832E+00
 .63868E+00  .86740E+00  .37388E+00  .22578E+00
 .41113E+00  .37388E+00  .48645E+00  .16674E+00
 .23832E+00  .22578E+00  .16674E+00  .29779E+00

STANDARDIZED VECTOR OF TEST STATISTICS (ESTIMATE/S.E.)
    1.99839     2.36363     1.15580      .41666

CORRELATION MATRIX
    1.00000      .65536      .56333      .41736
     .65536     1.00000      .57558      .44425
     .56333      .57558     1.00000      .43810
     .41736      .44425      .43810     1.00000

LINEAR COMBINATIONS OF TEST STATISTICS

          WEIGHTS                STATISTIC      VARIANCE        Z         P
(1,1,...,1)                      .5326E+01     .6856E+01     2.034      .021
RECIPROCALS OF THE VARIANCES     .6868E+01     .1765E+02     1.635      .051
(1,1,...,1) X SIGMA INVERSE      .1321E+01     .3868E+01      .672      .251

OMNIBUS CHI-SQUARE TEST STATISTIC =    6.715     DF=4     P=  .152
```