

Extension of the Carter-Yang polynomial growth curve model to allow unique times of measurement for subjects

Elizabeth A. Mauger^a, Charles J. Kowalski^b, Emet D. Schneiderman*^c,
Stephen M. Willis^c

^aDepartment of Biostatistics,

^bDepartment of Biologic and Materials Sciences, The University of Michigan, Ann Arbor, MI 48109, USA

^cDepartment of Oral and Maxillofacial Surgery and Pharmacology, Baylor College of Dentistry, PO Box 660677, Dallas, TX 75266-0677, USA

Received 16 November 1993; accepted 20 December 1993

Abstract

A PC program extending the procedure due to Carter and Yang (*Commun Stat: Theory Methods*, 8 (1986) 2507–2526) to allow unique times of measurement for subjects is described, illustrated and made available. Given longitudinal observations on each of N subjects comprising a single group, this program determines the lowest degree polynomial in time adequate to fit the average growth curve (AGC); estimates this curve and provides confidence bands for the AGC, and confidence intervals for the corresponding polynomial regression coefficients; and so-called prediction intervals which, with a given level of confidence, will contain the growth curve of a 'new' subject from the same population of which the N subjects constitute a random sample. Two kinds of missing data are accommodated. First, in the context of studies planned so that subjects will be measured at identical times and, second, in unstructured studies where subjects may present with their own, unique times of measurement.

Keywords: Longitudinal data; Missing values; Unbalanced designs; PC program

1. Introduction

A longitudinal study which is planned so that subjects will be measured at the same set of time points is said to be balanced (or to have a balanced design). If there are no missing data, the resulting data set is said to be complete. In an earlier paper [1], we described, illustrated, and made available a menu-driven PC program implementing Rao's [2]

so-called two-stage, or random coefficients, polynomial growth curve model. Given a data set which is both balanced and complete, this program can be used to (a) determine the degree of the polynomial adequate to fit the average growth curve (AGC) of the sample, and (b) estimate and compute confidence bands for the AGC. More recently [3], we implemented the Carter-Yang [4] extension of this procedure which accommodates missing data, but still requires a balanced design. This program produces output similar to that in

* Corresponding author.

[1], with the addition of (c) prediction intervals which, with a given level of confidence, will contain the growth curve for a 'new' subject from the same population as the sample used in the analysis. In this paper, we describe and illustrate a program extending [3] to unbalanced designs: subjects may have their own, unique times of measurement.

2. The method

The method is based on Carter and Yang [4]. In our notation [1,3], we assume that the observations at the T_i times of measurement for the i th subject ($i = 1, 2, \dots, N$) have the structure

$$x_i \sim MVN(W_i\tau, W_i\Delta W_i' + \sigma^2 I_{T_i}) \tag{1}$$

i.e., x_i has a multivariate normal distribution with mean vector $W_i\tau$ and covariance matrix $W_i\Delta W_i' + \sigma^2 I_{T_i}$. In (1), the time design matrix, W_i , consists either of powers of the times of measurement t_1, t_2, \dots, t_{T_i} for the i th subject, or the values of orthogonal or orthonormal polynomials [5,6]; I_{T_i} is the $T_i \times T_i$ identity matrix; and τ is the $P \times 1$ vector of regression coefficients for the AGC. The regression coefficients specific to subject i , τ_i , are assumed to have the distribution

$$\tau_i \sim MVN(\tau, \Lambda) \tag{2}$$

and σ^2 measures the variability of subjects about their individual growth curves (error variance). The subscripts on W and T emphasize the fact that subjects need not be measured at the same times nor have the same total number of observations.

To illustrate the structure of the model, suppose the first of 2 subjects had observations 25.2, 29.0, 33.6, and 35.8 at times 1, 2, 3 and 5; and the second observations 27.3, 32.1 and 41.8 at times 1.2, 2.3 and 3.8. These correspond to the first 2 subjects from the data set considered in [1,3] in which $T = 5$ measurements, $t = 1(1)5$, were taken on each of $N = 12$ subjects; here we have discarded some of the data, and altered the times of measurement for the second subject. We have $T_1 = 4$, $T_2 = 3$; and

$$x_1 = \begin{bmatrix} 25.2 \\ 29.0 \\ 33.6 \\ 35.8 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 27.3 \\ 32.1 \\ 41.8 \end{bmatrix}$$

Then, if polynomials of degree $D = 2$ are fit to these subjects, and the successive-powers-of- t form of the time design matrix is used,

$$W_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 5 & 25 \end{bmatrix} \quad \text{and} \quad W_2 = \begin{bmatrix} 1 & 1.2 & 1.2^2 \\ 1 & 2.3 & 2.3^2 \\ 1 & 3.8 & 3.8^2 \end{bmatrix}$$

3. Estimation

The parameters of the model are estimated by (all summations are from $i = 1$ to N)

$$\hat{\tau}_i = (W_i' W_i)^{-1} W_i' x_i \tag{3}$$

$$\hat{\sigma}^2 = \frac{\sum (T_i - P) s_i}{\sum (T_i - P)} \tag{4}$$

and

$$\hat{\Lambda} = S_\tau - \frac{\hat{\sigma}^2}{N} \sum (W_i' W_i)^{-1} \tag{5}$$

or

$$\hat{\Lambda} = S_\tau - \frac{(\hat{\gamma} - 1/a)}{N} \sum (W_i' W_i) \tag{6}$$

according to whether $\hat{\gamma} > \hat{\sigma}^2$ or $\hat{\gamma} \leq \hat{\sigma}^2$, respectively, where $\hat{\gamma}$ is the smallest root of the determinantal equation

$$\left| S_\tau - \frac{\hat{\gamma}}{N} \sum (W_i' W_i)^{-1} \right| = 0 \tag{7}$$

and

$$a = \Sigma T_i + \text{trace} \left[\sum (W_i' W_i)^{-1} \right] \tag{8}$$

$$s_i = \frac{1}{T_i - P} x_i' [I_{T_i} - W_i(W_i' W_i)^{-1} W_i'] x_i \tag{9}$$

and

$$S_\tau = \frac{1}{N-1} \sum (\hat{\tau}_i - \hat{\tau})(\hat{\tau}_i - \hat{\tau})' \quad (10)$$

It will be noted that, in the above, P is fixed (a $D = P - 1$ degree polynomial is fit to each subject's curve), and that S_τ is the sample covariance matrix of the $\hat{\tau}_i$.

The coefficients of the AGC are then estimated by

$$\hat{\tau} = \left(\frac{1}{N} \sum A_i \right)^{-1} \left(\frac{1}{N} \sum A_i \hat{\tau}_i \right) \quad (11)$$

where

$$A_i = [Var(\hat{\tau}_i)]^{-1} = [\hat{\Lambda} + \hat{\sigma}^2(W_i'W_i)^{-1}]^{-1} \quad (12)$$

4. Tests of hypotheses

Hypotheses of the form $H: L\tau = \tau_0$, where L is a specified $L \times P$ matrix and τ_0 is a vector of known constants, can be tested using the statistic

$$T^2 = N(\hat{\tau} - \tau)'L'(L\hat{\Omega}L')^{-1}L(\hat{\tau} - \tau) \quad (13)$$

where

$$\hat{\Omega} = \left(\frac{1}{N} \sum A_i \right)^{-1} \quad (14)$$

When $\min\{T_i\}$ is large, and H is true, (13) has Hotelling's T^2 distribution with $N - 1$ degrees of freedom, so that

$$F^* = \frac{N-L}{L(N-1)} T^2 \sim F(L, N-1) \quad (15)$$

where $F(L, N - 1)$ is the F -distribution with L and $N - 1$ degrees of freedom.

Alternatively, if N is large,

$$T^2 \sim X^2(L) \quad (16)$$

where $X^2(L)$ represents the chi-squared distribution with L degrees of freedom.

In our program, tests for specification (i.e., for the adequacy of a given value of D) are based on

these results. We begin by asking the user to specify the degrees, FD and RD , of the so-called full and reduced models, respectively [1,3]. $FD < T_{\min} - 1$ is a degree which the user feels should be more than adequate to model the AGC, and RD the smallest degree which may be contemplated (often the user will choose $RD = 1$). We then, using the level of significance supplied by the user, perform a series of step-up goodness-of-fit tests, testing first that RD is adequate then, if necessary, that $RD + 1$ is adequate, etc., up to degree FD which is used if no smaller D passes the test for goodness-of-fit.

If, e.g., the user specifies $RD = 1$ and $FD = 3$ in the previous example, τ is estimated by

$$\hat{\tau} = \begin{bmatrix} \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\tau}_3 \\ \hat{\tau}_4 \end{bmatrix}$$

and the test for linearity corresponds to the hypothesis $H: \tau_3 = \tau_4 = 0$. This is of the form $H: L\tau = \tau_0$ with

$$L = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \tau_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In our program we compute and print the P -values corresponding to both (15) and (16). Following [4], who choose to be conservative, the P -value corresponding to (15), the F -test, is compared to the user-specified level of significance to determine whether a linear equation is adequate, i.e., if one accepts $H: \tau_3 = \tau_4 = 0$.

5. Confidence intervals and bands

We also compute confidence intervals for the elements of τ , and confidence bands for the AGC either at the planned times of measurement or at times specified by the user when times of measurement differ among individuals. The user specifies the level of confidence to be used (e.g., 0.95). Let w be a $1 \times P$ vector. Then approximate (for large values of N or $\min\{T_i\}$), $(1 - \alpha) \times 100\%$ confidence intervals for the elements of τ are given by

$$\mathbf{w}\hat{t} \pm \left(\frac{1}{N} \mathbf{w}\Omega\mathbf{w}'\right)^{1/2} t_{1-\alpha/2}(N-1) \quad (17)$$

by taking, in turn,

$$\mathbf{w}_1 = [1,0,0,\dots,0,0]$$

$$\mathbf{w}_2 = [0,1,0,\dots,0,0]$$

...

$$\mathbf{w}_p = [0,0,0,\dots,0,1]$$

In (17), $t_{1-\alpha/2}(N-1)$ denotes the $(1-\alpha/2) \times 100$ th percentile of the t -distribution with $N-1$ degrees of freedom.

To generate confidence bands for the AGC, identify \mathbf{w} with the rows of \mathbf{W} , and let $\hat{\mathbf{x}} = \mathbf{w}\hat{t}$, the corresponding fitted value. Then marginal confidence bands are computed using

$$\hat{\mathbf{x}} \pm \left(\frac{1}{N} \mathbf{w}\Omega\mathbf{w}'\right)^{1/2} t_{1-\alpha/2}(N-1) \quad (18)$$

Simultaneous confidence bands are of the form

$$\hat{\mathbf{x}} \pm \left[\frac{(N-1)P}{N(N-P)} \mathbf{w}\Omega\mathbf{w}' F_{1-\alpha}(P, N-P) \right]^{1/2} \quad (19)$$

where $F_{1-\alpha}(P, N-P)$ is the F -distribution with P and $N-P$ degrees of freedom.

We also compute prediction intervals [3,4]. These are of the form

$$\hat{\mathbf{x}}_f \pm \left[\frac{1}{N} \mathbf{w}_f \Omega \mathbf{w}_f' + \mathbf{w}_f \Lambda \mathbf{w}_f' + \hat{\sigma}^2 \right]^{1/2} t_{1-\alpha/2}(N-1) \quad (20)$$

Marginal confidence bands for the AGC are appropriate for use when the investigator is interested in bounding the AGC at a single, preselected value of t . Simultaneous bands, as suggested by their name, provide a joint confidence statement for all values of t . As such, these intervals will be wider than their marginal (one-at-a-time) counterparts. Prediction intervals may be used to bound the growth curve of a new subject from the same population as the sample used to construct the AGC. These must take into account both the variability of the individual curves about the AGC ($\mathbf{w}\Lambda\mathbf{w}' + \sigma^2$) and the variability (uncertainty) associated with the fact that the AGC was

estimated (Ω). They will therefore be the widest of the 3 kinds of confidence bands computed by our program. They are, however, useful in diagnostic contexts where one may wish to judge, e.g., whether a given subject is 'growing normally,' where normal is defined in terms of the population from which the sample used to construct the AGC is presumed to constitute a random sample.

6. The program

The program is invoked by issuing the command *gsruni cyu*. The user is first prompted as to whether the data represent COMMON or UNIQUE times of measurement. If COMMON, the program follows the format outlined in [3]. In this case, it is assumed that the study was planned to have common times of measurement, say t_1, t_2, \dots, t_T . The user is then prompted for the name and location of the (ASCII or GAUSS) file containing the measurements; the columns containing the first and last response variables (these must be in consecutive columns); the degrees of the reduced (RD) and full (FD) models [1,3]; the times of measurement; the confidence coefficient (e.g., 0.95) to be used in constructing confidence intervals and bands; the level of significance (e.g., 0.05) to be used in determining the degree of the polynomial to be fit to the AGCs in the groups; and the missing data code. Periods '.' may be used to represent missing data, as well as (user-specified) numerical values such as 9999. Note that in this case, the times of measurement are not part of the data set; they are entered by the user while running the program. In terms of our earlier example, assuming the planned common times of measurement are 1, 2, 3, 4 and 5, and that the 3 measurements for the second case were made at times 1, 2 and 4, the first 2 lines of the data set might read

```
25.2 29.0 33.6 9999 35.8
27.3 32.1 9999 41.8 9999
```

or

```
25.2 29.0 33.6 . 35.8
27.3 32.1 . 41.8 .
```

If UNIQUE times are indicated, the prompts are similar, but now it is assumed that the times of measurement are part of the data set: observations and times appear on alternating lines. In the context of the earlier example, the first 4 lines of the data set might read

```
25.2 29.0 33.6 35.8 .
1 2 3 5 .
27.3 32.1 41.8 . .
1.2 2.3 3.8 . .
```

Here the actual ages are used, and there is no real need for a missing data code. Note, however, due to the structure of GAUSS, all lines in the data set must be of the same length, and it is necessary to 'fill-up' the data set with 'missing data' so that each line has length T_{\max} , where T_{\max} denotes the maximum number of observations for individuals in the data set. (In the example, $T_{\max} = 5$.)

7. An example

We consider an example based on the first 10 cases of the data set given in Goldstein [7]. The observations are the heights of children in the Harpenden growth study conducted by Tanner et al. [8]. For convenience, these data are reproduced below, in the format required by our program in the case of UNIQUE times of measurement, alternating lines of observations and times of measurement:

```
110.0 117.3 123.3 127.7 132.6
5.055 6.164 7.071 8.030 9.082
113.1 120.4 126.2 133.7 140.2
5.022 6.091 7.008 7.992 9.134
113.6 120.7 127.1 133.0 140.0
5.030 6.085 7.008 7.986 9.087
109.6 115.3 121.4 127.8 134.4
5.041 6.071 7.052 7.997 9.030
104.3 111.0 116.9 122.7 130.5
5.161 6.084 6.964 7.961 9.013
108.6 116.2 122.4 127.7 133.7
5.000 6.038 7.008 7.972 9.024
101.5 107.4 114.0 119.5 125.9
5.006 6.025 6.981 7.959 9.006
95.5 101.2 105.6 110.6 117.0
```

```
102.3 106.5 112.2 117.1 122.2
5.011 6.394 7.383 7.967 8.981
112.8 121.8 128.2 134.5 142.1
5.025 6.178 7.107 8.008 9.197
```

After indicating that UNIQUE times of measurement are present, and the name and location of the (GAUSS or ASCII) data set, the user is prompted for the number of subjects ($N = 10$) and the maximum number of time points ($T_{\max} = 5$). The user then indicates whether the observations are in the first or second row (we require alternating rows of observations and time points, but these can be in any order), and chooses between 4 forms of the time design matrix [1,3,5,6]. For purposes of this example, we choose powers of the original time scores. Taking the degrees of the reduced and full models, respectively to be $RD = 1$ and $FD = 3$ (this causes the program to first test for the adequacy of a line, then, if necessary, for a quadratic. If a quadratic is inadequate, a cubic is used), and 0.95 for the level of confidence to be used in constructing intervals and bands, and 0.05 the level of significance for the tests of fit of the models, we find that a line is adequate to fit the data ($P = 0.4581$ for the F -test; $P = 0.3833$ for the chi-squared test), and that the estimated regression coefficients for each individual are

```
82.48 5.607
79.76 6.661
81.14 6.499
77.58 6.268
70.00 6.690
78.39 6.182
70.78 6.129
70.24 5.080
75.01 5.183
78.10 7.008
```

The estimated error variance is $\hat{\sigma}^2 = 0.476$ and

$$\hat{\Lambda} = \begin{bmatrix} 18.764 & 1.041 \\ 1.041 & 0.374 \end{bmatrix}$$

The generalized least squares estimator of the coefficients of the AGC is

Table 1
95% confidence intervals for the coefficients of the linear equation fit to the AGC when UNIQUE times are used

Degree	Coeff	Std error	Conf interval	Half interval
0	76.351	1.458	73.053-79.649	3.298
1	6.130	0.205	5.666-6.595	0.465

$$\hat{\tau} = \begin{bmatrix} 76.35 \\ 6.130 \end{bmatrix}$$

Confidence intervals of 95% for these coefficients are presented as shown in Table 1.

At this point, since unique times of measurement are allowed, and since subjects may present with considerably different, even non-overlapping periods of observation, it is necessary for the user to indicate the interval, and the points within this interval, at which the confidence bands should be computed. To help in this choice, the program prints

Min time value = 5.000
 Max time value = 9.211
 Max number of observations for an individual = 5.

The user is then asked to either input a 'typical time vector' or use the default which places 5 points equally spaced from 5.000 to 9.211. In this example, the choice of the typical vector is clear, and the user enters

5
 6
 7
 8
 9.

Table 2
Marginal 95% confidence bands for the AGC when UNIQUE times of measurement are used

Time point	Fitted value	Std error	Marginal conf interval	Half interval
5	107.00	1.970	102.546-111.459	4.457
6	113.13	2.118	108.432-117.924	4.791
7	119.26	2.275	114.117-124.409	5.146
8	125.39	2.439	119.877-130.911	5.517
9	131.52	2.609	125.623-137.425	5.901

Table 3
Simultaneous 95% confidence bands for the AGC when UNIQUE times of measurement are used

Time point	Fitted value	Std error	Simultaneous conf interval	Half interval
5	107.00	2.955	100.710-113.295	6.292
6	113.13	3.177	106.368-119.897	6.765
7	119.26	3.412	111.998-126.529	7.265
8	125.39	3.658	117.604-133.183	7.789
9	131.52	3.913	123.192-139.856	8.332

We then output the fitted AGC, the corresponding standard errors and 95% confidence bands for the AGC at each of these time points. There are 3 sets of output of this type corresponding to (i) marginal confidence bands for the AGC, (ii) simultaneous confidence bands for the AGC and (iii) prediction intervals. These are shown in turn in Tables 2, 3, and 4.

8. Discussion

The example considered above was chosen because it lends itself nicely to comparisons of alternate analysis strategies which might be considered in this situation. Two other approaches which suggest themselves are to (1) ignore the (in this example) slight departures from the target ages, and analyze the data as if everyone was measured at $t = 5(1)9$; and (2) adjust the observations to the target ages before performing the analysis.

In the first situation, we simply run the program using COMMON times of measurement (this is equivalent to our earlier program [3]) to estimate and compute confidence intervals and bands for

Table 4
95% confidence bands for prediction intervals when UNIQUE times of measurement are used

Time point	Fitted value	Std error	Prediction interval	Half interval
5	107.00	6.548	92.190-121.815	14.812
6	113.13	7.048	97.190-129.076	15.943
7	119.26	7.570	102.139-136.387	17.124
8	125.39	8.110	107.049-143.739	18.345
9	131.52	8.664	111.925-151.123	19.599

the AGC. Assuming all measurements were made at the target ages $t = 5(1)9$, we again find that a line is adequate to fit the data ($P = 0.07624$ for the F -test; $P = 0.02010$ for the chi-squared test), and now the estimated regression coefficients for each individual are

83.26 5.560
 79.47 6.750
 81.31 6.510
 78.23 6.210
 72.21 6.410
 78.53 6.170
 71.03 6.090
 69.30 5.240
 76.78 5.040
 78.97 7.130

The estimated error variance is now $\hat{\sigma}^2 = 0.371$ and

$$\hat{\Lambda} = \begin{bmatrix} 18.800 & 1.002 \\ 1.002 & 0.398 \end{bmatrix}$$

The generalized least squares estimator of the coefficients of the AGC is

$$\hat{\tau} = \begin{bmatrix} 76.81 \\ 6.111 \end{bmatrix}$$

Confidence intervals of 95% for these coefficients are given in Table 5.

It is seen that, in this example, the estimates of the model parameters are not greatly affected when one ignores the departures from the planned times of measurement. While not shown, the confidence bands for the AGC and prediction intervals are also quite similar (ignoring departures from the target ages results in slightly wider bands). One

Table 5
 95% confidence intervals for the coefficients of the linear equation fit to the AGC when differences from target ages are ignored

Degree	Coeff	Std error	Conf interval	Half interval
0	76.809	1.438	73.555–80.063	3.254
1	6.111	0.209	5.639–6.583	0.472

point of difference in the results is the test for the adequacy of $D = 1$ in fitting the AGC. When UNIQUE times are taken into consideration, the F -test (the test used by the program in deciding whether or not to step-up) yields $P = 0.4581$, so the specification $D = 1$ is not called into question. When COMMON times are used, the corresponding $P = 0.07624$ which might cause some to consider using a higher degree polynomial.

In the second approach, one might use a method like that described by Goldstein [7] to adjust the measurements to the target ages $t = 5(1)9$ before doing the analysis. The adjusted measurements are as follows [7]:

109.64 116.22 122.89 127.56 132.22
 112.95 119.80 126.14 133.75 139.44
 113.40 120.12 127.05 133.09 139.45
 109.37 114.88 121.06 127.82 134.21
 103.13 110.41 117.13 122.96 130.40
 108.60 115.94 122.35 127.86 133.56
 101.47 107.24 114.12 119.74 125.86
 95.37 101.12 105.72 110.84 115.93
 102.27 104.86 109.60 117.33 122.30
 112.60 120.49 127.46 134.45 140.84

When our program, with COMMON times, is used to analyze these data, again assuming the times of measurement are $t = 5, 6, 7, 8, 9$, once again we find that a line is adequate to fit the data ($P = 0.5701$ for the F -test; $P = 0.5101$ for the chi-squared test), and the estimated regression coefficients for each individual are now given by

82.16 5.650
 79.66 6.683
 81.07 6.507
 77.63 6.262
 69.84 6.709
 78.37 6.184
 70.79 6.128
 70.21 5.084
 74.50 5.253
 77.86 7.044

The estimated error variance is now $\hat{\sigma}^2 = 0.488$ and

$$\hat{\Lambda} = \begin{bmatrix} 18.529 & 1.097 \\ 1.097 & 0.365 \end{bmatrix}$$

Table 6
95% confidence intervals for the coefficients of the linear equation fit to the AGC when measurements are adjusted to target times

Degree	Coeff	Std error	Conf interval	Half interval
0	76.209	1.450	72.930–79.489	3.280
1	6.150	0.204	5.690–6.611	0.460

The generalized least squares estimator of the coefficients of the AGC is

$$\hat{\tau} = \begin{bmatrix} 76.21 \\ 6.150 \end{bmatrix}$$

and 95% confidence intervals for these coefficients are shown in Table 6.

The values in Table 6 agree quite closely with the results when UNIQUE times (which we view as the 'gold standard') are used (Table 1). The confidence bands and prediction intervals are also quite similar, showing that, in this example at least, Goldstein's adjustments to the target ages was effective.

Acknowledgement

Supported by DE08730 from the National Institute of Dental Research.

Appendix

A full set of PC programs for longitudinal data analysis, including this program, can be obtained on 5.25" or 3.5" diskettes (please request type) by sending \$25 to defray the cost of handling and licensing fees. These programs require a 80386- or 80486-based personal computer (PC) running the MS-DOS operating system (version 5.0 or higher is recommended, although versions as low as 3.3 will suffice). 80386 computers must also be equipped with a 80387 math coprocessor. At least 4 MB of memory is required, and must be available to

GAUSS386i, i.e., not in use by memory resident programs such as Windows. EGA or VGA graphic capabilities are required to display the color graphics; VGA or SVGA is suggested to display optimally the graphic results. Runtime modules are supplied with the programs so that no additional software (i.e., compiler or interpreter) is required to run these programs. One can create and edit ASCII data sets for use by these programs using the full screen editor supplied with MS-DOS version 5.0. The programs are written and compiled using GAUSS386i, version 3.0, require no additional installation or modification, and are run with a single command. When requesting the programs, address inquiries to the corresponding author and make checks payable to Baylor College of Dentistry.

References

- 1 Ten Have TR, Kowalski CJ and Schneiderman ED: PC program for analyzing one-sample longitudinal data sets which satisfy the two-stage polynomial growth curve model, *Am J Hum Biol*, 3 (1991) 269–279.
- 2 Rao CR: The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, *Biometrika*, 52 (1965) 447–458.
- 3 Furey AM, Ten Have TR, Kowalski CJ, Schneiderman ED and Willis SM: PC program extending the two-stage polynomial growth curve model to allow missing data, *Int J Biomed Comput*, 33 (1993) 287–296.
- 4 Carter RL and Yang MCK: Large-sample inference in random coefficient regression models, *Commun Stat: Theory Methods*, 8 (1986) 2507–2562.
- 5 Ten Have TR, Kowalski CJ and Schneiderman ED: PC program for obtaining orthogonal polynomial regression coefficients for use in longitudinal data analysis, *Am J Hum Biol*, 4 (1992) 403–416.
- 6 Ten Have TR, Kowalski CJ, Schneiderman ED and Willis SM: A PC program for performing multigroup longitudinal comparisons using the Potthoff-Roy analysis and orthogonal polynomials, *Int J Biomed Comput*, 30 (1992) 103–112.
- 7 Goldstein H: Some graphical procedures for the preliminary processing of longitudinal data. In *Interpreting Multivariate Data*, (Eds: V Barnett), Wiley, New York, 1981.
- 8 Tanner JM, Whitehouse RH, Marubini E and Resele LF: The adolescent growth spurt for boys and girls in the Harpenden growth study, *Ann Hum Biol*, 3 (1976) 109–126.