

Analysis of Longitudinal Data in Craniofacial Research: Some Strategies

Emet D. Schneiderman

Department of Oral and Maxillofacial Surgery, Baylor College of Dentistry, Dallas, TX 75266-0677

Charles J. Kowalski

Department of Biologic and Materials Sciences, School of Dentistry, University of Michigan, Ann Arbor, MI 48109

ABSTRACT: Although it is generally acknowledged that longitudinal data provide the most information on growth and development and other time-dependent phenomena, such data are often analyzed by conventional (cross-sectional) statistical methods. This widespread practice ignores the distinctive characteristics (e.g., covariance structure) of longitudinal data and may yield misleading results. The purpose of this article is to present some strategies and make available computer programs for the appropriate analysis of longitudinal data. User-friendly PC programs for the estimation of average growth curves, computation of tracking indices, prediction of future values, diagnosis, classification, clustering, estimation of missing values, and testing hypotheses concerning individual and group differences are presented. Benefits of these methods over the usual techniques are illustrated with the example of maxillary growth in the rhesus monkey.

KEY WORDS: growth curve analysis, biostatistics, craniofacial growth and development, PC programs, serial data, randomization tests.

I. INTRODUCTION

It has long been recognized that the best approach for studying craniofacial growth, and growth and development in general, involves the observation of individuals at a number of times throughout the maturation process. Thus numerous techniques, such as cephalometric radiography, vital staining of bone, and various imaging methodologies have been developed and perfected to collect data on growing individuals in order to gain information on the morphological changes that occur during ontogeny. Despite great increases in the sophistication in methodology for collecting longitudinal data, practical statistical methods for analyzing them appropriately have lagged far behind. Recent papers by Vacek *et al.* (1989), Cullis and McGilchrist (1990), and Stanek (1990) all point to the dual needs of developing appropriate software and informing researchers of its existence. They also suggest that biologists often use simple techniques because they are simple,

without regard to their validity and the quality of the information provided. We believe that both clinical and basic scientists are capable of understanding the logical bases of more complex statistical procedures provided only that they are clearly explained and illustrated and will use them when user-friendly software implementing such procedures becomes available. It is the latter of these notions that concerns us here. The authors and collaborators Stephen M. Willis, Thomas R. Ten Have, Amy Furey, Ingrid Y. Guo, and Elizabeth A. Mauger have developed (and are continuing to develop) an integrated set of easy-to-use programs for longitudinal data analysis and wish to make them available to the dental research community. Thus, the purpose of this article is to describe and illustrate these stand-alone PC programs for performing various types of longitudinal data analyses. The first analyses that we programmed are best suited to highly structured data sets (i.e., as might issue from a planned experiment), whereas the more recent ones accommodate the irregular

data sets (e.g., with missing data or different times of measurement) that often arise in human, clinical studies. These predominantly menu-driven programs are written and compiled using GAUSS (Edlefson and Jones, 1985; GAUSS, 1992), obviating the need for a compiler and additional software. Details for acquiring and using these programs are given in the Appendix.

A. Why Are Special Procedures and Programs Needed?

This question is addressed best by way of a real data analysis problem in the study of craniofacial growth. In the early 1980s, one of us (EDS) collected a large number of longitudinal cephalometric observations on the rhesus monkey, spanning the life of this long-lived primate. Data for a typical variable in this data set are shown in Table 1. These are measurements of maxillary length taken from tracings of lateral cephalograms made at half-year intervals from 0.5 to 10 years of age. The first column contain group indicators, 1 for male and 2 for female. Note the number of missing observations and the sparseness of data for the females. Research goals at the time of data collection were to establish sex-specific average growth curves (AGCs), explore hypotheses about changes in growth rates with maturation, and test for sex differences in growth. In addition to describing central tendencies in growth, it was also desired to determine the variability associated with the AGCs (Schneiderman, 1985).

The analytical goal at that time was to improve on the practice of calculating a conventional mean, standard deviation, and perhaps a confidence interval at each time point. This has been the usual way of summarizing longitudinal data, for example, in the large-scale study on craniofacial growth in normal children (Riolo *et al.*, 1974) and on cranial vault and facial growth in children with clefts (Friede *et al.*, 1986). While this approach gives us a general idea of what average size is at a given time, it has several serious faults:

1. The covariance structure (the pattern of correlations between all pairs of serial mea-

surements) and hence the longitudinal nature of the data is ignored.

2. The probability of making type I errors (the α rate) is substantially inflated when testing hypotheses about the growth curve coefficients (a type I error occurs when one falsely rejects the null hypothesis of no difference or no change).
3. The form of the true average growth curve cannot be determined merely by connecting these means.
4. Velocities and accelerations cannot be determined from these means.
5. One cannot conclude that 95% of the population falls within ± 2 SDs of the estimated means, as is commonly done (e.g., Riolo *et al.*, 1974: p. 10). Also, conventional univariate confidence intervals computed from these summary statistics are spuriously narrow; they imply greater precision than is possible based on the sample.

While it has long been known that intercorrelations in a longitudinal data set will inflate the α rate, this fact has been often ignored by those faced with analyzing longitudinal data; investigators have had to do so because of limitations in available software. The simulation experiments of Hoel (1964) allow us to explore the effects of intercorrelation of serial observations and to illustrate the severity of this problem. The null hypothesis addressed specifically by these experiments is that the regression coefficients that specify the AGC are equal to some specified constants (e.g., zero); Hoel also notes that the same kind of error will arise when testing the equality of two sets of regression coefficients. He demonstrated that if the correlation between observations one time unit apart was r , that between observations two units apart was r^2 , etc., and if $r = 0.80$, then with a nominal α rate of 0.05, the actual α rate would be 0.92. The simulations producing this result involved 25 time points and samples of 1000 from a multivariate normal distribution. In our example (complete data blocks in Table 1), all but one of the values for r are somewhat greater than those in Hoel's example (0.64, 0.90, 0.93, and 0.97). This is alarming, in that it indicates that if one attempted to determine the polynomial specification of the AGC (e.g., quadratic or cubic)

TABLE 1
Maxillary Length Measurements, Full Data Set Spanning from 0.5 to 10 Years of Age. Column 1
Contains Group Indicators, Where 1 = Male and 2 = Female. Note Blocks of Complete Data Set
in Boldface

	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
1	.	.	34.5	37.5	39.0	44.0	48.5	51.5	54.2	55.4	56.0	56.6	57.2	57.4	57.4	58.2	58.6	.	.	.
1	37.5	46.4	49.1	50.4	52.3	53.1	.	53.3	.	56.7	56.5	56.5	.	57.2	57.5
1	.	.	34.6	37.2	39.7	41.9	43.7	47.4	49.9	52.4	54.0	56.3	57.1	56.9	57.9	57.9	58.4	58.8	58.8	.
1	.	.	.	36.7	40.9	45.9	49.7	53.0	53.1	57.5	59.4	60.7	61.8	62.8	62.9	63.8	63.8	64.2	64.2	64.3
1	.	.	.	35.8	39.1	46.5	48.2	50.3	51.4	52.9	54.9	56.6	56.4	57.3	57.2
1	.	.	31.4	35.8	38.6	40.7	43.6	46.5	51.1	55.1	56.1	57.0	58.0	58.9	59.9
1	.	.	.	36.4	38.4	40.9	43.2	46.3	49.9	52.9	57.5	59.0	59.7	60.1	60.9	61.8
1	.	.	31.2	33.3	37.1	40.1	42.6	47.4	49.8	51.6	.	55.4	56.1
1	21.9	26.1	28.7	32.6	35.7	40.3	44.5	46.9	48.4	49.3	50.1	51.4	54.2	56.4	57.1
1	24.3	28.3	.	35.0	37.9	40.5	43.7	47.5	47.9	49.4	50.9	52.1	52.8
1	.	28.2	32.6	37.7	40.8	42.7	45.8	49.2	51.0	51.4	.	53.8	56.2
1	.	29.0	32.5	35.3	38.0	41.3	44.3	44.3	48.2	50.9	.	53.5	55.6
1	24.9	28.9	30.8	32.3	37.5	42.6	45.2	47.4	51.0	53.4	55.1	55.6
1	.	.	31.7	34.8	37.3
1	40.1	43.7	45.5
1	.	.	.	38.8	42.8
1	.	.	35.7	37.1	39.0	41.5	43.0
1	.	.	35.7	40.5	41.9
1	.	29.5	32.8	36.6	38.3	40.1
1	.	29.4	33.0	37.1
2	.	.	.	39.5	40.7	41.5	41.9	43.8	44.4	44.6	45.5	45.8	46.7	47.8
2	40.7	41.2	43.1	43.4	44.3
2	23.6	28.8	.	34.8	37.2	41.0	43.6	44.1	45.2	46.3	47.3
2	24.8	30.0	31.6	36.4	38.5	40.6	44.0	44.6	45.9	47.3	48.0	50.0
2	.	29.0	34.0	37.2	39.9	41.4	44.1	44.9	45.9	47.4
2	24.0	.	32.0	36.0	38.7	41.8	41.1	41.2	43.0	44.4	.	47.2
2	.	.	32.3	34.7	.	40.2	42.7	45.6	46.6	47.3	48.8	.	50.5
2	48.4	48.6	49.9
2	41.8	43.0	44.3
2	38.8	39.8	42.6
2	46.8	47.2
2	47.1	47.4
2	.	.	36.9	37.1
2	49.6	50.0
2	48.1	48.8

using ordinary least-squares (OLS) regression, he or she would run a greater than 92% chance of being wrong, not the nominal 5% chance.

Generating reliable estimates of growth patterns is not merely of academic interest — knowing whether a growth curve is quadratic rather than linear in form may tell us whether a disease or therapy is having an accelerating effect on growth or not; for example, in the case of children afflicted with craniosynostosis, this knowledge may be key to the prevention of brain damage or development of gross disfiguration in head shape, or in the case of a child with a cleft palate, the development of a deformed midface.

It is common practice for investigators to test more complex hypotheses in longitudinal data (e.g., tests for between group differences) using a

battery of t-tests, one-way or repeated measures ANOVA. Riolo *et al.* (1974), for example, used t-tests at each of 11 ages to test for sex differences. The problem of multiplicity when performing 11 tests is substantial, even under the unlikely condition that the data are uncorrelated; in this example, the overall or experimentwise error rate would be $1 - (1 - 0.05)^{11} = 0.431$. The chance of concluding the presence of sex differences when there are none (type I error) for this variable is 43%, not the nominal 5%. Thus, it is clear that both the problems of intercorrelations and multiplicity in longitudinal data analysis can contribute significantly to the inflation of the type I error rate.

While on the surface, ANOVA or repeated-measures ANOVA would appear to be more ap-

appropriate than the t-tests; they too require assumptions that are seldom fulfilled with longitudinal data sets. Univariate repeated measures ANOVA assumes that all correlations between pairs of measurements are equal (Fleiss, 1986; Schneiderman *et al.*, 1991; Kowalski, 1993), whereas in longitudinal data we normally expect observations made close together in time will be more highly correlated than those far apart.

Recognizing that the software available in 1983 was inadequate for properly analyzing craniofacial growth data and that there was no empirical basis for choosing among various longitudinal methods outlined in the biostatistical literature, we began studying and implementing methods for analyzing longitudinal data. In 1988 we obtained an NIDR grant for implementing and evaluating such methods. We hope with this article to share our results with the dental research community. We believe that the availability of these tools is critical to improving our understanding of normal and abnormal craniofacial growth.

Implementation is the first step in developing the knowledge base that will provide researchers with a rational basis for choosing among methods for longitudinal data analysis. Until the advent of reliable matrix programming languages, a variety of theoretically validated methods of growth curve analysis had languished in the biostatistical literature. We began by implementing Rao's (1959) classic single-sample polynomial growth curve model (Schneiderman and Kowalski, 1985). This permitted the estimation of AGCs and 95% simultaneous confidence bands, as well as the polynomial regression coefficients and corresponding confidence intervals. First programmed on a mainframe in SAS, it was later extended to accept unequal time intervals and reprogrammed in GAUSS to run on PCs, greatly enhancing its usability (Schneiderman *et al.*, 1991). Because this method is based on generalized least-squares rather than OLS, it explicitly incorporates the covariance structure of the data into the model; the investigator can make statements about the whole course of growth covered by the analysis, that is, the shape of the AGC and the width of the confidence band at any particular time. Below we present a concrete data analysis in which these and several related longitudinal models are illustrated and then outline the features of these and 26

other longitudinal procedures and the PC programs implementing them. The names of the programs are shown in uppercase.

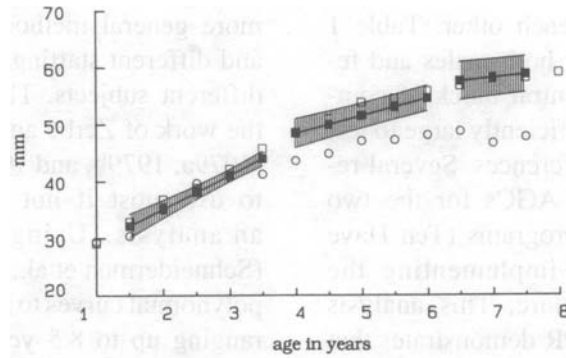
B. An Example — Growth in Maxillary Length

The graphic results of the analysis of maxillary length using Rao's model is shown in Figure 1. Using the related Hills' model (1968), also implemented as a GAUSS program (Schneiderman and Kowalski, 1989), velocities and accelerations were computed. These male data were fit adequately with linear equations. This and approximately 30 other craniofacial growth measurements for the rhesus monkey were similarly analyzed with early versions of the RAO and HILLS programs (Schneiderman, 1985) and published as a book (Schneiderman, 1992), making it the first to contain longitudinal descriptions of lower facial growth for any primate species based on an adequate sample. The average values at the five time points and simultaneous confidence intervals, as generated by RAO and a related program 2STG, are shown in Table 2. Also shown are conventional means and non-simultaneous confidence intervals. One can see that the conventional confidence intervals are 24 to 43% narrower than those generated by RAO. This apparent gain in precision provided by the conventional method is spurious, an artifact of ignoring the intercorrelated nature of this data. It is, however, desirable to minimize these intervals, that is, maximize the precision of the estimate of the AGC. Therefore, we implemented Rao's two-stage model (1965), which will give the best unbiased estimates of the AGC when its assumptions are fulfilled; when the individual curves and the AGC can be fit with the same degree polynomial this model will yield tighter confidence bands. The confidence intervals issuing 2STG (Ten Have *et al.*, 1991) are approximately 18 to 22% narrower than those issuing from RAO, providing real gains in precision with this particular data set (these conditions are by no means fulfilled with all data sets).

Despite breaking new ground, these methods and computer programs have several limitations. Note the three disjointed segments of the growth curve (Figure 1). This was necessitated by the

MAXILLARY LENGTH

A. Total Length



B. Change in Length

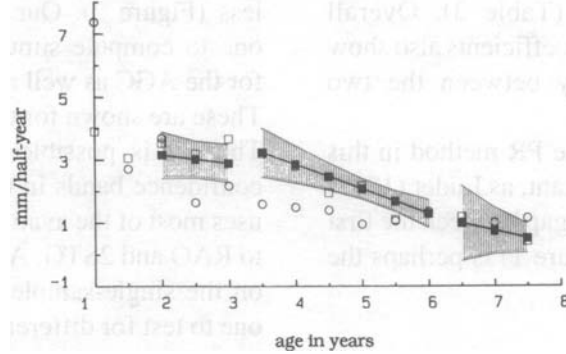


FIGURE 1. Average growth curves and simultaneous confidence bands for maxillary length in males as computed using Rao's Single Sample Polynomial Growth Curve Model. Conventional means for males and females are also shown as open symbols (squares and circles, respectively). (A) Attained growth curve; (B) growth velocity curve. (Reproduced with permission from Princeton University Press from Schneiderman (1992).)

TABLE 2
Conventional Means and Their 95% Confidence Intervals

T	Age	Conventional		RAO		2STG	
		\bar{X}	CI	AGC	CI	AGC	CI
1	3	41.93	1.49	42.48	2.26	42.41	1.81
2	3.5	45.36	1.37	45.15	2.00	45.07	1.62
3	4	48.22	1.42	47.83	1.88	47.73	1.52
4	4.5	50.49	1.10	50.50	1.94	50.39	1.54
5	5	52.66	1.41	53.17	2.15	53.05	1.67

Note: Fitted values of the AGC and 95% simultaneous confidence intervals for each from the Rao and two-stage analyses; maxillary length in millimeters, males only (N = 13); ages are in years.

requirement of the Rao and Hills procedures for complete data; three separate blocks of time, for which all subjects had all observations, were thus analyzed independently of each other. Table 1 contains all observations for both males and females; note that only the central block of complete data (in boldface) is sufficiently large to test the hypothesis of no sex differences. Several related hypotheses about the AGCs for the two sexes are tested using our programs (Ten Have *et al.*, 1992a and 1992b) implementing the Potthoff-Roy (1964) procedure. This analysis performed by our program PR demonstrates that from 3 to 5 years of age there is a significant difference in overall AGCs (test of coincidence) as well as in parallelism (Table 3). Overall univariate tests made on the coefficients also show them to differ significantly between the two groups.

While these results of the PR method in this example are valid and important, as Luder (1987) has pointed out, the apparent gap between the first and second data blocks (Figure 1) is perhaps the

most interesting in terms of the pubertal growth spurt, and, by implication, sex differences in growth patterns. Recently, we have implemented more general methods that allow missing data and different starting and finishing times for the different subjects. These methods are based on the work of Zerbe and Walker (1977) and Zerbe (1979a, 1979b, and 1979c) allow the investigator to use most if not all the available data for an analysis. Using our program ZCOEFF (Schneiderman *et al.*, 1993d), we interactively fit polynomial curves to subjects' data for time spans ranging up to 8.5 years. Twenty-two of the 35 potential subjects are used in the analysis; these subjects are fit with polynomials of degree 3 or less (Figure 2). Our program ZBANDS allows one to compute simultaneous confidence bands for the AGC as well as for the individual subject. These are shown for the male sample in Figure 3. Thus, it is possible to view the AGC and its confidence bands in their entirety; this approach uses most of the available data in marked contrast to RAO and 2STG. A related program that builds on the single-sample model in ZCOEFF allows one to test for differences in AGCs among two or more groups. This program, ZRTA, uses approximate randomization tests for hypothesis testing and allows one to delimit the analysis to any time span of interest. As one might expect, the results for 3- to 5-year time span above are similar to those yielded by PR — the difference is also significant ($p = 0.001$). When the total span between infancy and adulthood is considered, the overall growth curves are also found to be highly different ($p = 0.0005$). However, when we focus on the more ambiguous and interesting interval between 0.5 and 3.5 years of age for which the female data is scant, we find no evidence for sexual dimorphism in maxillary growth ($p = 0.227$). If we next confine the analysis to the interval between 3 and 4 years, the sex differences become significant ($p = 0.002$), confirming statistically an earlier speculation (Schneiderman, 1992) that growth patterns diverge between the sexes during this period. These computer programs have the advantage over the earlier ones in that they allow us to use much more of the available data in the analysis. In this example, we are able to use about 92% (250) of the 271 observations shown in Table 1 in the Randomization

TABLE 3
Key Results of Potthoff-Roy Analysis

Centered AGC Coefficients

Sex	N	α_0	α_1
F	7	43.55	2.354
M	13	47.42	5.033

Manova tests

For parallelism among average group curves
 $F(1.000, 18.00) = 30.63$ PROB > F = 2.966e-05
 For overall group differences
 $F(2.000, 17.00) = 22.80$ PROB > F = 1.544e-05

Univariate tests for overall group differences

For coefficient of degree 0
 $F(1.000, 18.00) = 26.51$ PROB > F = 6.740e-05
 For coefficient of degree 1
 $F(1.000, 18.00) = 30.63$ PROB > F = 2.966e-05

Group means of fitted (predicted) values

Sex	Age				
	3	3.5	4	4.5	5
F	41.19	42.37	43.55	44.73	45.90
M	42.38	44.90	47.42	49.93	52.45

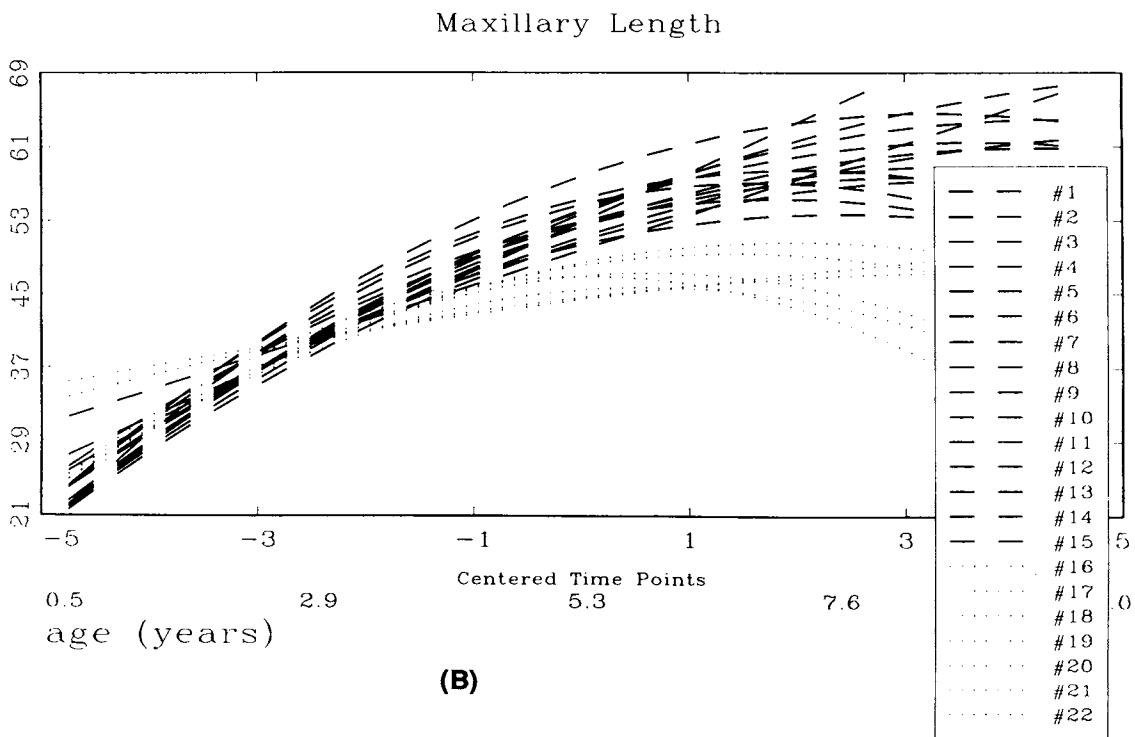
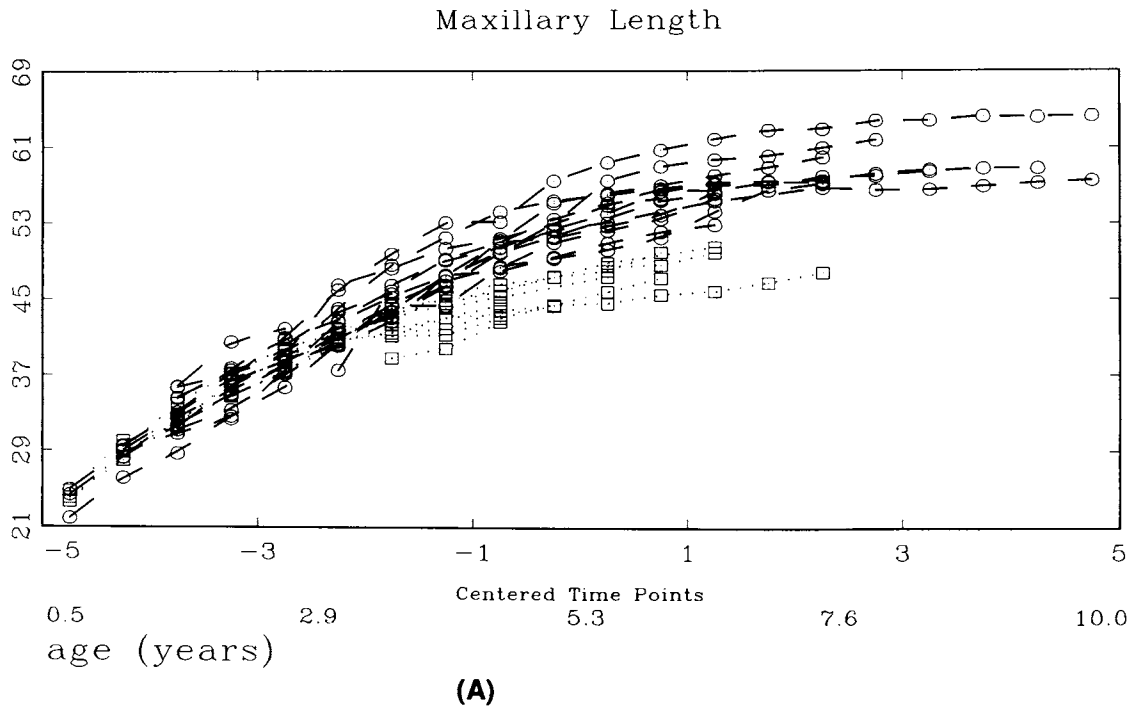


FIGURE 2. (A) Growth profiles for all original observations ($N = 35$) on maxillary length and (B) fitted polynomial growth curves for 22 of these subjects — all plotted with the program ZCOEFF.

AGC with Confid. Bands and Pred. Intrvls. (Scheffe)

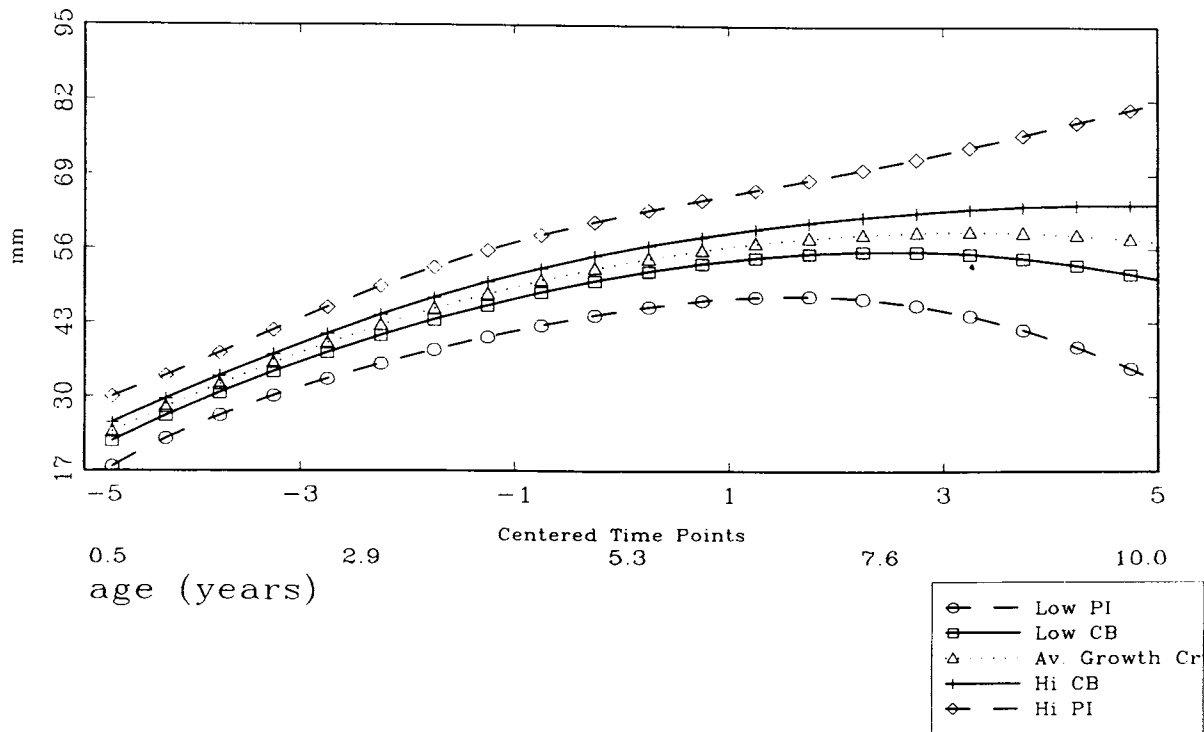


FIGURE 3. Plot of the 95% confidence bands for the AGC and for the individual (prediction interval) for males. All bands are of the Scheffé type and were generated by the program ZBANDS.

Tests, whereas only 37% (100) of the observations could be used in the RAO and other highly structured analyses.

Other questions of interest when the monkey craniofacial data were first collected were (1) how stable are growth patterns within and between populations, that is, how well do they track? and the related question, (2) can we predict later growth outcomes from earlier growth behavior? We were also interested in being able to (3) assess an individual's growth performance relative to the rest of a sample, with the goal of being able to diagnose growth problems. As no computer programs existed for these activities in the early 1980s, we also set out to implement methods that would address these questions.

Below is an overview of our current collection of GAUSS386i programs and the sorts of longitudinal data analyses one can accomplish with them. For greater detail regarding these programs and the underlying methods, we recommend that the reader consult our earlier papers.

These methods are categorized under three major headings, One-sample, G-sample (multisample), and Miscellaneous procedures, each with several subheadings describing particular techniques. One-sample procedures are appropriate in the situation in which the investigator has a single sample of N individuals, each measured at times that may be equally or unequally spaced. G-sample procedures are used for evaluating two or more groups of subjects in a single analysis, most often to contrast them with one another. Miscellaneous procedures include several "general purpose" methods that do not fit under the first two rubrics.

II. ONE SAMPLE PROCEDURES

A. Estimation of the Average Growth Curve (AGC)

We have several programs that, given a one-sample longitudinal data set, determine the

lowest degree polynomial adequate to fit the AGC, fit the corresponding polynomial, and provide confidence intervals for the coefficients of the polynomial and confidence bands for the AGC. There are several programs, because each is appropriate for a different set of assumptions concerning the covariance (correlation) structure of the longitudinal observations.

1. Rao's 1959 Model

When this structure is considered to be arbitrary, we have implemented (Schneiderman and Kowalski, 1985; Schneiderman *et al.*, 1991) a procedure due to Rao (1959). Our program RAO (formerly GCA7) finds the best polynomial specification for the AGC by using step-up goodness of fit tests, computes 95% simultaneous confidence intervals or bands, and reports the polynomial regression coefficients and the corresponding confidence intervals.

2. Hills Procedure

A second procedure due to Hills (1968) that, while formally equivalent to Rao (1959), provides explicit estimates of growth velocities and accelerations for each case and for the sample as a whole (Schneiderman and Kowalski, 1989). Our program HILLS and the underlying procedure using divided differences is intuitively appealing due to its simplicity.

3. Two-Stage Model

When the covariance structure of the repeated measurements may be assumed to have a particular form, known as "Rao's simple structure," an alternative procedure may be used (Ten Have *et al.*, 1991). This structure arises in the context of the so-called two-stage or random coefficients polynomial growth curve model (Rao, 1965) that, when the underlying assumptions are tenable, can be expected to produce sharper results than when the covariance structure is arbitrary. This procedure is implemented in the program 2STG.

4. Carter and Yang Model

An extension of the two-stage model due to Carter and Yang (1986) that accommodates missing data has also been implemented (Furey *et al.*, 1993b). In data sets where a small number of observations are missing at random, this program, CY, greatly extends the utility of the two-stage model. We have extended this model further to accommodate unique times of measurement for subjects with the program CYU (Mauger *et al.*, 1994). We are currently testing a program (VC) that further extends this model, after the work of Vonesh and Carter (1987), to accommodate a group indicator variable as well as unique times of measurement for subjects.

B. Tracking

Another set of one-sample analysis programs has to do with the computation of tracking indices. Tracking refers to the "regular behavior" of growth profiles, for example, do individuals who are "small" at one age tend to stay "small" as time progresses? More generally, tracking can be defined as the tendency of an individual to maintain a particular course of growth over time relative to other individuals. There are a number of ways in which "regular behavior" can be defined, and we have programs corresponding to several of these ways.

1. Cohen's Kappa

The first is based on the notion that individuals who are tracking tend to stay in the same quantile (track) of the distribution of the measurement whose growth is being monitored as that distribution changes over time. A chance-corrected measure of this tendency, based on Cohen's (1960) kappa statistic, is provided in Schneiderman *et al.* (1990). The statistic computed by our program KTRK is based on the numbers of times each individual is in each of the tracks; these numbers being compared with those that would be expected if individuals were assigned to tracks by chance at each time point.

2. Foulkes and Davis' Models

Another way of looking at tracking is in terms of pairs of individuals maintaining relative rank, that is, when one of the individuals is always smaller or larger than the other. Our program FDTRK (Schneiderman *et al.*, 1992a), following Foulkes and Davis (1981), estimates the probability that two randomly selected growth profiles will not intersect and provides a confidence interval for this quantity. A second program FDTRK2 (Schneiderman *et al.*, 1993a) uses Rao's (1959) arbitrary covariance structure for fitting polynomials to the growth curves and then computes a tracking index based on the crossings of these fitted curves.

3. McMahan's and Goldstein's Models

Yet another way to quantify tracking was given by McMahan (1981), who said tracking occurs to the extent that individuals are the same numbers of standard deviations away from the mean at each time of measurement. McMahan (1981) incorporated polynomial smoothing (using the two-stage model) in his procedure. This is implemented in our program MC (Guo *et al.*, 1994a). Goldstein (1981) simply compared the sum of squares of the standardized scores between individuals with the total sum of squares. This and related indices are computed by our program GTRACK (Furey *et al.*, 1994). A detailed discussion and comparison of these tracking methods is presented in Kowalski and Schneiderman (1992). Some of the differences in operational properties of these methods were considered in Porter *et al.* (1993).

C. Growth Prediction

Tracking is a useful concept in its own right, but it is also related to our ability to predict growth. We also have programs to perform simple types of growth prediction.

1. Short-Term

We have N individuals comprising a normative sample measured at T points of time and a "new" individual measured at the first $T - 1$ time

points, whose value at time T is to be predicted. The program predicts the value at time T and the confidence interval for that quantity (Schneiderman *et al.*, 1992c). Our program PREDICT also employs the leave-one-out (LOO) method, which can be used to further assess the accuracy with which the predictions based on this data set are made. In brief, the LOO method involves repeating the analysis N times, leaving one subject out of the standardizing sample each time, and predicting his or her value at time T . Because the true values at time T are known, the actual and predicted values may be compared, providing valuable insight into the predictability of the measurement under consideration.

2. Long-Term

This can also be extended to more points in time (Schneiderman *et al.*, 1993b). Suppose, for example, that the new individual has been measured at but T^* points in time and it is desired to predict the remaining $R = T - T^*$ values. The program LONGTERM predicts these values and their confidence intervals when the covariance structure is arbitrary.

3. Ware and Wu's Model

This program, WW (Guo *et al.*, 1994b), is similar to PREDICT above, except that it uses the two-stage polynomial growth curve model (Ware and Wu, 1981). If this model provides an adequate fit for the data, then one might expect tighter confidence intervals and bands than if the covariance structure is assumed to be arbitrary (as in PREDICT).

D. Diagnostic Testing

We have implemented two procedures that may prove useful in diagnostic contexts involving longitudinal measurements. The first is due to Zerbe (1979a). Given normative longitudinal data on N individuals, he shows how to compute the distances between the polynomial growth curves fit to these individuals and a "new" individual

(the one we wish to diagnose). The new individual is then assessed by computing his or her percentile rank in the distribution of these distances. We determine the proportion, γ , of these distances that are less than or equal to the average distance for the new individual. If this exceeds, say, 95%, we may consider this new individual to be "abnormal". This rule has the property that $100 \times \gamma$ % of the individuals most central in the normal sample would be correctly diagnosed. This technique can be used to construct normal ranges to classify an individual's growth, velocity, and acceleration of growth over a time interval, selected by the user for its biological importance. The procedure is nearly assumptionless and missing data are allowed. Our program, ZDIAG (Schneiderman *et al.*, 1994a) implements this procedure.

The second diagnostic procedure was used by Dawson, Todorov, and Elston (1980) in a study of the development of head circumference in achondroplastic children. In addition to developing estimates of and confidence bands for the AGC, they showed how one can obtain confidence bands for individual growth curves. The former indicates the accuracy with which the AGC has been estimated; the latter can be used to assess the likelihood that an individual is from the same population as the one from which the standardizing sample was drawn (i.e., for diagnostic purposes). Missing data are allowed. Our program, ZBANDS (Schneiderman *et al.*, 1994b), computes and plots the confidence bands for both the AGC and individual curves.

III. G-SAMPLE PROCEDURES

A. Structured Models

1. The Potthoff-Roy (PR) Analysis

Potthoff and Roy (1964) approached the problem of growth curve analysis by generalizing the standard multivariate analysis of variance (MANOVA) model into a form that accommodates fitting polynomial growth curves, constructing confidence bands about these curves, and testing hypotheses in the context of more complex experimental designs than had been considered

previously. We have implemented the PR procedure in both SAS (Ten Have *et al.*, 1992b) and GAUSS (Ten Have *et al.*, 1992a) for the analysis of G-sample data sets as defined in the Introduction. Our programs determine the polynomial of lowest degree adequate to fit the AGCs in each of the groups, fits polynomials of this degree to the AGCs, tests the equality of the resulting regression coefficients, and performs pairwise comparisons among the groups both with respect to all the coefficients simultaneously (MANOVA) and one coefficient at a time (ANOVA). Two forms of the multivariate tests are included: the first is for parallelism of the growth curves (all coefficients except for the intercept equal), the second for coincidence (all coefficients, including the intercept, equal). Several different forms of these tests are computed, viz., those based on the Hotelling-Lawley trace, Pillai's trace, Wilks' criterion, and Roy's maximum root criterion (see, Kshirsagar, 1972, p. 331, for more details concerning these statistics).

In order to perform the PR analysis, one must specify the form of a matrix containing the weights to be used in the weighted least squares estimation of the polynomial regression coefficients. In our program, the user may choose between the identity matrix, the sample covariance matrix, or the matrix can be input by the user. Timm (1975, p. 495) discusses the first two of these possibilities. Potthoff and Roy (1964) illustrate how the third option might arise in practice. It should be noted that the PR method remains valid regardless of the choice of this matrix.

We also allow four choices of the within-individual or time design matrix in the PR model. The advantages of using orthogonal and orthonormal time scores were detailed by Ten Have *et al.* (1992a, 1993); however, more traditionally popular time scores are also available, for example, the successive-powers- of -t form used by Schneiderman and Kowalski (1985).

2. Rao's Time Metameter

Rao's (1958) analysis is based on a transformation of the time axis with respect to which the growth curves in the G groups are approximately linear, regardless of the shapes of the curves on

the t-axis. This new time scale, t^* , is referred to as a "time metameter," and its structure can be estimated from the data in hand. Because the growth curves are linear with respect to t^* , the $N \times T$ data matrix can be reduced to a $N \times 2$ matrix containing the initial value of the measurement for each of the N individuals and estimates of their growth rates. The average growth rates in each of the groups can now be compared using a simple univariate analysis of variance or covariance, adjusting for differences in the initial mean values if necessary. Our SAS program (Ten Have *et al.*, 1992b) does both analyses, the user being free to select the one most appropriate for his or her situation. For additional information on Rao's time metameter see Rao (1961), Hoel (1964), and Goldstein (1979, p. 89).

3. Comparison of Tracking Indices

We have extended our program for computing the tracking index based on the kappa statistic (Schneiderman *et al.*, 1990) to allow the comparison of the kappas computed within each of G independent groups (Schneiderman *et al.*, 1992b). When the hypothesis of no difference between the G kappas is not contradicted by the data, the program KTRK computes an estimate of the overall value of kappa and provides a confidence interval for this quantity.

B. Randomization Tests

The G -sample procedures considered above require that individuals be measured at identical times; that polynomials of the same degree can be used to model the AGCs in each of the groups; and that the distribution of the variable whose growth is being monitored be multivariate normal. Zerbe (1979b) and Zerbe and Walker (1977) developed a class of procedures that make none of the above assumptions and thus are of great value to those who must cope with all of the problems inherent in collecting and analyzing longitudinal data. These procedures are based on randomization tests. For a good general account see Edgington (1987).

In our programs (Schneiderman *et al.*, 1994c) distances are computed between growth curves

for all pairs of individuals within each of the groups. These distances are averaged within each group and these averages are summed over the groups. We then recompute this sum for each of the possible assignments of N individuals to G groups (exact tests) or a randomly selected subset of all possible assignments (approximate tests). The p -value for the hypothesis of no group differences is then the proportion of these sums that are less than or equal to the sum of the average distances for the original data. The program ZRTA implements the approximate randomization tests and is suitable for any sample size; the program ZRTE performs exact randomization tests and is therefore suitable only for small problems.

IV. MISCELLANEOUS PROCEDURES

A. Computation of Orthogonal Polynomials

Orthogonal polynomials are useful in a number of contexts, including longitudinal data analysis. In order to obtain them you have had either to use tables (equally spaced time points) or to do laborious computations (unequally spaced points). We have a program ORPOL that generates them for you whether the points are equally spaced or not (Ten Have *et al.*, 1993).

B. Construction of Design Matrices

This program DESIGN may be used to evaluate and compare alternate choices of the design matrix in the general linear model, including polynomial regression (Furey *et al.*, 1993a). Measures of the "stability" of the design matrix are computed and influential data points are identified.

C. Imputation of Missing Data

The program MISSING may be used to estimate randomly missing observations in longitudinal data sets (Schneiderman *et al.*, 1993c). This is an extension of the theory of growth prediction allowing "predictions" to be made at any point along the time scale (not just at the end). Random

residuals are added to the imputed values to guard against oversmoothing the growth curves.

D. Computation of Polynomial Regression Coefficients

Zerbe (1979a) proposed a method for analyzing longitudinal data in which data are missing. Our program implementing Zerbe's procedure permits the estimation of polynomial growth, velocity, and acceleration curves (Schneiderman *et al.*, 1993d). This program ZCOEFF allows the interactive fitting of individual curves using step-down goodness-of-fit tests to arrive at the lowest degree adequate to fit them. The polynomial regression coefficients can then be saved for subsequent analysis by other programs (e.g., the MANOVA procedures provided in most commercial packages).

After fitting a set of growth curves in this fashion, it is then possible to compute confidence bands for the AGC and individual growth curves using our program ZBANDS (Schneiderman *et al.*, 1994b). Both Scheffé and Bonferroni type bands are computed.

E. Classification/Clustering of Growth Curves

Zerbe's (1979a) approach to diagnostic problems can also be used to classify a "new" individual into one of several existing groups. It is possible to determine the new individual's average distance from members of the various groups and assign this individual to the group in which he or she is "most central." One can compute $\gamma_1, \gamma_2, \dots, \gamma_G$, the percentile ranks of the new individual in each of the groups and assign the individual to the group for which this rank is minimal. Our program ZCLASS (Schneiderman *et al.*, 1994d) performs these computations.

Clustering is converse of classification. In the classification (sometimes called discrimination) problem, we wish to assign a new individual to one of G preexisting groups. Clustering, on the other hand, refers to the situation in which we have a sample of N individuals and wish to determine whether an underlying group structure can be discerned from the measurements in hand. Kendall (1980) gives a good description of both

procedures — and the differences between them. If we are given N individuals, the search for a group structure begins with a distance matrix with elements d_{ik} , the distances between each pair of individuals. Clustering algorithms then proceed to form groups by putting together that pair of individuals for which this distance is minimized, that is, the pair of individuals who are most similar. This pair is said to form the first "cluster." Then, using a user-specified measure of the distance between individuals and clusters, individuals or clusters are further linked together to form new clusters. This process continues until the user is satisfied that further agglomeration is not warranted, that is, that a balance between the within-cluster and between-cluster distances has been achieved. Anderberg (1973) provides a readable account of the entire procedure. Our program ZDIST (Schneiderman *et al.*, 1993e) allows the user to save the distance matrix for input into other clustering programs.

V. DISCUSSION

We have described, illustrated, and made available a number of user-friendly PC programs for longitudinal data analysis that we hope will strengthen the analysis of longitudinal data sets in dental research contexts (see the Appendix for details on acquiring copies). The interactive plotting routines of the Z programs (i.e., those based on Zerbe's methods) should prove especially valuable for exploring longitudinal data as well as performing analyses. These efficient plotting routines accept missing data, allow one to view directly unfitted and fitted curves (i.e., the correspondence between the original data points and their fitted counterparts); they also allow one to examine the implications of fitting a particular polynomial to an individual's observations. Additionally, the user has complete control over the display of a legend, symbols used for individual observations, and the color and pattern and thickness of lines used for plotting curves. While these features may seem cosmetic or trivial, they are actually essential to ensuring that one analyzes and interprets his or her data correctly. These features enable the investigator to immediately establish the identity of an outlier. With these program features, one can avoid the pitfalls of

running a statistical analysis blindly — this is not always done easily with commercial statistical packages, especially with larger data sets.

This article is limited to brief descriptions of the programs and is essentially an announcement of their availability. Space limitations preclude us from detailing the methods, interpretation of results, the situations in which particular procedures may be applicable, and all of the assumptions that must be met before each technique may be validly employed. The reader must refer to the publications cited for this important information. We are concerned not only that longitudinal methods be used to analyze longitudinal data sets; we also want to encourage the use of the proper longitudinal analysis in particular situations. We feel it is important that the dental research community have this information; however, it should be emphasized that intelligent use of these now-available tools for longitudinal analysis will require further study of them.

ACKNOWLEDGMENT

This work was supported in part by grant DE 08730 from the National Institute of Dental Research.

REFERENCES

- Anderberg, M. R.: *Cluster Analysis for Applications*, Academic Press, New York (1973).
- Carter, R. L. and M. C. K. Yang: Large-Sample Inference in Random-Coefficient Regression Models. *Comm. Stat.* **8**:2507–2526 (1986).
- Cohen, J.: A Coefficient of Agreement for Nominal Scales. *Educ. Psych. Measur.* **20**:37–46 (1960).
- Cullis, B. R. and C. A. McGilchrist: A Model for the Analysis of Growth Data from Designed Experiments. *Biometrics* **46**:131–142 (1990).
- Dawson, D. V., A. B. Todorov, and R. C. Elston: Confidence Bands for the Growth of Head Circumference in Achondroplastic Children during the First Year of Life. *Am. J. Med. Genet.* **7**:529–536 (1980).
- Edgington, E. S.: *Randomization Tests*, Marcel Dekker, New York (1987).
- Edlefsen, L. E. and S. D. Jones: GAUSS (Computer Program), Applied Technical Systems, Kent, WA (1985).
- Fleiss, J. L.: *The Design and Analysis of Clinical Experiments*, Wiley, New York (1986).
- Foulkes, M. A. and C. E. Davis: An Index of Tracking for Longitudinal Data. *Biometrics* **37**:439–446 (1981).
- Friede, H., A. A. Figueroa, M. L. Neagele, H. Gould, N. Kay, and H. Aduss: Craniofacial Growth Data for Cleft Lip Patients Infancy to 6 Years of Age: Potential Applications. *Am. J. Orthodont. Dentofacial Orthoped.* **90**:388–409 (1986).
- Furey, A. M., C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: A PC Program to Aid in the Choice of the Design Matrix in Multiple Linear Regression. *Int. J. Biomed. Comput.* **33**:1–23 (1993a).
- Furey, A. M., T. R. Ten Have, C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: A PC Program Extending the Two-Stage Polynomial Growth Curve Model to Allow Missing Data. *Int. J. Biomed. Comput.* **33**:287–296 (1993b).
- Furey, A. M., C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: GTRACK: A PC Program for Computing Goldstein's Growth Constancy Index and an Alternative Measure of Tracking. *Int. J. Biomed. Comput.* in press (1994).
- GAUSS (Computer Program): The GAUSS System Version 3.0. Aptech Systems, Inc., Maple Valley, WA (1992).
- Goldstein, H.: *The Design and Analysis of Longitudinal Studies*, Academic Press, New York (1979).
- Goldstein, H.: Measuring the Stability of Individual Growth Patterns. *Ann. Hum. Biol.* **8**:549–557 (1981).
- Guo, I. Y., T. R. Ten Have, C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: A PC Program for Computing McMahan's Tracking Indices from One-Sample Longitudinal Data Sets. *Int. J. Biomed. Comput.* **35**:255–266 (1994a).
- Guo, I. Y., E. D. Schneiderman, C. J. Kowalski, and S. M. Willis: PC Program for Growth Prediction in the Two-Stage Polynomial Growth Curve Model. *Int. J. Biomed. Comput.* **35**:39–46 (1994b).
- Hills, M.: A Note on the Analysis of Growth Curves. *Biometrics* **24**:189–196 (1968).
- Hoel, P. G.: Methods for Comparing Growth Type Curves. *Biometrics* **20**:859–872 (1964).
- Kendall, M. G.: *Multivariate Analysis*, Griffin, London (1980).
- Kowalski, C. J.: Data Analysis in Craniofacial Biology with Special Emphasis on Longitudinal Studies. *Cleft-Palate Craniofacial J.* **30**:111–120 (1993).
- Kowalski, C. J. and E. D. Schneiderman: Tracking: Concepts, Methods and Tools. *Int. J. Anthropol.* **7**:33–50 (1992).
- Kshirsagar, A. M.: *Multivariate Analysis*, Marcel Dekker, New York (1972).
- Luder, H. U.: Evidence for a Pubertal Spurt in Mandibular Condylar Growth of Nonhuman Primates. In: *Craniofacial Growth During Adolescence*. pp. 49–68. (Carlson, D. and K. A. Ribbens, Eds.) Center for Human Growth and Development, Ann Arbor (1987).
- Mauger, E. A., C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: Extension of the Carter-Yang Growth Curve Model to Allow Unique Times of Measurement for Subjects. *Int. J. Biomed. Comput.* in press, (1994).
- McMahan, C. A.: An Index of Tracking. *Biometrics* **37**:447–455 (1981).

- Porter, C. L., P. H. Buschang, and E. D. Schneiderman: Studying Mandibular Growth with the κ Tracking Statistic: Effects of Sample Size and Number of Tracks. *J. Dent. Res.* **72**(SI):340 (1993).
- Potthoff, R. F. and S. N. Roy: A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems. *Biometrika* **51**:313–326 (1964).
- Rao, C. R.: Some Statistical Methods for the Comparison of Growth Curves. *Biometrics* **14**:1–17 (1958).
- Rao, C. R.: Some Problems Involving Linear Hypotheses in Multivariate Analysis. *Biometrika* **46**:49–58 (1959).
- Rao, C. R.: Some Observations on Multivariate Statistical Methods in Anthropological Research. *Bull. Int. Stat. Inst.* **38**:99–109 (1961).
- Rao, C. R.: The Theory of Least Squares When the Parameters are Stochastic and its Application to the Analysis of Growth Curves. *Biometrika* **52**:447–458 (1965).
- Riolo, M. L., R. E. Moyers, J. M. McNamara, and W. S. Hunter: *An Atlas of Craniofacial Growth*, Monograph No. 2, Craniofacial Growth Series, Center for Human Growth and Development, University of Michigan, Ann Arbor (1974).
- Schneiderman, E. D.: A Longitudinal Cephalometric Analysis of Craniofacial Growth in the Rhesus Monkey (*Macaca mulatta*), Univ. Michigan Doctoral Dissertation, University Microfilms, Ann Arbor (1985).
- Schneiderman, E. D.: *Facial Growth in the Rhesus Monkey: A Longitudinal Cephalometric Study*. Princeton Univ. Press, Princeton, New Jersey (1992).
- Schneiderman, E. D. and C. J. Kowalski: Implementation of Rao's One-Sample Polynomial Growth Curve Model Using SAS. *Am. J. Phys. Anthropol.* **67**:323–333 (1985).
- Schneiderman, E. D. and C. J. Kowalski: Implementation of Hills' Growth Curve Analysis for Unequal-Time Intervals using GAUSS. *Am. J. Hum. Biol.* **1**:31–42 (1989).
- Schneiderman, E. D., C. J. Kowalski, and T. R. Ten Have: A GAUSS Program for Computing an Index of Tracking from Longitudinal Observations. *Am. J. Hum. Biol.* **2**:475–490 (1990).
- Schneiderman, E. D., S. M. Willis, T. R. Ten Have, and C. J. Kowalski: Rao's Polynomial Growth Curve Model for Unequal-Time Intervals: A Menu-Driven GAUSS Program. *Int. J. Biomed. Comput.* **29**:235–244 (1991).
- Schneiderman, E. D., C. J. Kowalski, T. R. Ten Have, and S. M. Willis: Computation of Foulkes and Davis' Non-parametric Tracking Index Using GAUSS. *Am. J. Hum. Biol.* **4**:417–420 (1992a).
- Schneiderman, E. D., S. M. Willis, C. J. Kowalski, and T. R. Ten Have: PC Program for Comparing Tracking Indices in Several Independent Groups. *Am. J. Hum. Biol.* **4**:399–401 (1992b).
- Schneiderman, E. D., S. M. Willis, C. J. Kowalski, and T. R. Ten Have: A PC Program for Growth Prediction in the Context of Rao's Polynomial Growth Curve Model. *Comput. Biol. Med.* **22**:181–188 (1992c).
- Schneiderman, E. D., C. J. Kowalski, T. R. Ten Have, and S. M. Willis: A PC Program for Computing the Foulkes-Davis Tracking Index for Polynomial Growth Curves. *Int. J. Biomed. Comput.* **32**:35–43 (1993a).
- Schneiderman, E. D., S. M. Willis, C. J. Kowalski, and T. R. Ten Have: Longer-Term Growth Prediction Using GAUSS. *Comput. Biol. Med.* **23**:149–154 (1993b).
- Schneiderman, E. D., C. J. Kowalski, and S. M. Willis: Regression Imputation of Missing Values in Longitudinal Data Sets. *Int. J. Biomed. Comput.* **32**:121–133 (1993c).
- Schneiderman, E. D., S. M. Willis, and C. J. Kowalski: PC Program for Estimating Polynomial Growth, Velocity, and Acceleration Curves When Subjects May Have Missing Data. *Int. J. Biomed. Comput.* **33**:249–265 (1993d).
- Schneiderman, E. D., C. J. Kowalski, and S. M. Willis: Clustering on the Basis of Longitudinal Observations. *Comput. Biol. Med.* **23**:399–406 (1993e).
- Schneiderman, E. D., S. M. Willis, C. J. Kowalski, and I. Y. Guo: A PC Program for Diagnosis on the Basis of Longitudinal Data. *Int. J. Biomed. Comput.* **35**:247–254 (1994a).
- Schneiderman, E. D., C. J. Kowalski, S. M. Willis, and I. Y. Guo: A PC Program for Computing Confidence Bands for Average and Individual Growth Curves. *Comput. Biol. Med.* **24**:119–127 (1994b).
- Schneiderman, E. D., S. M. Willis, C. J. Kowalski, and I. Y. Guo: Implementation of Exact and Approximate Randomization Tests for Polynomial Growth Curves. *Int. J. Biomed. Comput.* **36**:187–192 (1994c).
- Schneiderman, E. D., S. M. Willis, and C. J. Kowalski: A PC Program for Classification Into One of Several Groups on the Basis of Longitudinal Data. *Comput. Biol. Med.* **24**(4): in press (1994d).
- Stanek, E. J.: A Two-Step Method for Understanding and Fitting Growth Curve Models. *Stat. Med.* **9**:841–851 (1990).
- Ten Have, T. R., C. J. Kowalski, and E. D. Schneiderman: PC Program for Analyzing One-Sample Longitudinal Data Sets Which Satisfy Rao's Two-Stage Polynomial Growth Curve Model. *Am. J. Hum. Biol.* **3**:269–279 (1991).
- Ten Have, T. R., C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: A PC Program for Performing Multigroup Longitudinal Comparisons Using the Potthoff-Roy Analysis and Orthogonal Polynomials. *Int. J. Biomed. Comp.* **30**:103–112 (1992a).
- Ten Have, T. R., C. J. Kowalski, E. D. Schneiderman, and S. M. Willis: Two SAS Programs for Performing Multigroup Longitudinal Analyses. *Am. J. Phys. Anthropol.* **88**:251–254 (1992b).
- Ten Have, T. R., C. J. Kowalski, and E. D. Schneiderman: A PC Program for Obtaining Orthogonal Polynomial Regression Coefficients for Use in Polynomial Regression Analysis. *Am. J. Hum. Biol.* **4**:403–416 (1993).
- Timm, N. H.: *Multivariate Analysis with Applications in Education and Psychology*, Brooks/Cole, Monterey, CA (1975).
- Vacek, P. M., R. M. Mickey, and D. Y. Bell: Application of a Two Stage Random Effects Model to Longitudinal Pulmonary Function Data from Sarcoidosis Patients. *Stat. Med.* **8**:189–200 (1989).

- Vonesh, E. F. and R. L. Carter: Efficient Inference for Random-Coefficient Growth Curve Models with Unbalanced Data. *Biometrics* 43:617-628 (1987).
- Ware, J. H. and M. C. Wu: Tracking: Prediction of Future Values from Serial Observations. *Biometrics* 37:424-437 (1981).
- Zerbe, G. O.: A New Nonparametric Technique for Constructing Percentiles and Normal Ranges for Growth Curves Determined from Longitudinal Data. *Growth* 43:263-272 (1979a).
- Zerbe, G. O.: Randomization Analysis of the Completely Randomized Design Extended to Growth and Response Curves. *J. Am. Stat. Assoc.* 74:215-221 (1979b).
- Zerbe, G. O.: Randomization Analysis of Randomized Blocks Extended to Growth and Response Curves. *Comm. Stat.* A8:191-205 (1979c).
- Zerbe, G. O. and S. H. Walker: A Randomization Test for Comparison of Groups of Growth Curves with Different Polynomial Design Matrices. *Biometrics* 33:653-657 (1977).

APPENDIX

Computer Implementation

The full set of PC programs for longitudinal data analysis 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*). Five mb of hard disk space are required to store the programs. 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.