Comput. Biol. Med. Vol. 22, No. 3, pp. 181–188, 1992
Printed in Great Britain

0010-4825/92 $5.00 + .00

A PC PROGRAM FOR GROWTH PREDICTION IN THE CONTEXT OF RAO'S POLYNOMIAL GROWTH CURVE MODEL

EMET D. SCHNEIDERMAN,*† STEPHEN M. WILLIS,† CHARLES J. KOWALSKI‡ and THOMAS R. TEN HAVE§

†Department of Oral and Maxillofacial Surgery, Baylor College of Dentistry, 3302 Gaston Avenue, Dallas, TX 75246 and the ‡Departments of Oral Biology and §Biostatistics, The University of Michigan, Ann Arbor, MI 48109, U.S.A.

(*Received* 26 *August* 1991; *received for publication* 24 *January* 1992)

Abstract—We consider the problem of growth prediction in the context of Rao's [1] one-sample polynomial growth curve model and provide a PC program, written in GAUSS, to perform the associated computations. Specifically, the problem considered is that of estimating the value of the measurement under consideration for a "new" individual at the T^{th} time point given measurements on that individual at $T-1$ previous points in time and the values of the measurement on N "similar" individuals at all T time points. The times of measurement $t_1, t_2, \ldots, t_T$ need not be equally spaced, but we assume that each of the N individuals comprising the normative sample were measured at these times. The method and the program are illustrated using the leave-one-out method on a sample of $N = 12$ male rhesus monkeys whose mandibular ramus height was measured five times at yearly intervals.

Longitudinal studies Growth curves Polynomials Prediction
PC program

INTRODUCTION

We have previously described Rao's [1] one-sample polynomial growth curve model [2] and provided GAUSS programs to perform the associated computations [3, 4]. Given a longitudinal data set consisting of the values of the measurement under consideration for N individuals at T time points, namely,

$$\underset{N\times T}{\boldsymbol{X}} = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1T} \\ x_{21} & x_{22} & \ldots & x_{2T} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N1} & x_{N2} & \ldots & x_{NT} \end{bmatrix} \tag{1}$$

and assuming that each row of $\boldsymbol{X}$ has a multivariate normal distribution with mean or expected value

$$E(\boldsymbol{x}_i) = \boldsymbol{W}\boldsymbol{\tau} \tag{2}$$

and (arbitrary) covariance matrix Σ, our programs can be used to:

(i) Find the lowest degree polynomial in time adequate to fit the average growth curve (AGC);

(ii) Estimate the coefficients of this polynomial and provide confidence intervals for them;

* Author to whom correspondence should be addressed.

(iii) Obtain confidence bands for the AGC; and

(iv) Plot the individual growth profiles and the AGC along with its associated confidence bands.

In (2), W is the within-individual (or time) design matrix and τ the vector of regression coefficients for the AGC, namely,

$$W = \begin{bmatrix} 1 & t_1 \ldots t_1^D \\ 1 & t_2 \ldots t_2^D \\ \vdots & \vdots \quad \vdots \\ 1 & t_T \ldots t_T^D \end{bmatrix} \quad \text{and} \quad \tau = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_P \end{bmatrix} \tag{3}$$

where D is the degree of the polynomial being fit and $P(=D+1)$ the corresponding number of polynomial regression coefficients. Thus for any $t = t_1, t_2, \ldots, t_T$

$$E(x_i|t) = \tau_1 + \tau_2 t + \tau_3 t^2 + \ldots + \tau_P t^D. \tag{4}$$

The purpose of the present paper is to extend this methodology—and our program—to accommodate a simple form of *growth prediction*, i.e. to allow the user to estimate the value of the measurement under consideration for a "new" individual at the T^{th} time point given measurements on that individual at $T-1$ previous points in time and the values of the measurement on N "similar" individuals at all T time points. The times of measurement $t_1, t_2, \ldots, t_T$ need not be equally spaced, but we assume that the time design matrix, W is the same for each of the $N+1$ individuals (the N individuals comprising the normative sample and the individual whose growth we wish to predict).

This method can be applied to longitudinal data sets, where one is interested in predicting future values for subjects. The most obvious areas for application are to child growth studies and the clinical practices of pediatrics and orthodontics. Often such investigators and practitioners are interested in predicting where a particular child will be, in terms of stature or weight, or facial dimensions, at some future time. The availability of an appropriate standardizing sample, i.e. one having subjects with characteristics similar to those of the subject under consideration, and observations at the time points of interest, is required for using this approach. For example, a pediatrician interested in predicting how tall a 10-year-old male achondroplastic patient will be at age 12 when treated with a particular hormone therapy, having 6 years of data on the patient, would need a comparable reference sample of achondroplastic boys treated in the same way, with complete data from 4 to 12 years. We demonstrate in this paper, with a real example, that it is possible to generate quite accurate estimates of future values, even on the basis of a small reference sample—12 subjects in this case. Application of this method to "filling-in" missing data in incomplete longitudinal data sets is also considered in this paper.

Formally, we may state the problem as follows: Given X and given the first $T-1$ entries of

$$x_v = \begin{bmatrix} x_{v1} \\ x_{v2} \\ \vdots \\ x_{v,T-1} \\ x_{vT} \end{bmatrix} \tag{5}$$

estimate the value of x_{vT}.

PREDICTION OF x_{vT}

The solution to this prediction problem is most easily described in terms of *partitioned matrices* [5]. We partition the vector x_v into its known and unknown parts, namely,

$$\boldsymbol{x}_\nu = \begin{bmatrix} x_{\nu 1} \\ \vdots \\ x_{\nu,T-1} \\ \hline x_{\nu T} \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_\nu^* \\ \hline x_{\nu T} \end{bmatrix} \tag{6}$$

so that $\boldsymbol{x}_\nu^*$ is $(T-1)\times 1$, the observed values for the ν^{th} individual, and $x_{\nu T}$ is the (scalar) quantity to be predicted. The time design matrix $\boldsymbol{W}$ is partitioned similarly into the $(T-1)\times P$ matrix $\boldsymbol{W}_1$ and the $1\times P$ matrix (vector) $\boldsymbol{W}_2$, namely,

$$\boldsymbol{W} = \begin{bmatrix} 1 & t_1 & \dots & t_1^D \\ \vdots & \vdots & \dots & \vdots \\ 1 & t_{T-1} & \dots & t_{T-1}^D \\ 1 & t_T & \dots & t_T^D \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_1 \\ \hline \boldsymbol{W}_2 \end{bmatrix}. \tag{7}$$

Finally, the covariance matrix $\boldsymbol{\Sigma}$ is written

$$\boldsymbol{\Sigma} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right] \tag{8}$$

where $\boldsymbol{\Sigma}_{11}$ is $(T-1)\times(T-1)$, $\boldsymbol{\Sigma}_{12}=\boldsymbol{\Sigma}_{21}'$ is $(T-1)\times 1$ and $\boldsymbol{\Sigma}_{22}$ is a scalar. Then from standard multivariate normal theory [6], the conditional mean and variance of $\boldsymbol{x}_{\nu T}$ given $\boldsymbol{x}_\nu^*$ are

$$E(\boldsymbol{x}_{\nu T}|\boldsymbol{x}_\nu^*) = \boldsymbol{W}_2\boldsymbol{\tau} + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_\nu^* - \boldsymbol{W}_1\boldsymbol{\tau}) \tag{9}$$

and

$$V(\boldsymbol{x}_{\nu T}|\boldsymbol{x}_\nu^*) = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}. \tag{10}$$

The so-called *empirical Bayes predictor* [7] of $\boldsymbol{x}_{\nu T}$ and its estimated variance are then obtained by substituting estimates of $\boldsymbol{\tau}$ and $\boldsymbol{\Sigma}$ in the above equations. That is, if $\boldsymbol{S}$ is the sample covariance matrix (partitioned analogously to Σ in (8)) and if $\bar{\boldsymbol{x}}$ is the $T\times 1$ vector of means at each time point, $\boldsymbol{\tau}$ is estimated by [2]

$$\hat{\boldsymbol{\tau}} = (\boldsymbol{W}'\boldsymbol{S}^{-1}\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{S}^{-1}\bar{\boldsymbol{x}} \tag{11}$$

and hence

$$\hat{E}(\boldsymbol{x}_{\nu T}|\boldsymbol{x}_\nu^*) = \boldsymbol{W}_w\hat{\boldsymbol{\tau}} + \boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}(\boldsymbol{x}_\nu^* - \boldsymbol{W}_1\hat{\boldsymbol{\tau}}) \tag{12}$$

and

$$\hat{V}(\boldsymbol{x}_{\nu T}|\boldsymbol{x}_\nu^*) = \boldsymbol{S}_{22} - \boldsymbol{S}_{21}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}. \tag{13}$$

An approximate 95% prediction interval for $\boldsymbol{x}_{\nu T}$ is then

$$\hat{E}(\boldsymbol{x}_{\nu T}|\boldsymbol{x}_\nu^*) \pm 2\sqrt{\hat{V}(\boldsymbol{x}_{\nu T}|\boldsymbol{x}_\nu^*)}. \tag{14}$$

The approach to prediction outlined above is entirely similar to that taken by Ware and Wu [7] in the context of the so-called two-stage (or random coefficients) polynomial growth curve model. The essential difference between their model and the one considered here is that in the two-stage model $\boldsymbol{\Sigma}$ is assumed to have a special structure. In

our development, Σ is an *arbitrary* positive-definite matrix. Also, Ware and Wu tie prediction to the concept of *tracking* [8]. This is perfectly natural in as much as accurate growth predictions are an indication of tracking behavior. Stated otherwise, one would not expect to be successful in predicting growth when the sample under consideration is tracking poorly. This point of view is given explicit consideration in the example to follow. We show how our program can be used to predict growth and provide the values of the corresponding tracking indices [8] for the sample under consideration. Before turning to this example, a few general remarks concerning the program and its use may be in order.

THE PROGRAM

A data set of the form (1) is assumed to be available. This may be in either GAUSS or ASCII format; the user is asked first to indicate which and to provide the name of the file. If the file is ASCII (A), the user must know the values of N and T to proceed: He/she is prompted for these values. If a GAUSS (G) data set is used, the value of N is determined by the program. The user is then asked whether or not the time points are equally spaced. If yes (Y), the user may select the default values $1, 2, \ldots, T$ or type in the actual time points (one per line). If no (N), the user is asked to enter the values of the time points. The user is also asked to specify the level of significance for the step-up goodness-of-fit tests used to determine the degree of the polynomial fit to the AGC. Finally, the user is asked to enter the values of x_v at the first $T-1$ points in time as in (5).

The output includes D, the smallest degree polynomial adequate to fit the data; the estimated values of the elements of τ and their corresponding 95% confidence intervals; the 95% confidence bands for the AGC at each time of measurement; the estimated value of x_{vT}; and an approximate 95% confidence interval for this quantity. The AGC and its confidence bands are then plotted and the predicted value for the first "new" individual is highlighted. The user is then asked whether or not another prediction is to be made. If yes, the user is prompted for the observed values of the second "new" individual at the first $T-1$ time points. The numerical output at this stage consists only of the predicted value and the prediction interval. The graphical output again highlights the predicted value for this individual against the backdrop of the AGC and its 95% confidence band. The program continues in this fashion until the user responds no (N) to the question concerning another individual's prediction. The user is then given the opportunity to save the original data set, X, augmented by the observed and predicted values for the new individual(s). That is, the $N \times T$ data matrix X may be expanded to $(N+n) \times T$ where n is the number of predictions made. This feature may prove useful when the investigator wishes to fill-in an incomplete data set due to dropouts at time t_T. The enlarged data set can be saved in ASCII format in a file named by the user and subsequently used in any of our (or others') programs.

Finally, as an option, the user may choose to apply the leave-one-out method to his/her data set to evaluate the accuracy of the predictions made. This method is described in the following section.

AN EXAMPLE

Our example is based on the data set previously considered in [2] consisting of mandibular ramus height measurements (in mm) for 12 male rhesus monkeys at $T=5$ yearly intervals (coded 1, 2, 3, 4, 5). A second degree ($D=2$) polynomial was found to fit the data adequately ($p=0.14$), the AGC being estimated by

$$x(t) = 18.56 + 8.819t - 0.8198t^2.$$

For a new individual with observations 25.89, 30.09, 35.30 and 37.86 at the first four times of measurement, the predicted value was 38.647 with prediction variance 0.3375, leading to the approximate 95% prediction interval of (37.49, 39.81). If this were the

only prediction made, the user might then opt to save the expanded 13×5 data matrix consisting of the original 12×5 matrix, X, augmented by a 13^{th} row with values

$$25.89 \quad 30.09 \quad 35.30 \quad 37.86 \quad 38.647.$$

If we choose to employ the leave-one-out (LOO) method, $N = 12$ additional predictions are made: we leave one monkey out of the computations involving the normative sample at each stage and predict *his* value at $T = 5$. Since the actual values at $T = 5$ are known for each monkey, a comparison of these values with the predicted values provides some insight into the accuracy with which predictions are being made. This method was used in growth prediction contexts by Rao [9, 10] and other applications were indicated in [11]. The results are shown below:

Monkey	$T = 5$ actual	$T = 5$ predicted
1	35.8	35.56
2	43.5	43.41
3	38.9	39.42
4	44.4	43.46
5	37.9	38.63
6	43.8	44.01
7	43.1	43.17
8	44.0	44.79
9	43.8	44.02
10	42.0	42.12
11	43.8	42.85
12	43.8	44.44

It is seen that the predictions are quite close for this data set. This occurs despite the fact that these monkeys do not track especially well as judged by the values of the tracking indices we have implemented, these being an index based on the kappa statistic [12], and two forms of the index developed by Foulkes and Davis [13], denoted here by FDI and FDII. In fact, their values and the 95% confidence intervals for the corresponding parameters are:

Kappa (with three tracks): 0.24242 ± 0.12990
FDI: 0.39394 ± 0.14902
FDII (with $D = 2$): 0.53030 ± 0.13018.

This is somewhat unexpected [7], perhaps reflecting the facts that tracking indices measure particular aspects of growth patterns, and small values do not *preclude* prediction. One can expect that prediction will be quite good when tracking is in evidence, but tracking is not a necessary condition for the ability to predict.

A measure of the overall accuracy of prediction which can be used to compare rival methods of prediction and/or the accuracy of a given method on several data sets is the root mean square error

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A - P)^2}. \tag{15}$$

In (15), A denotes the actual value and P the predicted value. this quantity is also computed when the user chooses to perform the LOO method on his/her data. The smaller the value of the RMSE, the more accurate the predictions which have been made. The value of this quantity for our example data set is $\text{RMSE} = 0.56$.

When this analysis is repeated on maxillary length measurements in the same monkeys, RMSE = 1.13 providing an indication that it is somewhat easier to predict ramus length than maxillary length in these monkeys. When done on ramus height in human males aged 8, 8.5, 9 and 9.5 years (these data have been analyzed, among others, by [14, 15]), RMSE = 0.64, suggesting that the accuracy of the predictions of ramus length measurements in male rhesus monkeys and humans are comparable, at least over the age ranges considered.

DISCUSSION

We have anticipated that some users will realize that this program can be used to "fill-in" longitudinal data sets that are incomplete (contain missing data) due to drop-outs. While several authors have proposed methods for analyzing incomplete repeated measurements data [16–19], use of these methods in practice awaits the development of appropriate software. And, while we are working at filling this void, it must be emphasized that whether one estimates the values of the missing observations or uses an analysis which can accommodate missing data, it is important to be sure that the drop-outs have occurred "at random," i.e. that the incomplete measurement sequences are not atypical. Diggle [20] gives a good discussion and outlines a test which may be used to check on this assumption. In any case, we suggest that users prudently limit the numbers of observations filled-in using the methods of this paper.

Having said this, as mentioned earlier, we do allow users to save the enlarged data set which results when several individuals' observations have been estimated. That is, if n predictions have been made, the original $N \times T$ data matrix, X, can be augmented to produce an $(N+n) \times T$ ASCII data set, named by the user, which can be read into any program requiring complete data. We suggest that n should be small relative to NT, certainly less than 10%, and note that the effect(s) of using the enlarged data set can be at least partially assessed by comparing the outputs from a given program obtained when the original and augmented data sets are used in turn. Finally, we suggest that the investigators report, in all subsequent uses of the augmented data set, the proportion of observations that have been estimated.

SUMMARY

We have considered growth prediction in the context of Rao's [1] one-sample polynomial growth curve model and provided a PC program to perform the associated computations. Specifically, the problem considered was that of estimating the value of the measurement under consideration for a "new" individual at the T^{th} time point given measurements on that individual at $T-1$ previous points in time and the values of the measurement on N "similar" individuals at all T time points. The times of measurement $t_1, t_2, \ldots, t_T$ were not assumed to be equally spaced, but we did assume that each of the N individuals comprising the normative sample were measured at these times. The method and the program were illustrated using a sample of $N = 12$ male rhesus monkeys whose mandibular ramus height was measured five times at yearly intervals.

This method has applications in the study of growth, development and treatment effects in humans and other species. For example, it can be used to predict, with a known level of confidence, the future height of a child, given the availability of a suitable normative sample. In addition to predicting the T^{th} value for one or several subjects, the program computes the root mean squared error, which reflects the accuracy of the predicted values. In the example considered, we demonstrated that it was possible to obtain accurate predictions even in an instance when the standardizing sample was small and did not track particularly well.

Finally, it was noted that while our program could be used to fill-in longitudinal data sets containing missing data due to drop-outs, care should be taken when doing so to ensure that the incomplete measurement sequences are not atypical.

Acknowledgement—This research was supported by Grant DE08730 from the National Institute of Dental Research.

REFERENCES

1. C. R. Rao, Some problems involving linear hypotheses in multivariate analysis, *Biometrika* **46**, 49 (1959).
2. E. D. Schneiderman and C. J. Kowalski, Implementation of Rao's one-sample polynomial growth curve model using SAS, *Am. J. Phys. Anthropol.* **67**, 323 (1985).
3. E. D. Schneiderman and D. J. Kowalski, Implementation of Hills' growth curve analysis for unequal-time intervals using GAUSS, *Am. J. Hum. Biol.* **1**, 31 (1985).
4. E. D. Schneiderman, 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 (1991).
5. F. A. Graybill, *Introduction to Matrices with Applications in Statistics*. Wadsworth, Belmont, CA (1969).
6. C. R. Rao, *Linear Statistical Inference and Its Applications*, p. 442. Wiley, New York (1965).
7. J. H. Ware and M. C. Wu, Tracking: Prediction of future values from serial measurements, *Biometrics* **37**, 427 (1981).
8. C. J. Kowalski and E. D. Schneiderman, Tracking: concepts, methods and tools, *Multivariate Methods in Bioanthropology*, G. N. van Vark, W. Schaafsma and R. S. Corruccini, Eds. II Sedecesimo, Florence, Italy (1992).
9. C. R. Rao, Prediction of future observations with special reference to linear models, *Multivariate Analysis IV*, P. R. Krishnaiah, Ed., p. 193. North-Holland, Amsterdam (1977).
10. C. R. Rao, Prediction of future observations in growth curve models (with discussion), *Stat. Sci.* **2**, 434 (1987).
11. P. A. Lachenbruch, *Discriminant Analysis*. Hafner, New York (1975).
12. E. D. Schneiderman, 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 (1990).
13. M. A. Foulkes and C. E. Davis, An index of tracking for longitudinal data, *Biometrics* **37**, 439 (1981).
14. H. Goldstein, *The Design and Analysis of Longitudinal Studies*. Academic Press, New York (1979).
15. R. C. Elston and J. E. Grizzle, Estimation of time–response curves and their confidence bands, *Biometrics* **18**, 148 (1962).
16. H. Crepeau, J. Kozial, N. Reid and Y. S. Yuh, Analysis of incomplete multivariate data from repeated measurement experiments, *Biometrics* **41**, 505 (1985).
17. R. H. Jones, Serial correlation in unbalanced mixed models, *Bull. Int. Stat. Inst.* **52**, 105 (1987).
18. M. G. Kenward, A method for comparing profiles of repeated measurements, *Appl. Stat.* **36**, 296 (1987).
19. G. O. Zerbe and S. H. Walker, A randomization test for comparison of groups of growth curves with different polynomial design matrices, *Biometrics* **33**, 653 (1977).
20. P. J. Diggle, Testing for random dropouts in repeated measurement data, *Biometrics* **45**, 1255 (1989).

APPENDIX. COMPUTER IMPLEMENTATION

This program can be obtained on a 5.25″ double density floppy disk by sending $10 to defray the cost of handling and licensing fees. The program will ruin on a IBM-PC/XT or AT compatible computer. The computer *must* be equipped with a numerical coprocessor from the 8087 family and 640 K of memory. The computer must be configured so that at least 430 K of memory is available, i.e. not tied up with memory resident programs such as *Windows*. EGA or VGA graphics capability is required to display the color graphics. No additional software is required (other than what one would normally use to enter a data set); run-time modules are supplied with the program so that no compiler or interpreter is necessary. The program, written in GAUSS, version 2.0, revision 20, requires no additional installation or modification, and is run with a single command. When requesting the program, address inquiries to EDS and make cheques payable to Baylor College of Dentistry.

About the Author—EMET DAN SCHNEIDERMAN received a B.A. and M.A. in Anthropology from Northwestern University in 1978, and the Ph.D. in Biological Anthropology from The University of Michigan in 1985. While at The University of Michigan he was affiliated with the Center or Human Growth and Development and began conducting research in the area of craniofacial growth. In collaboration with Joseph Mudar, Schneiderman developed an integrated software system for the analysis of cephalometric radiographs (X-rays of the head). While on the orthodontics faculty of the University of Detroit School of Dentistry from 1985 to 1988, Dr Schneiderman created a computerized cephalometry laboratory. In 1988 Dr Schneiderman went to the Baylor College of Dentistry in Dallas where he is the Director of Research for the Department of Oral and Maxillofacial Surgery. Dr Schneiderman and co-investigator Dr Charles Kowalski have been funded by NIH/NIDR from 1988 to 1993 to conduct this biostatistical research from which this paper issued. Dr Schneiderman has 41 publications including two chapters and the monograph, *Facial Growth in the Rhesus Monkey*, to be published shortly by Princeton University Press.

About the Author—CHARLES J. KOWALSKI received the B.S. in Mathematics from Roosevelt University in Chicago in 1962, the M.S. in Statistics from Michigan State University in 1965, and the Ph.D. in Biostatistics from The University of Michigan in 1968. Dr Kowalski then joined the faculty of the Department of Oral Biology at the University of Michigan School of Dentistry. Dr Kowalski served as the assistant director of the university's Statistical Research Laboratory from 1971 to 1978 and research scientist at the Dental Research Institute from 1978 to the present and directed the institute's biometrics laboratory. At various times Dr Kowalski has served as a

consultant to the National Football League, Park, Davis and Co., Nymegen University, Lanchester Cleft Palate Clinic, the Department of Antiquities of the University of Alexandria in Cairo, the U.S. Veterans Administration and the Eastman Dental Center. Dr Kowalski has published more than 175 scientific papers, including numerous chapters and the book *A Mixed-Longitudinal Interdisciplinary Study of Growth and Development* published by Academic Press in 1979. Dr Kowalski's research has focused on the application of statistical methods to dental and oral research with special emphasis on measurement processes, their validity, reliability and calibration. Longitudinal data analysis and the computer implementation of polynomial growth curve models have also been and continue to be a major thrust of his research. Drs Kowalski and Schneiderman have been funded by NIH/NIDR from 1988 to 1993 to study and implement biostatistical methods for the analysis of longitudinal data in the form of user-friendly microcomputer programs.

About the Author—STEPHEN M. WILLIS received the B.S. degree in Mathematics from the University of Texas at Arlington in 1987. Mr Willis has over 15 years of experience in clinical toxicology and is currently operations manager of a regional toxicology laboratory in Dallas. Mr Willis is also the lead programmer/systems analyst for the NIH/NIDR grant on longitudinal statistical methods with Drs Kowalski and Schneiderman. Mr Willis has played a major role in the development of user-friendly interfaces for programs that have broad applications in the biomedical sciences.

About the Author—THOMAS R. TEN HAVE received the B.A. in Statistics in 1981, an M.P.H. in Biostatistics in 1982, and the Ph.D. in Biostatistics in 1991, all from The University of Michigan. Dr Ten Have has served as a biostatistician at the Center for Human Growth and Development and the Statistical Research Laboratory of The University of Michigan since 1985. Dr Ten Have has recently joined the faculty of the School of Public Health at The University of Michigan. In addition to the analysis of longitudinal craniofacial growth and development data, Dr Ten Have is also conducting research on the longitudinal analysis of categorical data. Dr Ten Have has co-authored more than 20 scientific papers.