LINEAR MODELS WITH NESTED ERROR STRUCTURE

IN PREDICTING VISION LOSS FOR PATIENTS

WITH SUBRETINAL NEOVASCULAR MEMBRANES

LINEAR MODELS WITH NESTED ERROR STRUCTURE IN PREDICTING VISION LOSS FOR PATIENTS WITH SUBRETINAL NEOVASCULAR MEMBRANES

BY

MEIYING HOU

A project

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

Aug. 1992

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Dr. A. Willan for his kindly guidance, assistance and encouragement throughout the length of this project. I also wish to thank Dr. P. Macdonald for his valuable comments and suggestions, and his help for my writting.

My heartful gratitude reserve to my parents who give me moral support throughout university career in Canada, to my brother and sister-in-law who provide financial assistance for my study.

finally I want to thank the Dept. of Chem. Eng. for its equipment of computer used in my project typing.

MASTER OF SCIENCE (1992)
(STATISTICS)

McMASTER UNIVERSITY
Hamilton, Ontario

Title: Linear models with nested error structure inpredicting vision loss for patients with subretinal neovascular membranes

Author: Meiying Hou, B.S. Computer Software Engineering

Supervisor: Dr. A. Willan

Number of Pages: v, 32

Table of Contents

LINEAR MODELS WITH NESTED ERROR STRUCTURE IN PREDICTING VISION LOSS FOR PATIENTS WITH SUBRETINAL NEOVASCULAR MEMBRANES

Abstract

Age-related macular degeneration (AMD)* and presumed ocular histoplasmosis (POHS) are common causes of macular degeneration. Both are major causes of blindness, with AMD being the leading cause of blindness in people over the age of 65. The major cause of visual loss in both categories is the presence of a subretinal neovascular membrane (NVM) in the macular. Sometimes these conditions can be treated successfully with laser therapy. Our task was to develop a regression model for predicting post-treatment vision as a function of time from treatment and baseline prognostic factors measured at diagnosis. The particular analysis of the model was to examine how patients' post-treatment vision is affected by baseline factors. A nested-error structure was used in a linear model.

*: Abbreviations see Table A.

I. Medical Background

Age-related macular degeneration (AMD) and presumed ocular histoplasmosis (POHS) are common causes of macular

degeneration. Both are major causes of blindness with AMD being the leading cause of blindness for those over 65 years of age. The major cause of visual loss in both conditions is the presence of a subretinal neovascular membrane (NVM) in the macular with resultant bleeding under and scarring of the retina. These conditions can be treated successfully with laser therapy in some cases. The goals of treatment are to obliterate completely the subretinal neovascular membrane without damaging the foveal ascular zone (FAZ).

Patients for this study were those recruited for the Canadian Ophthalmology Study Group trial. All patients had a suspected subretinal neovascular membrane associated with age-related macular degeneration or presumed ocular histoplasmosis.

Visual acuity (vision) was measured following refraction using the Early Treatment Diabetic Retinopathy Study Chart. Vision was recorded as the total number of letters the patient can read. Patients had their vision measured just prior to treatment and again at 3 months, 6 months, 12 months, 18 months, 24 months, 30 months and 36 months following treatment.

Laser therapy was performed with either Aryon-green or Krypton-red photocoagulators according to a standard protocol for each instrument. Treatment will continue until the

membrane is completely closed or until it grows to involve the centre of the fovea.

II. Purpose of Analysis

The object of this study is to determine which factors predicted visual loss, and to develop a regression model for predicting vision as a function of baseline prognostic factors and time from treatment.

The prognostic variables to be considered are diagnostic category (AMD or POHS), diameter of the NVM, distance from the foveal edge of the NVM to the centre of the FAZ, duration measured in days between first symptom and diagnosis, and time measured in years between baseline vision and follow-up vision. Patients belonged in one of two categories, and distance ranged from 200 to 2500 microns. Distribution of number and average duration of patients in each category and each distance range is shown in Table 1.

We wish to examine how patients' post-treatment vision is affected by baseline factors and the time from treatment. Two methods were used to achieve this. In the first method, the slope of the regression line of vision on time were calculated

for each patient. Then the effect of the baseline factors on the slope was examined. In the second method all the observations from all patients were used in a regression analysis using a nested error structure. A nested error structure was used to account for the correlation between observations on the same patient. The error structure for the ith observation on the jth subject was assumed to be $s_j + e_{ij}$, where s_j follows a normal distribution with mean zero and variance σ^2_s, and e_{ij} follows a normal distribution with mean zero and variance σ^2_e.

TABLE - 1.

	AMD		POHS		TOTAL	
200 -- 500	89	52.37	49	54.59	138	53.16
501 -- 1000	73	58.25	58	31.38	131	46.35
1001 -- 1500	13	71.08	13	22.46	26	46.77
1501 -- 2000	10	54.90	2	29.50	12	50.67
2001 -- 2500	4	59.75	7	64.00	11	61.45
TOTAL	189	56.22	129	41.04	318	50.06

Number of patients and average duration (days) between first symptom and diagnosis in each category and distance group. First entry is number of patient and second is average duration.

III. Theoretical Considerations

1. Method of slopes.

The first method considered for this study was to calculate the slope for the regression of vision on time, and the variance of slope for each patient, as follows

$$V_{ij} = b_{0i} + b_i t_{ij} + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0, \sigma^2)$$

$$\hat{b}_i = \frac{\Sigma^{n}_{j=1}(V_{ij} - \bar{V}_i)(t_{ij} - \bar{t}_i)}{\Sigma^{ni}_{j=1}(t_{ij} - \bar{t}_i)^2}$$

$$\text{var}(\hat{b}_i) = \frac{\Sigma^{n}_{j=1}(V_{ij} - \bar{V}_i)^2}{(n_i - 2)\Sigma^{ni}_{j=1}(t_{ij} - \bar{t}_i)^2}$$

The slope, b_i, was calculated from the regression of vision versus time, where vision is measured by the number of letters the patient can read and time is the number of years between baseline vision and observation vision.

We used the method of least squares to fit general linear models.

Let $B = X\Theta + \varepsilon$ with $E(B) = X\Theta$

where

$$B = \begin{pmatrix} b_1 \\ b_2 \\ : \\ : \\ b_n \end{pmatrix} \qquad X = \begin{pmatrix} x_{10} & x_{11} & .. & .. & .. & x_{1p} \\ x_{20} & x_{21} & .. & .. & .. & x_{2p} \\ .. & .. & & & .. & \\ .. & .. & & & .. & \\ x_{n0} & x_{n1} & .. & .. & .. & x_{np} \end{pmatrix}$$

$$\Theta = \begin{pmatrix} \Theta_0 \\ \Theta_1 \\ : \\ : \\ \Theta_p \end{pmatrix} \qquad \varepsilon = \begin{pmatrix} e_1 \\ e_2 \\ : \\ : \\ e_n \end{pmatrix}$$

where b_i is the slope of the ith patient and $(x_{i0}, x_{i1}, \ldots, x_{ip})$ is the vector of predictors.

The ordinary least squares used involves choosing Θ as the value of Θ which minimizes the sum of squares of deviations of the observations from their expected values, i.e., choose Θ as that Θ which minimizes

$$\Sigma^{n}_{i=1}[b_i - E(b_i)]^2 = (B - X\Theta)'(B - X\Theta)$$

The resulting estimator is

$$\hat{\Theta} = (X'X)^{-1}X'B$$

However, since the weighted least squares estimates are best linear unbiased estimators if the weights for the observations are proportional to the reciprocals of the error variances, we adopted the BLUE to get the estimates of Θ. A weighted residual sum of squares $\Sigma^{n}_{i=1}W_i(b_i - E(b_i))^2$ is minimized, where

$$W = 1/(\text{variance of slope})$$

calculated from data. So the weighted normal equations used are:

$$\hat{\Theta} = (X'WX)^{-1}X'WB$$

This $\hat{\Theta}$ is the best linear unbiased estimator. To test the null hypothesis H_0: $\Theta_j = 0$, we define the p-value to be twice the area the right of $|t|$ under the curve of the t-distribution having (n-p) degrees of freedom. If the inference assumptions are satisfied, we can reject H_0: $\Theta_j = 0$ in favour of H_1: $\Theta_j \neq 0$ by setting the p-value of type I error equal to α if and only if following condition hold:

p-value < α

We use this method to choose our final model for which the factors are significant.

2. Regression with nested error structure

Frequently data arise from the random selection of "individuals" on which several "measurements" are made. So a sample of patients may be selected and vision measurements taken for the individuals over several years in the study of the relationship between vision and time. This "nesting" pattern by which the data are generated has a significant bearing on the statistical model that is appropriate for valid analyses of the data.

In the presentation of the statistical model for the analysis of observations that arise in a one-fold nested structure, we denote the variable under study by the letter y with two subscripts. The first subscript distinguishes the individual (patient) in the sample, and the second subscript distinguishes the measurement (observation) for the particular individual. We assume the N individuals are selected at random with eligibility criteria and that n_i measurements are made on the ith individual. The linear model is expressed as

$$y_{ij} = \Sigma^{p}_{k=1} x_{ijk}\beta_k + u_{ij} \qquad (3.2.1)$$

$i = 1,2,\ldots,N$

$j = 1,2,\ldots,n_i$

and

$$u_{ij} = s_i + e_{ij} \qquad (3.2.2)$$

where,

y_{ij} denotes the value of the jth measurement for the ith individual.

x_{ijk}, k=1,2,...,p, denotes the levels of the p predictor variables at which the observation y_{ij} is obtained.

β_k, k=1,2,...,p, denotes the unknown regression coefficient to be estimated.

and

u_{ij}, the random error associated with y_{ij}, is assumed the sum of the random effect associated with ith sample individual (s_i) and the random effect associated with the jth measurement for the ith individual in the sample (e_{ij}).

The random errors s_i and e_{ij} are assumed to be independently normal distributed with means zero and variances σ^2_s and σ^2_e respectively, where $\sigma^2_s \geq 0$ and $\sigma^2_e > 0$. The covariance structure for the random errors u_{ij} is thus expressed by

$$E(u_{ij}u_{i'j'}) = \sigma^2_\varepsilon + \sigma^2_e \quad \text{if } i=i',\ j=j'$$
$$= \sigma^2_\varepsilon \quad \text{if } i=i',\ j\neq j'$$
$$= 0 \quad \text{if } i\neq i'$$

See Fuller & Battese (1973). As proposed by Fuller & Battse (1973) we transform (3.2.1) into the regression equation

$$y_{ij} - \alpha_i\bar{y}_{i.} = \Sigma^p_{k=1}(x_{ijk} - \alpha_i\bar{x}_{i.k})\beta_k + u^*_{ij} \qquad (3.2.3)$$

where

$$\alpha_i = 1-[\sigma^2_e/(\sigma^2_e+n_i\sigma^2_\varepsilon]^{\frac{1}{2}} \qquad (3.2.4)$$

and $\bar{y}_{i.},\bar{x}_{i.k}$, $k=1,2,\ldots,p$, denote the averages of the n_i y-and x-measurements on the ith individual. The errors, u^*_{ij}, are uncorrelated with the variances σ^2_e, and the ß parameters in (3.2.3) are identical to these in (3.2.1).

We write the linear model (3.2.1) as

$$Y_{ij} = \Sigma X_{ijk}\beta_k + u^*_{ij} \qquad (3.2.5)$$

where,

$$E(u^*u^{*\prime}) = var(u^*) = I\sigma^2_e$$

Since the variance components σ^2_s and σ^2_e are unknown, the values of the transformation factors α_i defined in (3.2.4) must be estimated from estimates of σ^2_s and σ^2_e. Fuller and

Battese (1973) used the "fitting-of-constants" method suggested by Henderson (1953) and discussed by Searle (1971)*. By the regressing the y-deviations, $y_{ij}-y_{i.}$, on the x-deviations, $x_{ijk}-x_{i.k}$, k=1,2,...,p, that are not identically zero, we obtain the unbiased estimator for σ^2_e

$$\hat{\sigma}^2_e = \hat{e}'\hat{e} \;/\; (N_1 - N - p + \phi_1)$$

where,

$N_1 = \Sigma^N_{i=1} n_i$

$\hat{e}'\hat{e}$ denotes the residual sum of squares obtained

from the regression and ϕ_1 is the number of x-variables which are a linear combination of the indicator variables for individuals. The variance component σ^2_s is unbiasedly estimated by

$$\hat{\sigma}^2_s = \frac{\hat{u}'\hat{u} - (N_1 - p)\hat{\sigma}^2_e}{N_1 - tr[(X'X)^{-1}\Sigma^N_{i=1} n^2_i \bar{x}_{i.}'\bar{x}_{i.}]}$$

where $\hat{u}'\hat{u}$ denotes the residual sum of squares from the regression of Y on X, and $\bar{x}_{i.}$ denotes the (1xp) vector having kth element $x_{i.k}$, k=1,2,...,p.

*: "fitting-of-constant" was presented by Searle (1971). The value of this method is to yields estimators of the variance components unaffected by the fixed effects.

3. Alternatives

An alternative approach of predicting a future measurement on an individual given the past measurements will be considered here. It was introduced by Rao (1987) who presented several papers on prediction of future observations from linear models. Statistical techniques have also been proposed by Barndorff-Nielsen (1981), Bock (1976), Geisser (1975), Hinkley (1979), Lee (1972) and Young (1977).

Rao (1987) gave some formulae for predicting future observations in a linear model, and compared different formulae by applying them on empirical data relating to biological growth. The method of principal components is used to estimate the coefficients of a linear model when the coefficients are not specified. Rao (1987) also assessed the efficiencies of different methods of prediction by cross-validation or leave-one-out technique. Bayesian and empirical Bayesian methods were used to estimate unknown parameters.

Rao's method can be summaried as follows:
Consider the linear model

$$Y = X\beta + E \qquad (3.3.1)$$

where Y and E are px1 vectors, X is a pxn matrix and ß is an

nx1 vector. Further, let y be a vector of k random variables with a Gauss-Markoff structure

$$y = x\beta + e \qquad (3.3.2)$$

where ß is the same parameter as above. The problem we consider is that of predicting y be a linear function of Y depending on the nature of information available on X, ß, E and e.

We note that the problem of finding an optimum predictor y under the loss function

$$(y - \hat{y})'G(y - \hat{y}), \qquad (3.3.3)$$

where G is a positive definite matrix, is equivalent to that of optimum prediction of each component of y under a quadratic loss function. Thus the solution under the loss function (3.3.3) is independent of G. We shall, therefore, consider y to be a single future observation to be predicted.

Let the dispersion matrix of (E,e) given ß be written in the partitioned form, apart from a multiplier σ^2,

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \qquad (3.3.4)$$

where V_{ij} and σ^2 do not depend on ß. Further let ß be a random variable with mean u and dispersion matrix $\sigma^2 F$. We shall assume that (V_{ij}), F and X are all full rank to avoid some complications.

When all parameters are known, the best linear predictor (BLP) of y under quadratic loss function is the regression of y on Y,

$$xu+(xFX'+V_{21})(XFX'+V_{11})^{-1}(Y-Xu), \qquad (3.3.5)$$

and the associated prediction mean square error is

$$\sigma^2[V+xFx'-(xFX'+V_{21})(XFX'+V_{11})^{-1}(XFx'+V_{12})] \quad (3.3.6)$$

Let us denote by $\hat{ß}$, the least squares estimator of ß from (3.3.1).

$$\hat{ß} = (X'V^{-1}{}_{11}X)^{-1}X'V^{-1}{}_{11}Y = UX'V^{-1}{}_{11}Y$$

where $U = (X'V^{-1}{}_{11}X)$. Then (3.3.5) can be written as the sum of three expressions:

$$x\hat{\beta}$$

$$-(x - V_{21}V^{-1}_{11}X)U(F + U)^{-1}(\hat{\beta} - u)$$

$$+ V_{21}V^{-1}_{11}(Y-X\hat{\beta}) \qquad (3.3.7)$$

The prediction mean square error (3.3.6) can also be written, apart from the multiplier σ^2, as

$$V_{22}-xUX'V^{-1}_{11}V_{12}-V_{21}V^{-1}_{11}XUx'+xUx'$$

$$-(x-V_{21}V^{-1}_{11}X)U(F+U)^{-1}U(x'-XV^{-1}_{11}V_{12})$$

$$-V_{21}(V^{-1}_{11}-V^{-1}_{11}XUX'V^{-1}_{11})V_{12} \qquad (3.3.8)$$

corresponding to the three terms in (3.3.7). The expression (3.3.5) can be also written as

$$xu + (xF+V_{21}V^{-1}_{11}XU)(F+U)^{-1}(\hat{\beta}-u) \qquad (3.3.9)$$

$$+ V_{21}V^{-1}_{11}(Y-X\beta) \qquad (3.3.10)$$

where (3.3.9) is the regression of y on ß and (3.3.10) is the regression of y on the residual (Y-Xß).

The BLP depends on all the parameters (V_{ij}), u, σ^2 and F whose values may not be known in any particular situation. If past data on the linear model (3.3.1) are available, it may be possible to estimate the unknowns, substitue the estimates for parameters in the formula for the BLP and thus obtain am empirical best linear predictor.

Rao (1987) suggested that in the absence of any information on

the stochastic process describing an individual's growth, a standard approach to the prediction problem is to consider the joint distribution of Y = W and $(Y_1,Y_2,\ldots,Y_p)$ = U over the individuals of the relevant population and derive the conditional distribution of W given U for use in prediction.

IV. Results and Discussion

We used two methods to analyze this study.

1. The method of slope.

There are two data files collected for the Canadian Ophthalmology Study Group trial. One is the patients data set, called PATIENT.dat. It has a record for each patient including patient identifier, category (AMD vs POHS), diameter, distance, duration, baseline vision, slope of regression line of vision on time, and weight (1/variance of slope).

The GLM procedure in SAS was employed to fit general linear models with slope as the dependent variable. The GLM procedure uses the method of least squares and allows many different analyses, such as simple regression, multiple regression, analysis of variance, analysis of covariance, weighted regression and so on. In addition, The GLM procedure allows the specification of any degree of interaction (cross effects)

and nested effects. It also provides for continuous-nesting effects. Through the concept of estimability, the GLM procedure can provide tests of hypotheses for the effects of a linear model. The GLM prints not only the sum of squares (SS) associated with each hypothesis tested but also upon request the form of the estimate function employed in the test. The GLM can produce the general form of all estimable functions.

We chose slope as the dependent variable and used category, diameter, distance, duration, and baseline vision as independent variables respectively. The program is described briefly as follows writing by SAS.

```
DATA = PATIENT.DAT
PROC GLM
MODEL SLOPE = CATEGORY
WEIGHT WGHT
;
MODEL SLOPE = DIAMETER
WEIGHT WGHT
;
MODEL SLOPE = DISTANCE
WEIGHT WGHT
;
```

:

:

Where, WGHT = 1/(variance of slope)

We set $\alpha = 0.05$ to test each hypothesis. If p-value of an independent variable was less than 0.05 (two tailed test), then we included it in the model. We used a step-foward approach to model building. In the first step each independent variable was entered in a model as a single factor. The most significant of these was category ($p < 0.0001$), see Table 2.

Through running this program with one independent variable at a time, we found only category, distance and duration are important factors, ie, the associated p-value is less than 0.05. The p-values of category, distance, and duration are 0.0001, 0.0079 and 0.0367, respectively. Therefore we have strong evidence that category, distance and duration are significantly related to Y, the slope.

In the next step we calculated the p-value of each independent variable in a model that included category. The only significant variable in this step was distance ($p = 0.0046$), see Table 2.

TABLE - 2

	Step1	Step2	Step3	Step4	Step5	Step6	Step7	Step8
CATE	.0001	*	.0001	.0001	*	*	.0001	*
DIAM	.2178	.9178	.1267	.2562	.6907	.9191	.1522	.6927
DIST	.0079	.0046	*	.0075	*	.0050	*	*
DURA	.0367	.3033	.0399	*	.3304	*	*	*
TRET	.7135	.8182	.8425	.6471	.9620	.7787	.7715	.9220
VISI	.0541	.9800	.1217	.9880	.6020	.9252	.2009	.5577

P-values at each step of independent variable.

The terms marked by * are included in the model with any other factors. In the first column of Table 2 the p-value of each variable entered by itself is given. In the second column the p-value given for each variable is the p-value with category already in the model. In the column 3 the p-value given for each variable is the p-value with distance already in the model, and so on.

The final best model is

$$E(\text{slope}) = \Theta_0 + \Theta_1 \text{category} + \Theta_2 \text{distance}$$

We conclude the category and distance are the only significant factors.

Category and distance are main effects and there is no interaction between them, i.e., the causes of visual loss with subretinal neovascularization and distance from the foveal edge to the centre of the foveal ascular zone are important components in this study, and the effect of distance is the same for all the causes of visual loss. Rate of visual loss was greater in patients with AMD and for patients whose NVM was closer to the FAZ.

2. Regression with nested error structure.

Let us analyze this study from another point of view. Since the data arose from a random selection of patients (189 AMD patients and 129 POHS), for which several vision measurement were taken over a period of up to 36 months, we can regard it as a linear model with nested error structure.

The second data file used is VISIT.dat which recorded several vision measurements for each of 318 patients for a total of 2823 observations. Each patient may be regarded as a cluster, providing different values of the dependent variable. There are 1702 observations with AMD and 1121 observations with POHS in this data file.

Model: $y_{ij} = \Sigma^{p}_{k=1} x_{ijk}\beta_k + u_{ij}$

Thus the intra-class correlation coefficient (i.c.c.c.) is given by

$$\text{i.c.c.c.} = \frac{\sigma^2_s}{\sigma^2_s + \sigma^2_e}$$

So we can test a null hypothesis H_0: $\beta_j = 0$ by defining the t-test statistics

$$t = \frac{\hat{\beta}_k}{s\sqrt{c_{kk}}} \qquad (4.2.3)$$

where, s is standard error, defined as

$$s = \sqrt{S^2} = \frac{\sqrt{SSE}}{\sqrt{n-p}} \qquad (4.2.4)$$

$$SSE = \Sigma^{ni}_{j=1}(y_{ij} - \bar{y}_{i.})$$

and c_{kk} is jth diagonal element of $(X'X)^{-1}$. If $|t| > t^{(n-p)}_{[\alpha/2]}$ holds, we can reject H_0: $\beta_k=0$ in favour of H_1: $\beta_k\neq0$ by setting the p-value of type I error equal to α.

The first model we considered is:

$$y = x\beta + u$$

i.e.,

$$\text{Vision} = \beta_0 + \beta_1\text{category} + \beta_2\text{distance} + \beta_3\text{time} + \varepsilon + e \qquad (4.2.5)$$

or,

$$y_{ij} = \beta_0 + \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ji3} + \varepsilon_i + e_{ij}$$

where, $y_{ij} = \text{vision}_{ij}$

$x_{ij1} = \text{category}_{ij}$ (CATE_{ij})

$x_{ij2} = \text{distance}_{ij}$ (DIST_{ij})

$x_{ij3} = \text{time}_{ij}$ (TIME_{ij})

$i=1,2,\ldots,N$ number of patients

$j=1,2,\ldots,n_i$ number of objects on patient i

The numbers of vision measurements taken are not same for all the patients. However,

$$\text{CATE}_{i1} = \text{CATE}_{i2} = \ldots = \text{CATE}_{ini}$$

$$\text{DIST}_{i1} = \text{DIST}_{i2} = \ldots = \text{DIST}_{ini}$$

$$\text{TIME}_{ij} \neq \text{TIME}_{ij'} \qquad \text{when } j \neq j'$$

The program SUPER CARP was used to fit this model to the data points defined by 2823 values of (y, x_1, x_2, x_3), where x_1,x_2 and x_3 denote the potential confounders category (0 = AMD, 1 = POHS), distance (microns), and time (year) respectively.

To assess the statistical significant of the independent variables x_1, x_2 and x_3, we inspect the estimated generalized least-squares coefficients, their associated standard errors and corresponding t-statistic, as presented in Table 3. The intra-class correlation coefficient is estimated as 0.579.

Table 3

VARIABLE	COEFFICIENT	STD. ERROR	t-STATISTIC
INTERCEPT	48.068	1.875	25.636
CATEGORY	17.106	1.981	8.634
DISTANCE	0.009	0.002	4.421
TIME	-5.947	0.297	-20.004

Since the t-statistic of category is 8.634, we conclude that the patients with AMD have significantly worse vision than patients with POHS, adjusting for distance and time.

Although the coefficient of distance, given by $ß_2 = 0.009$, is very small, it is significant in this model since $t=4.421>1.96$ with 2814 the degrees of freedom. So we can reject H_0: $ß_2=0$ in favour of H_1: $ß_2\neq0$, and conclude that patients with smaller distance have poorer followup vision. The t-value indicates that time is by far the most important predictor of vision Y. Since the coefficient of time, $ß_3 = -5.947$, is negative, we

conclude that the visual acuity decreases as time increases. We want to know that the relationship of vision and time is presented by straight line or by curve, so we tried next model:

$$E(y_{ij}) = \beta_0 + \beta_1 CATE + \beta_2 DIST + \beta_3 TIME + \beta_4 TIME^2$$

The corresponding results is given by Table 4.

Table 4

VARIABLE	COEFFICIENT	STD. ERROR	t-STATISTIC
INTECEPT	49.316	1.889	26.104
CATEGORY	17.051	1.983	8.599
DISTANCE	0.009	0.002	4.440
TIME	-11.112	0.956	-11.636
TIMESQUA	1.933	0.339	5.689

$P < 0.05$

From Table 4, we found t-value of time^2 is 5.689, ie, the effect of time^2 is significant in this model. The loss of vision with increasing time is represented by a concave-up curve.

We use a forward stepwise approach, considering all the interactions between category, distance and time, adding the most significant terms, one at a time. This leads to the following model:

$$E(VISION) = \beta_0 + \beta_1 CATE + \beta_2 DIST + \beta_3 TIME + \beta_4 TIME^2 + \beta_5 TIME*DIST + \beta_6 TIME*CATE + \beta_7 TIME^2*CATE + \beta_8 TIME^2*DIST \qquad (4.2.6)$$

The result is shown in Table 5. We conclude that any cross-product terms on category, distance and time are highly significant. The cause of vision loss depends on the causes of macular degeneration, distance from the centre of the foveal avascular zone and years between baseline and observations.

For AMD patients (category = 0), using the coefficients from Table 5, equation 4.2.6 becomes

$$E(VISION)=(55.0+0.00575*DIST)+(-23.6+0.00829*DIST)*TIME + (4.94-0.00237*DIST)*TIME^2$$

For POHS patients (category = 1), equation 4.2.6 becomes

$$E(VISION)=(63.7+0.00575*DIST)+(-6.99+0.00829*DIST)*TIME + (1.56-0.00237*DIST)*TIME^2$$

TABLE - 5

EFFECT	COEFFICIENT	T-STATISTIC	P-VALUE
INTECEPT	54.98889	28.07291	<0.00001
CATEGORY	8.76783	4.20944	0.0001
DISTANCE	0.00575	2.67756	0.001
TIME	-23.64002	-13.56672	<0.00001
TIMESQUA	4.93554	7.87168	<0.00001
TIMEDIST	0.00829	4.56533	<0.00001
CATETIME	16.65519	8.90217	<0.00001
TISQCATE	-3.36927	-5.07531	<0.00001
TISQDIST	-0.00237	-3.71227	0.001

Thus we can see that for both categories VISION is a quadratic in TIME where the coefficients of the quadratic, while dependent on category, are affected by distance in the same way.

These relationships between VISION, DISTANCE, CATEGORY and TIME are illustrated in Figure 1.

V. CONCLUSION

The method of slope is used to analyze the relation between vision and time. Using slope as the dependent variable in the model showed the vision was affected not only by baseline factors also by time.

Using methods proposed by Fuller and Battese (1973), a model was developed by considering a time-squared term to allow for curvature and all possible interactions in a forward stepwise procedure. The final model includes terms for time, time-squared, patient category, distance, time by distance and category interaction and a time-squared by distance and category interaction. The model allows us to conclude that patients with AMD have poorer vision, although vision deteriorates at about the same rate in both categories; that the slope of the final model is flatter as the distance goes up; that the vision of patients whose subretinal neovascular membrane (NVM) is close to the foveal ascular zone (FAZ) deteriorates at a fast rate, with the rate of deterioration declining over time; and that the vision in patients whose NVM is far from the FAZ deteriorates at a slow but constant rate. By observing the quadratic relationship between vision and time in the second analysis, we found method of slope was invalidated and linear-relation turned to non-linear.

Because the model was supposed as linear regression and the coefficient of time-square was very significant, so the relation between VISION and TIME on POHS is not deeply down, whereas flatter up when the distances go up (greater than 1000). See figure 2. It means the model we supposed does not fit very well. We may try to use other regression model, such exponetial as following:

$$\text{Vision} = f_1(t) + f_2(t)e^{-f_3 t} + E$$

to fit data later.

TABLE A

1. AMD: age-related macular degeneration
2. POHS: presumed ocular histoplasmosis
3. FAZ: foveal ascular zone
4. NVM: neovascular membrane

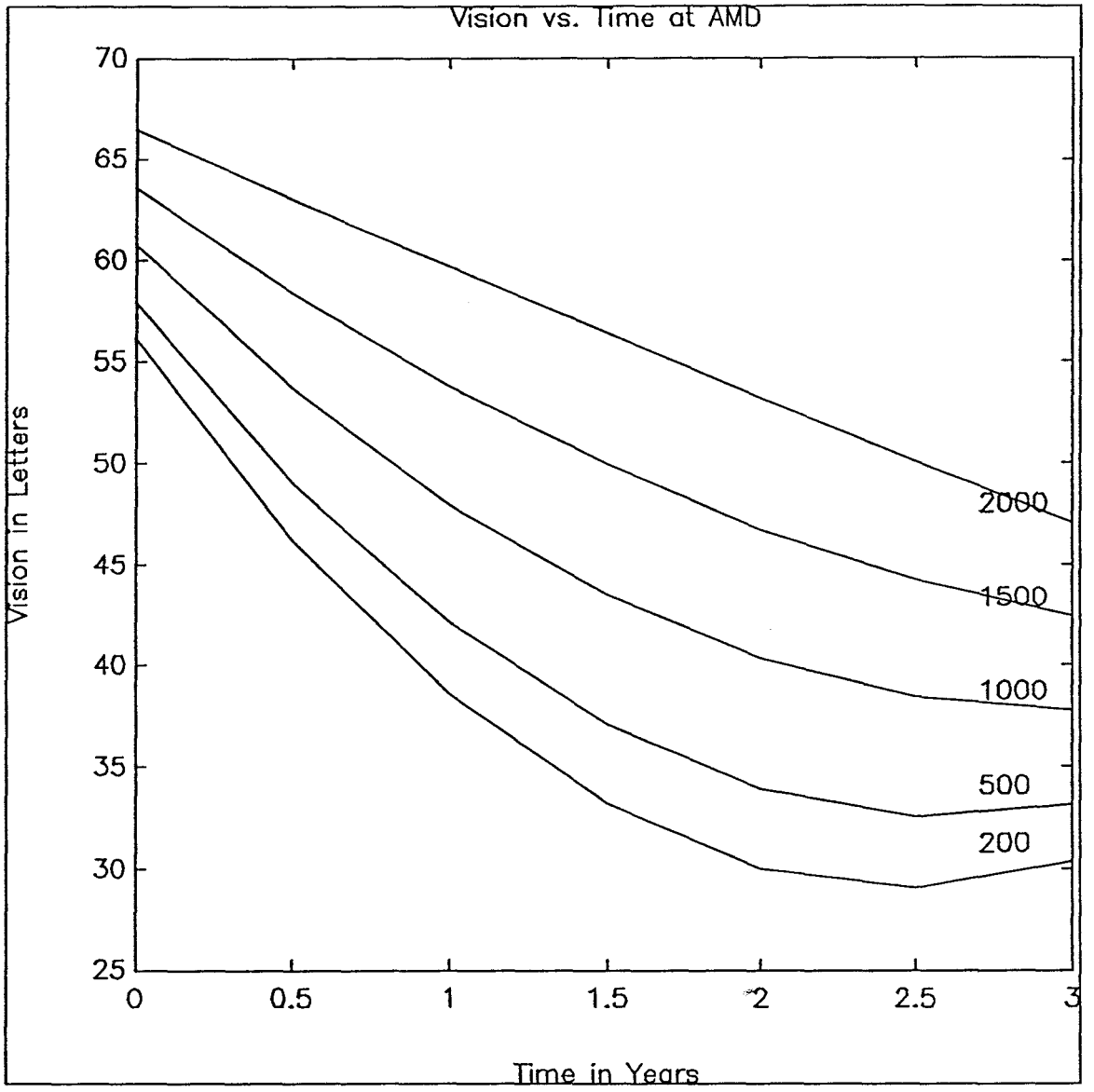
Vision vs. Time at AMD
Vision in Letters
Time in Years
70
65
60
55
50
45
40
35
30
25
0
0.5
1
1.5
2
2.5
3
2000
1500
1000
500
200

figure 1.a

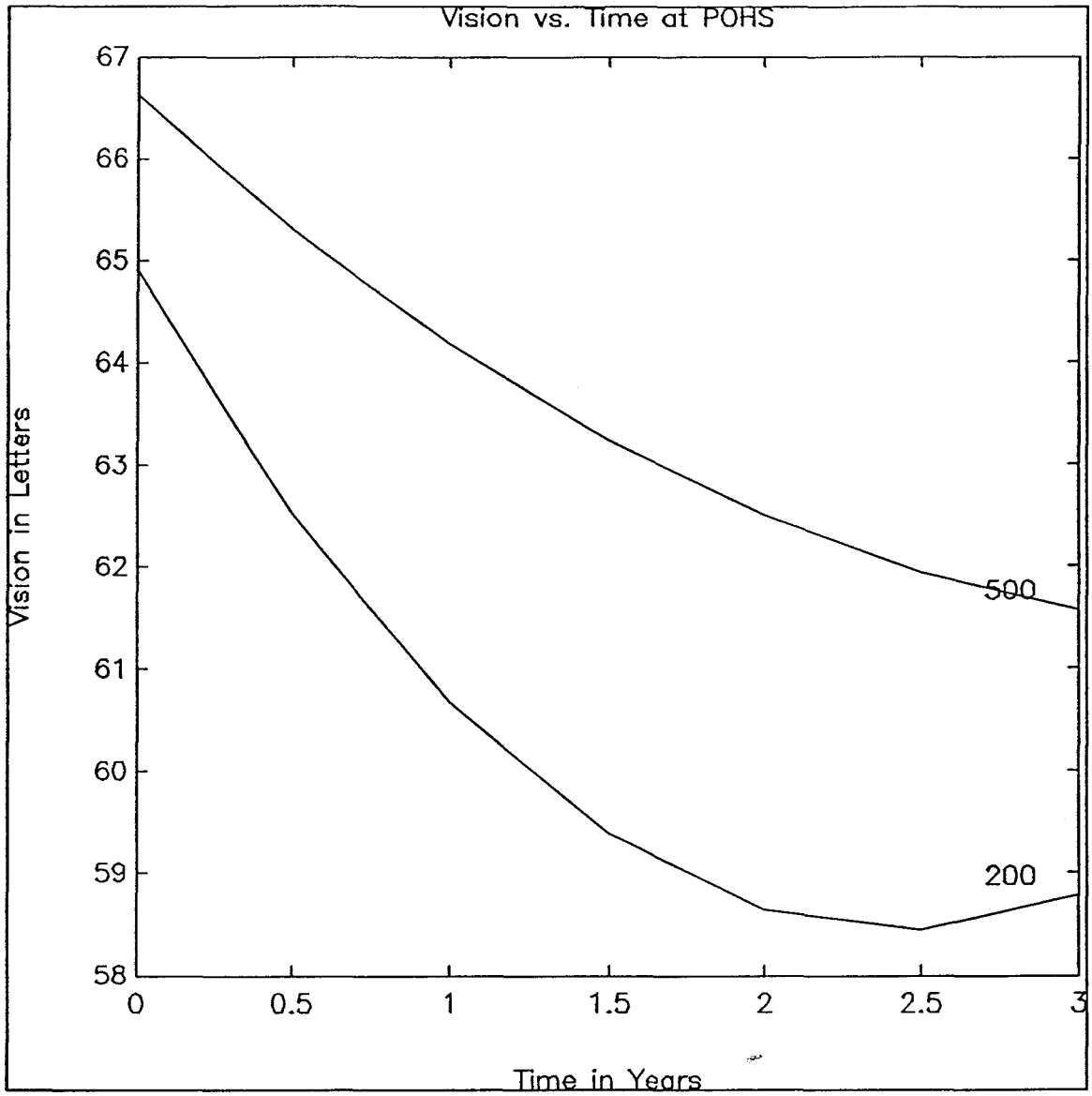
Vision vs. Time at POHS
Vision in Letters
67
66
65
64
63
62
61
60
59
58
0
0.5
1
1.5
2
2.5
3
Time in Years
500
200

figure 1.b

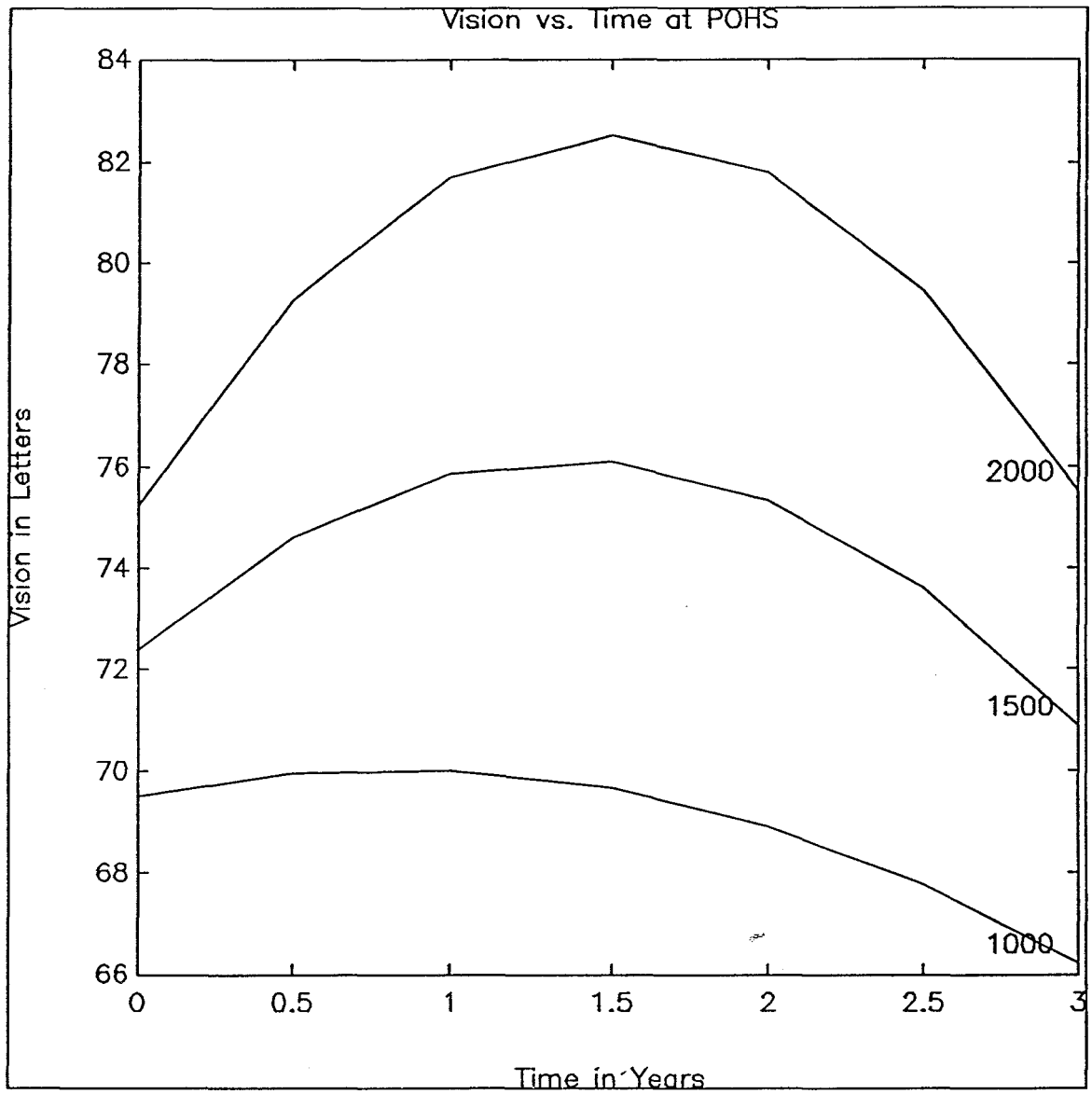
Vision vs. Time at POHS
84
82
80
78
76
74
72
70
68
66
0
0.5
1
1.5
2
2.5
3
Vision in Letters
Time in Years
2000
1500
1000

figure 2

REFERENCE

Bentler, P.M., Some Contributions to Efficient Statistics in Structural Models: Specification and estimation of moment structures. Psychometrika, Vol. 48, 493-517, 1983.

Donner, A., A Regression Approach to the Analysis of Data Arising from Cluster Randomization, International Journal of Epidemiology, Vol.14, No.2, 322-326, 1985.

Draper, N. and Smith, H., Applied Regression Analysis, 2nd ed. New York, John Wiley, 1981.

Elston, R.C. and Grizzle, .E., Estimation of Time-Response Curves and Confidence bands. Biometrics, Vol.18, 148-159, 1962.

Fuller, W.A. and Battese, G.E., Transformations for Estimation of Linear Models with Nested-Error Structure, Journal of the American Statistical Association, Vol. 68, 626-632, 1973.

Graybill, F.A., An Introduction to Linear Statistical Models, Vol. 1, New York, McGraw Hill Book Co., 1961.

Green, W.R., McDonnell, P. and Yeo, J.H., Pathologic Features of Senile Macular Degeneration, Ophthalmology, Vol.92, No.5,

615-627, 1985.

Grizzle, J.E., and Allen, D.M., Analysis of Growth and Dose Response Curves. Biometrics, Vol.25, 357-382, 1969.

Henderson, C.R., Estimation of Variance and Covariance Components, Biometrics, Vol.9, 226-252, 1953.

Hocking, R.R., The Analysis and Selectoin of Variables in Linear Regression, Biometrics, Vol.32, 1-49, 1976.

Kempthorne, O., Design and Analysis of Experiments, New York: John Wiley and Sons, Inc., 1952.

Maguire, J.I., Benson, W.E. and Brown, G.C., Treatment of Foveal Pigment Epithelial Detachments with Contiguous Extrafoveal Choroidal Neovascular Membranes, American Journal of Ophthalmology, Vol.109, 523-529, 1990.

Rao, C.R., Prediction of Future Observations in Growth Curve Models, Statistical Science, Vol2, No.4, 434-471, 1987.

SAS Software Programming, version 6.1.

Scott, A.J. and Holt, B., The Effect of Two-Stage Sampling on Ordinary Least Squares Methods, Joural of the American

Statistical Association, Vol. 77, 848-854,1982.

Searle, S.R., Topics in Variance Component Estimation, Biometrics, Vol.39, 1-76, 1971.

Sorbom, D., A General Method for Studying Differences in Factor Means and Factor Structure Between Groups, British J. Math. Statist. Psych, Vol.27, 229-239, 1974.

The Canadian Ophthalmology Study Group, Argon Green versus Krypton Red Laser in the Treatment of Age-Related Macular Degeneration and Presumed Ocular Histoplasmosis, June, 1987.