MODELLING HEARING THRESHOLDS IN THE ELDERLY

CHRISTOPHER H. MORRELL
Loyola College in Maryland, Mathematical Sciences Department, 4501 North Charles Street, Baltimore, MD 21210, U.S.A.

AND

LARRY J. BRANT
National Institute on Aging, Gerontology Research Center, 4940 Eastern Avenue, Baltimore, MD 21224, U.S.A.

SUMMARY

This paper concerns a linear mixed-effects repeated measures model in the analysis of a large data set with over 17,000 observations in a longitudinal study of pure-tone hearing perception in the elderly. The repeated measurements are described by fixed and random components in the model. The fixed effects include the age at entry, time of follow-up, a quadratic component in natural logarithm of frequency, a component to allow for participants with hearing impairments, as well as interaction terms between age and frequency and between impairment and frequency. The random factors include a term for subject, a time component and a frequency component. The analysis shows that hearing impaired individuals have similar patterns of hearing loss over time but, on average, have higher hearing thresholds than normal individuals. Estimation of the random effects in the model by restricted maximum likelihood (REML) using the Newton–Raphson method made possible the analysis of this large data set with speed and efficiency.

1. INTRODUCTION

The characteristics of age-related hearing loss or presbycusis, like other age-related changes in the human body, involve different patterns of loss among individuals. Not only do individuals age differently, but the patterns of ageing for a given individual can differ among different organ systems. Longitudinal studies are the only method to study individual change directly and to identify factors associated with that change. Recent models for longitudinal data analysis have contributed greatly to the study of individual patterns of change by providing flexibility, allowing for unbalanced designs and incomplete data, and modelling the correlation between repeated observations for a subject. For example, the basis for the linear mixed-effects model is the assumption that the population is heterogeneous with respect to some set of latent or unknown characteristics that may affect the outcome over time. Thus it is now possible to study situations where individuals have different patterns of change and to identify factors that may affect the pattern of change.

Three recent papers demonstrate the use of the linear mixed-effects model proposed by Rao and further developed by Harville and by Laird and Ware to study longitudinal data. Two of them study pulmonary function and the third studies the effect of blood lead concentration on cognitive development. While the first two papers consider individual vari-
ability by including random effects in their models, only the third uses the estimates of the random effects to detect outlying individuals. In addition these papers, like most published papers using the linear mixed-effects model, do not consider large data sets. For example, Diem and Liukkonen\textsuperscript{6} had 223 usable cases which resulted in a data set with at most 2007 observations; Vacek et al.\textsuperscript{7} used 95 observations; and Wateraux et al.\textsuperscript{8} had 603 observations in their study.

For studies with greater numbers of observations, which may have numerous parameters to estimate and/or an increased number of random effects components, the Laird–Ware approach using the EM algorithm to obtain restricted maximum likelihood estimates is time consuming and requires many iterations. Diem and Liukkonen\textsuperscript{6} consider it a major drawback that they required many iterations to obtain convergence to the final estimates. Lindstrom and Bates\textsuperscript{11} have shown that the number of iterations increases with the number of random effects in the model. They demonstrate that the Newton–Raphson method is more efficient than the EM algorithm in computing parameter estimates for models with increased numbers of random effects and data sets with increasing numbers of observations. They present examples that show that the Newton–Raphson method requires fewer iterations than the EM algorithm for moderate sized data sets (308 and 804 observations).

In this paper we use the linear mixed-effects model with the Newton–Raphson method to obtain restricted maximum likelihood estimates for a model with 10 fixed effects and 5 random effects for a data set with 17,236 observations on 268 individuals. We use estimates of the random effects that represent individual variability to find outlying data and to describe more accurately individual patterns in the dependent variable. The dependent variable in this study is hearing threshold level recorded in decibels (dB) as measured by a pure-tone audiogram at 11 different sound frequency levels. For each individual, hearing thresholds at the 11 different frequencies are obtained for both right and left ears. The result is a data set with a large number of observations, since there is an average of 3.5 such repeated hearing tests for the 268 persons.

Brant and Fozard\textsuperscript{12} described age changes in pure-tone hearing thresholds among a population of 20- to 95-year-old males followed longitudinally over a 25-year period. They found that males aged under 70 years who complained of hearing problems, and had a medical history of otological disease or diagnosis indicative of possible hearing problems at the time of testing, had only slight differences in threshold levels compared with the remainder of males of similar age without hearing problems or impairment. For males over the age of 70, they found noticeable differences in thresholds between those identified with and without impairment. Older males with hearing impairment were thus excluded in their study of pure-tone hearing loss.

The present paper examines hearing loss in an elderly male population who participated in a longitudinal study of normal human ageing. In addition to the effect of hearing impairment on hearing threshold level, other factors considered in the model are: age at entry into the study, number of years or time in the study, listener's ears, and frequency of the tone. With use of a frequency-intensity function developed by Brant and Fozard,\textsuperscript{12} we described the relationship between hearing threshold level and frequency with a quadratic relationship in the natural logarithm of frequency. The parameters of this quadratic relationship change with age and by participant and allow an economical way to summarize the large mass of data. Accordingly, in addition to demonstration of how one can apply a linear mixed-effects model to a very large data set with many random effects (used to account for individual differences in outcomes and to detect outlying subjects), it provides previously unknown comparisons of age-related hearing loss between hearing impaired and not impaired elderly individuals. Finally, because the model contains both age at entry and time in the study as factors, we examine both cross-sectional as well as longitudinal differences within the framework of this analysis.
2. DESCRIPTION OF DATA

The Baltimore Longitudinal Study of Ageing (BLSA) is an ongoing multidisciplinary study begun in 1958 to study normal human ageing. The study collected data from several hundred male volunteers on hearing thresholds of continuous pure tones that varied from 125 to 8000 hertz (Hz) with use of a sound-proof chamber and a Bekesy audiometer. In our study we consider 268 elderly male BLSA participants whose first visit was at about 70 years of age and older. The distribution of ages at first visit ('age1' in years) is

<table>
<thead>
<tr>
<th>Age at first visit:</th>
<th>&lt; 74·9</th>
<th>75–79·9</th>
<th>80–84·9</th>
<th>85–89·9</th>
<th>90+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number:</td>
<td>122</td>
<td>88</td>
<td>44</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>

with a median age at first visit of 76·2 years.

Participants in the study returned at various intervals. These visits are not equally spaced as the participants return at their convenience. There is usually about two years between visits. A total of 71 participants had only 1 visit, while 4 participants had as many as 12 visits. The distribution of the number of visits is

<table>
<thead>
<tr>
<th>Number of visits:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants:</td>
<td>71</td>
<td>79</td>
<td>28</td>
<td>16</td>
<td>13</td>
<td>14</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

At each visit, threshold sound pressure levels (SPLs in dB) were measured at 11 different frequencies on both ears, giving a maximum of 22 observations per visit. For many participants, however, there are missing thresholds at some frequencies for one or both ears. Hence, methods used to model the data must account for unbalanced data, unequal intervals and missing observations.

Initial analysis showed that the hearing thresholds for the different ears did not differ significantly. Hence we do not consider this factor further, and we consider the measurements on each ear as replicates at each frequency for each visit. There were 16 participants with a diagnosed hearing disease or impairment, which we take into account in the analysis by use of an indicator variable 'impair'. Impaired subjects included those with chronic otitis media, perforated/punctured eardrum, Menière's disease, scarred tympanic membrane, otosclerosis, trauma/noise injury from military service, chronic noise exposure and stapedectomy.

3. MIXED-EFFECTS MODEL FOR REPEATED MEASURED DATA

In this paper we use the linear mixed-effects model to analyse these longitudinal data. Using similar notation to Lindstrom and Bates, the model for each individual in the study is

$$ Y_i = X_i \beta + Z_i b_i + e_i, \quad i = 1, \ldots, M, $$

where

- $Y_i$ is an $n_i \times 1$ vector of observations for individual $i$
- $\beta$ is a $p \times 1$ vector of parameters for the fixed effects
- $X_i$ is an $n_i \times p$ design matrix of independent variables for the fixed effects
- $b_i$ is a $q \times 1$ vector of random effects for individual $i$
- $Z_i$ is an $n_i \times q$ design matrix of independent variables linking the random effects $b_i$ to $y_i$
- $e_i$ is an $n_i \times 1$ vector of errors for each individual
- $M$ is the number of subjects in the study.
We assume the errors $e_i$ are independent and have multivariate normal distributions $N(0, \sigma^2 I)$, where $I$ is an $n_i \times n_i$ identity matrix. The random effects $b_i$ are multivariate normal $N_q(0, \sigma^2 D)$, independently of each other and of the $e_i$, where $\sigma^2 D$ is a positive definite covariance matrix. Note that the Lindstrom and Bates formulation uses $\sigma^2 D$ for the covariance matrix of $b_i$ for computational convenience. The random factors allowed in this model are persons and the interactions of persons with the other factors (see Laird et al.\(^\text{14}\)). We used software written by Mary Lindstrom (described in Lindstrom and Bates\(^\text{11}\)) to find the restricted maximum likelihood (REML) estimators (and maximum likelihood (ML) estimators) of the variance components in the model, $\sigma^2$ and the $q(q+1)/2$ elements of $D$. We used REML estimators because maximum likelihood estimators of variance components do not take into account the loss in degrees of freedom from the estimation of $\beta$, and will be biased downward. The REML method corrects for this problem by maximizing the log-likelihood of $N-p$ linearly independent error contrasts, where $N = \sum_{i=1}^{M} n_i$, the total number of observations from all subjects in the study. This log-likelihood was first developed by Harville.\(^\text{15}\) Given estimates of $\sigma^2$ and $D$, we estimate the $p$ elements of $\beta$ by generalized least squares, and the random effects $b_i$, $i = 1, \ldots, M$, by their posterior mean. The computer program uses the Newton–Raphson method to obtain the REML estimates with a relative offset orthogonality convergence criterion.\(^\text{16}\) Using this criterion, at convergence, the residual vector will have zero projection onto the tangent plane of the surface maximized.\(^\text{11}\) The criterion used with the EM algorithm is usually the size of change in the parameter estimates or likelihood between two iterations. This is a measure of lack of progress as opposed to actual convergence and is another drawback of the EM algorithm.

We also computed maximum likelihood estimates of the variance components to compare with the REML estimates. We expect the ML estimates to be close to the REML estimates since the data set is large.

The design matrix for the fixed effects $X_i$ has ten columns as follows: 1, age1, follow-up time (time), time\(^2\), ln(freq), ln\(^2\)(freq), age $\times$ ln(freq), age $\times$ ln\(^2\)(freq), impair $\times$ ln\(^2\)(freq), impair.

The design matrix for the random effects $Z_i$ has five columns: 1, time, time\(^2\), ln(freq), ln\(^2\)(freq).

Thus if $y_{ij}$ is the $j$th hearing threshold of participant $i$, the model for each observation is

$$y_{ij} = (\beta_1 + b_{i1}) + \beta_2 \text{ age1} + (\beta_3 + b_{i2}) \text{ time} + (\beta_4 + b_{i4}) \text{ time}^2 + [\beta_5 + b_{i5} + \beta_7 \text{(age)}] \ln(\text{freq})$$

$$+ [\beta_6 + b_{i6} + \beta_9 \text{(age)} + \beta_9 \text{(impair)}] \ln^2(\text{freq}) + \beta_{10} \text{ impair} + e_{ij},$$

$$i = 1, \ldots, M, \quad j = 1, \ldots, n_i. \quad (2)$$

The age terms (in years) model cross-sectional differences in hearing thresholds of subjects of different ages, while the time terms (in years) model the longitudinal age changes in the thresholds of each subject. Note that age = age1 + time. We modelled the relationship between frequency and threshold with use of a quadratic function in ln(freq) (developed by Brant and Fozard\(^\text{12}\)) that allows for both J-shaped and U-shaped curves which the audiograms typically follow, thus efficiently summarizing the large mass of data. The shape of the curve may change with age for each participant regardless of whether the participant has a hearing impairment. We initially included an impair $\times$ ln(freq) term, but we found it statistically non-significant. Usually, polynomials that exclude hierarchically inferior terms are not well formulated.\(^\text{17,18}\) There is, however, already a component in ln(freq) in the model for all participants. The term impair $\times$ ln(freq) would allow the 16 impaired participants to have a fixed component of ln(freq) that deviates from the remainder of the participants. Since this term is not statistically significant, the frequency effect on hearing thresholds is the same for all subjects in the ln(freq) term and differs for impaired subjects only in the impair $\times$ ln\(^2\)(freq) term.
MODELLING HEARING THRESHOLDS IN THE ELDERLY

The random terms in (2) allow for different estimates of the intercept and partial slopes for each subject. Thus the elements of $\mathbf{b}$ are the ‘average’ intercept and slopes, while the elements of $\mathbf{b}_i$ are the deviations from these ‘averages’ for participant $i$. In addition, our model has random shape (frequency) components ($b_{i4}, b_{i5}$) and random growth curve (time) components ($b_{i2}, b_{i3}$).

We estimated the covariance matrix of the $\mathbf{b}_i, \sigma^2 D$, by REML. The first diagonal element of this matrix yields the estimate of variance between subjects. The other diagonal terms measure the variance of subject x factor interactions, while the off-diagonal terms give the covariances between the random factors.

To decide upon the significance of individual fixed effects in the model, we computed $z$-ratios for each fixed-effects parameter estimate. These are

$$z = \frac{\hat{\beta}_i}{\left[ \text{var}(\hat{\beta}_i) \right]^{1/2}},$$

where $\text{var}(\hat{\beta}_i)$ is the $i$th diagonal term of the variance–covariance matrix of $\hat{\beta}$, and $\left[ \text{var}(\hat{\beta}_i) \right]^{1/2}$ is the standard error of $\hat{\beta}_i$. We can compare these $z$-values to standard normal critical values to determine their significance.

Since we estimated the variance components by REML in model (2), we are only able to obtain approximate tests for the random factors in the model. This is because the log-likelihood for REML contains an extra term when compared with the usual likelihood. To obtain these approximate tests for the random components in the model, we fitted reduced models leaving out columns of $\mathbf{Z}_f$ the design matrix of the independent variables for the random effects. To evaluate whether we may include additional terms, we added appropriate columns to $\mathbf{Z}_f$. Then, to determine whether we could drop terms from or add terms to the model, we computed $-2$ (difference in the log-likelihoods). We compare these values to the $\chi^2$ distribution with degrees of freedom (d.f.) given by the difference in the number of parameters between the two models. If $q$ is the number of random effects in the model under consideration, the covariance matrix has $q(q+1)/2$ terms to be estimated. If one random effect is dropped, only $(q-1)q/2$ terms remain. Thus if we drop one random term, d.f. = $q$, while if we add one random term, d.f. = $q + 1$. We compare the results of these approximate tests to the likelihood ratio test using the ML estimates. Once more, we expect similar results since the sample size is large.

4. RESULTS

We used REML to estimate the variance components in model (2), generalized least squares to estimate the fixed effects $\mathbf{b}$, and the posterior mean to estimate the individual random effects $\mathbf{b}_i$, for the $M = 268$ participants from the BLSA. There were $N = 17,236$ hearing threshold for these 268 participants. We fitted model (2) to the data. To check the assumptions of the analysis, we first look at the $\mathbf{b}_i$ to identify outlying individuals. Bivariate plots of pairs of $\mathbf{b}_i$ show that subject number 74 (represented by a *) lies outside the plot for the time and time' random components (Figure 1). Thus, subject 74 warrants closer investigation. A plot of the hearing thresholds for this participant against frequency for each ear (Figure 2) shows erratic values for the fourth visit. (We attempted to identify why there was a problem, but we could not locate the original audiogram tracing for the fourth visit that corresponded to the computerized data.) Thus we excluded the data from the fourth visit from the analysis, leaving $N = 17,217$ hearing thresholds, and we re-estimated the model. Figure 3 contains the bivariate plots of all pairs of random effects. Subject 74 is no longer an outlier. Table I contains estimates of the elements of the parameter vector $\mathbf{b}$ and their standard errors. All the fixed effects in the model are significant at the 5 per cent level.

The estimate of the error variance is $\hat{\sigma}^2 = 155.334$. The REML and ML estimates of the components of the variance–covariance matrix of the random effects and their correlations
Figure 1. Plot of time and time$^2$ estimated random effects for each subject showing outlying individual ($\ast$)

Figure 2. Hearing thresholds by frequency for outlying individual, participant 74, for left and right ears for visits 1 ($\ast$), 2 ($\circ$), 3 ($+$) and 4 ($\times$)
Table 1. Estimates of the fixed effects in model (2), their standard errors and z-ratios

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>z-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$: constant</td>
<td>-113.115</td>
<td>9.791</td>
<td>-11.55</td>
</tr>
<tr>
<td>$\beta_2$: age</td>
<td>1.8051</td>
<td>0.1276</td>
<td>14.15</td>
</tr>
<tr>
<td>$\beta_3$: time</td>
<td>1.0542</td>
<td>0.2123</td>
<td>4.97</td>
</tr>
<tr>
<td>$\beta_4$: time^2</td>
<td>0.07518</td>
<td>0.01923</td>
<td>3.91</td>
</tr>
<tr>
<td>$\beta_5$: ln(freq)</td>
<td>-15.7972</td>
<td>1.9546</td>
<td>-8.06</td>
</tr>
<tr>
<td>$\beta_6$: ln^2(freq)</td>
<td>25.0025</td>
<td>1.5523</td>
<td>16.11</td>
</tr>
<tr>
<td>$\beta_7$: age x ln(freq)</td>
<td>0.3473</td>
<td>0.02468</td>
<td>14.07</td>
</tr>
<tr>
<td>$\beta_8$: age x ln^2(freq)</td>
<td>-0.2457</td>
<td>0.01968</td>
<td>-12.48</td>
</tr>
<tr>
<td>$\beta_9$: impair x ln(freq)</td>
<td>-1.9736</td>
<td>0.8639</td>
<td>-2.28</td>
</tr>
<tr>
<td>$\beta_{10}$: impair</td>
<td>15.561</td>
<td>3.667</td>
<td>4.24</td>
</tr>
</tbody>
</table>

Figure 3. Estimated random effects for each subject for all pairs of random effects. Participant 74 is plotted as a •
appear in Table II(a) and (b). The REML and ML estimates are close. We restrict our discussion to REML estimates. The between-persons variance $\hat{\sigma}_p^2 = 212.164$ indicates a large difference between participants in the study. The correlation between the time and time$^2$ random terms is $-0.9667$. This is also evident from the plot of the estimates $b_i$, which shows highly negatively correlated values. Thus if an individual's coefficient for time is above $\hat{\beta}_3$, the coefficient for time$^2$ is below $\hat{\beta}_4$.

To determine whether we require all the random factors in the model to describe adequately the data, we omitted various columns from the design matrix $Z_i$ and performed approximate likelihood ratio tests. We excluded only the squared terms for time and ln(freq) so that these random terms remain hierarchically well formulated. Also, we included (separately) extra random terms, age1 and impair, to determine whether we require these random terms in the model. The results appear in Table III along with $\chi^2$ values for both REML and ML estimation. The results are all close to one another and the conclusions remain the same. No random factors may be omitted, and the two extra terms considered are not necessary to describe these data. Thus, even though the random effects of time and time$^2$ are highly correlated, they are both required in the model to describe the data adequately.

Figure 4 contains plots of the expected hearing thresholds against frequency for normal and impaired participants. As anticipated, the fitted curves for the impaired subjects are above those for normal subjects. The curves for the cross-sectional changes are more equally spaced than those for the longitudinal changes. This is because the curves for cross-sectional change are linear in age at each frequency, whereas those for longitudinal change have a quadratic component in
**MODELLING HEARING THRESHOLDS IN THE ELDERLY**

Table II.

(a) REML estimates of the random effects variance–covariance matrix ($\sigma^2 D$)*

<table>
<thead>
<tr>
<th>Random effect</th>
<th>ln(freq)</th>
<th>ln²(freq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>212.164</td>
<td>-32.4934</td>
</tr>
<tr>
<td>time</td>
<td>-0.1157</td>
<td>-1.5153</td>
</tr>
<tr>
<td>time²</td>
<td>0.0961</td>
<td>0.1043</td>
</tr>
<tr>
<td>ln(freq)</td>
<td>0.2704</td>
<td>1.4998</td>
</tr>
<tr>
<td>ln²(freq)</td>
<td>-0.7124</td>
<td>9.8060</td>
</tr>
</tbody>
</table>

* The diagonal terms and values above the diagonal are the elements of the variance–covariance matrix. The values below the main diagonal are the correlations between the random effects.

(b) ML estimates of the random effects variance–covariance matrix ($\sigma^2 D$)

<table>
<thead>
<tr>
<th>Random effect</th>
<th>ln(freq)</th>
<th>ln²(freq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>210.137</td>
<td>-32.1827</td>
</tr>
<tr>
<td>time</td>
<td>-0.1158</td>
<td>-1.5053</td>
</tr>
<tr>
<td>time²</td>
<td>0.0964</td>
<td>0.1031</td>
</tr>
<tr>
<td>ln(freq)</td>
<td>0.2712</td>
<td>1.4978</td>
</tr>
<tr>
<td>ln²(freq)</td>
<td>-0.7127</td>
<td>9.7046</td>
</tr>
</tbody>
</table>

Table III. Approximate likelihood-ratio tests for the random effects*

<table>
<thead>
<tr>
<th>Factor</th>
<th>$\chi^2_{\text{REML}}$</th>
<th>d.f.</th>
<th>$\chi^2_{\text{ML}}$</th>
<th>$p$-value$_{\text{REML}}$</th>
<th>$\chi^2_{\text{ML}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclude:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln²(freq)</td>
<td>- 53957:600</td>
<td>5</td>
<td>1435:17</td>
<td>0.000</td>
<td>1431:95</td>
</tr>
<tr>
<td>time²</td>
<td>- 53303:530</td>
<td>5</td>
<td>127:03</td>
<td>0.000</td>
<td>125:74</td>
</tr>
<tr>
<td>Add:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>impair</td>
<td>- 53234:400</td>
<td>6</td>
<td>11:23</td>
<td>0.081</td>
<td>11:13</td>
</tr>
<tr>
<td>age</td>
<td>- 53238:517</td>
<td>6</td>
<td>3:00</td>
<td>0.809</td>
<td>2:96</td>
</tr>
</tbody>
</table>

* For the final model, $q = 5$ and $\ln L_{\text{REML}} = - 53240:017$.

Thus the curves on the two graphs of the upper portion of Figure 4 represent cross-sectional differences, while those of the lower portion of the figure give longitudinal changes of the subjects in the study. At 1 kHz there is a cross-sectional loss of 1.8 dB per year, whereas the longitudinal loss is 1.1 dB per year at the beginning of follow-up and increases to 2.6 dB per year after ten years. Thus cross-sectional studies may produce biased estimates of true rates of hearing loss.

Of interest to investigators studying hearing is whether subjects with hearing impairments lose their hearing at a different rate than those without an impairment. The fixed factors, time $\times$ impair and time² $\times$ impair terms allow tests for different rates of hearing loss for impaired subjects. The $z$-values for these terms, however, were 0.144 and 0.225 respectively, which are both far from statistical significance. Thus, while impaired participants have a higher threshold than
normal subjects ($\beta_{10} \pm SE = 15.56 \pm 3.67$), we cannot conclude that the impairment results in a different rate of hearing loss after age 70.

5 CONCLUSIONS

The Newton–Raphson method for estimating the parameters in the linear mixed-effects model was very useful in this longitudinal study of hearing loss with an extremely large data set and a relatively large number of random effects. To compare the algorithms, we use both the Newton–Raphson and EM methods to fit the model. We choose a diagonal matrix of starting estimates of $\sigma^2D$ close to the final values from our final fitted model (2), $\sigma^2D = \text{diag}(200, 10, 0.1, 20, 20)$. The Newton–Raphson method converged in 10 iterations on a VAX 11/785 computer using a relative offset orthogonality convergence criterion. The EM algorithm took 124 iterations to converge on the same computer. The linear mixed-effects model enabled us to detect an outlying individual and hence eliminate faulty data. We could also account more adequately for individual differences in the data since the model allowed for each individual to have a different intercept, time and time$^2$ parameters as well as different ln(freq) and ln$^2$(freq) parameters. The fixed effects entered into the model included age at entry to the study, a quadratic in time of follow-up and a quadratic in ln(freq), the parameters of which changed with age, and whether or not
not a participant had a hearing impairment. These fixed effects provide the ‘average’ curve for the population, whereas the random effects give the deviation from the ‘average’ for each participant.

Many investigators have had concern with presbycusis (that is, the amount of hearing loss that results because of normal human ageing in individuals).19–21 These papers use cross-sectional data to show that hearing loss occurs with increasing age, especially at the higher frequencies such as 8000 Hz. Möller22 studied hearing in the elderly in a Swedish study of 70- and 75-year-olds. She considered cross-sectional and longitudinal changes in hearing thresholds but was restricted to only two visits for data collection. She found no differences in pure-tone thresholds for men between the beginning and end of the five-year period. Moscicki et al.23 use data from the Framingham Heart Study to investigate hearing loss in the elderly at a single time point. In addition to age having a significant impact upon hearing loss, other factors such as illness, history of hearing loss and noise exposure were also significant predictors. Since the data are cross-sectional, they used logistic regression to build a model that could predict the proportion of ‘events’ expected for various combinations of risk factors.

This paper provides the largest set of longitudinal data to date for examination of hearing loss between hearing impaired and unimpaired elderly males. Since there are only 16 impaired individuals for 8 types of impairment, it is not meaningful to estimate the effect due to individual types of impairment. The results from our analysis show that, on average, impaired subjects have a hearing threshold of 15.6 dB above a normal subject at 1 kHz, but at higher frequencies the difference decreases. We did not find the rate of decline for impaired subjects to be statistically significantly different from normal participants. Thus impaired individuals appear to have patterns of ageing similar to those of unimpaired individuals. Further, at 1 kHz there is a cross-sectional loss of 1.8 dB per year, whereas the longitudinal loss is 1.1 dB per year at the beginning of follow-up and increases to 2.6 dB per year after ten years. These findings have particular pertinence since hearing loss constitutes an important determinant of function in the elderly.24

**ACKNOWLEDGEMENTS**

We thank Mary Lindstrom for supplying the Newton–Raphson and EM algorithm program for the linear mixed-effects model. We also thank Scott Zeger for his helpful comments that improved the presentation.

**REFERENCES**