STATISTICS IN MEDICINE, VOL. 16, 2475–2488 (1997)

CONSTRUCTION OF HEARING PERCENTILES IN WOMEN WITH NON-CONSTANT VARIANCE FROM THE LINEAR MIXED-EFFECTS MODEL

CHRISTOPHER H. MORRELL[1,*], JAY D. PEARSON[2], LARRY J. BRANT[2]
AND SANDRA GORDON-SALANT[3]

[1]*Mathematical Sciences Department, Loyola College, 4501 North Charles Street, Baltimore, MD 21210 U.S.A.*
[2]*Longitudinal Studies Branch, Gerontology Research Center, National Institute on Aging, Baltimore, MD 21224, U.S.A.*
[3]*Hearing and Speech Sciences, University of Maryland, College Park, MD 20742, U.S.A.*

SUMMARY

Current age-specific reference standards for adult hearing thresholds are primarily cross-sectional in nature and vary in the degree of screening of the reference sample for noise-induced hearing loss and other hearing problems. We develop methods to construct age-specific percentiles for longitudinal data that have been modelled using the linear mixed-effects model. We apply these methods to construct percentiles of hearing level using data from a carefully screened sample of women from the Baltimore Longitudinal Study of Aging. However, the variation in the residuals and random effects from the linear mixed-effects model does not remain constant with age and frequency of the stimulus tone. In addition, the distribution of the hearing levels is not symmetric about the mean. We develop a number of methods to use the output from the linear mixed-effects model to construct percentiles that do not have constant variance. We use a transformation of the hearing levels to provide for skewness in the final percentile curves. The change in the variation of the residuals and random effects is modelled as a function of beginning age and frequency and we use this variance function to construct the hearing percentiles. We present a number of approaches. First, we use the absolute values of the population residuals to model the total deviation about the mean as a function of beginning age and frequency. Second, we model the standard deviation in the person-specific (cluster) residuals as well as the standard deviation in the estimated random effects. Finally, we use weighted least squares with the regressions on the absolute cluster residuals and absolute estimated random effects where the weights are the reciprocal of the standard deviations of their estimates.

Statist. Med., **16**, 2475–2488 (1997)
No. of Figures: 6   No. of Tables: 2   No. of References: 17

INTRODUCTION

Reference ranges or norms have wide usage in medicine. When the single variable of interest follows a normal distribution, one usually computes the norms as the mean plus or minus

* Correspondence to: C. H. Morrell, Mathematical Sciences Department, Loyola College, 4501 North Charles Street, Baltimore, MD 21210 U.S.A.

Contract grant sponsor: National Institute on Deafness and Other Communication Disorders, National Institutes of Health
Contract grant number: 1 R03 DC 02566-01

CCC 0277–6715/97/212475–14$17.50   *Received January 1996*
 Revised January 1997

a percentile of the standard normal distribution times the standard deviation. Frequently, reference ranges are desired at each age. If the mean of the variable of interest can be modelled as some function of age, one can construct percentiles as this mean plus or minus multiples of the standard deviation. However, the standard deviation may not remain the same at all ages and meaningful percentile curves must reflect this non-constant variance.

Royston[1] and Royston and Matthews[2] discuss the construction of reference centiles using various models to describe the relationship between the response variable and the explanatory variables. Cole and Green[3] and Green and Silverman[4] use a penalized likelihood method to obtain percentiles when one models the relationship between the response and explanatory variables using cubic splines with changing variation and skewness. Altman[5] models the absolute residuals using linear regression and obtains an estimate of the standard deviation that changes with the value of an explanatory variable.

In a previous study, Pearson *et al.*[6] described longitudinal patterns of change in hearing thresholds from 416 women (age 20 to 90 years) followed for up to 13 years who had been screened for otologic disorders, unilateral hearing loss, and evidence of noise-induced hearing loss. In this paper, we begin with the output from the previous linear mixed-effects analysis[7] to construct hearing percentiles at various frequencies for these women. However, the distribution of hearing levels about the mean is not symmetric and the amount of variability in the hearing levels changes with age and frequency of stimulus tone. To overcome these problems, we transform the data and develop methods to reflect this changing variability in the percentile curves.

We apply a logarithmic transformation to the hearing levels that have had a constant added to ensure all of the values are positive. We refit these transformed longitudinal data using the linear mixed-effects model. We then model the absolute residuals and random effects from the model for the transformed data as functions of beginning age and frequency to obtain non-constant standard deviation functions.

The principal goal of this paper is to illustrate how to calculate percentiles from longitudinal data when the variance is not constant. We apply the methods developed to hearing threshold data on a sample of women from a longitudinal study.

BACKGROUND

Study Population

The participants are female volunteers in the Baltimore Longitudinal Study of Aging (BLSA), an open-panel multidisciplinary study of normal human ageing which began studying women in 1978 and which is conducted by the intramural research programme of the National Institute on Aging.[8] Participants in the study are predominantly white (95 per cent), well-educated (over 75 per cent have a bachelor's degree or higher), and financially comfortable (82 per cent) volunteers. The participants are scheduled to visit the Gerontology Research Center in Baltimore at approximately 2-year intervals where they stay for $2\frac{1}{2}$ days of evaluation and testing that includes audiologic testing.

The present analyses exclude data from participants with otologic disease, unilateral hearing loss, or evidence of noise-induced hearing loss.[6] Due to these exclusions, the final study group consists of 416 women (71 per cent of the original group) evaluated at the BLSA between 1978 and 1991 with age at entry between 18 and 86 years. These women produced a total of 1331 audiograms with a mean of 3·2 visits and 5·2 years of follow-up (maximum of 12·9 years).

Approximately 48 per cent of the women have 5 or more years of follow-up and 17 per cent have 10 or more years of follow-up.

Apparatus and Procedures

As part of the BLSA testing, participants completed continuous followed by pulsed pure tone audiologic testing. In this paper we report the hearing threshold levels determined for 9 frequencies (500, 750, 1000, 1500, 2000, 3000, 4000, 6000, 8000 Hz) separately in each ear from a Bekesy audiogram obtained using a pulsed pure tone generated by a Grason–Stadler audiometer (Model 1701). All thresholds are expressed in dB HL using the ANSI[9] standards.

Statistical Methods

The linear mixed-effects (LME) model has the form[7,10,11]

$$\boldsymbol{y}_i = X_i\boldsymbol{\beta} + Z_i\boldsymbol{b}_l + \boldsymbol{e}_i, \quad i = 1, \ldots, M$$

where $\boldsymbol{y}_i$ is the $n_i \times 1$ vector of observations for individual i, X_i is the design matrix of independent variables for the fixed effects for individual i, $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed-effects parameters, Z_i is the design matrix of independent variables for the random effects, $\boldsymbol{b}_i$ is the $q \times 1$ vector of random effects for individual i, $\boldsymbol{e}_i$ is the $\boldsymbol{n}_i \times 1$ random error vector, and M is the number of subjects in the study. It is usually assumed that $\boldsymbol{e}_i \sim \mathrm{N}(0, \sigma^2 I)$ and $\boldsymbol{b}_i \sim \mathrm{N}(0, D)$ and that $\boldsymbol{e}_i$ and $\boldsymbol{b}_i$ are independent. The random component in the model, $\boldsymbol{b}_i$, allows for the natural heterogeneity in intercept and partial slopes among individuals in the study so that each individual has her own pattern of change over the time and frequency domains that differ from those of other subjects in the study.

The fixed-effects parameters in the model, $\boldsymbol{\beta}$, are estimated via generalized least squares, and the variance components, σ^2 and D, may be estimated either by maximum likelihood or restricted maximum likelihood.[10,12] In this paper we estimate the variance components by restricted maximum likelihood.

Mixed-effects models allow estimation of the average hearing level curve for the population and also allow each subject's estimated longitudinal change, inter-aural difference, and audiometric curve to deviate from the group average. The fixed effects estimate the average intercept and rates of change for the independent variables, while the random effects represent the deviation for each individual from the average intercept and slope terms. Thus, the random effects account for natural heterogeneity in initial level, ear, patterns of longitudinal change, and audiometric shape among the individuals in the study. Mixed-effects models provide a marginal covariance structure among repeated measures within individuals that may adequately model the correlation within subjects and allow the analysis of unbalanced data where individuals have differing numbers of observations taken at varying intervals between the observations.

Arranging the terms in the mixed-effects model to see how the longitudinal change depends upon first age and frequency, the initial model entertained for a hearing level observation on the ith woman at time j and frequency k is

$$\begin{aligned} y_{ijk} = {} & (\boldsymbol{\beta}_0 + b_{i0}) + (\boldsymbol{\beta}_1 + b_{il})\mathrm{ear}_i + (\boldsymbol{\beta}_2 + b_{i2})\mathrm{time}_{ij} + (\boldsymbol{\beta}_3 + b_{i3})\ln(\mathrm{freq})_{ik} \\ & + (\boldsymbol{\beta}_4 + b_{i4})\ln^2(\mathrm{freq})_{ik} + (\boldsymbol{\beta}_5 + b_{i5})\ln^3(\mathrm{freq})_{ik} + \boldsymbol{\beta}_6\mathrm{fage}_i + \boldsymbol{\beta}_7\mathrm{fage}_i^2 + \boldsymbol{\beta}_8\mathrm{fage}_i^3 \\ & + \boldsymbol{\beta}_9\mathrm{visit1}_i + [\boldsymbol{\beta}_{10}\mathrm{fage}_i + \boldsymbol{\beta}_{11}\mathrm{fage}_i^2 + \boldsymbol{\beta}_{12}\mathrm{fage}_i^3 \end{aligned}$$

$$+ (\boldsymbol{\beta}_{13} + \boldsymbol{\beta}_{14}\text{fage}_i + \boldsymbol{\beta}_{15}\text{fage}_i^2 + \boldsymbol{\beta}_{16}\text{fage}_i^3)\ln(\text{freq})_{ik}$$

$$+ (\boldsymbol{\beta}_{17} + \boldsymbol{\beta}_{18}\text{fage}_i + \boldsymbol{\beta}_{19}\text{fage}_i^2 + \boldsymbol{\beta}_{20}\text{fage}_i^3)\ln^2(\text{freq})_{ik}$$

$$+ (\boldsymbol{\beta}_{21} + \boldsymbol{\beta}_{22}\text{fage}_i + \boldsymbol{\beta}_{23}\text{fage}_i^2 + \boldsymbol{\beta}_{24}\text{fage}_i^3)\ln^3(\text{freq})_i]\,\text{time}_{ij}$$

$$+ [\boldsymbol{\beta}_{25}\text{fage}_i + \boldsymbol{\beta}_{26}\text{fage}_i^2 + \boldsymbol{\beta}_{27}\text{fage}_i^3 + \boldsymbol{\beta}_{28}\text{visit1}_i]\ln(\text{freq})_{ik}$$

$$+ [\boldsymbol{\beta}_{29}\text{fage}_i + \boldsymbol{\beta}_{30}\text{fage}_i^2 + \boldsymbol{\beta}_{31}\text{fage}_i^3 + \boldsymbol{\beta}_{32}\text{visit1}_i]\ln^2(\text{freq})_{ik}$$

$$+ [\boldsymbol{\beta}_{33}\text{fage}_i + \boldsymbol{\beta}_{34}\text{fage}_i^2 + \boldsymbol{\beta}_{35}\text{fage}_i^3 + \boldsymbol{\beta}_{36}\text{visit1}_i]\ln^3(\text{freq})_{ik} + e_{ijk} \qquad (1)$$

where longitudinal change is represented by follow-up time (time), cross-sectional age differences are represented by polynomial terms for age at first visit (fage, fage^2, and fage^3), audiogram shape is represented by polynomial terms for the natural logarithm of the frequency in kilohertz (ln(freq), $\ln^2$(freq), and $\ln^3$(freq)), interaural differences are represented by the indicator variable (ear), learning effects are represented by a contrast between first visit and subsequent visits using an indicator variable (visit1), and e_{ijk} represents the statistical error term. Since only 38·5 per cent of the women had more than three visits, the only longitudinal terms included in the model were time and visit1 (that is, no time^2 terms were included). Previous analyses have shown that the use of polynomials in ln(freq) is an efficient and flexible method of modelling the audiogram frequency-intensity function.[13,14] Interaction terms are included that allow the longitudinal patterns of change to differ with age at entry (fage × time), allow the audiogram shape to change longitudinally (ln(freq) × time) and with age at entry (ln(freq) × fage). The visit1 × ln(freq) interaction terms allow the frequency intensity function at the first visit to differ from those at later visits so that the learning effect may differ at the various frequencies. Three-way interactions between fage, time and ln(freq) are included to model differences in rate of change in thresholds at different ages and frequencies.

We include six random-effect terms ($b_{i0}, b_{i1}, b_{i2}, b_{i3}, b_{i4}$, and b_{i5}) in the initial model to account for natural heterogeneity among individuals with respect to hearing level (intercept), interaural difference (ear), longitudinal pattern of change (time), and audiogram shape (ln(freq), $\ln^2$(freq), and $\ln^3$(freq)). Thus, each person's hearing thresholds may have a level and audiometric shape that deviates from the overall average, each person's longitudinal pattern of change may deviate from the overall average, and the degree of each person's interaural symmetry/asymmetry may vary from the overall average.

To reduce the multicollinearity among the polynomial terms, we centre the follow-up time and first age variables on the mean follow-up time and age at first visit by subtracting the corresponding means of 5 and 53 years, respectively, from time and fage. The most parsimonious well-formulated model[15] is obtained by backward elimination of the highest-order non-significant polynomial and cross-product terms.

Results

The final model for describing female hearing levels contained 24 fixed-effects variables (omitting fixed-effects terms numbered 8, 12, 16, 20, 21, 22, 23, 24, 27, 31, 34, 35, 36).[6] Longitudinally, at higher frequencies and ages, hearing sensitivity declines, although hearing levels at 1000, 2000 and 4000 Hz improve slightly ($<$ 2·0 dB per decade) for women under age 60. Hearing levels worsen at all ages for 500 Hz, after age 50 for 1000 and 2000 Hz, and after age 40 for 4000 Hz. The decline in hearing sensitivity accelerates at approximately age 40–50. The cross-sectional and longitudinal changes in hearing level do not differ appreciably.[6] Thus, there is little difference in the percentile plots whether we use first age or age.

The women exhibited a statistically significant learning effect from the first visit to subsequent visits. The estimated improvement in hearing levels ranges from 0·1 dB at 500 Hz to 1·6 dB at 8000 Hz. Thus, there was little meaningful difference in the magnitude of the learning effect at different frequencies. As has been found in previous studies, hearing levels are slightly poorer on average for the left compared to the right ear (0·4 dB).

Examination of graphs of the estimated residuals and random effects indicates that the amount of variation increases slightly with age and the amount of spread in the residuals differs among frequencies. This changing variation is not so dramatic as to cast doubt on the linear mixed-effects analysis. In any case, the estimate of the fixed effects parameters remains consistent[11,16] and hence we estimate the mean curve correctly.

CONSTRUCTING PERCENTILES

Initial Percentiles

The marginal distribution of the vector **y** is[7]

$$\boldsymbol{y} \sim \mathrm{N}(X\boldsymbol{\beta}, \sigma^2 I + ZDZ^{\mathrm{T}}).$$

We first use this marginal distribution to construct the hearing percentiles by replacing the parameters by their estimates. In particular, the pth percentile is

$$X\hat{\boldsymbol{\beta}} + z_p\sqrt{(\hat{\sigma}^2 I + Z\hat{D}Z^{\mathrm{T}})} \tag{2}$$

where z_p is the pth percentile of a standard normal distribution. These percentiles are calculated for first age from 25 to 80 years for a number of frequencies (see Figure 1). While we construct the plots at first age, we set visit1 to 0 so that we obtain percentiles of hearing levels that correct for the learning effects. Note that at a particular frequency the percentiles have constant deviation from the mean curve at all ages and are symmetric about the mean. This is clearly not a realistic representation as the amount of spread in hearing level increases with age and the distribution of hearing level is not symmetric about the mean but is skewed to the right. The amount of spread at different frequencies differs due to the random ln(freq) terms in the mixed-effects model. To serve as a comparison, we plot the observed hearing levels for the four selected frequencies at first visit against age at first visit (Figure 2). These observed data show the increasing spread with age as well as the skewness in the data.

To account for the asymmetry, we transformed the data by adding 15 to the hearing level and taking logarithms (the smallest hearing levels for these women is -13). We then refit the linear mixed-effects model (1) eliminating statistically non-significant terms using backward elimination as before. The final model for the transformed data contains 30 fixed effects (omitting fixed-effects terms numbered 12, 16, 20, 22, 23, 24, 36 in (1)) and the same 6 random effects. Application of (2) to the transformed data and transformation back to the original scale produces Figure 3. This graph

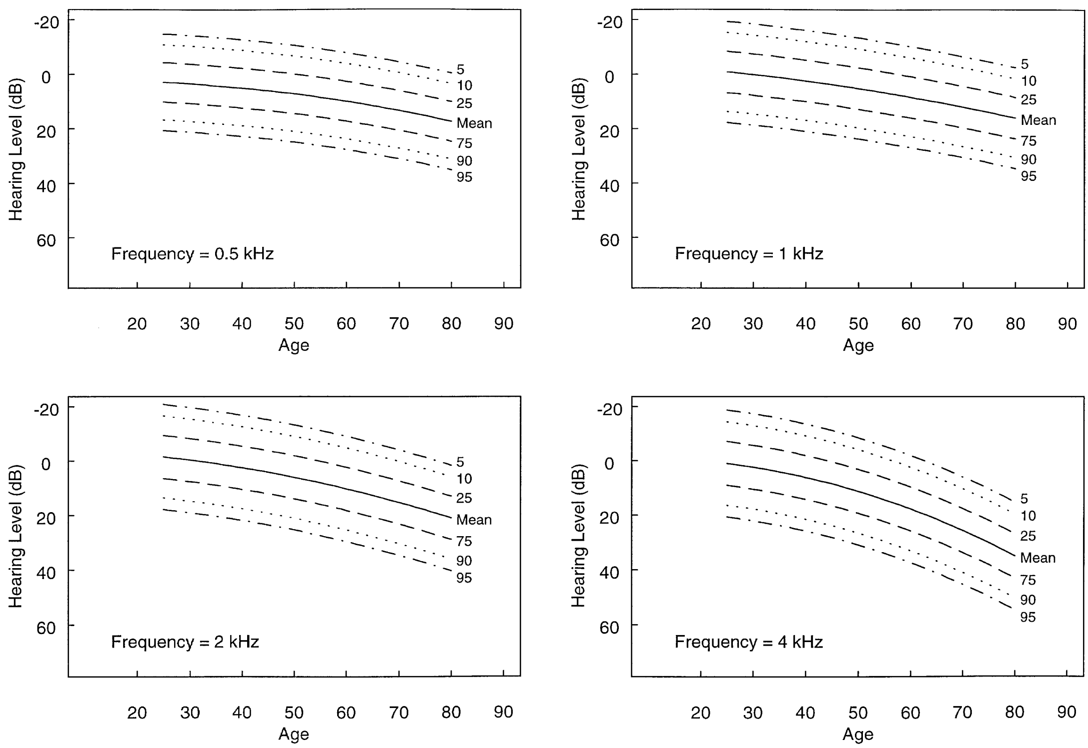

Figure 1. Percentile curves computed from the linear mixed-effects model with constant variance in residuals and random effects

shows skewness and non-constant variation due to the transformation. In addition, the median curve is almost identical to the mean curve from the untransformed data. However, the output from the model again exhibits non-constant variation in the residuals with first age and frequency and in the random effects with first age. We use these estimated residuals and random effects from the model with the transformed data to construct further percentile curves.

Modelling the Non-constant Variance

We relax the restriction that the variance of the error term, σ^2, and the covariance matrix of the random effects, D, are constant. This is achieved by modelling the amount of variation in the residuals as a function of first age and frequency, and the variation in the random effects as a function of first age. This provides more realistic and useful norms or reference standards.

Altman[5] demonstrated how to construct percentiles with non-constant variance from the linear regression model. Suppose that the residuals plotted against age exhibit changing variation. Altman's method models the absolute residuals as a function of age (usually some low-order polynomial). The mean of the absolute residuals times $(\pi/2)^{1/2}$ is an estimate of the standard deviation of the residuals.

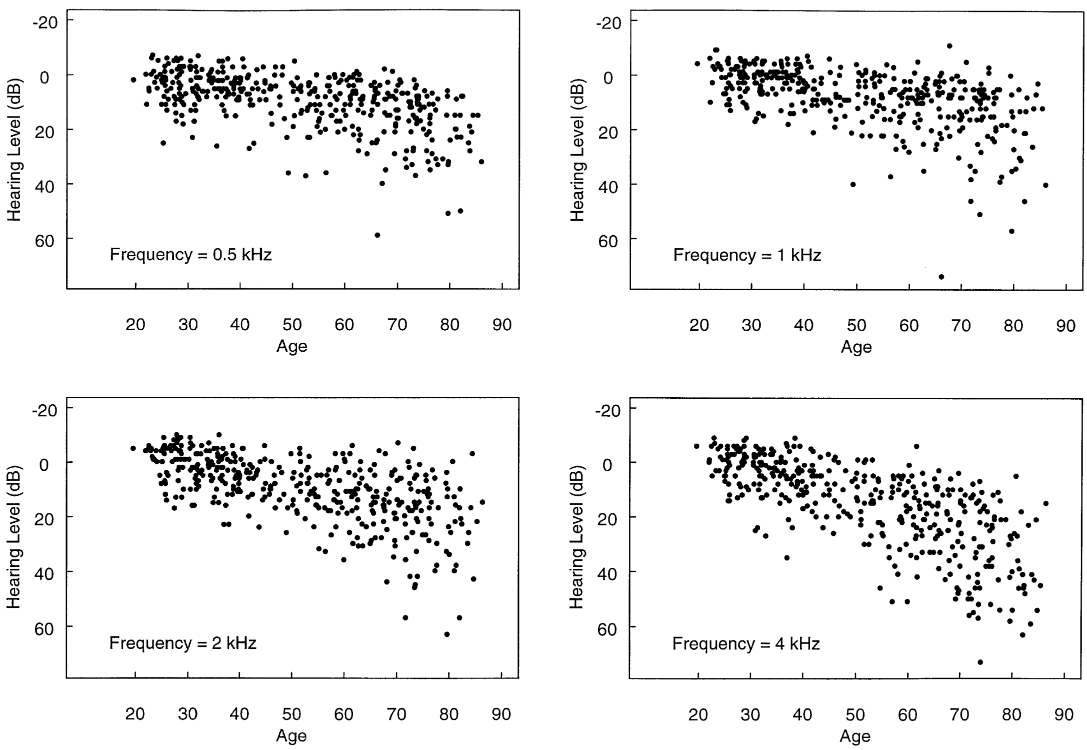

Figure 2. Observed hearing levels at first visit

In this paper we use three approaches to model the non-constant variance using the absolute residuals and absolute random effects from the linear mixed-effects model on the transformed data.

Non-Constant Variance in Population Residuals

As a simple approach for addressing the non-constant variance, we apply Altman's method to the population residuals

$$\hat{\boldsymbol{e}}_{Pi} = \boldsymbol{y}_i - X_i\hat{\boldsymbol{\beta}}.$$

We construct percentiles around the fixed-effects part of the model, $\hat{\boldsymbol{y}} = X\hat{\boldsymbol{\beta}}$, with the modified variation. We use multiple linear regression to model the absolute population residuals as a function of first age and frequency. We consider polynomial terms up to first age^3 and $\ln^3$(freq), as well as all the cross products of these terms. Backward elimination of statistically non-significant terms (while constraining the model to remain well-formulated) yields the model that describes the changing standard deviation of the population residuals as a function of first age and ln(freq). First age is again centred by 53 and frequency is expressed in kilohertz. The

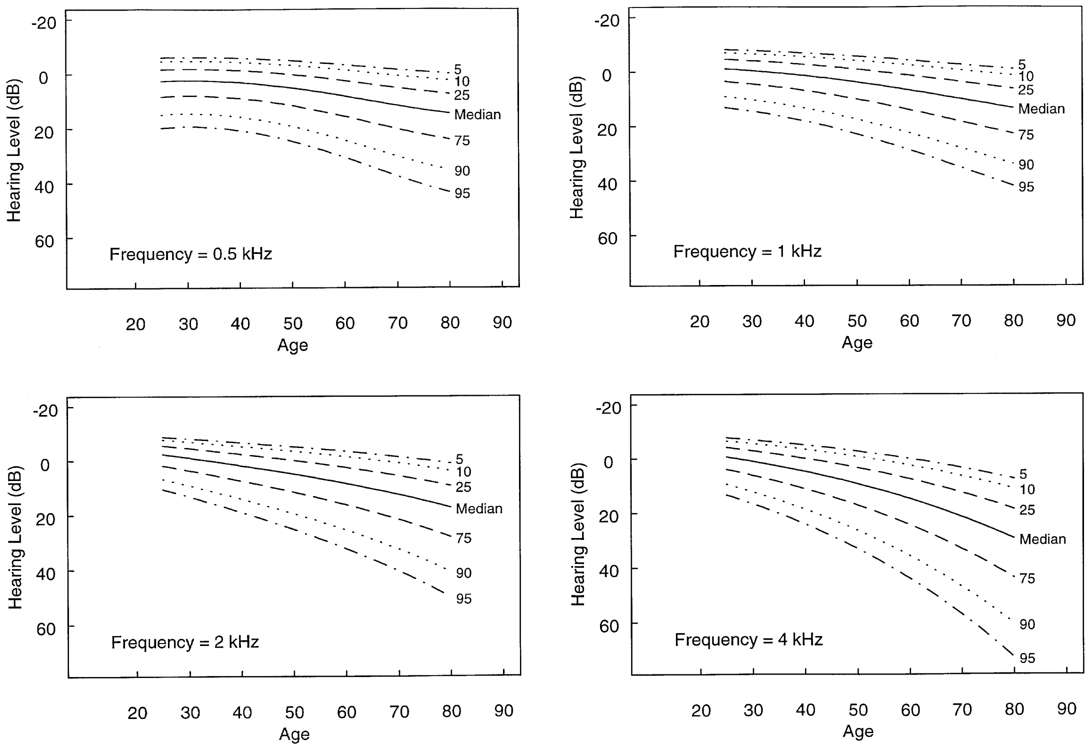

Figure 3. Percentile curves computed from the linear mixed-effects model for transformed data with constant variance in residuals and random effects

parameters of the resulting function appear in the first column of Table I. The percentiles for the transformed data are then:

$$X\hat{\boldsymbol{\beta}} + z_p \times \hat{\sigma}_P(\text{fage}, \ln(\text{freq})).$$

Figure 4 shows the percentiles for the untransformed data. The percentiles are now wider, especially at higher ages. However, the population residuals associated with the data obtained from each woman are not independent. Thus the results of the regression analysis on these population residuals may not be valid.

Non-Constant Variance in Cluster Residuals and Random Effects

As a second approach, we use the cluster (person specific) residual

$$\hat{\boldsymbol{e}}_i = \boldsymbol{y}_i - X_i\hat{\boldsymbol{\beta}} - Z_i\hat{\boldsymbol{b}}_i. \tag{3}$$

The LME model assumes that $\text{var}(\boldsymbol{e}_i) = \sigma^2 I$, so that the errors for an individual's measurements are independent. While the cluster residuals are not independent, they do have very small correlations so that the independence assumption of linear regression is not violated adversely.

Table I. Parameter estimates of the standard deviation function of the residuals. The parameters for the regressions on absolute residuals must be multiplied by $(\pi/2)^{1/2}$ to obtain the estimated standard deviation function

Variable	Regression on absolute population residuals	Regression on absolute cluster residuals	WLS regression on absolute cluster residuals
Intercept	0·2980543	0·1768215	0·1764454
fage	0·00070201	− 0·000925408	− 0·000954173
fage^2	0·0000047865	− 0·0000158289	− 0·000015658
fage^3	− 0·0000002004	− 0·0000003126	− 0·0000002509
ln(freq)	0·01587456	0·017246286	0·01733571
$\ln^2(\text{freq})$	− 0·017806145	− 0·0026337024	− 0·002667801
$\ln^3(\text{freq})$	0·00223408	− 0·006379758	− 0·006341108
$\text{fage} \times \ln(\text{freq})$	0·000114931	− 0·000520406	− 0·0005274326
$\text{fage} \times \ln^2(\text{freq})$	− 0·000239387	− 0·000136155	− 0·0001302225
$\text{fage} \times \ln^3(\text{freq})$	0·0001810754	0·000013536	0·0001348578
$\text{fage}^2 \times \ln(\text{freq})$	0·00005394	0·0000164459	0·0000166894
$\text{fage}^2 \times \ln^2(\text{freq})$	0·0000118212	0·0000071263	0·0000067064
$\text{fage}^2 \times \ln^3(\text{freq})$	− 0·0000285415	− 0·0000072157	− 0·0000071340
$\text{fage}^3 \times \ln(\text{freq})$	− 0·0000002949		
$\text{fage}^3 \times \ln^2(\text{freq})$	− 0·0000011018		

We model the variance in the cluster residual as a function of first age and ln(freq) as in the previous section to obtain $\sigma(\text{fage}, \ln(\text{freq}))$.

The estimates of the random effects are empirical Bayes' estimates,[7] $\hat{\boldsymbol{b}}_i = \hat{D} Z_i^{\mathrm{T}} W_i (\boldsymbol{y}_i - X_i \hat{\boldsymbol{\beta}})$ where $W_i = (\sigma^2 I + Z_i \hat{D} Z_i^{\mathrm{T}})^{-1}$. We model the non-constant variance in these random effects in a similar fashion. We model each of the absolute estimated random effects as a function of first age (a polynomial of order at most 3). We assume that the correlation structure remains constant over the age span. The covariance matrix is decomposed as

$$D(\text{fage}) = \text{cov}(\boldsymbol{b}_i) = V^{1/2} \text{corr}(\boldsymbol{b}_i) V^{1/2} = S\, \text{corr}(\boldsymbol{b}_i) S \tag{4}$$

where $V^{1/2} = S = \text{diag}(s_0(\text{fage}), s_1(\text{fage}), \ldots, s_{q-1}(\text{fage}))$ is a $q \times q$ diagonal matrix whose diagonal elements are the standard deviations of the q random effects as functions of first age. The variance of the marginal distribution is then calculated as:

$$\text{var}(\boldsymbol{y}) = \hat{\sigma}(\text{fage}, \ln(\text{freq}))^2 I + Z\hat{D}(\text{fage}) Z^{\mathrm{T}}$$

where $\hat{D}(\text{fage})$ is decomposed as in (4). The second column of Table I gives the estimated coefficients of the standard deviation function for the cluster residuals and the first column of Table II gives the estimated coefficients of the standard deviation functions for each of the random effects. The estimated standard deviations for the random effects intercept and $\ln^2(\text{freq})$ increase with first age, remain constant for time and $\ln^3(\text{freq})$, and decrease for ear. While for some of these random effects the variance remains constant or even decreases on the transformed scale, once we transform the data back to the original scale they will exhibit increasing variation since the hearing level increases with age. The percentiles on the transformed scale are

$$X\hat{\boldsymbol{\beta}} + z_P \times \sqrt{\{\hat{\sigma}(\text{fage}, \ln(\text{freq}))^2 I + Z\hat{D}(\text{fage}) Z^{\mathrm{T}}\}}.$$

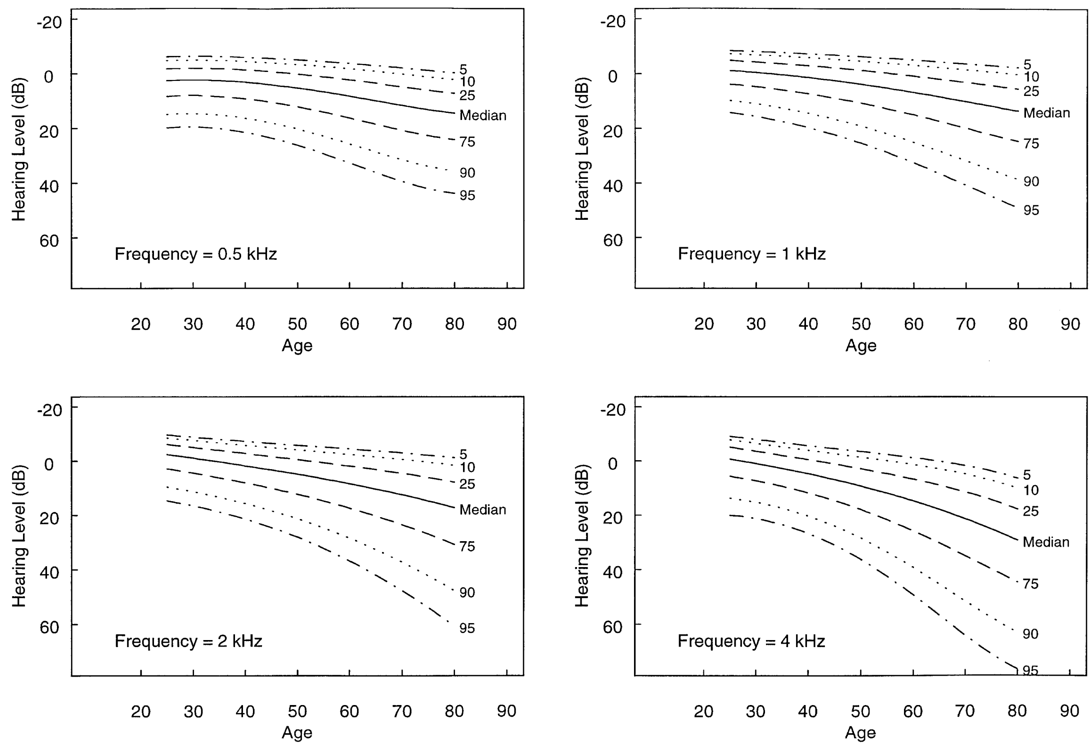

Figure 4. Percentile curves computed from the non-constant standard deviation function in the population residuals for transformed data computed from regressions using the absolute values of the population residuals

Table II. Parameter estimates of the standard deviation functions of the random effects. The parameters for the regressions on absolute random effects must be multiplied by $(\pi/2)^{1/2}$ to obtain the estimated standard deviation functions

Random effect	Variable	Regression on absolute random effects	WLS regression on absolute random effects
b_0-Intercept	Intercept	0·21071335	0·21192842
	fage	0·00192444	0·00206399
b_1-Ear	Intercept	0·063431	0·06338411
	fage	− 0·00045033	− 0·00041733
b_2-Time	Intercept	0·01790994	0·0165208456
	fage		− 0·0001125164
b_3-ln(freq)	Intercept	0·159129	0·15962
	fage		
b_4-$\ln^2$(freq)	Intercept	0·07547885	0·07576808
	fage	0·00046646	0·00044405
b_5-$\ln^3$(freq)	Intercept	0·04363397	0·04474316
	fage		

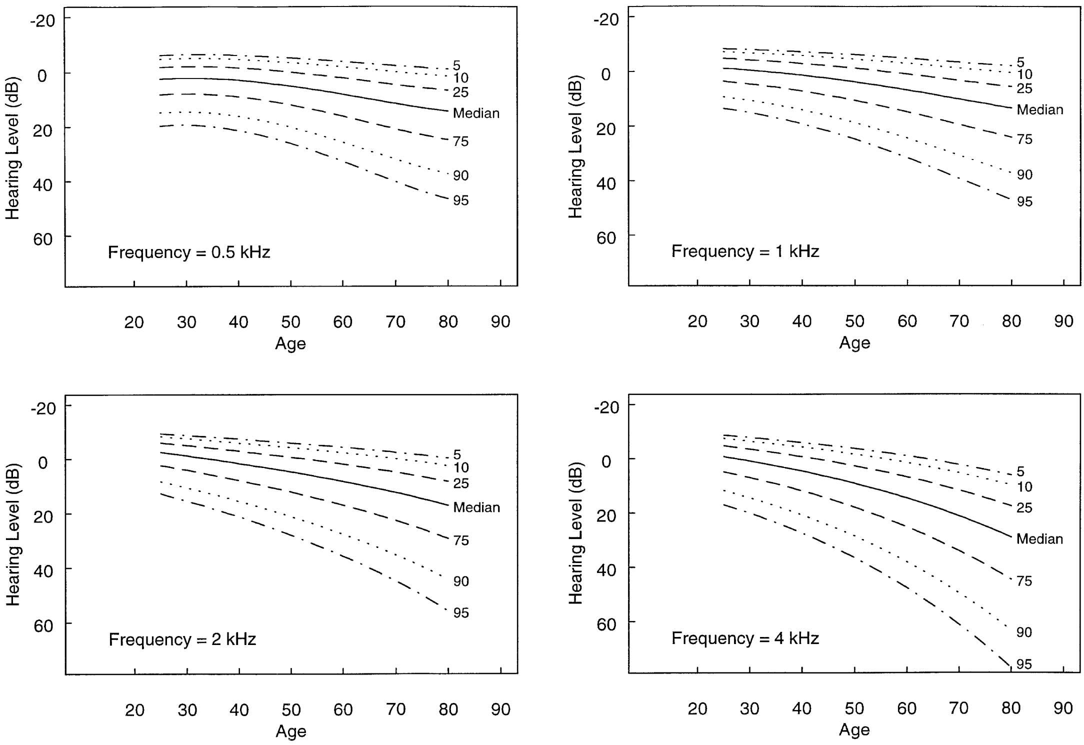

Figure 5. Percentile curves computed from the non-constant standard deviation functions for the cluster residuals and random effects for transformed data computed from regressions using the absolute values of the cluster residuals and random effects

Figure 5 shows these percentiles on the original scale which appear very similar to those in Figure 4.

Non-Constant Variance in Cluster Residuals and Random Effects: Weighted Least Squares

This approach is similar to the method in the previous section but uses weighted least squares to perform the regressions. The weights are the reciprocal of the estimated standard deviations of the cluster residuals and random effects. To assess the error in the estimation of the random effects, Laird and Ware[7] suggest the use of

$$\operatorname{cov}(\hat{\boldsymbol{b}}_i - \boldsymbol{b}_i) = \hat{D} - \hat{D}Z_i^{\mathrm{T}} \left\{ W_i - W_i X_i \left(\sum_{i=1}^{M} X_i^{\mathrm{T}} W_i X_i \right)^{-1} X_i^{\mathrm{T}} W_i \right\} Z_i \hat{D}.$$

We can write the cluster residual in (3) as

$$\hat{\boldsymbol{e}}_i = (\boldsymbol{y}_i - X_i\hat{\boldsymbol{\beta}}) - Z_i\hat{D}Z_i^{\mathrm{T}} W_i(\boldsymbol{y}_i - X_i\hat{\boldsymbol{\beta}}) = (I - Z_i\hat{D}Z_i^{\mathrm{T}} W_i)(\boldsymbol{y}_i - X_i\hat{\boldsymbol{\beta}})$$

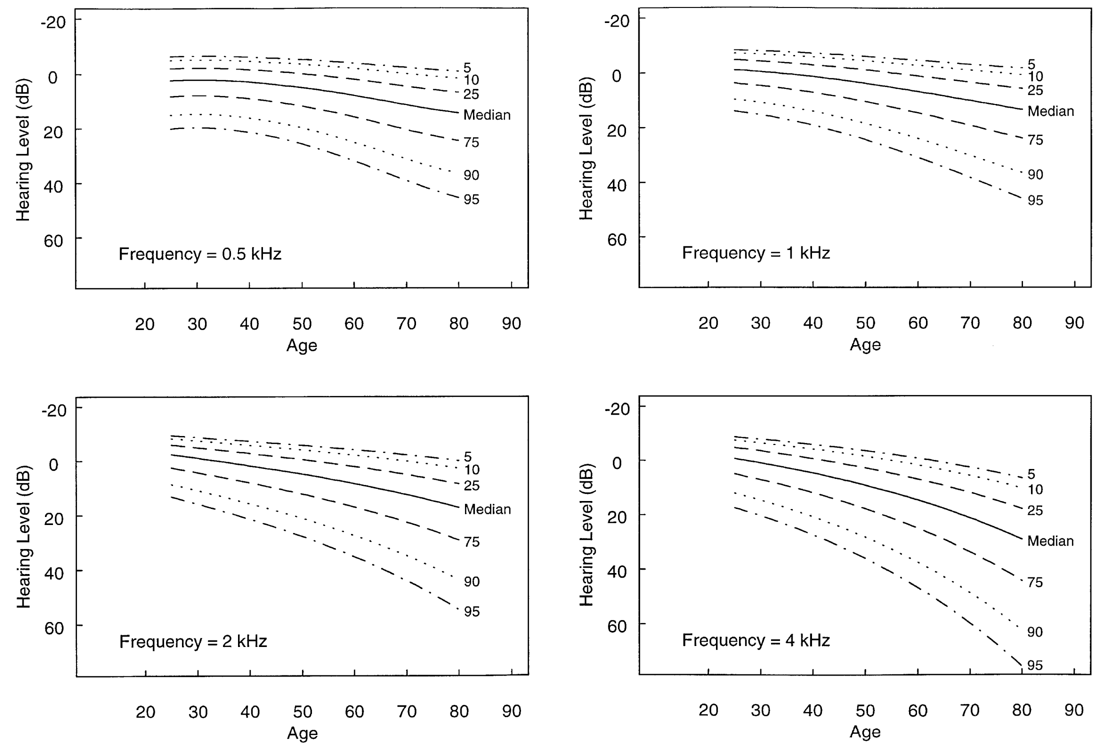

Figure 6. Percentile curves computed from the non-constant standard deviation functions for the cluster residuals and random effects for transformed data computed from regressions using the absolute values of the cluster residuals using weighted least squares

so that

$$\operatorname{cov}(\hat{\boldsymbol{e}}_i) = (I - Z_i\hat{D}Z_i^{\mathrm{T}}W_i)(W_i^{-1} - X_i\operatorname{cov}(\hat{\boldsymbol{\beta}})X_i^{\mathrm{T}})(I - Z_i\hat{D}Z_i^{\mathrm{T}}W_i)^{\mathrm{T}}.$$

We use the reciprocals of the square root of the variances from the diagonals of these matrices to provide the weights for the weighted least squares. The coefficients of these regressions are in column 3 of Table I and column 2 of Table II. The results are almost identical to the unweighted least squares calculations except that the regression for the random effect for time decreases with first age. Figure 6 shows these percentiles which are almost identical to those in Figure 5 at the frequencies shown. Thus, in this example, use of weighted least squares has little effect on the estimates of the standard deviation functions.

CONCLUSIONS

We have presented several methods of constructing percentiles based on the output from the linear mixed-effects model when the amount of variability is not constant. We investigated both non-constant variability in the population residuals as well as in the cluster residuals and random effects. The population residuals are not independent for data associated with an individual and so we violate the assumptions of linear regression. Our study included subjects with diverse numbers of observations, but the extra effort needed to compute the variances of the cluster residuals and random effects to perform a weighted least squares did not produce percentiles that differed much from the least squares results. Based on these comments, we recommend modelling the absolute cluster residuals and absolute random effects using least squares linear regression to construct the percentile plots. In addition, the percentiles constructed using this method (Figure 5) match well the original first visit data (Figure 2).

Proc Mixed in SAS[17] may also be used to fit the linear mixed-effects model. Proc Mixed allows for a number of covariance structures for the random effects and for the error term. However, these covariance structures are not flexible enough to allow for the modelling of the changing variability in the random effects and errors as a function of age and ln(freq) while keeping the correlation structure between the random effects the same over the age and frequency domains.

ACKNOWLEDGEMENTS

The work of Christopher Morrell was supported by research grant number 1 R03 DC 02566-01 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We thank the referees and editor for their useful comments.

REFERENCES

1. Royston, P. 'Constructing time-specific reference ranges', *Statistics in Medicine*, **10**, 675–690 (1991).
2. Royston, P. and Matthews, J. N. S. 'Estimation of reference ranges from normal samples', *Statistics in Medicine*, **10**, 691–695 (1991).
3. Cole, T. J. and Green, P. J. 'Smoothing reference centile curves: the LMS method and penalized likelihood', *Statistics in Medicine*, **11**, 1305–1319 (1992).
4. Green, P. J. and Silverman, B. W. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman and Hall, London, 1994.
5. Altman, D. G. 'Construction of age-related reference centiles using absolute residuals', *Statistics in Medicine*, **12**, 917–924 (1993).
6. Pearson, J. D., Morrell, C. H., Gordon-Salant, S., Brant, L. J., Metter, E. J., Klein, L. L. and Fozard, J. L. 'Gender differences in a longitudinal study of age-associated hearing loss', *Journal of the Acoustical Society of America*, **97**, 1196–1205 (1995).
7. Laird, N. M. and Ware, J. H. 'Random effects models for longitudinal data', *Biometrics*, **38**, 963–974 (1982).
8. Shock, N. W., Greulich, R. C., Andres, R., Lakatta, E. G., Arenberg, D. and Tobin, J. D. *Normal Human Aging: The Baltimore Longitudinal Study of Aging*, NIH Publication No. 84-2450, U.S. Government Printing Office, Washington, D.C. 1984.
9. ANSI S3·6-1989 *American National Standard Specification for Audiometers*, American National Standards Institute, New York, 1989.
10. Lindstrom, M. J. and Bates, D. M. 'Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data', *Journal of the American Statistical Association*, **83**, 1014–1022 (1989).
11. Longford, N. T. *Random Coefficient Models*, Clarendon Press, Oxford, 1993.
12. Laird, N., Lange, N. and Stram, D. 'Maximum likelihood computations with repeated measures: application of the EM algorithm', *Journal of the American Statistical Association*, **82**, 97–105 (1987).

13. Brant, L. J. and Fozard, J. L. 'Age change in pure-tone hearing thresholds in a longitudinal study of normal human aging', *Journal of the Acoustical Society of America*, **88**, 813–820 (1990).
14. Morrell, C. H. and Brant, L. J. 'Modeling hearing thresholds in the elderly', *Statistics in Medicine*, **10**, 1453–1464 (1991).
15. Peixoto, J. L. 'A property of well-formulated polynomial regression models', *American Statistician*, **44**, 26–30 (1990).
16. Verbeke, G. N. M. 'The linear mixed model. A critical investigation in the context of longitudinal data', unpublished Ph.D. thesis, Catholic University of Leuven, 1995.
17. SAS Institute Inc. *SAS Technical Report P-229, SAS/Stat Software: Changes and Enhancements, Release 6·07*, SAS Institute Inc., Cary, NC, 1992.

