A log-normal distribution model of the effect of bacteria and ear fenestration on hearing loss: a Bayesian approach

Byron J Gajewski, *University of Kansas Medical Center*
Jack D Sedwick, *University of Florida*
Patrick J Antonelli, *University of Florida*
A log-normal distribution model of the effect of bacteria and ear fenestration on hearing loss: a Bayesian approach

Byron J. Gajewski1,∗,†,‡; Jack D. Sedwick2 and Patrick J. Antonelli2

1Schools of Allied Health and Nursing, The University of Kansas, Kansas City, KS 66160, U.S.A.
2Department of Otolaryngology, University of Florida, Gainesville, FL 32610, U.S.A.

SUMMARY

Chronic ear infection is a potentially life-threatening illness that medical doctors typically treat with ear surgery. Despite the success of this treatment, complications can occur due to bacteria infection. Surgeons believe that this infection causes the patient to have clinically significant hearing damage. In order to understand such complications, surgeons must quantify the effect of bacteria, their toxins and ear surgery on hearing loss. To this end, the other two authors of this paper performed two experiments on guinea pigs to measure hearing thresholds following a bacterial infection and surgery of the inner ear. The response variable in these experiments is hearing thresholds measured in decibels (dB). The problem in analysing such experiments is that the hearing threshold observations often suffer from missing data and censoring mechanisms of various types. Additionally, the distribution of hearing thresholds has heavy tails and is peaked. In order to account for the above statistical issues, we present a Bayesian method with a location-shifted log-normal distribution. The method accounts for the uncertainty in the data collection mechanism and the parameters associated with a location-shifted log-normal distribution. We refer to one of the parameters as the ‘location-shift’ parameter. The Bayesian approach provides a posterior distribution of the location-shift parameter that we compare with values estimated in previously published studies. The immediate goal of our proposed method was to quantify the effects of ear surgery and bacteria infection on hearing loss. Thus, we present the merits of the method in the form of a case study, and report posterior distributions of mean hearing loss, probability of clinically significant hearing loss and relative risk. The results show that surgeon 2, using the surgical procedure ‘oval window’, poses a greater than 40 per cent chance of a 15 dB hearing loss regardless of injection of bacteria or not. However, surgeon 1, using the surgical procedure ‘semicircular canal’, does not pose a significantly greater than 40 per cent chance of a 15 dB hearing loss unless there is a Pseudomonas aeruginosa-induced infection. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: Gibbs sampler; censored data; hierarchical model; relative risk; data augmentation; MCMC

1. INTRODUCTION

Chronic ear infection is a potentially life-threatening illness [1] that afflicts up to 5 per cent of some populations [2]. Treatment typically requires middle ear and mastoid surgery. While
most surgical treatments are safe and effective, serious complications such as deafness can occur [3]. Deafness following surgical treatment commonly occurs after injury to the inner ear [4]. While the inner, non-inflamed ear can be manipulated without compromising hearing [5], bacterial infection has been postulated to contribute to the origin and the development of disease infections leading to surgically induced hearing loss [6]. In order to prevent such complications, the effects of bacteria, their toxins and ear surgery on hearing loss must be clarified. To this end, two experiments were performed on guinea pigs to measure hearing thresholds following induced bacterial infection and surgical violation of the inner ear. Hearing thresholds were measured in decibels (dB). See Reference [7] for a practical explanation on dB-measured thresholds.

Auditory thresholds were determined with a measuring technique called electrocochleography. Ball-ended electrodes were implanted immediately prior to ear canal manipulation. Hearing thresholds were measured using clicks. Stimuli were generated using an auditory workstation and were introduced into the external auditory canal with an insert earphone tube. Thresholds were determined by decreasing sound intensities in 5 dB increments from a maximum until the waveform was lost, then raising the stimulus by 5 dB increments until a reproducible waveform was restored. This restored waveform corresponds to the recorded hearing threshold. The accuracy of this method was to within 5 dB, and the maximum intensity was 120 dB.

Our main interest was in the effects of the bacteria types \textit{P. aeruginosa} (PA) and \textit{Streptococcus pneumoniae} (SP) on long-term, post-surgical hearing loss. In order to determine effect over time, hearing thresholds of each subject were measured following bacteria injection but prior to surgery, immediately after surgery and 1-week later. For each subject, hearing thresholds in both ears were measured, regardless of intervention, resulting in a total of six measurements.

A different surgeon performed each of the two experiments, using a different ear surgery approach, thus potentially increasing the variation in experimental results. Surgeon 1 used the ‘semi-circular canal’ technique and surgeon 2 used the ‘oval window’ technique. Thus, surgeon effect is confounded with surgical technique. Note that confounding is not overly harmful because the main interest was in comparing the bacteria. Each experiment had a control group consisting of guinea pigs receiving surgery on one randomly selected ear per subject. The other ear received no surgery. No bacterium was induced in the control group subjects. In the first experiment (surgeon 1), each subject in the comparison group was injected with PA in both ears, and then received surgery on one randomly selected ear. In the second experiment (surgeon 2), this procedure was repeated twice; once using a PA group, and once using an SP group. No guinea pigs were used more than once as each was sacrificed following their six measurements.

The key scientific questions of interest all centre on mean differences and probabilities of certain hearing threshold changes. Specifically, what was the probability that a guinea pig would have lost 15 dB of hearing in a surgically compromised ear, and how does this probability change as a function of bacteria group? From a statistical point of view, we desired a method that answered these questions with a parametric model and simultaneously accounted for the uncertainty in:

1. usual parameters associated with a parametric model;
(2) an added parameter that models extra heavy tails and peakness not found in the normal distribution and
(3) the nature of the data collection which included missing, grouped and right-censored hearing threshold observations.

To answer the scientific questions, as well as account for the list of uncertainties, we present a Bayesian method for analysing the data from these experiments. Our approach also allows us to quantify uncertainty in the mean differences and probabilities by accounting for uncertainty in parameters conditional on data. A review of the current methods is presented in Section 2. We propose a repeated measures log-normal model with additional location-shift parameter, and report the prior distributions for model parameters in Section 3. In Section 4, we provide a discussion of the multifaceted computational issues related to our Bayesian application. The resulting posterior distributions, their relation to parameter estimates previously reported in the literature and a discussion of the scientific implications are presented in Section 5. Finally, we present our conclusions in Section 6.

2. GENERAL MODELLING AND DATA

2.1. Approaches

The quantification of hearing threshold is measured in decibels (dB). Several hearing threshold research studies employ a log-normal distribution with an added parameter, \( \alpha \), to model hearing thresholds [8, 9]. This parameter accounts for longer tails than those found in the traditional log-normal distribution [8], and is introduced into the probability model by adding it to the response, \( y \), and taking the natural logarithm, or \( z = \log(y + \alpha) \). The parameter \( \alpha \) prevents any problems associated with taking the logarithm of zero. To demonstrate the effect of the parameter \( \alpha \), consider the distribution of \( y \) written

\[
 f(y | \mu, \sigma^2, \alpha) = \exp\left\{ \frac{-(\log(y + \alpha) - \mu)^2}{2\sigma^2} \right\} \sqrt{\frac{2\pi \sigma}{y + \alpha}},
\]

where \( y + \alpha \) is between 0 and \( \infty \), \( -\infty < \mu < \infty \), \( \sigma > 0 \) and \( \alpha > 0 \). Thus, the pdf of \( y \) consists of three unknown parameters \( \mu \), \( \sigma^2 \) and \( \alpha \). Several versions of the pdf are presented in Figure 1, all with \( E(y) = 30 \) dB, but \( \alpha \) varying from 0, 10 to 20. The increase in \( \alpha \) results in heavier right and left tails and a lower peak. We refer to \( \alpha \) as the location-shift parameter.

In general, the location-shift parameter is unknown. Bowater et al. [8] use maximum likelihood estimation (MLE) to estimate the location-shift parameter. They report the location-shift parameter estimates based on age for an analysis of hearing threshold in humans. The resulting estimates for subjects 51–60 years of age and 71–80 years of age were \( \hat{\alpha} = 9.10 \) (standard deviation = 0.98) and \( \hat{\alpha} = 28.20 \) (standard deviation = 8.21), respectively. Longford [9] fixes the location-shift parameter at \( \alpha = 20 \). In Reference [10], the authors model thresholds without a location-shift parameter. Estimation of the location-shift parameter for our data set is an important modelling consideration, but the nature of the data collection scheme presents analysis challenges.

Unique to our modelling solution is the incorporation of missing data and censored data directly into the modelling. To account for censoring we utilize Bayesian data augmentation. See Reference [11] for an example. In addition to addressing data censoring, the Bayesian perspective provides a posterior distribution of the location-shift parameter, as well as other parameters, rather than the point estimates provided in the previously cited studies. Our Bayesian
approach to parameter estimation in the hearing data problem is presented using first principles found in recent Bayesian literature [12]. Ours is the first application of Bayesian methods to hearing threshold studies.

2.2. Data

A key aspect of our Bayesian approach is that it allows us to easily incorporate censoring into the estimation process. This is of particular importance in our case as censoring is commonly found in hearing threshold studies. As implied in the introduction, the three types of censored data we encounter are interval censored, right censored and missing data. Interval-censored data are responses between two finite values. In this case, finite values are in increments of 5dB (e.g. 0–5, 5–10, 10–15, etc.). Therefore, in the interval censored cases the observed values, $y_{OBS}^{ij}$, are the lower end of the interval. Right-censored data are responses between 120 dB and infinity. Finally, as the name implies, missing data occur when no data are available and thus are treated as being within the range of all feasible values (i.e. 0-infinity).

Of the 50 guinea pigs by six hearing thresholds, totalling 300 measurements, 19 were missing due to death of six subjects and seven were right censored. To quantify the long-term effect of surgery and bacterial injection on hearing loss, we compare the difference between the before surgery measurements and the 1-week post-operative measurements. The immediate post-operative measurements were included in the modelling to help estimate the variability of hearing thresholds. To illustrate the long-term effects, we refer the reader to the vertical lines.
Figure 2. Here is a presentation of the raw data summary statistics for hearing thresholds on animals. The top of each vertical line is the mean plus one standard deviation, and the bottom of the line is the mean minus one standard deviation. On the horizontal line the measures are: (1) before surgery (no surgery ear); (2) immediately after surgery (no surgery ear); (3) one week after surgery (no surgery ear); (4) before surgery (surgery ear); (5) immediately after surgery (surgery ear) and (6) one week after surgery (surgery ear).

in Figure 2. The raw data are not yet manipulated and so are in the form of interval censored, right censored and missing. The missing values are dropped for the purpose of calculating the summary statistics. The difference, before minus after surgery, in the sample average hearing loss for this group is 33.1 minus 40.6 dB, resulting in a loss of a sample average of 7.5 dB.
We compare this difference to the PA-induced surgical sample average for surgeon 2, which is 38 minus 78 dB, resulting in a much more drastic sample hearing loss of 40 dB. The results shown in Figure 2 indicate that surgeon 1 has a greater overall impact on hearing loss than surgeon 2. However, surgeon effect is confounded by technique. This descriptive analysis is extended with our formal method built with a location-shifted log-normal linear mixed model, that accounts for the interval and right censoring.

3. SPECIFIC MODEL AND PRIOR DISTRIBUTIONS

3.1. General model

Consider our model for the two experiments described previously, which involve six measurements on $N = 50$ guinea pigs. Define log as a vectorized function so $Z_i = \log(Y_i + \alpha \mathbf{1})$ where $Z_i$ and $Y_i$ are column vectors of length six, $\mathbf{1}$ is a column vector of six ones and $\alpha$ is the location-shift parameter. The elements of the column vector are indexed by $j = 1, 2, 3, 4, 5, 6$ so that $y_{ij}$, the uncensored hearing threshold from the $j$th measurement on the $i$th guinea pig, represents the $j$th element of $Y_i$. The main model becomes

$$Z_i = X_i \beta + 1d_i + F_i + e_i, \quad i = 1, 2, 3, \ldots, 50$$

where:

(i) $\beta$ is the vector of fixed-effects parameters of length 30;
(ii) $d_i$ is a scalar that represents the guinea pig random effect independently distributed as $N(0, \sigma_d^2)$;
(iii) $F_i$ is a 6-vector with the first three elements equal to $f_{i1}$ and the second three elements equal to $f_{i2}$. The $f_{i1}$ are the two ear random effects for guinea pig $i$ independently distributed $N(0, \sigma_f^2)$; and
(iv) $e_i$ is the 6-vector of residuals $e_{ij}$ each independently distributed as $N(0, \sigma_e^2)$.

The distributions in (ii)–(iv) are independent. The natural log transformation is used to relate the hearing thresholds to the log-normal distribution, while additionally accounting for location-shift using the parameter $\alpha$. The $X_i$ matrix, $6 \times 30$, indicates which group and measurement the $i$th subject belongs to. Using the Kronecker product notation, $X_i = M_i \otimes I_6$, where the first three rows of the identity matrix $I_6$ maps the fixed effects for three measurements in order of time measured for the non-surgical ear, and the second three rows map the three measurements in order of time measured for the surgical ear. The row vector, $M_i$, maps the group containing subject $i$. All possible values for $M_i$ are $M_i = [1, 0, 0, 0, 0, 0]$ for control surgeon 1; $M_i = [0, 1, 0, 0, 0]$ for PA surgeon 1; $M_i = [0, 0, 1, 0, 0]$ for control surgeon 2; $M_i = [0, 0, 0, 1, 0]$ for SP surgeon 2; and $M_i = [0, 0, 0, 0, 1]$ for PA surgeon 2. For illustration, all subjects in control surgeon 1 have $X_i = [1, 0, 0, 0, 0, 0] \otimes I_6 = [I_6, \mathbf{0}_6, \mathbf{0}_6, \mathbf{0}_6]$. The notation for $X_i$ defines $\beta = [\beta_1' \beta_2' \beta_3' \beta_4' \beta_5']'$, where the first six elements are the fixed effects, placed in the $\beta_1$ vector, from control surgeon 1, the second six, $\beta_2$, are from PA surgeon 1, the third six, $\beta_3$, are from control surgeon 1, the fourth six, $\beta_4$, are from SP surgeon 2 and the last six elements, $\beta_5$, are from PA surgeon 2.
In closing this subsection, we note that the model can be written in compact form as

\[ \mathbf{Z}_i \mid \mathbf{\beta}, \sigma_\beta^2, \sigma_d^2, \sigma_e^2, \mathbf{X}_i, \alpha \sim \text{MVN}(\mathbf{X}_i, \mathbf{\beta}, \Sigma) \]

where ‘MVN’ stands for multivariate normal distribution and \( \Sigma \) is made up of \( \sigma_d^2 + \sigma_e^2 \) in the diagonal, \( \sigma_d^2 + \sigma_e^2 \) is the covariance for measures within ear and \( \sigma_d^2 \) is the covariance between ears.

### 3.2. Priors

Except for the location-shift parameter, our Bayesian framework models the unknown parameters using prior distributions that are all of a semi-conjugate form [12]. The prior distribution of the location-shift parameter is not semi-conjugate. The variance parameters are modelled with fairly diffuse inverse gamma distributions, specifically \( \sigma_d^2 \sim \text{IG}(1, 0.1) \), \( \sigma_e^2 \sim \text{IG}(1, 0.1) \) and \( \sigma_f^2 \sim \text{IG}(1, 0.1) \).

The fixed-effects parameters, surgery crossed with bacteria, are modelled with \( \mathbf{\beta} \sim \text{MVN}(\mathbf{\eta}, \Sigma_0) \). The elements of the \( \mathbf{\eta} = [\eta_1, \eta_2, \eta_3, \eta_4, \eta_5]' \) vector have the same indexes as \( \mathbf{\beta} \). The elements of the \( \mathbf{\eta} \) vector are set to 0, representing the mean prior response of SP from surgeon 2; \( \eta_1 = \eta_3 = \mathbf{B}_1 \), represents the mean prior vector for control from surgeons 1 and 2, respectively; and \( \eta_2 = \eta_5 = \mathbf{B}_2 \), represents the mean prior vector for PA from surgeons 1 and 2, respectively. The prior distributions for \( \mathbf{B}_1 \) and \( \mathbf{B}_2 \) are multivariate normal with a mean of \( \mathbf{0} \) and variance \( 10^3 \mathbf{I} \). The matrix, \( \Sigma_0 \), is diagonal and has \( \sigma_i^2 \) in each of its first 18 diagonal elements, \( 10^3 \) in the next six diagonal elements and \( \sigma_i^2 \) in the last six elements. Because \( \sigma_i^2 \mathbf{I}_6 \) is the variance of the fixed parameters, it represents a compromise in the borrowed strength between the experiments performed by the two surgeons. That is, if \( \sigma_i^2 \) was zero, the mean parameter for surgeons in PA (control) would be identical. Specifically, \( \beta_1 = \beta_3 \) and \( \beta_2 = \beta_5 \). Conversely, if \( \sigma_i^2 \) was infinite, then the mean parameters for the two surgeons are independent. To this end, we treat this variance component as a random variable; specifically \( \sigma_i^2 \sim \text{IG}(1, 0.1) \).

The prior parameters can be placed into two categories, both interpreted on the log of the response plus \( \alpha \) scale. The first category holds the mean parameters, the elements of \( \mathbf{\beta} \), and the second category holds the variance components, \( \sigma_r^2, \sigma_d^2, \sigma_e^2 \) and \( \sigma_f^2 \). When comparing the variances of our priors to the estimated values in Bowater et al, one can see that, in fact, we have diffuse priors. For the mean parameter, Bowater et al. quote values in the range of \( 3-4 \), well within the range of our prior. For the variance component, Bowater et al. quote values from 0.09 to 0.25, which is the 33 and the 67 percentile of the \( \text{IG}(1, 0.1) \).

Finally, we chose to model the location-shift parameter, \( \alpha \), as a uniform distribution from 1 to 30. We chose this prior because previous studies indicate that location-shift parameters outside of this range are unlikely [8,9].

### 3.3. Modelling censored data

For the rest of the analysis, we assume that the data are missing completely at random (MCAR). Thus, the probability of a censored value, like interval, right and missing, does not depend on the observed data or the unobserved data. To account for censored data in this case, one simply specifies the region where the response must have fallen.
Based on the observed data, which are stored in the matrix $Y_{OBS}$, we define the theoretical values, stored in $Y$, and account for its uncertainty. Specifically, we model the actual value, $y_{ij}$, to be between $c_{1ij}$ and $c_{2ij}$, where $c_{1ij}$ is the lower end of the interval, and $c_{2ij}$ is the upper end of the interval. The three cases of these are

$$
c_{1ij} = \begin{cases} 
y_{OBSij} & \text{if interval censored} \\
120 & \text{if right censored} \\
0 & \text{if missing} 
\end{cases}
$$

and

$$
c_{2ij} = \begin{cases} 
y_{OBSij} + 5 & \text{if interval censored} \\
\infty & \text{if right censored} \\
\infty & \text{if missing} 
\end{cases}
$$

Therefore, in the interval-censored cases the observed values, $y_{OBSij}$, are the lower end of the interval. Consequently, the calculation $\int_{c_{1ij}}^{c_{2ij}} f(y_{ij} \mid \theta) \, dy_{ij}$, where $\theta$ is the vector of unknown parameters, is the likelihood contribution for $y_{ij}$. We incorporate this calculation into our Bayesian method. This general treatment of the endpoints is useful for programming purposes since we can treat each observation, $y_{ij}$, as bounded between $c_{1ij}$ and $c_{2ij}$. This replacement of the observed data with the ‘actual’ data is an example of data augmentation. For a recent paper on data augmentation, see Reference [13].

Our data augmentation method uses a simulation that we explain assuming that the model parameters are known. First, a uniform random variable between $\mathcal{U}(\log(c_{1ij} + \alpha) \mid \mu_j, \sigma^2)$ and $\mathcal{U}(\log(c_{2ij} + \alpha) \mid \mu_j, \sigma^2)$ is generated, which we call $u_{ij}$. The value $\Phi(\bullet \mid \mu_j, \sigma^2)$ is the normal cumulative density with mean parameter $\mu_j$ and variance parameter $\sigma^2$. This random variable is transformed back to the hearing threshold scale with the equation $y_{ij} = \exp\{\Phi^{-1}(u_{ij} \mid \mu_j, \sigma^2)\} - \alpha$.

In our approach to data augmentation using a Bayesian framework, the nature of the location-shifted log-normal model provides non-uniform imputations within each of the defined bounds. The Bayesian approach is set up in a similar manner to a frequentist parametric approach through a likelihood function. But rather than an MLE, we calculate a posterior distribution of the model parameters that simultaneously accounts for uncertainty associated with the observed data. As such, we obtain posterior distributions from which adequate summaries such as percentiles, means and variances can be computed.

4. MOMENT CALCULATION AND COMPUTATION

4.1. Moments and probability calculations

For the remainder of our discussion, we make statistical inference on the mean and probability function of the values from the sampling distribution $Y$. With help from properties of the log-normal distribution the mean is

$$
E(y_{ij} + \alpha \mid \beta, \sigma_j^2, \sigma_d^2, \sigma_e^2, X_i, \alpha) = \exp\{(X_i \beta) + (\sigma_j^2 + \sigma_d^2 + \sigma_e^2)/2\}
$$

Notice that we use the $i$ subscript to indicate means associated with a certain measure within a group. Therefore, the mean represents a prediction for future guinea pigs and not the ones already in the data set. We can then calculate $E(y_{ij} - y_{ij'} \mid \beta, \sigma_j^2, \sigma_d^2, \sigma_e^2, X_i, \alpha)$, where
measurement \( j \) typically represents the index before surgery, and \( j’ \) represents 1-week after surgery or immediately after surgery. We also wish to report our results as a probability in the form of \( \Pr(y_{ij} - y_{ij'} < -15\,\text{dB} \mid \beta, \sigma_f^2, \sigma_d^2, \sigma_e^2, X_i, \alpha) \) because a hearing loss less than \(-15\,\text{dB}\) is ‘highly clinically significant’ [6]. However, this probability is difficult, if not impossible, to calculate in closed form. Given this challenge, conditional on the parameters, we generate 10,000 simulated values to calculate the probability by recording the proportion of times; the difference between appropriate values is less than \(-15\,\text{dB}\).

### 4.2. Computational discussion

We use Gibbs sampling to calculate the posterior distributions of all unknown model parameters. We use MATLAB, a matrix-based program, to calculate posterior distributions. This program requires complete conditionals [12] to be written into the code, allowing the user to call simulation functions. Our MATLAB code addresses censoring and non-standard distributions simultaneously in a straightforward manner. Thus, despite programming challenges, we find MATLAB a viable choice. Fortunately for us, most of the complete conditionals [12] are of semi-conjugate [12] form and are ‘usual’ distributions, thus easy to simulate. The complete conditionals for this program are provided in the appendix, and the MATLAB code used here can be obtained by contacting the first author.

Thus, the posterior distribution of the parameters from a Gibbs sampler, supplies the posterior distribution for \( E(y_{ij} - y_{ij'} \mid \beta, \sigma_f^2, \sigma_d^2, \sigma_e^2, X_i, \alpha) \) and for the measure \( \Pr(y_{ij} - y_{ij'} < -15\,\text{dB} \mid \beta, \sigma_f^2, \sigma_d^2, \sigma_e^2, X_i, \alpha) \). From the posterior distribution we can take the 5 percentile and the 95 percentile of the model parameters as well as both \( E(y_{ij} - y_{ij'} \mid \beta, \sigma_f^2, \sigma_d^2, \sigma_e^2, X_i, \alpha) \) and \( \Pr(y_{ij} - y_{ij'} < -15\,\text{dB} \mid \beta, \sigma_f^2, \sigma_d^2, \sigma_e^2, X_i, \alpha) \) to obtain 90 per cent credible regions for parameters, mean differences and probabilities associated with hearing threshold changes. Throughout Section 5, we extract 90 per cent Bayes credible regions for the various functions described in Section 4.1.

### 5. RESULTS

#### 5.1. The estimation of model parameters

We estimate the distributions of the model parameters using a Gibbs sampler. Inspection of the time and autocorrelation plots for all model parameters including the variance components and the location-shift parameter for various starting values indicate quick convergence to the stationary distribution. The distribution for the analysis uses a burn-in period [14] of 50,000 iterations, with inferences based on the second 50,000 iterations. For the purpose of this paper, we discuss summaries of the posterior distributions of the model parameters, starting with the location-shift parameter (\( \alpha \)).

Previous hearing threshold studies have used either estimation from the data or pre-selection of a fixed location-shift parameter. Our Bayesian method allows for a range of values for the location-shift parameter in the form of a posterior distribution. The posterior mean of \( \alpha \) is 4.64 (standard deviation = 2.89), similar to the MLE result found in Reference [8] of \( \hat{\alpha} = 9.10 \), but different than the Longford [9] result of \( \alpha = 20 \).
Consider exploring the posterior distribution for certain contrasts of the $\beta$ vector. The contrast $\beta_4 - \beta_6$ represents long-term hearing loss in the surgical ear for control surgeon 1, PA surgeon 1 is $\beta_{10} - \beta_{12}$, control surgeon 2 is $\beta_{16} - \beta_{18}$, SP surgeon 2 is $\beta_{22} - \beta_{24}$ and PA surgeon 2 is $\beta_{28} - \beta_{30}$. The largest drop in mean difference among these five contrasts is $\beta_{10} - \beta_{12}$, with a posterior mean of $-0.47$ (standard deviation $= 0.15$). Among the ten parameters that define the above contrasts, the largest posterior mean belongs to the parameter in the long-term PA surgeon 2, $\beta_{30}$, which has a posterior mean of 4.43 (standard deviation $= 0.17$), and the smallest posterior mean belongs to the fixed parameter in the before surgery control surgeon 2, $\beta_{16}$, which has a posterior mean of 3.37 (standard deviation $= 0.19$).

5.2. Model checking

We demonstrate model adequacy by comparing the imputed data generated from the observed data and posterior distribution of the model parameters to predictive data generated from the posterior distribution of the model parameters. Each imputed value, $y_{ij}$, is forced into the bound $(c_{1ij}, c_{2ij})$. Each predictive value, $y_{Pij}$, is not forced into the bound $(c_{1ij}, c_{2ij})$. Specifically, $Z_{pi} \sim \text{MVN}(X_{pi}, \Sigma)$ and $Y_{pi} = \log(Z_{pi} + z_1)$, thus predicted values do not use the same random effects from the original guinea pigs. Our comparison between the imputed values and the predicted values was performed with a $\chi^2$ discrepancy function, described in Reference [12, p. 172], that we calculate for both $Y$ and $Y_P$. The $\chi^2$ discrepancy function is defined as $\chi^2 = \sum_{i=1}^{N} (Z_{i} - X_{i}\beta)'\Sigma^{-1}(Z_{i} - X_{i}\beta)$ and the predictive discrepancy function is defined as $\chi^2_p = \sum_{i=1}^{N} (Z_{pi} - X_{pi}\beta)'\Sigma^{-1}(Z_{pi} - X_{pi}\beta)$. From these two $\chi^2$ discrepancy functions, we compute a Bayes $p$-value, described in Reference [12], by counting the proportion of times the $\chi^2$ discrepancy from the imputed data is larger than the $\chi^2$ discrepancy associated with the predicted data. We consider a $p$-value $<0.01$ or $>0.99$ to indicate a poor model fit. In our fully Bayesian approach we obtain a $p$-value $= 0.0344$, indicating that the model provides an adequate reflection of the data.

5.3. Addressing the scientific questions

The effect of bacteria infection and surgery on both short- and long-term hearing loss was the scientific question motivating our statistical analysis. We make inference by calculating $E(y_{ij} - y_{ij'} \mid \beta, \sigma_j^2, \sigma_d^2, X_i, z)$ as we describe in Section 4.1. We also make inference by calculating $Pr(y_{ij} - y_{ij'} < -15 \text{dB} \mid \beta, \sigma_j^2, \sigma_d^2, X_i, z)$. Specifically, this is the posterior distribution of the probability of the difference measures less than $-15 \text{dB}$, a clinically important drop in hearing threshold that we refer to as ‘highly clinically significant’ [6].

Figure 3 summarizes the mean of differences in hearing loss from before surgery relative to after, with 90 per cent Bayes credible regions (CR) [12]. Specifically these are regions of $E(y_{ij} - y_{ij'} \mid \beta, \sigma_j^2, \sigma_d^2, X_i, z)$ based on surgeons 1 and 2 that (1) show short- and long-term differences for control, SP and PA, and (2) attributes the differences to either the surgery or non-surgery ears. These Bayes credible regions provide an opportunity to decide whether the results in the hearing loss are considered somewhat clinically significant (90 per cent CR < 0 dB), clinically significant (90 per cent CR < $-10 \text{dB}$) or highly clinically significant (90 per cent CR $<-15 \text{dB}$). Examining the mean values of the surgery ears for surgeon 2, we immediately notice that the 90 per cent CR for long-term hearing loss for control, SP and PA are $<-15 \text{dB}$. The 90 per cent CR for the short-term effect of control is $<-10 \text{dB}$. Since a $15 \text{dB}$ drop in any group is considered highly clinically significant, this drop in the long-
A BAYESIAN ANALYSIS OF HEARING DATA 503

Figure 3. The following presents 90 per cent Bayes credible regions for $E(y_{ij} - y_{ij}' | \beta, \sigma^2, \sigma^2, X_i, x)$, where $y_{ij}$ is the hearing threshold before surgery and $y_{ij}'$ is the threshold after surgery, for each combination of surgeon, surgery and bacteria. The end of the lines represent the 5 and the 95 percentile. The horizontal lines represent 0, $-10$ and $-15$ dB changes.

term control group specifically points to the surgical technique of surgery from surgeon 2 as a significant cause of hearing loss. This is not an indictment of ‘oval window’ in the context of the experimental design surgeon that is confounded by technique. By comparison, surgery from surgeon 1 shows a 90 per cent CR for the long-term control group, which covers 0 dB. The 90 per cent CR for long-term damage in the non-surgical ears exposed to PA for surgeon 2 is $<0$ dB, showing evidence that the virulent nature of PA can cause hearing loss, independent of surgery.

The CR for the mean is informative, allowing us to find interesting and significant patterns in the data. However, we also report the probability of a 15 dB hearing loss and use this measure to draw conclusions about the data. The top part of Figure 4 reports a 90 per cent CR for the probability of a 15 dB hearing loss. This measure is a probabilistic way to explore the relative effect of surgery. For example, in the long-term surgical ear for both surgeons, the 90 per cent CR is above 0.40 for PA and SP. This is also the case for the control group for surgeon 2. In contrast, we see that all the long-term hearing losses for the non-surgical ears for surgeon 2 have 90 per cent CR covering 0.40, whereas the 90 per cent CR for control surgeon 1 is completely below 0.40.

By taking the ratio of two probabilities at each posterior sample, we can use the Gibbs sampler to quantify the relative risk, within surgeon, of bacteria infection on control. As an example calculation, take the probability of a 15 dB loss for the PA-infected surgery ear for
Figure 4. The following is a summary of the posterior distributions of the probability of a 15 dB hearing loss and the log-relative risk. Log-relative risk is the natural log of the probability of a 15 dB hearing loss for one effect divided by the probability of a 15 dB hearing loss for another effect. The distributions are summarized using a 90 per cent Bayes credible region. The log relative risk plot also has a horizontal display of 0 which represents one relative risk.

surgeon 1, at each iteration, and divide this by the probability for control surgery ear for surgeon 1. This conversion results in a posterior distribution of relative risk and we obtain a 90 per cent CR relative risk of 1.25–6.69. For visualization purposes, log-relative risk is reported in the bottom part of Figure 4 so that a relative risk of 1 is shown as a log-relative risk of 0. This indicates to us that the surgery risk in surgeon 1 significantly changes from control relative to PA. Using the same calculation technique with SP and control for surgeon 2, we find a 90 per cent CR relative risk of 0.55–1.11 (covering 1, insignificant). Conversion
of PA and control results in surgeon 2 is 0.77–1.25 (covering 1, insignificant). Based on these calculations, we conclude that the bacteria-induced surgery risk with surgeon 2 is insignificant relative to control surgery.

The results of the posterior distributions allow us to define priors for future studies. Assuming that the future study has the same parameters as the current study we calculate the prior parameters using a method of moments on the posterior. That is we took the mean and variance from the posterior and plugged them into a normal and inverse gamma distribution. Consider treating the prior of the fixed effect as a normal distribution, which from our analysis is $\beta \sim N(4.06,0.32^2)$. The variance component priors are the following the error variance is $\sigma^2_e \sim IG(7.75,0.61)$, the variance between guinea pigs is $\sigma^2_d \sim IG(6.63,0.25)$, and the variance between ears is $\sigma^2_f \sim IG(4.40,0.35)$, where $IG(\cdot,\cdot)$ is the inverse gamma distribution. Finally, the posterior distribution from our study results in a prior for the location-shift parameter to be $\alpha \sim N(4.63,2.89^2)$.

6. CONCLUSION

In summary, surgeon 2, using the surgical procedure ‘oval window’, poses a greater than 40 per cent chance of a 15 dB hearing loss regardless of injection of bacteria or not. However, surgeon 1, using the surgical procedure ‘semicircular canal’, does not pose a significantly greater than 40 per cent chance of a 15 dB hearing loss, unless there is a PA-induced infection and the PA probability relative to the control probability is significant. But because of the current experimental design, we require further experiments to separate the effects of surgery and technique. The main point of the results is to demonstrate the flexibility of these types of inferences provided by the Bayesian methodology, which include inferences from mean differences and probabilities of certain hearing threshold changes.

In general, practitioners appreciate a methodology that is accessible, straightforward and easy to use. This accounts for the popularity of point and click software packages like SAS analyst and SPSS. However, our custom-built method appeals to those who want to fully account for uncertainty (e.g. censoring), and who want direct evidence for the scientific question posed. That is, in addition to intervals about a mean for transformed data, we obtain intervals for functions of the model parameters such as a probability and relative risk.

To do this we consider treating the model parameters as random variables. We make inference by calculating posterior distributions using both the raw data as well as prior distributions of the model parameters, a Bayesian approach. It seems reasonable to predict that Bayesian analyses will become more frequent as the approach becomes commonplace in research and practice. In our future studies using a Bayesian approach, it will benefit us to use informative priors based on the posterior distributions calculated in this paper.

The method we present in this paper is of great interest to us as future hearing threshold studies become available. As illustrated by our analysis of the effect of bacteria and surgery on hearing loss, the Bayesian approach can flexibly answer questions posed by researchers, with a posterior distribution of a probability of hearing loss and a posterior distribution of relative risks. Among immediate future challenges is the fact that our results are based exclusively on hearing thresholds for a hearing tone level called ‘clicks’. Other tone levels can also be measured (4, 8, 12 kHz, …) and simultaneously analysed as multidimensional data.
APPENDIX

We use Gibbs sampling to obtain posterior distributions of all parameters in the \( \theta \) vector [12, 14]. We update each element with a chain conditional on the previously generated parameters. For example, at the \( m \)th iteration of the Gibbs sampler for the \( k \)th parameter, we generate the random variable \( \theta_k^{m+1} | \theta_1^m, \ldots, \theta_{k-1}^m, \theta_{k+1}^m, \ldots \). In our paper, the vector \( \mathbf{u}_i = \mathbf{X} \beta + \mathbf{1}d_i + \mathbf{F}_i \) is updated after every update of \( \beta, d_i \) or \( \mathbf{F}_i \).

After chain initialization, we begin the set of simulations by generating the location-shift parameter. Its complete conditional is

\[
f(\alpha | \mathbf{Y}, \mathbf{u}, \sigma_c^2) \propto \prod_{i=1}^{N} \prod_{j=1}^{6} \{ \exp(-\log(y_{ij} + \alpha) - \mu_{ij})^2/(2\sigma_c^2) \}/(y_{ij} + \alpha) \} I_{(1,30)}(\alpha)
\]

where \( I_{(1,30)}(\alpha) \) indicates that the probability density for \( \alpha \) is between 1 and 30. Since equation (A1) is not a usual distribution, we generate this random variable with a grid sampling method [12, p. 302]. We compute the target density \( f(\alpha | \mathbf{Y}, \mathbf{u}, \sigma_c^2) \) at a set of evenly spaced values \( \alpha_1, \alpha_2, \ldots, \alpha_M \) that are between 1 and 30, then approximate \( f(\alpha | \mathbf{Y}, \mathbf{u}, \sigma_c^2) \) with the discrete density at \( \alpha_1, \ldots, \alpha_M \), with probabilities \( f(\alpha_k | \mathbf{Y}, \mathbf{u}, \sigma_c^2)/\sum_{k=1}^{M} f(\alpha_k | \mathbf{Y}, \mathbf{u}, \sigma_c^2) \). Then a discrete cumulative distribution method simulates \( \alpha \).

The remaining complete conditionals are calculated with normal distributions and inverse gamma distributions. Continuing with the notation from Section 3, let \( \mathbf{Z}^{*} = \mathbf{Z} - 1d_i - \mathbf{F}_i \) and \( \tilde{\mathbf{Z}}^{*}(g) = \sum_{i \in \text{group}} \mathbf{Z}_i/n_g \) be the mean 6-vector for the \( g \)th treatment group, where \( n_g \) is the number of subjects in the \( g \)th group. The parameter \( \beta_g \) is the 6-vector from the \( g \)th group and \( n_g \) and \( \Sigma_{0g} \) are the prior mean 6-vector and \( 6 \times 6 \) covariance matrix. The hyper parameter \( n_g \) is the sample size in the \( g \)th group. The first complete conditional normal distribution is

\[
\beta_g | \mathbf{Y}, \mathbf{u}, \Sigma_0, \mathbf{d}, \mathbf{F} \sim \text{MVN}(\theta_g, \Sigma_{\theta_g}) \tag{A2}
\]

where the 6-vectors are \( \mathbf{A}_g = \Sigma_{0g}^{-1} \eta_g + \tilde{\mathbf{Z}}^{*}(g)/n_g/\sigma_c^2 \), and \( \mathbf{B}_g = \Sigma_{0g}^{-1}1 + (n_g/\sigma_c^2)\mathbf{I} \), and elements of which vectors are \( \theta_{gij} = \mathbf{A}_{gij}/\mathbf{B}_{gij} \) and \( \sigma_{\theta_g}^2 = 1/\mathbf{B}_{gij} \). The matrix \( \Sigma_{\theta_g} \) is diagonal. Two of the variance components are

\[
\sigma_d^2 | \mathbf{d} \sim \text{IG} \left( N/2 + 1, \sum_{i=1}^{N} d_i^2/2 + 0.1 \right) \tag{A3}
\]

and

\[
\sigma_c^2 | \mathbf{Z}, \mathbf{u} \sim \text{IG} \left( 3N + 1, \sum_{i=1}^{N} \sum_{j=1}^{6} (z_{ij} - \mu_{ij})^2/2 + 0.1 \right) \tag{A4}
\]

One type of random effect is

\[
d_i | \mathbf{z}, \mathbf{X}, \beta, \sigma_c^2, \sigma_d^2 \sim \text{N}((\mathbf{z}_i - C_i)/(6 + \sigma_c^2/\sigma_d^2), \sigma_d^2\sigma_c^2/(6\sigma_d^2 + \sigma_c^2)) \tag{A5}
\]

where \( C_i = \mathbf{X}_i \beta + \mathbf{F}_i \). The regression hyperparameters are

\[
\mathbf{B}_1 | \beta \sim \text{MVN}(((\beta_1 + \beta_3)/\sigma_c^2)/(1/10^3 + 2/\sigma_c^2), 1/(1/10^3 + 2/\sigma_c^2)\mathbf{I})
\]
and

\[ \mathbf{B}_2 \mid \beta \sim \text{MVN}((\mathbf{(\beta_2 + B_5)/\sigma^2_5)} / (1/10^3 + 2/\sigma^2_5), 1 / (1/10^3 + 2/\sigma^2_5) \mathbf{I}) \]  \hspace{1cm} (A6)

The hierarchical scale parameter is

\[ \sigma^2_c \mid \ldots \sim \text{IG}(13, \frac{1}{2} D + 0.1) \]  \hspace{1cm} (A7)

where

\[ D = \mathbf{1}^\top[(\mathbf{\beta}_1 - \mathbf{B}_1) + (\mathbf{\beta}_3 - \mathbf{B}_1) + (\mathbf{\beta}_2 - \mathbf{B}_2) + (\mathbf{\beta}_5 - \mathbf{B}_2)] + 1 \]

Let \( \omega_i = \mathbf{z}_i - \mathbf{X}_i \mathbf{\beta}_i - \mathbf{1} d_i \), then the updates for the other random effects

\[ f_{1i} \mid \ldots \sim \mathcal{N}\left( \frac{3}{\mathbf{\sum}_{j=1}^{3} \omega_{ij} \sigma_j^2 / (3\sigma_j^2 + \sigma_c^2), \sigma_c^2 \sigma_j^2 / (3\sigma_j^2 + \sigma_c^2) \right) \]  \hspace{1cm} (A8)

and

\[ f_{2i} \mid \ldots \sim \mathcal{N}\left( \frac{6}{\mathbf{\sum}_{j=4}^{6} \omega_{ij} \sigma_j^2 / (3\sigma_j^2 + \sigma_c^2), \sigma_c^2 \sigma_j^2 / (3\sigma_j^2 + \sigma_c^2) \right) \]  \hspace{1cm} (A9)

The variance of this random variable is updated using

\[ \sigma^2_f \mid \ldots \sim \text{IG}\left( N + 1, \sum_{i=1}^{N} \{ f_{1i}^2 + f_{2i}^2 \} / 2 + 1 \right) \]  \hspace{1cm} (A10)

The last step in the entire process is to generate the data within the intervals set by the type of censoring. The first line generates a random number on the log-normal cdf scale from the interval-censored values. The last two lines convert this back to the log-normal scale

\[ u_{ij} \sim \mathcal{U}(\Phi(\log(c_{1ij} + \alpha) \mid \mu_{ij}, \sigma_c), \Phi(\log(c_{2ij} + \alpha) \mid \mu_{ij}, \sigma_c)) \]

\[ z_{ij} = \Phi^{-1}(u_{ij}) \sigma_c + \mu_{ij} \]

and

\[ y_{ij} = \exp(z_{ij}) - \alpha \]  \hspace{1cm} (A11)

Note that \( \Phi(x \mid \mu, \sigma^2) \) is the normal CDF and \( \Phi^{-1}(\cdot) \) is the inverse standard normal CDF.

ACKNOWLEDGEMENTS

We thank Mary Benson Gajewski for editorial assistance in the preparation of this article. We thank C. Shane Reese and Mary Richardson for their reviews, and Scott Berry for inspiration and direction. We also thank two anonymous referees and editors for their patient reviews and suggestions that greatly improved this manuscript.
REFERENCES


