Bayesian Estimation of the Probability of Asbestos Exposure from Lung Fiber Counts

Scott Weichenthal,1 Lawrence Joseph,2,3,∗ Patrick Bélisle,2 and André Dufresne4

1Department of Epidemiology, Biostatistics and Occupational Health, McGill University, 1020 Pine Avenue West, Montreal, Quebec, H3A 1A2, Canada
2Division of Clinical Epidemiology, McGill University Health Center, 687 Pine Avenue West, V Building, Room V2.10, Montreal, Quebec, H3A 1A1, Canada
3Department of Epidemiology and Biostatistics, McGill University, 1020 Pine Avenue West, Montreal, Quebec, H3A 1A2, Canada
4Department of Occupational Health, McGill University, 1130 Pine Avenue West, Montreal, Quebec, H3A 1A3, Canada
∗email: Lawrence.Joseph@mcgill.ca

1. Introduction

Occupational asbestos exposure is a well-recognized risk factor for the development of serious lung diseases including malignant mesothelioma. Reaching a conclusive decision on the association between asbestos exposure and lung diseases for the purpose of occupation related compensation is a process that may become arduous when a reliable work history is unavailable. Nevertheless, a decision must be made, and the impacts of such decisions can be substantial, affecting the financial welfare of families.

When limited information is available concerning the likelihood of asbestos exposure, compensating agencies rely on lung fiber counting and analysis as an exposure assessment tool. This practice is supported by a number of studies which have related lung fiber burden to asbestos exposure (Gylseth et al., 1981; Mowé et al., 1985; Dufresne et al., 1996). Several benchmarks of lung fiber retention have been proposed. For example, a lung fiber concentration of more than one million per gram of dried lung tissue is thought to be indicative of occupational exposure to asbestos (Gylseth et al., 1981).

Short asbestos lung fibers are less than 5 micrometers (μm) in length, whereas long fibers are greater than or equal to 5 μm in length. Long fibers are better predictors of diseases such as mesothelioma compared to short fibers, but short fibers are usually found in greater numbers. Although the association between short fibers and disease is less strong, short fibers may also expose an occupational history, as, like long fibers, they are found in higher concentrations in exposed individuals compared to a reference population. Asbestos bodies are asbestos fibers that have been coated with ferritin by macrophages.

Although data on all three types of lung fibers can provide useful information regarding the likelihood of occupational asbestos exposure, the numbers of short or long lung fibers or asbestos bodies counted can vary substantially between tissue samples both within and between individuals. This is an important concern for inferences about exposure because decisions based on lung fiber analysis based on a single block of lung tissue may lead to misclassification of a truly exposed or unexposed individual. In the past, evaluation boards relied on the analysis of only a single block of lung tissue in classifying occupational asbestos exposures. In recent years, however, it has become more common to examine three or four blocks of lung tissue to decrease the probability of exposure misclassification.

© 2009, The International Biometric Society
Whether using one or more tissue blocks, very little has been published about the properties of lung fiber counts as diagnostic tests for asbestos exposure history. Here we develop Bayesian latent class models to evaluate and compare the test properties of asbestos exposure classification from all three types of fiber counts discussed above. We consider each fiber type used singly, as well as when information from all three types of data are used in combination. We compare the test properties when only a single block of lung tissue is analyzed relative to when data from two to four blocks of lung tissue are available per subject. The problem is rendered especially difficult because, in the absence of a strong work history, there is no gold standard method by which to classify each subject into truly exposed versus truly nonexposed categories. An additional complication is that the data provided by the three types of fiber counts do not follow standard distributions, in large part because there is always a value below which tests are not sensitive, leaving a mass of probability at that value. Further, the data within the different types of fiber counts may be correlated within individuals, indicating lack of independence conditional between blocks within each of the three tests. Although past literature has been divided about the importance of accounting for such dependence among diagnostic tests (Dendukuri and Joseph, 2001; Gustafson, 2005), we derive inferences from dependent as well as independent models, and compare results. Our model thus accommodates all important features of the data, and summarizes the inferences in several useful ways. In addition to standard diagnostic testing properties, such as sensitivity and specificity leading to receiver operating characteristic (ROC) curves, we provide methods for calculating the individual level probabilities of exposure given the test values on one or more tests and using data from one or more tissue blocks.

Bayesian latent class models for dichotomous diagnostic data arising from laboratory tests in the absence of a gold standard have been discussed by many, including Gastwirth, Johnson, and Reneau (1991); Joseph, Gyorkos, and Conpal (1995); Demissie et al. (1998); Johnson, Gastwirth, and Pearson (2001); and Gustafson (2005). Similar models for continuous data and ROC curves have been investigated by Zou and O’Malley (2005), although they assumed availability of a gold standard test result for each patient. Scott et al. (2008); considered normally distributed diagnostic test data in the absence of a gold standard, but considered only a single continuous test result from each subject, and did not account for the possibility of discrete probability masses in the test result distribution. Choi, Johnson, and Thurmond (2006) estimated disease prevalence and predictive probabilities for individual level continuous test results in the absence of a gold standard, but only considered results from a single test, and assumed the availability of a training sample of known positive and negative test results to estimate the distributions of truly diseased and truly disease-negative subjects. Erkanli et al. (2006) proposed a nonparametric analysis using mixtures of Dirichlet processes to model distributions within diseased and non-diseased subjects, leading to nonparametric ROC curve estimation, but again assumed a gold standard test. Because it is often the case in practice that results are available on more than one test per subject, including both repeated measures of the same test and results from different tests, the models developed extend the methodology available to date in important directions.

Section 2 describes the study setting, whereas Section 3 presents our models, discusses estimation of the test properties, and calculation of the probability of asbestos exposure. The results of applying our models to our fiber count data are in Section 4, and we conclude with a discussion.

2. Study Setting and the Source of Data

Occupational exposure to asbestos typically involves a mixture of mineral fibers of amphibole (banned since the early 1980s) or serpentine origin. Between 1996 and 2000, lung fiber retention analyses were conducted for 78 Quebec workers who had died of lung diseases potentially caused by occupational asbestos exposure. As previously described (Dubreuil et al., 1996), lung retention data for the numbers of long and short fibers and asbestos bodies per milligram of dry lung tissue were collected. Seventy-five of these cases were men, and in total, 35 workers had three and 43 workers had four blocks of lung tissue examined. We also were able to obtain data on a limited number (N = 41) of controls. These data were from lung tissue of either accidental death or death caused by acute myocardial infarction in men autopsied between 1990 and 1992. These results will help estimate distributional parameters for our tests for nonexposed individuals.

3. Statistical Methods

Examining the available data for our three types of fiber counts revealed clear patterns. For short fibers, data ranged from 35 fibers per milligram (f/mg) to over 164,000 f/mg. The smallest values among our data set of controls was 70 f/mg, but in the data set of possible cases the lower limit was 35 f/mg, the next lowest value being over 100 f/mg. In each case, the lower limits (i.e., 35 or 70 f/mg) were reported for many subjects, making it clear that this was the minimum possible value. We therefore modeled these lowest values as a probability mass, and considered a lump sum probability representing all values equal to or below the relevant minimum possible value. Conditional on the latent exposure status, we first assumed that the counts from each type of fiber were independent from each other both within and between individuals. In a second similar model, we allow for within-individual correlations across tissue blocks for...
each test. We also fit a hierarchical model that assumed conditional independence both between and within subjects, but only after further conditioning on a distinct mean value for each fiber type within each subject, these means assumed to be drawn from a common distribution between subjects. This model is a Bayesian version of a random effects model. We now provide the details of our models, starting with a model for each test used singly, and then show how the model can be extended when more than one type of fiber count is used, and/or there are several repeated values from each subject.

### 3.1 One Mixed Discrete/Continuous Diagnostic Test

Let $\theta$ be the true asbestos exposure rate in the population of interest. Let $Z_i$ represent the latent true exposure history for individual $i$, $i = 1, 2, \ldots, n$. That is, $Z_i = 1$ if individual $i$ was truly exposed to asbestos, and $Z_i = 0$ otherwise. Let $X_{i1}, \ldots, X_{in}$ represent the logarithms of the fiber counts across the $n$ individuals in the study, and let $C$ represent the cutoff value representing the minimum value for the test. For the remaining values above this cutoff, let $\mu_E$ and $\sigma_E$ represent the mean and standard deviation (SD) of the fiber count distribution conditional on a positive true (latent) asbestos exposure status, and let $\mu_{NE}$ and $\sigma_{NE}$ represent the same quantities among the truly not asbestos exposed. Finally, let $p_E$ and $p_{NE}$ denote the probabilities of obtaining the minimum possible value in the exposed and nonexposed groups, respectively.

Let $\mathbf{X}$ and $\mathbf{Z}$ be vectors of fiber counts and latent exposure history on $n$ subjects, respectively. Assuming for the moment

\[
P(\text{exposed} \mid x, \theta, \mu_E, \sigma_E, \mu_{NE}, \sigma_{NE}, p_E, p_{NE})
\]

that just a single observation is available per subject, the likelihood function of the data $\mathbf{X}$ and augmented data $\mathbf{Z}$ for a model using results from any one fiber test is

\[
f(\mathbf{X}, \mathbf{Z} \mid \theta, \mu_E, \sigma_E^2, \mu_{NE}, \sigma_{NE}^2, p_E, p_{NE}) = \prod_{i=1}^{n} \left( \frac{1}{\sqrt{2\pi\sigma_E^2}} \exp \left\{ -\frac{1}{2} \left( x_i - \mu_E \right)^2 / \sigma_E^2 \right\} \right)^{I_{\{x_i > C\}}} \times \left( \theta(1 - p_E) \frac{1}{\sqrt{2\pi\sigma_E^2}} \exp \left\{ -\frac{1}{2} \left( x_i - \mu_E \right)^2 / \sigma_E^2 \right\} \right)^{I_{\{x_i \leq C\}}} \times \left( \theta p_E \frac{1}{\sqrt{2\pi\sigma_{NE}^2}} \right)^{I_{\{x_i > C\}}} \times \left( (1 - \theta) (1 - p_{NE}) \frac{1}{\sqrt{2\pi\sigma_{NE}^2}} \right)^{I_{\{x_i \leq C\}}} \times \left( (1 - \theta) p_{NE} \right)^{I_{\{x_i > C\}}}
\]

where $I(\cdot)$ is the indicator function.

The estimation problem is similar to that of fitting a normal mixture model, with the extra complication of the point masses, which adds two parameters to be estimated, $p_E$ and $p_{NE}$. We will use Bayesian methods, and thus require prior distributions over the parameter space $(\theta, \mu_E, \sigma_E^2, \mu_{NE}, \sigma_{NE}^2, p_E, p_{NE})$. We will use beta prior distributions for $\theta, p_E$ and $p_{NE}$, which includes the beta(1,1) or uniform distribution as a special case, useful when little prior information will be incorporated into the analysis. Normal prior distributions are used for $\mu_E$ and $\mu_{NE}$, and uniform distributions over a suitably chosen finite range for $\sigma_E$ and $\sigma_{NE}$. As there is no closed form solution for the posterior distributions, we will use the Gibbs sampler as implemented in WinBUGS version 1.4.1 (Lunn et al., 2000). We used 5000 iterations for burn-in and ran a further 20,000 iterations for use in inferences. Convergence of the Gibbs sampler algorithm was assessed by examining history plots across all iterations and by running the sampler several times from different starting values.

Of course, once all of the above parameters are estimated, one can estimate any function of these parameters, including the sensitivity and specificity of the fiber count test for any given cutpoint for positivity, which in turn leads to ROC curves (Hanley, 1996), although these ROC curves are not quite of the usual form because of the point masses. ROC curves plot the false positive rate (1 – specificity) versus the true positive rate (sensitivity) across the range of possible cutoff values for the test. In addition, it is of great clinical importance to know the probability of being exposed given any fiber count value, $x$. Once $(\theta, \mu_E, \sigma_E^2, \mu_{NE}, \sigma_{NE}^2, p_E, p_{NE})$ have been estimated, the posterior probabilities $P(\text{exposed} \mid x, \theta, \mu_E, \sigma_E^2, \mu_{NE}, \sigma_{NE}^2, p_E, p_{NE})$ are obtainable via the formula

\[
P(\text{exposed} \mid x, \theta, \mu_E, \sigma_E^2, \mu_{NE}, \sigma_{NE}^2, p_E, p_{NE}) = \frac{\theta}{\sqrt{2\pi\sigma_E^2}} \exp \left\{ -\frac{1}{2} \left( x - \mu_E \right)^2 / \sigma_E^2 \right\} + \frac{1 - \theta}{\sqrt{2\pi\sigma_{NE}^2}} \exp \left\{ -\frac{1}{2} \left( x - \mu_{NE} \right)^2 / \sigma_{NE}^2 \right\}.
\]
606

Biometrics, June 2010

\[
\begin{aligned}
&\exp \left\{ \frac{1}{2}(x_{ij} - \mu_E)^2/\sigma_E^2 \right\} I_{x_{ij} \leq C} \\
&\times (\theta_{PE})_{x_{ij} \leq C} \\
&\times \left( 1 - \theta \right)^{1 - p_N E} \frac{1}{\sqrt{2\pi}\sigma_{NE}} \\
&\times \exp \left\{ \frac{1}{2}(x_{ij} - \mu_{NE})^2/\sigma_{NE}^2 \right\} I_{x_{ij} \leq C} \\
&(1 - \theta)(p_N E)_{x_{ij} \leq C},
\end{aligned}
\]

where \( \theta = (\theta, \mu_E, \sigma_E^2, \mu_{NE}, \sigma_{NE}^2, \rho, \mu_{PE}, \sigma_{PE}^2, \mu_{PE}, \sigma_{PE}^2, \rho) \), and \( j \) indicates the \( j \)th test, \( j = 1, 2, 3 \). This model assumes conditional independence of results across three tests within each individual. Conditional independence is a much weaker condition compared to unconditional independence, in that results are assumed independent conditional on the (latent) true disease state for each individual. In Section 3.3 below, we describe two different models that account for correlations between repeated measurements of the same test within individuals, but we chose not to also account for any possible correlations between different tests, for two reasons: First, correlations between different tests seemed much lower than those within the same test within each individual. Second, to account for both types of correlations within the same model requires inverting a nine by nine (or larger) matrix, where each entry of the matrix is a very complex function of parameters of the model. Thus, the model becomes very unwieldy, making it difficult to estimate parameters, even using Monte Carlo techniques.

Although there are now many more parameters, the forms of the prior distributions can be chosen to be identical to the case of a single test. Of course, similar models can be created when any two tests are used rather than all three tests.

3.3 More than One Measurement Per Test for Each Individual

The above models assume just a single observation from each of the three tests is analyzed from each individual. Assuming independence of the observations both between and within individuals, conditional on the latent true state for each subject and given the values for all unknown parameters relating to the normal distributions and probabilities of values below the cutoffs, within subject observations for repeated observations on a single test can easily be accommodated. The likelihood function given in equation (2) can be used, the only addition being a further product term over the numbers of observations from each individual.

In our data set, some subjects had three observations for each fiber count type (i.e., three different tissue samples were taken), whereas others had four. As discussed in the introduction, it is of great interest to compensation boards to compare test properties as larger numbers of tissue samples are taken. In Section 4, we report on the evolving probabilities of exposure as more tissue samples are available to be analyzed.

All of the models discussed so far assume that the data across different blocks within each subject are conditionally independent. That is, given the true exposure status, there is no information about the results from one block given the results of the other blocks. However, this assumption likely does not hold for these data, as we found within subject across-block correlations above 0.3 within all three types of fiber counts. These approximate correlations were calculated after dividing the subjects into exposure classes “by eye,” except for the 41 subjects who were a priori known to be nonexposed.

Some have argued that correlations may not always have a strong effect on final inferences (Denulukuri and Joseph, 2001; Gustafson, 2005), and sometimes a simpler model not including correlations may perform better. To investigate this issue, we created two distinct models that explicitly accommodate such correlations, and compared inferences from these models to those assuming independence. One model added specific correlation parameters within a multivariate normal distribution, whereas the other model handled correlations implicitly, via addition of a hierarchical component. As discussed in the next section, results were very similar between the two types of models accommodating correlations, but there were substantial differences between correlated and uncorrelated inferences. Given the observed correlations, the multivariate normal and hierarchical models are more plausible. Therefore, although we report results from both independent and correlated models, those from the independent models are included mostly for comparison purposes. Given the large number of different models run, and because results from our multivariate normal and hierarchical models are so similar, in Section 4 we present detailed results only from the hierarchical model.

The likelihood function for our correlated model using data from each of the three tests is similar to that given by equation (1) of Section 3.1, modified by considering \( m \) observations per subject from the test via an \( m \)-dimensional multivariate normal distribution, with a distinct correlation parameter for each test. Within each test, however, we assumed the same correlation parameter between each pair of measures, because these represent different samples from the same lung. Thus, the multivariate normal distribution used a single mean parameter across blocks, with variance-covariance matrix having just a single covariance parameter for all off-diagonal elements (compound symmetry). For the \( j \)th subject, and for data across blocks \( j, k = 1, \ldots, m \), and test \( k \) we have

\[ x_{ik} \sim MVN \left( \mu_k, \Sigma \right) \]

where the \( . \) represents \( j = 1, \ldots, m \), and where \( \Sigma \) is an \( m \times m \) variance-covariance matrix with \( \sigma^2 \) on the diagonal entries, and all off-diagonal entries equal to \( \rho \sigma^2 \), where \( \rho \) is the between-block correlation parameter for test \( k \).

In our hierarchical model, all observations, both within and between subjects were again considered as independent, but conditioned on distinct individual level mean parameters for the logarithms of each type of fiber count. These parameters, in turn, are tied together through a hierarchical normal distribution across individuals for each fiber type. This allows for the correlations that occur within individuals if only a single overall mean is used. Therefore, at the first level of our hierarchical model, the test results within each block for each test type are considered as independently normally distributed, but now each subject has their own unique mean value within each test. In turn, at the second level of the
Bayesian Estimation of Probability of Asbestos Exposure

4. Results

4.1 Descriptive Statistics and Fiber Count Distributions

Table 1 contains fiber count distributional results from our hierarchical latent class model including all of the data (i.e., three or four samples for subjects having more than one observation). For long fibers, there is a clear separation of the distributions in exposed versus unexposed subjects. Although under 4% (1 − 0.966 = 0.034) of exposed subjects have values at the lower limit of detection, almost 60% of unexposed subjects are at this value. For subjects above this threshold, the means on the log scale are also well separated (7.72 − 5.44 = 2.28, see also Figure 4). In addition, exposed subjects have larger variation compared to unexposed subjects (SD = 2.1 versus 0.3). The area under the ROC curve (AUC) for long fibers is 0.92, although with a relatively wide credible interval, because unlike the other estimates in this table, this value is based only on one block of data. ROC curves are not well defined when two or more blocks of data are used, because the choice of cutoff to use is not unique for two or more dimensional data.

Short fibers may be slightly less diagnostic than long fibers, with only 32% of unexposed subjects estimated to have values at the lower limit, compared to almost 60% in the case of long fibers. However, those above the lower threshold are widely separated, with a larger mean difference compared to long fibers. On the other hand, asbestos bodies have larger estimated SDs, and so the distributions of the exposed and unexposed will have larger overlap compared to long fibers. Therefore, we might expect the sensitivities and specificities values are very low in information compared to the information in our data set. Varying the prior distributions produced no noticeable changes in results, so we do not report further on these robustness checks here.
Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Posterior median</th>
<th>95% credible interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long fibers (on a log scale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean for exposed subjects</td>
<td>7.72</td>
<td>(7.23, 8.15)</td>
</tr>
<tr>
<td>SD for exposed subjects</td>
<td>1.42</td>
<td>(1.14, 1.79)</td>
</tr>
<tr>
<td>P (above lowest value) in exposed subjects</td>
<td>0.966</td>
<td>(0.928, 0.991)</td>
</tr>
<tr>
<td>Mean for unexposed subjects</td>
<td>5.44</td>
<td>(5.18, 5.85)</td>
</tr>
<tr>
<td>SD for unexposed subjects</td>
<td>0.30</td>
<td>(0.03, 0.76)</td>
</tr>
<tr>
<td>P (above lowest value) in unexposed subjects</td>
<td>0.423</td>
<td>(0.324, 0.517)</td>
</tr>
<tr>
<td>Area under the ROC curve</td>
<td>0.92</td>
<td>(0.76, 0.98)</td>
</tr>
<tr>
<td>Short fibers (on a log scale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean for exposed subjects</td>
<td>8.78</td>
<td>(8.38, 9.17)</td>
</tr>
<tr>
<td>SD for exposed subjects</td>
<td>1.33</td>
<td>(1.07, 1.68)</td>
</tr>
<tr>
<td>P (above lowest value) in exposed subjects</td>
<td>0.973</td>
<td>(0.937, 0.995)</td>
</tr>
<tr>
<td>Mean for unexposed subjects</td>
<td>6.30</td>
<td>(5.98, 6.65)</td>
</tr>
<tr>
<td>SD for unexposed subjects</td>
<td>0.43</td>
<td>(0.02, 0.90)</td>
</tr>
<tr>
<td>P (above lowest value) in unexposed subjects</td>
<td>0.680</td>
<td>(0.595, 0.756)</td>
</tr>
<tr>
<td>Area under the ROC curve</td>
<td>0.86</td>
<td>(0.74, 0.96)</td>
</tr>
<tr>
<td>Asbestos bodies (on a log scale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean for exposed subjects</td>
<td>8.44</td>
<td>(7.95, 8.95)</td>
</tr>
<tr>
<td>SD for exposed subjects</td>
<td>1.73</td>
<td>(1.40, 2.18)</td>
</tr>
<tr>
<td>P (above lowest value) in exposed subjects</td>
<td>0.958</td>
<td>(0.910, 0.995)</td>
</tr>
<tr>
<td>Mean for unexposed subjects</td>
<td>6.23</td>
<td>(5.68, 6.72)</td>
</tr>
<tr>
<td>SD for unexposed subjects</td>
<td>1.24</td>
<td>(0.87, 1.74)</td>
</tr>
<tr>
<td>P (above lowest value) in unexposed subjects</td>
<td>0.512</td>
<td>(0.425, 0.600)</td>
</tr>
<tr>
<td>Area under the ROC curve</td>
<td>0.88</td>
<td>(0.76, 0.97)</td>
</tr>
</tbody>
</table>

for short fibers and asbestos bodies to be slightly lower compared to long fibers.

4.2 ROC Curves

Figure 1 displays ROC curves for short fibers, long fibers, and asbestos bodies, based on data from block 1 only. ROC curves are not well defined when two or more blocks of data are used, because the cutoff to choose for each subject when they provide two or more values is not unique. The small jumps arise from the masses that represent the probability of being at the lower limits of detection for each test.

The AUC represents the probability that a randomly selected truly positive subject and a randomly selected truly negative subject will be correctly ordered by the continuous test. For long fibers the AUC is 0.92, although with a relatively wide credible interval (see Table 1), in part because this estimate is based only on one block of data. The AUC for short fibers is 0.86, slightly less than the point estimate for long fibers; however, the very wide credible intervals leave much uncertainty. The AUC for asbestos bodies is 0.88, intermediate to the AUCs from the other two tests.

4.3 Individual Probabilities of Asbestos Exposure

Of most interest to compensation boards is the probability of exposure, given data from any subject. A byproduct of running any of the models described in this article are estimates of the probabilities of exposure across all subjects. We will provide examples of how these probabilities vary between subjects by selecting three prototypic subjects and examining their exposure probabilities across the full range of models we have developed.

Subject #1 has a mix of values both at and above the lower limits of detection. Subject #2 has high values for each test, with no observations at the lower limit. Subject #3 has moderate values across the three tests, again with none at the lower limit. The data (on a log scale) from all three tests and
Figure 2. Posterior median estimates of the distributions for nonexposed and exposed subjects for short fibers, long fibers, and asbestos bodies. Superimposed on each graph are the data values for each of our three subjects, as discussed in Section 4.3.

Figure 3 presents the probability of exposure for all three subjects across all models and using increasing numbers of tissue blocks. If one only looks at the results from subject #1 short fibers (labeled as SF in the figure), one can say very little about the probability of exposure, as even with all four tissue blocks included in the analysis, the credible interval for this probability ranges from close to zero to almost 0.9. Similarly, little can be concluded if only a single block of data is used from long fibers or asbestos bodies for this subject. However, as soon as data from two or more blocks are included, the probability of exposure is concentrated very near to 0, indicating no exposure. Only one block of data is required if one combined data from all three tests, again indicating a probability of exposure close to zero. For subject #1, there is little difference between a model that assumed independent observations within subjects, versus the hierarchical model.

Subject #2 had generally high values across all three tests, with no values hitting the lower threshold, $C$. Here, even a single observation from any test is sufficient to classify this subject as exposed, and these high probabilities remain near one with very high probability regardless of which model is run or how many data points are used. Note that the lower limit of the $y$ axis of this graph is at 0.992, so that even a single short fibers observation provides very strong evidence of exposure.
In contrast to subjects #1 and #2, where clear decisions are possible, the data from subject #3 shows that this is not always the case, even when all available data are used. Results range from near certainty of exposure if one or two long fiber observations are examined, to certainty of nonexposure if all data are used, and the model assumes independence between all observations. However, all other models indicate great uncertainty about the probability of exposure, and because the independence model is probably not valid, one must admit that the data are not sufficient to make a strong recommendation. This example also clearly shows the danger of relying on only a single type of fiber count rather than all the data, and illustrates that a model that assumes independence can provide very different estimates from our correlated data models, even though exactly the same data are input into both models.

The contrasting results between correlated and noncorrelated models from subject #3 raises the issue of goodness-of-fit or model selection procedures. Indeed, some authors (for example Black and Craig, 2002) have performed formal model selection procedures or averaged the results over several models. Here, however, we have very strong reasons to doubt models that do not incorporate dependence between blocks within subjects, and both models that account for these correlations provide virtually identical estimates. This is not surprising, as correlations arise because data for each test within subjects are more similar compared to data between subjects, and both of our models account for this.

Figure 4 displays the mean exposed minus nonexposed mean differences, $\mu_E - \mu_{NE}$ across all models, which is interesting for several reasons. First, this parameter is of importance by itself, because the distance between exposed and nonexposed distributions is a marker for the usefulness of a continuous diagnostic test. Second, we can clearly see how accuracy for this parameter is affected by the number of data blocks used. For asbestos bodies, there is not much increase in accuracy, as judged by the length of the credible interval, as more data blocks are added. For both short and long fibers, there is an increase in accuracy going from $m = 1$ to $m = 2$ tests, but not much improvement after that. We can also see smaller mean differences in the hierarchical model as compared to a model that assumes independence of data blocks within subjects, perhaps an indicator that the amount of information is somewhat exaggerated in the independence model. We can also see similar sized credible intervals for $m = 2$, 3, or 4 in the hierarchical model, indicating that there is extra variability accounted for here which does not substantially decrease with increasing $m$.

5. Discussion

We have developed a series of Bayesian latent class models for mixed continuous/discrete diagnostic test data, and applied these models to determine the probability of asbestos exposure from lung fiber count data. We have shown that incorrect inferences may be made if only a single block of data is analyzed, and that for many subjects a clear decision is possible using a model that uses all possible data. However, for some subjects, even 12 data points are not sufficient for a definitive assessment.

There is no substitute for a detailed work history in determining the likelihood of occupational asbestos exposure. This
is particularly true for occupational exposure to chrysotile fibers, which do not accumulate as readily in the lungs, and for which the number of fiber-years of exposure is the best indicator of lung fiber burden. Unfortunately, detailed occupational exposure histories are not available for all suspected cases of asbestos-related lung disease, and in such circumstances, lung fiber retention analysis is one alternative method. Our models provide direct estimates of the probability of exposure, given all data collected, and have shown that collecting more than one data block per subject improves these estimates, resulting in better decisions concerning compensation.

REFERENCES


Received June 2008. Revised February 2009. Accepted February 2009.