

MODELLING OCCUPATIONAL EXPOSURE USING A RANDOM EFFECTS MODEL: A BAYESIAN APPROACH

Justin Harvey and Abrie van der Merwe***

**Centre for Statistical Consultation, University of Stellenbosch*

***University of the Free State*

ABSTRACT

In this article we develop a similar method to Krishnamoorthy and Mathew (2002), who used a generalized confidence interval and p-value approach for the occupational exposure limit (OEL), using Bayesian methodology for the one-way random effects model in order to determine the performance of different non-informative prior distributions.

Keywords: Occupational exposure limit (OEL); one-way random effects model; Bayes; non-informative prior.

INTRODUCTION

Lognormally distributed data are found in many settings including occupational health. It is not just confidence intervals that one is interested in though. Many standard statistical models can be adapted to facilitate the use of lognormal data. One such setting has been proposed by Krishnamoorthy and Mathew (2002), whereby they applied a one way random effects model to balanced, lognormally distributed data. This was done as a means to assess occupational exposure. The parameter of interest in these cases was the occupational exposure limit (OEL). The reader is referred to the original articles by Krishnamoorthy and Mathew (2002) for a more complete description of the medical applications of this method as well as the texts by Rappaport, Kromhout and Symanski (1993), Heederik and Hurley (1994) and Lyles, Kupper and Rappaport (1997).

Krishnamoorthy and Mathew (2002), in almost an extension of their other work on lognormally distributed data, attempted to analyse the data using generalized confidence intervals and generalized p-values in order to test hypotheses by means of confidence intervals on the overall mean exposure limits. The intention here is not to further describe the application to medical exposure limits (although a brief description of the setting will be given in sections that follow), but merely to develop a similar method using Bayesian methodology in order to improve the method. The prior distribution will be non-informative, since we have no subjective prior opinion as to the distribution of the data.

Although this article is primarily concerned with modeling the arithmetic mean occupational exposure limits, it is a simple procedure to extend this analysis to the geometric mean. In many instances the geometric mean can minimize the influence of outliers in the data and in this setting also has implications for the computational simplicity of the procedures. The interested reader is referred to work by Harvey and van der Merwe (2012).

Using a simulation study this article will illustrate the flexibility of the Bayesian methods described here. The primary advantage of the Bayesian approach is that while the methods proposed by Krishnamoorthy and Mathew (2002) cannot incorporate individual worker means (the overall sample mean is used and no accounting of individual worker means is made), the Bayesian approach enables the researcher to model mean exposure for each individual worker by using information specific to an individual. In so doing, truly individual worker means can be modeled. This will be illustrated by means of hypothetical example, similar to that proposed by Krishnamoorthy and Mathew (2002).

SETTING DESCRIPTION

One of these methods proposed to model OEL data has been to model the log-transformed shift-long exposures to attempt to incorporate the between- and within-worker sources of variability. In this case we

are interested in workers or groups of workers for whom the long-term exposure levels exceed the specific OEL.

The analysis of problems in this area is rather different from the traditional problems associated with the analysis of mixed models. Deriving exact tests is perhaps a slightly more complicated procedure since the analysis involves all the regular parameters of the traditional mixed model analysis, namely the mean, as well as the two variance components, that is the within- and between-subject variability.

As mentioned in the Introduction, Krishnamoorthy and Mathew proposed a generalized confidence interval and generalized p-value method to the analysis of these problems. Lyles et al. (1997) proposed large sample tests, such as Wald, likelihood ratio and score-type tests.

To represent this data situation consider the following example. There are n exposure measurements available on each of k workers. The exposure measurements X_{ij} ($i = 1, \dots, k; j = 1, \dots, n$) are lognormally distributed and therefore $Y_{ij} = \ln(X_{ij})$ is distributed normally. The assumed one-way random effects model is:

$$Y_{ij} = \mu + \tau_i + e_{ij}, \quad i = 1, \dots, k; j = 1, \dots, n. \quad (1)$$

where μ is the general mean, $\tau_i \sim N(0, \sigma_\tau^2)$ and $e_{ij} \sim N(0, \sigma_e^2)$. All the random variables are independent of each other and here τ_i represents the random effect due to the i -th worker.

According to Krishnamoorthy and Mathew (2002) let

$$\mu_{x_i} = E(X_{ij}|\tau_i) = E(\exp[Y_{ij}]|\tau_i) = \exp\left(\mu + \tau_i + \frac{\sigma_e^2}{2}\right) \quad (2)$$

and μ_{x_i} is the mean exposure for the i -th worker. Let θ denote the probability that μ_{x_i} exceeds the OEL. Thus,

$$\theta = P(\mu_{x_i} > OEL) = P\left(\ln(\mu_{x_i}) > \ln(OEL)\right) = 1 - \Phi\left(\frac{\ln(OEL) - \mu - \frac{\sigma_e^2}{2}}{\sigma_\tau}\right) \quad (3)$$

where $\Phi(\cdot)$ denotes the c.d.f of the standard normal distribution. The kinds of hypothesis that are going to be considered here are: $H_0: \theta \geq A$ vs $H_1: \theta < A$, where A is a specific quantity that is usually small, according to Krishnamoorthy and Mathew (2002).

An example data set of simulated “styrene exposures” that will serve as a basis for discussion in this article is given in Harvey (2012 – unpublished Ph.D thesis). However, the summary statistics from this data set are as follows:

$$k = 15; n = 10; v_1 = k(n - 1); v_2 = k - 1;$$

The vector of worker means is

$$[\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_k]' = [4.28 \ 4.65 \ 4.79 \ 4.04 \ 4.42 \ 5.05 \ 4.29 \ 4.42 \ 4.83 \ 4.80 \ 4.16 \ 4.09 \ 4.54 \ 4.77 \ 4.71]'$$

$$\bar{Y}_{..} = \frac{1}{kn} \sum_{i=1}^k \sum_{j=1}^n Y_{ij} = 4.5228;$$

$$SS_\tau = v_2 m_2 = n \sum_{i=1}^k (\bar{Y}_i - \bar{Y}_{..})^2 = 13.2283 = \text{"between workers sum of squares"};$$

$$SS_e = v_1 m_1 = \sum_{i=1}^k \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2 = 11.3897 = \text{"within workers sum of squares"};$$

BAYESIAN METHODOLOGY

The basis for analyzing any situation from a Bayesian perspective is the following relationship, a well-known result of Bayes' theorem: *Posterior* \propto *Likelihood* \times *Prior*. The likelihood function (in matrix form) is given by:

$$L(\mu, \sigma_\tau^2, \sigma_e^2, \mathbf{u}|\mathbf{Y}) \propto \left(\frac{1}{\sigma_e^2}\right)^{\frac{1}{2}kn} \exp\left\{-\frac{1}{2\sigma_e^2}(\mathbf{Y} - \mu\mathbf{1} - \mathbf{Z}\mathbf{u})'(\mathbf{Y} - \mu\mathbf{1} - \mathbf{Z}\mathbf{u})\right\} \times \left(\frac{1}{\sigma_\tau^2}\right)^{\frac{1}{2}k} \exp\left\{-\frac{1}{2\sigma_\tau^2} \mathbf{u}'\mathbf{u}\right\} \quad (4)$$

$\mathbf{Y} = [Y_{11} \dots Y_{1n} \ Y_{21} \dots Y_{2n} \dots Y_{kn}]'$ is the data vector, $\mathbf{e} = [e_{11} \dots e_{1n} \ e_{21} \dots e_{2n} \dots e_{kn}]'$ is the error or residual vector and $\mathbf{e} \sim N(\mathbf{0}, I_{kn}\sigma_e^2)$. $\mathbf{u} = [\tau_1 \ \tau_2 \dots \tau_k]'$ $\sim N(\mathbf{0}, I_k\sigma_\tau^2)$ and $\mathbf{1}$ is a $k \times n$ of ones and $\mathbf{Z} = I_k \otimes \mathbf{1}_n$ is the design matrix denoted by the aforementioned Kronecker product. The following non-informative prior distribution will be considered:

$$p(\mu, \sigma_\tau^2, \sigma_e^2) = p(\mu)p(\sigma_\tau^2, \sigma_e^2) \propto \frac{1}{\sigma_e^2} \frac{1}{(\sigma_e^2 + n\sigma_\tau^2)} \quad (5)$$

The above prior distribution is the non-informative distribution that has been discussed at length by Box and Tiao (1973) a full criticism can be found in this text. By combining (4) and (5) we arrive at the posterior distribution of $\mu, \sigma_\tau^2, \sigma_e^2$ and \mathbf{u} .

This can then be integrated over \mathbf{u} by completing the square with respect to \mathbf{u} . Therefore, the joint posterior density function of $(\mu, \sigma_\tau^2, \sigma_e^2)$ can be written as:

$$p(\mu, \sigma_\tau^2, \sigma_e^2 | \mathbf{Y}) \propto (\sigma_e^2)^{-\frac{1}{2}(v_1+2)} (\sigma_e^2 + n\sigma_\tau^2)^{-\frac{1}{2}(v_2+3)} \exp \left\{ -\frac{1}{2} \left[\frac{kn(\bar{Y} - \mu)^2}{\sigma_e^2 + n\sigma_\tau^2} + \frac{v_2 m_2}{\sigma_e^2 + n\sigma_\tau^2} + \frac{v_1 m_1}{\sigma_e^2} \right] \right\} \quad (6)$$

Now, to obtain the joint posterior distribution of $(\sigma_\tau^2, \sigma_e^2)$ we integrate (6) with respect to μ and arrive at the desired result:

$$p(\sigma_\tau^2, \sigma_e^2 | \mathbf{Y}) \propto (\sigma_e^2)^{-\frac{1}{2}(v_1+2)} (\sigma_e^2 + n\sigma_\tau^2)^{-\frac{1}{2}(v_2+2)} \times \exp \left\{ -\frac{1}{2} \left[\frac{v_2 m_2}{\sigma_e^2 + n\sigma_\tau^2} + \frac{v_1 m_1}{\sigma_e^2} \right] \right\} \quad (7)$$

Note that $\sigma_\tau^2 > 0$.

SIMULATION STUDY

A simulation study was performed using the results obtained in the previous section. However, there was a departure from the analysis that Krishnamoorthy and Mathew (2002) performed. In particular Krishnamoorthy and Mathew (2002) applied to the generalized confidence limits and p-values to the following quantity:

$$\mu_{x_i} = e^{\left(\mu + \tau_i + \frac{\sigma_e^2}{2} \right)}$$

which is the mean exposure of the i -th worker as defined in equation (2). However, as it turns out, the “mean exposure of the i -th worker did not include the actual row mean (mean for that specific worker) anywhere in the simulation, but rather the simulation was based on the overall mean and the prior distribution of τ_i , instead of the prior distribution of $\mu + \tau_i + \frac{\sigma_e^2}{2}$. So in actual fact, Krishnamoorthy and Mathew had more a “general” or “overall” mean exposure for each worker. In other words, the results derived by Krishnamoorthy and Mathew (2002) are for any worker (new or unknown) and not for a specific worker in the sample. In this section though we suggest an enhancement of this technique, whereby the mean for the i -th worker is estimated (from a Bayesian ideology) whereby the actual row means also influence the mean exposure level for that specific worker. So this section in fact presents various simulation studies in the following order:

1. Mean exposure of the i -th worker accounting for the row mean (enhancement on Krishnamoorthy and Mathew technique).
2. Mean exposure of the i -th worker not accounting for the row mean (comparable technique).
3. Overall mean exposure.

Simulation 1 – Individual Worker Mean

The intention is to simulate the mean exposure per worker from the posterior distribution using the prior mentioned by Box and Tiao (1973). Let

$$\mu_{x_i} = e^{\left(\mu + \tau_i + \frac{\sigma_e^2}{2} \right)}$$

represent the mean exposure level of the i -th worker. For each worker the probability that the mean exposure exceeds a certain pre-defined limit can be simulated from the posterior distribution as follows:

1. Simulate λ from a $\chi_{v_1}^2$ distribution where $v_1 = k(n-1)$ and using this calculate: $\sigma_e^2 = \frac{SSE}{\lambda}$
2. Simulate δ from a $\chi_{v_2}^2$ distribution where $v_2 = k-1$ and using this calculate: $\sigma_e^2 + n\sigma_\tau^2 = \frac{SS\tau}{\delta} = \sigma_{12}^2$
3. Calculate $\sigma_\tau^2 = \frac{\sigma_{12}^2 - \sigma_e^2}{n}$. This implies that we have simulated σ_e^2 and σ_τ^2 from their joint posterior distribution as given in (7).
4. If a negative value is obtained in step 3 we disregard both σ_e^2 and σ_τ^2 and repeat steps 1 - 3 until we find a pair where both are positive. This is somewhat different to Krishnamoorthy and

Mathew's technique, whereby in their method negative estimates were simply set equal to zero and not totally disregarded.

5. For each pair of $(\sigma_e^2, \sigma_\tau^2)$ simulate $\zeta^* \sim N(I, II)$ (the posterior distribution of $\mu + \tau_i + \frac{\sigma_e^2}{2}$ given the variance components has this normal distribution) where

- a. $I = \frac{n\sigma_\tau^2}{\sigma_e^2 + n\sigma_\tau^2} \bar{Y}_{i.} + \frac{\sigma_e^2}{\sigma_e^2 + n\sigma_\tau^2} \bar{Y}_{..} + \frac{\sigma_e^2}{2}$ and

- b. $II = \frac{\sigma_e^2}{\sigma_e^2 + n\sigma_\tau^2} \left\{ \frac{\sigma_e^2}{nk} + \sigma_\tau^2 \right\}$.

The posterior distribution of $\mu + \tau_i + \frac{\sigma_e^2}{2}$ given the variance components follows from the fact that $\mu + \tau_i | \sigma_\tau^2, \sigma_e^2, Y \sim N \left(\frac{n\sigma_\tau^2}{\sigma_e^2 + n\sigma_\tau^2}; \frac{\sigma_e^2}{\sigma_e^2 + n\sigma_\tau^2} \left\{ \frac{\sigma_e^2}{nk} + \sigma_\tau^2 \right\} \right)$. For further details see van der Merwe and Bekker (2007).

6. Calculate $\mu_{x_i} = e^{(\zeta^*)}$
7. Repeat steps 1 - 6 l ($= 10000$) times for each of the 15 workers. The result is a $15 \times l$ matrix of simulated observations.

Using the data from steps 1 - 7 above the following can be calculated:

1. For each worker (i.e. each row of the $15 \times l$ matrix of simulated observations mentioned in step 7 above):
 - a. Draw up a histogram
 - b. Calculate $P(\mu_{x_i} > OEL) = \frac{\# \text{ Simulated Values} > OEL}{l}$ where OEL is a pre-defined value of "clinical" interest.

Assume, for the purposes of illustration that $OEL = 130$ and furthermore take $l = 10\ 000$. Histograms of individual worker can also be obtained, but for brevity this is not illustrated here. From the above simulations the following descriptive statistics and Bayesian credibility intervals (CI) were calculated (only three specific workers' results are given for illustrative purposes; complete results for all 15 workers can found in Harvey (2012 – unpublished Ph.D thesis)):

Table 1: Simulation Summary Results

Worker	$P(\mu_{exposure} > 130)$	90% CI		95%CI		Mean	Median	Mode
		Low	High	Low	High			
Worker 2	0.0185	93.85262	125.052	91.19376	128.7124	108.6394	108.2715	108.25
Worker 9	0.4020	110.4973	147.6873	107.3347	152.1216	127.9384	127.2589	124.25
Worker 15	0.0698	98.45825	131.9492	95.61997	136.1275	114.4943	113.9588	110.75

As mentioned previously, the results here are rather different from the results obtained Krishnamoorthy and Mathew (2002). Using their method it does not seem possible to investigate the mean exposure levels of individual workers. For this reason, they found the probability (or tested the hypothesis) that $(100 \times A)\%$ of these workers had exposure levels in excess of a certain pre-defined threshold. However, using the Bayesian methodology as presented above, it is clear that one can examine the probability that a *specific* worker's exposure levels exceed a pre-defined threshold. The above table also provides credibility intervals for each worker's mean exposure levels. For example, it is evident that the mean exposure level for Worker 9, specifically, is higher than that of his/her fellow workers and the probability that his/her exposure levels are in excess of 130 is in excess of 0.402. So using this Bayesian framework we are able to examine the exposure of each worker.

Simulation 2 –Overall Mean Exposures

As mentioned previously, the method given by Krishnamoorthy and Mathew (2002) is an overall representation of the mean exposure level for each worker and since it does not incorporate the actual row means from the data it is somewhat different to the simulation mentioned previously, which offers the opportunity to look specifically at an individual worker. In this simulation the Bayesian equivalent of this "overall" method is described:

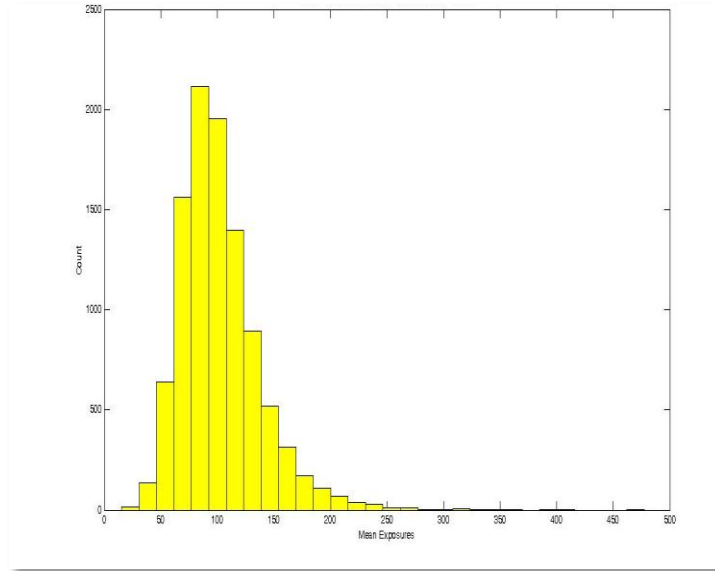
The quantity we are interested in simulating is again

$$\mu_{x_i} = e^{\left(\mu + \tau_i + \frac{\sigma_e^2}{2}\right)}$$

1. Let $\tilde{\lambda} = \mu + \tau_i + \frac{\sigma_e^2}{2}$
2. The conditional distribution of $\tilde{\lambda}$ is as follows: $\tilde{\lambda}|\mathbf{Y}, \tau_i, \sigma_e^2, \sigma_\tau^2 \sim N\left(\bar{Y}_{..} + \tau_i + \frac{\sigma_e^2}{2}, \frac{\sigma_e^2 + n\sigma_\tau^2}{nk}\right)$.
3. To simulate a value from the unconditional distribution, $\tilde{\lambda}|\mathbf{Y}$, we do the following:
 - a. Simulate a pair of $(\sigma_e^2, \sigma_\tau^2)$ values from their joint posterior distribution as described in the previous simulation.
 - b. Given σ_τ^2 , simulate a τ_i observation. Since $\tau_i \sim N(0, \sigma_\tau^2)$ this distribution will be the same for all the workers and therefore we have a “general” mean exposure over all workers, as in Krishnamoorthy and Mathew (2002).
 - c. Using the values simulated in a) and b), i.e. $\tau_i, \sigma_e^2, \sigma_\tau^2$, simulate an observation ($\tilde{\lambda}$) from the normal distribution described in Step 2.
 - d. Calculate $\mu_{x_i} = e^{(\tilde{\lambda})}$.
 - e. Repeat #a to #d l ($= 10000$) times. So now we will have a $1 \times l$ matrix representing an overall picture for all workers.

The $P(\mu_x > OEL)$ is obtained in a manner similar to previously described. In this case, the following results were obtained:

Fig 1 - Overall Mean Exposure per Worker



The following table represents the results from the above graph:

Table 2
Simulation Summary Results

Worker	$P(\mu_{exposure} > 130)$	90% CI		95%CI		Mean	Median	Mode
		Low	High	Low	High			
All Workers	0.1762	56.5422	166.8839	49.8624	189.4470	102.2833	96.4365	83.25

We can see that the credibility intervals for the overall mean exposure per worker are rather wide, even though the mean values are well below the OEL.

Krishnamoorthy and Mathew (2002) also simulate the following statistic (and for the purposes of comparison will be analysed using the Bayesian methodology developed previously):

$$T = \mu + Z_{1-A}\sigma_\tau + \frac{1}{2}\sigma_e^2$$

where A is a suitably chosen parameter between 0 and 1 and Z is the cumulative distribution function of the standard normal distribution. Using a specific value of OEL the following hypothesis can be tested:

$$H_0: \mu + Z_{1-A}\sigma_\tau + \frac{1}{2}\sigma_e^2 \geq \ln(OEL)$$

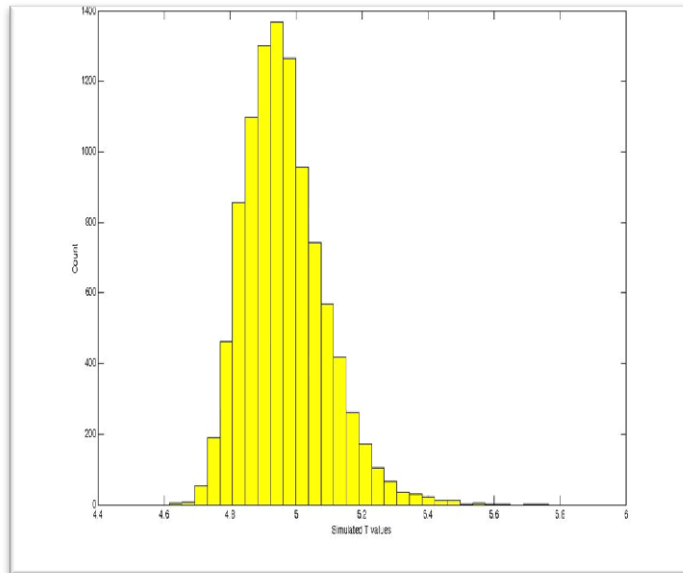
against the complementary alternative hypothesis. For example, if our choice of A is 0.05 then essentially we are testing (one-sided) whether at least 5% of the workers have mean exposure levels in excess of the chosen OEL. The OEL is chosen to be a clinically relevant value. The specific choice of OEL is not the primary concern of this research, but primarily a demonstration of the Bayesian methodology.

In order to replicate the methodology of Krishnamoorthy and Mathew from a Bayesian perspective the following simulation study was undertaken for a range of both OEL and A values:

1. Let $T = \mu + Z_{1-A}\sigma_\tau + \frac{1}{2}\sigma_e^2$
2. To simulate a value of T we do the following:
 - a. Simulate a pair of $(\sigma_e^2, \sigma_\tau^2)$ values from their joint posterior distribution as described in the previous simulation.
 - b. Using the values simulated in a) i.e. $\sigma_e^2, \sigma_\tau^2$, simulate μ from the following normal distribution $\mu|Y, \sigma_e^2, \sigma_\tau^2 \sim N\left(\bar{Y}_{..}, \frac{\sigma_e^2 + n\sigma_\tau^2}{nk}\right)$
 - c. Calculate $T = \mu + Z_{1-A}\sigma_\tau + \frac{1}{2}\sigma_e^2$, where Z is a standard normal variable and Z_{1-A} is the inverse of the cumulative distribution function.
 - d. Repeat #a to #c l ($= 10000$) times. So now we will have a $1 \times l$ matrix.
3. For the data (i.e. the $1 \times l$ matrix of simulated observations mentioned in sub-step d) above):
 - a. Order the observations from smallest to largest.
 - b. Taking $\alpha = 0.05$ find the $100(1 - \alpha)$ th percentile.
 - c. For a specific choice of OEL determine whether $\ln(OEL)$ falls in the critical region (i.e. test the hypothesis described earlier).

This procedure was performed for several choices of OEL ($= [130; 140; 150; 160; 170; 180]$) and for several choices of A ($= [0.1; 0.05; 0.025; 0.001]$) For the purposes of illustration only the combination of OEL = 130 and $A = 0.1$ will be displayed.

Fig 2: Simulated T-values for $A = 0.1$



* $100(1 - \alpha)$ th percentile = 5.1807

* $\ln(130) = 4.8675$

So for the simulated “Styrene Exposures” we can, by way of example arrive at the following: for Fig 2 we can say that 10% (or more) of the workers had exposure values in excess of 170 (95th percentile = 5.1907 and $\ln(170) = 5.1358$) whereas we cannot say that 10% (or more) of the workers had exposure values in excess of 180 ($\ln(180) = 5.193$) at a 5% significance level.

Simulation 3 – Mean Exposure for All Workers

We now shift our attention to the overall mean exposure, i.e. the mean exposure for all the workers. So we now define the overall mean exposure as:

$$\mu_x = e^{\left(\mu + \frac{\sigma_e^2 + \sigma_t^2}{2}\right)}$$

To simulate mean overall exposure levels we do the following:

1. Define $\beta = \mu + \frac{\sigma_e^2 + \sigma_t^2}{2}$.
2. Now, $\beta | Y, \sigma_e^2, \sigma_t^2 \sim N\left(\bar{Y}_{..} + \frac{\sigma_e^2 + \sigma_t^2}{2}, \frac{\sigma_e^2 + n\sigma_t^2}{nk}\right)$. This is the conditional posterior distribution.
3. To simulate an observation from the unconditional posterior distribution, $\beta | Y$, we can do the following:
 - a. Simulate a pair of (σ_e^2, σ_t^2) values from their joint posterior distribution as described in the previous simulation.
 - b. Using this pair of (σ_e^2, σ_t^2) values and the data simulate a $\beta | Y$. Since we have the variance components we can plug them in and simulate the values.
 - c. Calculate $\mu_x = e^{(\beta)}$.
 - d. Repeat sub-steps a) to c) $l (= 10000)$ times. So now we will have a $1 \times l$ matrix representing simulated overall mean exposure levels.
4. For the data (i.e. the $1 \times l$ matrix of simulated observations mentioned in sub-step d) above):
 - a. Draw up a histogram - so there will be only a single histogram representing all simulated mean exposure levels.
 - b. Calculate $P(e^{(\beta)} > OEL) = \frac{\# \text{ Simulated Values} > OEL}{l}$

In this case, the following results were obtained:

Fig 3 - Overall Mean Exposure

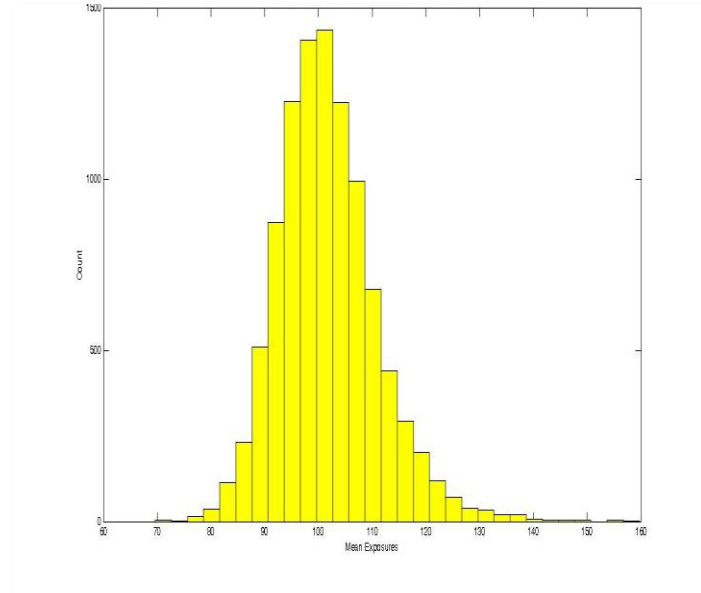


Table 3
Simulation Summary Results

Worker	$P(\mu_{\text{exposure}} > 130)$	90% CI		95% CI		Mean	Median	Mode
		Low	High	Low	High			
All Workers	0.0091	88.3341	117.9463	85.8763	122.439	101.6887	100.9088	97.75

Thus, the overall mean exposure can be easily simulated. The above distribution is substantially narrower than the mean for a new or unknown worker as discussed in the relevant simulation study.

CONCLUSION

In this article the usefulness of the Bayesian methodology to the proposed setting of occupational exposure data was examined. This can however, be extended to many other settings including insurance portfolio claims. The one-way random effects model for balanced data was developed using the chosen prior distribution. In addition to replicating the results achieved using generalized confidence intervals and p-values the proposed method is able to specifically model the occupational exposure of an individual worker. In addition to this significant improvement to the potential analysis of this setting, the advantages of the Bayesian paradigm are apparent. In this article a single non-informative prior has been proposed. However, the derivation and application of other non-informative priors, such as the reference and probability-matching priors, could be used to refine the analysis and improve performance. Ultimately, if subjective prior information is available this could lead to significant improvements in prediction of future exposure for individual workers as well as for groups of workers.

REFERENCES

1. Berger, J.O., Bernardo, J.M. (1992). "On the development of the reference prior method," *Bayesian Statistics 4: Proceedings of the Fourth Valencia Meeting*. Eds. Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. Oxford, U.K., Clarendon Press, pp. 35 – 60.
2. Box, G.E.P., Tiao, G.C. (1973). "Bayesian inference in statistical analysis," *Reading, MA*, Addison-Wesley.
3. Datta, G.S., Ghosh, J.K. (1995). "On priors providing frequentist validity of Bayesian inference," *Biometrika*, Vol 82, pp. 37 – 45.
4. Harvey, J. (2012). "Bayesian inference for the lognormal distribution," *Unpublished Ph.D dissertation, Bloemfontein, South Africa*.
5. Heerderik, D., Hurley, F. (1994). "Occupational exposure assessment: Investigating why exposure measurements vary," *Applied Occupational and Environmental Hygiene*, Vol 9, pp. 71-71.
6. Krishnamoorthy, K., Mathew, T. (2002). "Assessing occupational exposure via the one-way random effects model with balanced data," *Journal of Agricultural, Biological and Environmental Statistics, American Statistical Association and International Biometric Society*, Vol 7 No.3 (Sep), pp. 440 – 451.
7. Lyles R.H., Kupper, L.L. (1996). "On strategies for comparing occupational exposure data to limits," *American Industrial Hygiene Association Journal*, Vol 57, pp. 6-15.
8. Lyles, R. H., Kupper, L. L., Rappaport, S. M. (1997a). "Assessing Regulatory Compliance of Occupational Exposures via the Balanced One-Way RandomEffects ANOVA Model," *Journal of Agricultural, Biological, and Environmental Statistics*, Vol 2, pp. 64–86.
8. Lyles, R. H., Kupper, L. L., Rappaport, S. M. (1997b). "A Lognormal Distribution Based Exposure Assessment Method for Unbalanced Data," *Annals of Occupational Hygiene*, Vol 41, pp. 63–76.
8. Rappaport, S. M., Kromhout, H., Symanski, E. (1993). "Variation of Exposure Between Workers in Homogeneous Exposure Groups," *American Industrial Hygiene Association Journal*, Vol 54, pp. 654–662.
9. van der Merwe, A.J., Bekker, K.N. (2007). "A computational Bayesian approach to the balanced Buhlmann credibility model," *South African Statistical Journal*, Vol 41, pp. 65-103.