

A Bayesian Approach for Calibrating Risk Assessment Models

Michael S. Williams¹, Eric D. Ebel¹,
Jennifer A. Hoeting² and James L. Withee³

¹*Food Safety and Inspection Service,*

United States Department of Agriculture, Fort Collins, Colorado

²*Department of Statistics, Colorado State University, Fort Collins, Colorado*

³*GigaYeast Inc., Belmont, California*

USA

1. Introduction

Monte Carlo simulation is a commonly used tool for constructing foodborne pathogen risk assessment models. Monte Carlo simulation enables an analyst to construct a probabilistic model of almost any desired complexity. It requires relatively little mathematical rigor and the models can be presented in an intuitive manner. It has some drawbacks, however. For example, Monte Carlo simulation requires that each parameter, as well as its uncertainty, be quantitatively described.

The models are typically used to make a projection of possible outcomes. In food-safety risk assessment applications, we typically construct a model to predict the number of human illnesses in the population. These calculations are based on the prevalence of contaminated production units and their microbial load. These are tracked through food production, consumer handling and consumption. The final step is converting predicted contamination into a human health impact via a dose-response model.

Foodborne illness is often the result of an acute microbial pathogen exposure. More than 75 countries have implemented surveillance systems to monitor occurrences of these illnesses (Allos et al., 2004; de Jong B & K., 2006; Herikstad et al., 2002). These surveillance systems do not capture every case of foodborne illness, so scaling factors are developed to estimate the total number of illnesses for the pathogen of interest (Ebel et al., 2012; Scallan et al., 2011). Additional scaling factors can be developed to extend these estimates to a specific product-pathogen pair (Hald et al., 2004).

A conundrum for risk assessors occurs when the illness estimates from a Monte Carlo-based risk assessment model do not match the estimates based on the surveillance data. When these two estimates do not match the risk assessment model must be calibrated. A common approach to calibrating the model is to adjust the parameters of the dose-response function to match predicted to observed illnesses. Alternative calibration approaches are to replace components of the model, such as changing the models used to calculate pathogen attenuation during cooking. The concern with these calibration approaches is their lack of objectivity and rigor.

Bayesian methods offer an alternative approach to the problem of calibration. Various Bayesian methods have been used in previous risk assessment applications (Albert et al., 2008; Hald et al., 2004; Parsons et al., 2005). The Bayesian models are similar in their structure to a Monte Carlo model. They do, however, offer the advantage that any data available at a stochastic node can be incorporated into the model in the form of a prior distribution. These methods allow the user to incorporate the data for observed human illnesses. The model then produces a Bayesian revision of the system's parameter estimates. A consequence of conditioning inferences on the human illness data is that the parameter distributions for each node of the model are shifted so that the predictions more closely match the observed illness data. These adjusted distributions can be thought of as posterior distributions, though some Bayesian methods use the terms pre- and post-model distributions (Givens, 1993; Raftery et al., 1995). The direction and degree to which each of the prior distributions are shifted is to a large extent determined by the relative degrees of uncertainty in the prior distributions (i.e., the parameters in the model that are highly uncertain will experience the largest degree of adjustment in the process of calibrating the model).

This chapter will focus on introducing Bayesian methods for use in food-safety risk assessment. The use of Bayesian methods requires first establishing a simple probabilistic model. Once a model is established, a number of different Bayesian techniques can be used for drawing inferences. We will introduce a relatively simple resampling algorithm that can be used to calibrate a food-safety risk assessment model. A number of examples of the application of this framework can be found in the literature (Williams & Ebel, 2012; Williams, Ebel & Hoeting, 2011; Williams, Ebel & Vose, 2011a) so we will present a rather unusual example where the probabilistic model and Bayesian resampling method are used to study the laboratory test sensitivity of a test for *Escherichia coli* O157:H7 in ground beef samples.

2. Probabilistic models for risk assessment

The probabilistic model we consider assumes that interest lies in modeling a count of events during a specified time period. For food-safety applications, this count will usually be the number of illnesses observed ($I_{observed}$) by a surveillance system during a single year.

We assume the count of sporadic illnesses detected by a surveillance system is reasonably modeled as Poisson random variable. Our model assumes that each food serving of the commodity of interest ($N_{servings}$) has a probability of causing illness of $P(ill)$. The product $N_{servings} \times P(ill)$ describes the rate parameter for a Poisson distribution that describes the total number of illnesses for the product-pathogen pair of interest.

Two factors relate the total number of illnesses for a single product-pathogen pair to the number of illness observed by the surveillance system. The first factor describes the proportion of illnesses, α , attributed to the product of interest. This attribution factor modifies a rate parameter for a specific product to describe the illness rate for the product-pathogen pairing of interest. The second factor describes the proportion of illnesses ρ that are reported by the surveillance system. These factors modify the product-pathogen rate parameter to describe the number of observed illnesses whose etiology is the pathogen. This leads to the basic model for observed illnesses being

$$I_{observed} \sim \text{Poisson}\left(\frac{\rho}{\alpha} N_{servings} P(ill)\right). \quad (1)$$

Other factors may be needed to relate the number of observed illnesses to the total number of illnesses. For example, the surveillance system may only cover a fraction of the population or a pathogen may be specific to a single product (e.g., BSE cases are associated with beef consumption so $\alpha = 1$). Thus, the inclusion of the adjustment terms will be specific to each surveillance system and product-pathogen pair. For this reason, the adjustment factor(s) will not be included for the remainder of the general model development and we note that $N_{servings}P(ill)$ is the rate parameter describing the illness rate in the population, denoted λ_{ill} . An extensive development of these scaling factors can be found in Williams & Ebel (2012) as well as Ebel et al. (2012).

In many applications, the objective of the risk assessment is to predict the change in the number of human illnesses that would occur if the production process were improved. This improvement is expected to reduce $P(ill)$ and the resulting reduction in illnesses can be modeled as

$$I_{avoided} \sim Poisson(N_{servings}(P(ill) - P_{new}(ill))), \tag{2}$$

where $P_{new}(ill)$ is the reduced probability of illness.

3. Partitioning $P(ill)$

The $P(ill)$ term in Equation 1 is one of the typical outputs of a Monte Carlo risk assessment model. Efforts to reduce the complexity of a risk assessment model begin with expanding this term and looking for biologically plausible situations where a simpler model is appropriate. Model simplifications of quantitative microbial risk assessments often begin from first principles: microbial contamination begets food exposure begets illness. In this approach, the interest is in determining the unconditional probability of illness among all servings.

It is also possible to derive these simplifications by applying Bayes Formula. In this case, the question to be answered is "what is the probability that an illness occurred given exposure to a contaminated food". Using Bayesian language, the answer to this question is termed a posterior distribution or $P(ill|exp)$. This is a conditional probability statement. From Bayes Theorem, we have

$$P(ill|exp)P(exp) = P(exp|ill)P(ill) \tag{3}$$

$$P(ill|exp) = \frac{P(exp|ill)P(ill)}{P(exp)}, \tag{4}$$

where $P(exp) = P(exp|ill)P(ill) + P(exp|ill^C)P(ill^C)$.

That the probability of exposure must be this sum can be appreciated from a simple Venn diagram in Figure 1 (i.e., the fraction of exposure servings includes those with and without illness)

This diagram also illustrates the triviality of the conditional probability $P(exp|ill)$. Because all illnesses result from exposure to a contaminated serving, the conditional probability of exposure given that a serving causes illnesses is unity. This conclusion generates a simpler calculation of the posterior probability that we are interested in, namely

$$P(ill|exp) = \frac{P(ill)}{P(exp)}. \tag{5}$$

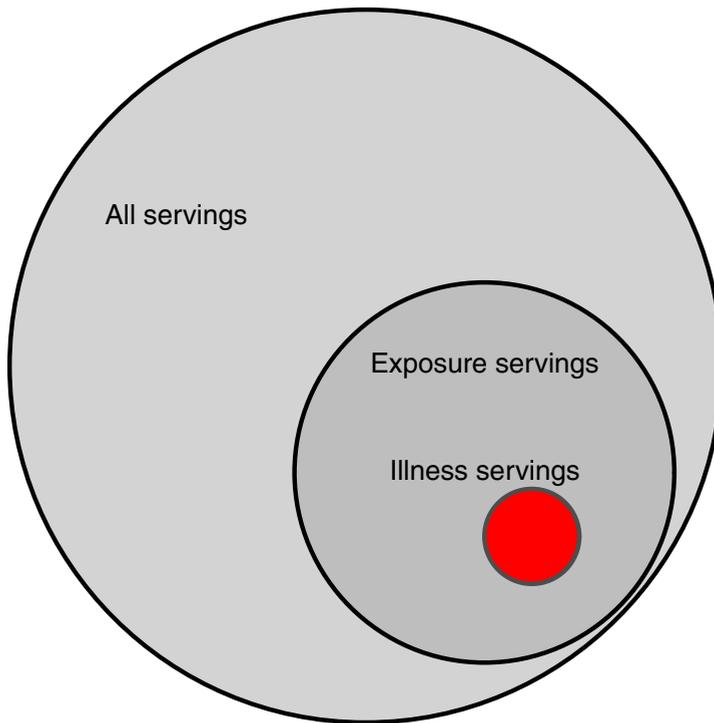


Fig. 1. Venn diagram describing the probability of exposure $P(exp)$ and the probability of illness for contaminated servings ($P(ill|exp)$).

A variation on this equation was employed by Bartholomew et al. (2005) to derive a linear risk model. In that example, the ratio of the estimated number of illnesses and number of exposures per annum derives a constant of proportionality that is ultimately used to project changes in illnesses from intentional changes in exposure. It should be noted that numbers of illnesses and exposures are simple transformations of $P(ill)$ and $P(exp)$ where each is multiplied by the number (or mass) of servings of a food consumed per year.

In microbial risk assessments, we usually have prior information about the number of illnesses per annum (i.e., $N_{servings} \times P(ill)$ where $N_{servings}$ is the number of exposure units), as well as prior information about the fraction of exposure units that are contaminated (i.e., $P(exp)$). This evidence can be used to solve

$$P(ill|exp) = \frac{P(ill)}{P(exp)} \quad (6)$$

using Monte Carlo methods. The result is a posterior distribution of this conditional probability.

Contemplating this calculation, however, highlights a potential problem in using the posterior distribution to make risk projections that might result from changes in exposure. Because this posterior is derived as a ratio of two random variables (each describing uncertainty about a

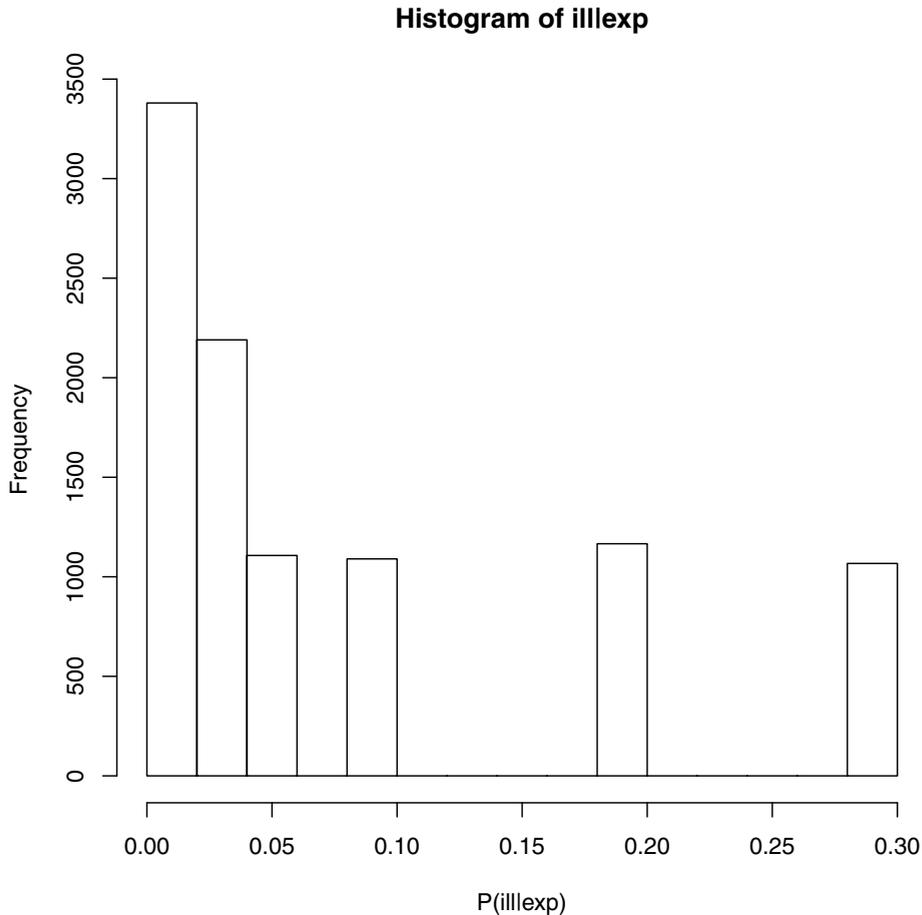


Fig. 2. Histogram for a simplistic Monte Carlo calculation of $P(ill|exp)$.

true parameter in nature), this distribution is necessarily informed by any covariance between these random variables. Furthermore, projections about future values of $P(ill)$ derived by $P_{new}(exp)P(ill|exp)$ need to account for the starting value of $P(exp)$.

A simple example may illustrate this point. Assume available prior evidence about numbers of illnesses (for a particular product-pathogen pair) is that 100, 200 or 300 cases occur per year with equal probability weights. Further, assume our understanding of exposures implies that 1%, 5% or 10% of 10,000 servings per year are contaminated with equal probability weights. A naive estimate of $P(ill|exp)$ would look like the histogram in Figure 2.

This ratio ranges from 0.01 (100 illnesses divided by 10,000 exposures) to 0.30 (300 illnesses divided by 1000 exposures). Furthermore, we can re-derive the uniform distribution of $P(ill)$ by simply multiplying the vector of $P(ill|exp)$ by the vector of $P(exp)$. But, if we consider

these two independent random variables, we will generate a distribution for $P(ill)$ that is not at all what the prior $P(ill)$ looked like.

Instead, the appropriate distribution for $P(ill|exp)$ is a distribution particular to the value of $P(exp)$. For example, $P(ill|exp)$ for $P(exp) = 0.01$ is a discrete Uniform distributions of values (100/1000, 200/1000, 300/1000).

If the analyst wants to predict the effect of a change in $P(exp)$, this dependence between $P(exp)$ and $P(ill|exp)$ should be borne in mind. Otherwise, incorrect representations of $P(ill)$ could result. Fortunately, the model simplifications developed in Williams, Ebel & Vose (2011a) avoid this trap because the $P(ill|exp)$ often cancels out of the equations and the change in illness occurrence can be estimated directly from changes in $P(exp)$. The term *prevalence-dependent* model is used to describe applications where this simplification is feasible.

4. Model simplification

A complete evaluation of the components of the model in Equation 1 can still be a complex task. Nevertheless, the factorization on $P(ill)$ into its exposure component ($P(exp)$) and hazard characterization component ($P(ill|exp)$) leads to situations where estimation of the number of illnesses can be greatly simplified. We outline two different models and describe methods for simplification.

4.1 Dose-dependent model

The first parameterization assumes that all servings have some level of contamination, where D describes the average number of pathogens in each serving. Note that when D describes an average concentration, it is possible for these concentration values to be much less than 1 unit per serving. Common examples are the description of pathogen levels in water. It may also be reasonable to model average concentrations for liquid and ground food products where no natural units exist. An exposure event from a particular food type will involve the ingestion of a random number of pathogenic organisms, where the distribution of organisms is described by the probability density $f(D)$. The lognormal distribution is a common and convenient choice (Limpert et al., 2001), so $f(D) \sim \text{Lognormal}(\mu_D, \sigma_D^2)$. The probability that a random person will become ill, given a microbial dose of size D , is $P(ill|D)$. Averaging across all possible doses yields the probability of a person becoming ill given exposure to the pathogen. When D describes an average dose, the probability of illness given exposures described by a continuous dose distribution is

$$P(ill) = \int_0^{\infty} P(ill|D)f(D)dD, \quad (7)$$

where $P(ill|D)$ is the dose-response function. The exponential and beta-Poisson dose-response functions are appropriate for continuous dose distributions. The term dose-dependent model will be used to denote this model.

The difficulty with this model is that data describing the dose at the point of consumption are not available. Instead, virtually all risk assessment models rely on a measurement of contamination, X , derived from data collected at a more convenient location in the farm-to-table continuum, such as during production or at retail. A typical risk assessment must rely on models of post-production activities to transform measurements of microbial

contamination following production (X) into measurements of human health risk at the point of consumption (D).

If it is assumed $X \sim \text{Lognormal}(\mu_X, \sigma_X^2)$ and the focus of a risk assessment is to determine how changes in the production process would lead to a change the number of illnesses, then the distribution of D can be derived from a single component Δ , that describes the cumulative change in average microbial level between production and consumption (i.e., it combines the effects of storage time and temperature as well as cooking and other process). Assuming that the cumulative change is distributed as $\Delta \sim \text{Lognormal}(\mu_\Delta, \sigma_\Delta^2)$, the distribution for D is $\text{Lognormal}(\mu_X + \mu_\Delta, \sqrt{\sigma_X^2 + \sigma_\Delta^2})$.

The dose-dependent model can be simplified by treating Δ as a latent variable, with its parameters $(\mu_\Delta, \sigma_\Delta^2)$ estimated during calibration. Williams, Ebel & Vose (2011a) provide an example based on *Campylobacter* contamination in chicken.

Additional simplifications of the model are possible in situations where pathogen numbers are uniformly low at the point of consumption (Williams, Ebel & Vose, 2011b).

4.2 Prevalence-dependent model

For the model in Equation 1, the number of illnesses avoided by reducing the prevalence of contaminated servings is readily predicted via Equation 2. Reduced prevalence of contamination might occur via changes in import practices or improved animal husbandry practices that reduce the occurrence of a pathogen among farms, herds, flocks or sheds. These changes are expected to reduce the prevalence of contaminated carcasses, but in a number of situations it is still reasonable be assumed that $P(\text{ill}|\text{exp})$ would remain essentially unchanged.

For example, suppose that a country, where a specific pathogen is endemic, will begin importing animal products from a country that is free from the disease. If the importation of uncontaminated carcasses is such that prevalence is reduced by $P_{\text{new}}(\text{exp}) = \delta P(\text{exp})$ where uncertainty about change in prevalence might be characterized as $\delta \sim \text{Beta}(a, b)$, and it is reasonable to assume that $P(\text{ill}|\text{exp})$ will remain unchanged, then the human health benefit is modeled as:

$$I_{\text{avoided}} \sim \text{Poisson}(N_{\text{servings}}(P(\text{exp})P(\text{ill}|\text{exp}) - P_{\text{new}}(\text{exp})P(\text{ill}|\text{exp}))) \tag{8}$$

$$\sim \text{Poisson}\left(\left(1 - \frac{P_{\text{new}}(\text{exp})}{P(\text{exp})}\right)\lambda_{\text{ill}}\right) \tag{9}$$

$$\sim \text{Poisson}((1 - \delta)\lambda_{\text{ill}}). \tag{10}$$

Note that this model relies only on the characterization of the number of illnesses (λ_{ill}) and the effect of the change in importation policy. Also note that measures of prevalence are not necessarily the prevalence of contaminated servings. Instead, one can argue that the prevalence of contaminated units at the point of data collection is proportional to the prevalence of contaminated servings. This constant of proportionality cancels out when $P(\text{exp})$ and $P_{\text{new}}(\text{exp})$ are measured at the same location in the farm-to-table continuum.

This formulation also obviates the need for modeling pathogen levels as well as eliminating the need to adjust for the difference between true and apparent prevalence. A linear relationship between contaminated carcass prevalence and human illnesses was

demonstrated for chicken and campylobacteriosis in a previous risk assessment (Bartholomew et al., 2005).

5. Bayesian methods

Recent research interest has focused on the replacement of Monte Carlo models with a Bayesian approach that uses Markov chain Monte Carlo (MCMC) methods (Albert et al., 2008; Hald et al., 2004; Parsons et al., 2005). Along with this proposed approach comes the inevitable suggestion that the models be built using MCMC packages such as WinBUGS (Lunn et al., 2009; 2000; Vose, 2008). These packages often rely on the Gibbs sampler, or similar algorithms, to obtain a set of random samples from the posterior probability distribution of the risk assessment model. While it is possible to use software packages such as WinBUGS (Williams, Ebel & Hoeting, 2011), personal experience suggests that convergence is difficult to achieve given the high degree of uncertainty in the parameters of even a highly simplified food-safety risk assessment model. The underlying problem can be understood by examining the mechanics of an MCMC algorithm.

Sampling and numerical search algorithms generally follow two approaches. MCMC algorithms generate a new realization of the model parameters at each iteration (following a burn-in period). When the model converges, each iteration is an element or observation from the posterior distribution.

An alternative approach to MCMC algorithms are algorithms that first generate a large number of candidate values for each parameter using Monte Carlo simulation. Bayesian logic combines the new evidence, denoted by E , with the Monte Carlo parameter estimates to select or reweight a subset of the Monte Carlo generated parameters. In this application, the new evidence will be the illness count from a public health surveillance system (i.e., $E = I_{observed}$).

The algorithm we employ is the sampling importance resampling (SIR) approach proposed by Rubin (1987). This method generates an unequal probability sample where the sample weights are determined by the degree of agreement between the prior information and the new sampling evidence.

To demonstrate, let θ represent a vector of model inputs. Examples of model inputs are parameters describing the contamination distribution and dose-response function. The input parameters are not fixed values, so uncertainty is represented by the distribution $p(\theta)$. This distribution is referred to as the prior or pre-model distribution for the inputs. Consider a process model that predicts, among other parameters, the rate parameter describing the number of human illnesses (Equation 1). Let this model be denoted by $M(\cdot)$. For a randomly sampled θ from $p(\theta)$, the observed output from the process model is $M(\theta)$. The $M(\theta)$ value will be compared to the observed number of illnesses from public health surveillance, which is denoted by $I_{observed}$.

The algorithm for implementing the SIR is:

1. Draw N samples $(\theta_1, \theta_2, \dots, \theta_N)$ from the prior distribution $p(\theta)$.
2. For each θ_i , use the model to determine $M(\theta_i)$.
3. Determine a weight for each $M(\theta_i)$. This weight describes the agreement between the model prediction and observed number of illnesses. For this application, the weight is $w_i = P(I_{observed} | M(\theta))$.

4. Draw an unequal probability with-replacement sample of size $m \ll N$ from $(\theta_1, \theta_2, \dots, \theta_N)$ using sample weights w_i .

As $N/m \rightarrow \infty$ the SIR algorithm produces an exact sample from the posterior distribution. Previous studies have found that values for N/m ranging from 20 to 40 are often sufficient (Rubin, 1987), but appropriate values must be considered on a case-by-case basis.

To illustrate the SIR algorithm, consider the problem of estimating the prevalence θ of a disease in a herd of animals when new sampling evidence is combined with prior information. Suppose the new evidence is a sample from the herd of size $n = 20$ of which $s = 4$ samples are positive. Suppose the prior evidence on the prevalence in the herd can be summarized by a beta distribution of the form $\theta \sim \text{Beta}(a = 1, b = 6)$. In this example the model, $M(\theta)$, uses the prior information on prevalence, and the number of additional test results, to predict the number of infected animals. Using the model $s \sim \text{Binomial}(n, \theta)$ and Bayes formula, it is known that the distribution the resulting posterior distribution $p(\theta|s, n) \sim \text{Beta}(s + a, n - s + b) = \text{Beta}(5, 22)$. The following R code (R Development Core Team, 2011) demonstrates the SIR algorithm and illustrates the equivalence of the SIR solution and the known posterior distribution.

```
# example of the beta-binomial by SIR

N=200000

# parameters for the beta prior
a=1
b=6

# evidence. 4-successes 20-trials
s=4
n=20

# draw a large sample from the prior (the "Sampling" step)
p.prior=rbeta(N,a,b)

# the "Importance" weight of each sample is determine by the likelihood
w=dbinom(s,n,p.prior)

# the "Resampling" step draws a weighted sample from the prior
m=round(N/40)
index=sample(seq(1,N),size=m,replace=T,prob=w)
p.posterior=p.prior[index]
```

6. Example

Previously published examples of the proposed probabilistic framework and Bayesian method (Williams & Ebel, 2012; Williams, Ebel & Vose, 2011a) focus on the prediction of changes in human illness in a farm-to-table model. The following example is a departure from the farm-to-table model and it is presented to highlight similarities and the utility of the proposed framework in applications where the data represent a surveillance system.

Surveillance sensitivity, in the context of testing for pathogens in food, is the probability that the pathogen is detected given that it exists in a sampled unit. Surveillance sensitivity is a function of test sensitivity in the sense that not only does a contaminated unit need to be sampled, but the results of testing must properly classify the unit as containing the pathogen of interest. When the sensitivity of a test is Se , and the number of units in the population is large in relation to the number of samples collected (n), and $p(S^+)$ is the proportion of the units that are contaminated. Then the surveillance system sensitivity is typically given by

$$P(\text{detecting one or more positives}) = 1 - (1 - Se \times p(S^+))^n. \quad (11)$$

Note that the role of test sensitivity is to modify the true prevalence term $p(S^+)$ to provide the apparent prevalence.

The concern with the standard approaches used in defining Se and equation 11 is that test sensitivity can decrease as the level of the pathogen drops, especially in cases where the average concentration is less than 1 cfu/tested unit (e.g., a test that uses 10 g of a 100 g food unit with only 1 organism will have an average concentration of 0.1 cfu/g).

In the testing for pathogens that occur at low levels, test sensitivity can be improved by employing enrichment techniques, increasing incubation time, and increasing the volume of material sampled. All of these methods increase the number of pathogens in the medium to be tested. The bonding and potential encapsulation of a microbe within fatty tissues, insufficient time for cells to leave a quiescent state during incubation, the possibility of cells entering a viable but nonculturable state (Oliver, 2005; Oliver et al., 2005), and the small volume of material tested all can lead to reductions in test sensitivity.

In this study, a situation is examined in which the prevalence of positive samples doubled over a one-year period. During this time, a minor modification was made to the enrichment technique used for testing. While the laboratory had performed testing to determine the equivalence of the new and old methodology, insufficient evidence exists to determine if the observed increase in prevalence is due to the change of the enrichment methodology or whether the change was due to an actual change in pathogen prevalence.

This study is predicated on the assumption that the observed increase in *E.coli* O157:H7 positive samples is the result of a change in enrichment media, rather than an actual increase in contamination. The analyses presented assess what the change in test sensitivity would be if this assumption were true. This initial analysis is used to specify both the required sample and the concentration of *E.coli* O157:H7 used to spike validation samples.

7. Data description

The Food Safety and Inspection Service of the United States Department of Agriculture has been collecting ground beef samples from all slaughter and grinding facilities producing ground beef products since the beginning of the year 2000. Nevertheless, the enumeration of positive samples was only begun in January 2007. This more limited dataset was used in this analysis. Each facility that produces ground beef for distribution is sampled on multiple occasions every year and the sample unit consists of approximately 900 g of ground beef collected at the end of production from approximately 5,000 kg lots.

The data used in this study spans the time period between January of 2007 and December 2009. The annual number of samples collected was 9401, 10760, and 10774 for each of the three years. The number of positive samples observed in each year was 26,43 and 35, with the change in the enrichment media occurring in January of 2008.

The recorded data values are the result of two tests. The first being a qualitative test that provides a positive or negative result for each sample. The sample for this qualitative PCR test that consists of a 325 g subsample of the original 900 g sample that is incubated for 24 hours using an enrichment broth. The size of the sample implies a detection limit of 0.003 cfu/g.

The second test is only performed on the qualitatively positive samples. This test provides an estimate of the number of organisms per gram, derived from an additional 33.3 g sample of the remaining portion of the original 900 g. The estimate is derived using the most probable number (MPN) method (Harrigan, 1998) using three dilutions and three tubes per dilution. The dilutions used for MPN testing were 10, 1, and 0.1 g. Thus, the 33.3 g sample is divided into 3 10-g, 3 1-g and 3 0.1-g subsamples, incubated in a test tube, and tested for *E. coli* O157:H7 using the same PCR technique.

The two tests have different levels of detection and there are three possible outcomes for any sample (i.e., -/-, +/-, +/+) (estimated number of organisms per g). Samples that are only positive on the first test will be referred to as qualitatively positive, while samples that were positive on both tests will be referred to as quantitatively positive.

The Food Safety Inspection Service (FSIS, 1998) provides additional descriptions of the testing methodology since the inception of the program.

8. Methods

8.1 Estimating the distribution for the average concentration of *E.coli* O157:H7

Ground beef at the point of production comprise no natural units, so it can be viewed as a homogeneously mixed product that mimics a viscous liquid. As such, the estimation of the levels of pathogens in ground beef follows a similar methodology as used for describing the distribution of contaminants in other fluids, such as water. The grinding process inevitably introduces some microscopic voids in the medium, so the unit of measurement is typically in grams rather than milliliters.

The sampling data are censored in the sense that the true number of organisms is not observed for all samples. This occurs because the test has a level of detection (*LOD*) at which the probability of a positive is low even though the original sample contained one or more viable organisms (Helsel, 2005).

Contamination generally occurs at very low levels for the vast majority of ground beef production. Nevertheless, there are situations where high contamination levels can occur. The commonly used biologically plausible model assumes that the average concentration of contamination varies according to a Lognormal distribution (i.e., $X \sim \text{Lognormal}(\mu, \sigma^2)$) (Haas et al., 1999).

The parameters of the contamination distribution can be estimated using maximum likelihood. To account for the different levels of detection, the total number of samples, N , can be broken down into $N = N_{neg} + N_{qual+} + N_{quan+}$. The two levels of detection are given by $LOD_{qual+} = 0.003$ and $LOD_{quan+} = 0.03$.

The likelihood is given by

$$L = [F(LOD_{qual+})]^{N_{neg}} \times [F(LOD_{quan+}) - F(LOD_{qual+})]^{N_{qual+}} \times \prod_{i=1}^{N_{quan+}} f(x_i), \quad (12)$$

where x_i is the estimated number of organisms from the MPN analysis for the i^{th} quantitatively positive sample, and F and f represent the cumulative and probability density function of the Lognormal distribution of contamination levels. Applying the specific LOD values for this application yields;

$$L = [F(0.003)]^{N_{neg}} \times [F(0.03) - F(0.003)]^{N_{qual+}} \times \prod_{i=1}^{N_{quan+}} f(x_i). \quad (13)$$

A nonlinear optimization routine (*optim* in R) was used to estimate $\hat{\mu}$ and $\hat{\sigma}^2$. This routine also provides the variance-covariance matrix so that $var[\hat{\mu}]$, $var[\hat{\sigma}^2]$ and $cov[\hat{\mu}, \hat{\sigma}^2]$ are available for further analyses. These estimates will be used as hyper-priors describing uncertainty in the estimated pathogen levels.

8.2 Defining test sensitivity as a function of pathogen level

The process under which a positive test occurs requires two sequential outcomes:

1. The original 325 g sample must contain one or more viable organisms.
2. The DNA in a cell must be successfully amplified by PCR for detection to occur.

Two factors can affect the probability of detection. The first is that material from the single cell undergoes successful amplification. The alternative is that the cell leaves its quiescent state during the enrichment phase and begins the process of exponential growth.

Suppose there are Z cells in the sample and for each cell there is a probability Q of detection. Given the probability of detection for an individual cell, and the number of cells, the probability of a positive test is defined as the compliment of non-detection for each of the Z organisms, therefore

$$P(T+ | Q = q, Z = z) = 1 - (1 - q)^z. \quad (14)$$

This probability assumes fixed values for Q and Z , though there is likely to be variation among organisms as well as variation in the number of organisms.

Suppose for each cell in the sample the probability, Q , of detection is distributed in the population as a Beta distribution with parameters α and β . Similarly, let Z represent the number of viable organisms per gram initially in the sample. Assuming that the organisms are uniformly distributed in the medium, Z can be modeled as a Poisson distribution. The rate parameter (X) for any given sample is determined by a draw from the Lognormal distribution describing *E. coli* O157:H7 levels previously described.

A two-step process is used to incorporate the variability in Q and Z . First, assume a fixed probability of detection for each organism. Applying the Theorem of Total Probability across the Poisson distributed organism counts, with rate parameter X , gives

$$\begin{aligned}
 P(T+ | Q = q, X = x) &= \sum_{z=0}^{\infty} P(T+ | Q = q, X = x, Z = z)P(X = x, Z = z) \quad (15) \\
 &= \sum_{z=0}^{\infty} (1 - (1 - q)^z) \frac{e^{-x}x^{-z}}{z!} \\
 &= 1 - e^{-x} \sum_{z=0}^{\infty} \frac{((1 - q)x)^{-z}}{z!} \\
 &= 1 - e^{-x} \times e^{(x-qx)} \\
 &= 1 - e^{-qx}.
 \end{aligned}$$

The probability of detection for each cell is likely to vary across the population due to factors such as genetic diversity and the level of damage the cell may have received during the production process due to the possible application of antimicrobials and steam to carcasses, as well as drying, freezing and storage times prior to sample collection and testing. Assuming that the detection probability is $Q \sim \text{Beta}(\alpha, \beta)$ and rate parameter X , the probability of detection is given by:

$$\begin{aligned}
 P(T+ | X = x) &= \int_0^1 P(T+ | Q = q, X = x)f(Q = q)dq \\
 &= \int_0^1 (1 - e^{-qx}) \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} q^{\alpha-1}(1 - q)^{\beta-1}dq. \quad (16)
 \end{aligned}$$

The integral in this expression describes a confluent hypergeometric function. Using the results of (Haas et al., 1999) the exact solution can be written as

$$P(T+ | X = x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \sum_{i=0}^{\infty} \frac{\Gamma(\alpha + i)}{\Gamma(\alpha + \beta + i)} \frac{(-1)^{i-1}x^i}{i!}. \quad (17)$$

The ultimate goal is to determine the probability of a positive test across the distribution of possible contamination levels. If $f_X(x)$ is the Lognormal probability density function describing contamination levels across the population of ground beef, then

$$P(T+) = \int_0^{\infty} P(T+ | X = x)f_X(x)dx, \quad (18)$$

where $P(T+ | X = x)$ is as defined in either 16 or 17. Note that the series expansion for the confluent hypergeometric in Equation 17 includes an alternating power series of x^i , and the range of integration for x is $(0, \infty)$. This causes numerical problems for levels of x greater than approximately 50 cfus and precludes the use of the numerical approximations based on the series expansion.

A solution to the problem is to combine 18 and 16 as

$$P(T+) = \int_0^{\infty} \int_0^1 (1 - e^{-qx}) \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} q^{\alpha-1}(1 - q)^{\beta-1}dqf_X(x)dqdx. \quad (19)$$

The evaluation of the double integral that results from combining 18 and 16 is numerically difficult and time consuming to reliably implement in a Monte Carlo simulation. One solution is to use computer hardware and software that accommodates parallel processing, such as the *Snowfall* package in *R* (Knaus et al., 2009; Rossini et al., 2007).

Another alternative comes from noting that the model in Equation 16 also appears in epidemiology and food-safety applications where it is functionally identical to a beta-Poisson dose-response function (Haas et al., 1999). In this setting the model is used to determine the probability of illness ($P(ill)$) with a Poisson distributed number of organisms with mean value X , with an individual dose-response

$$P(ill|D = d) = 1 - (1 - q)^d. \quad (20)$$

In this equation $P(ill|D = d)$ is the individual probability for illness, d the individual dose, and q an individual measure of susceptibility that is distributed in the population as a beta distribution with parameters α and β . Note that this expression is equivalent to the probability of illness given exposure ($P(ill|exp)$) defined previously.

An approximation to the exact solution in 18 has been derived for the beta-Poisson dose-response equation. The approximation is

$$P(T+ | X = x) = 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha}. \quad (21)$$

Furumoto & Mickey (1967) show, via a Taylor's series expansion, that this approximation is sufficiently accurate for $\beta > \alpha$ and x reasonably small. Prior experience indicates that the number of organisms is low for this application, however, the relationship between β and α is not known *a priori*. This model produces a sigmoidal curve in log-space and the interpretation of the two parameters is that α controls the slope of the curve while β controls the location.

The ultimate goal is to determine the probability of a positive test across the distribution of possible contamination levels. If $f_X(x)$ is the Lognormal probability density function describing contamination levels across the population of ground beef, then

$$P(T+) = \int_0^{\infty} 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha} f_X(x) dx. \quad (22)$$

Two different probabilities of a positive test are of interest. The first is as described in Equation 22, which is assumed to be the probability of a positive test prior to the change in enrichment methodology. The sensitivity of the test is greatly improved if the organism has sufficient time to undergo the process of replication. If, as postulated, the increase in the number of positives samples is a side-effect of an improvement in enrichment media, the probability of detection for each organism would be greater. The magnitude of this increase would be captured in the parameters that describe the probability of detection. Let the parameters α_{new} and β_{new} represent the postulated increase in test sensitivity associated with the new enrichment technique, then the probability of a positive test is

$$P_{new}(T+) = \int_0^{\infty} 1 - \left(1 + \frac{x}{\beta_{new}}\right)^{-\alpha_{new}} f_X(x) dx \quad (23)$$

8.3 Description of surveillance data

The goal of the analysis is to quantify the parameters $\alpha, \beta, \alpha_{new}, \beta_{new}$. The differences in the resulting test sensitivity models can be used to inform additional analyses related to determining the necessary sample size to test for equivalence of the two enrichment media.

The enumerated values from the surveillance data are used to characterize the distribution of contamination in ground beef (i.e., $f_X(x)$). These surveillance data from 2000-2006 represent what is thought to be a time period of relative stability in the level of contamination and are characterized by annual rate of positives samples of approximately 1 positive for every 500 samples (0.2%).

In the following three years there were $T_{2007} = 26$, $T_{2008} = 43$ and $T_{2009} = 35$ positive tests observed. The number of tests performed was $N_{2007} = 9,401$, $N_{2008} = 10,760$ and $N_{2009} = 10,774$. The 2007 data represent the last year during which no changes were made to the laboratory protocols. The 2008 and 2009 data represent the first years in which the new enrichment technique was implemented and there was an immediate increase in the number of positives samples observed, with the percent positive rate jumping from a long-term average of approximately 0.20 to 0.36%.

8.4 Bayesian model

Application of the SIR routine requires the definition of the model ($M(\theta_i)$) and the likelihood equation used to determine the sample weights (i.e., $w_i = P(E|M(\theta))$).

In this application, the model generates estimates of the number of positive tests during the 2007 and the combined time period 2008-2009. Define these two predictions as $M_{2007}(\theta_{2007}) = N_{2007}P(T+)$ and $M_{2008-9}(\theta_{2008-9}) = (N_{2008} + N_{2009})P_{new}(T+)$. The parameter vectors consist of $\theta_{2007} = (\mu, \sigma^2, \alpha, \beta)$ and $\theta_{2008-9} = (\mu, \sigma^2, \alpha_{new}, \beta_{new})$. Uncertainty in the μ, σ^2 parameters is modeled by a multivariate Normal distribution where the mean values and variance-covariance matrix were derived from the maximum likelihood solution of Equation 13. Uniform priors were used for the α and β parameters of the test sensitivity models. Equations 22 and 23 were used to calculate $P(T+)$ and $(P_{new}(T+))$, respectively.

There is no evidence of plant-level clustering amongst the positive tests collected across the year 2000-2009, so it is reasonable to assume that the number of positive tests can be modeled as a Poisson distribution, giving

$$T_{2007} \sim \text{Poisson}(N_{2007}P(T+)) \quad (24)$$

and

$$T_{2008-9} \sim \text{Poisson}((N_{2008} + N_{2009})P_{new}(T+)). \quad (25)$$

Each of these expressions serves as a likelihood function in a Bayesian model where uncertainty in $P(T+)$ and $P_{new}(T+)$ provide extra Poisson variability. The model was solved using the SIR algorithm where the weights w_i were the product of the two likelihoods, so

$$w_i = P(T_{2007}|N_{2007}P(T+)) \times P(T_{2008-9}|(N_{2008} + N_{2009})P_{new}(T+)). \quad (26)$$

The linkage between the two likelihoods is the common distribution explaining the level of contamination. The sample sizes used in the SIR algorithm, were $N = 1,000,000$ and $m = 25,000$.

9. Results

The test sensitivity models are essentially equivalent for extremely low concentrations and for concentrations substantially greater than the LOD, with the probability of detection being near zero at low levels and essentially one at higher concentrations (Figure 3). Any laboratory testing to determine the equivalence between the two enrichment methods could require an enormous sample size to detect a significant difference if the concentration were too high because essentially no difference exists between the two methods when samples contain more than four or five viable cells. However, there exists a range over which the models no longer overlap Figure. 3.

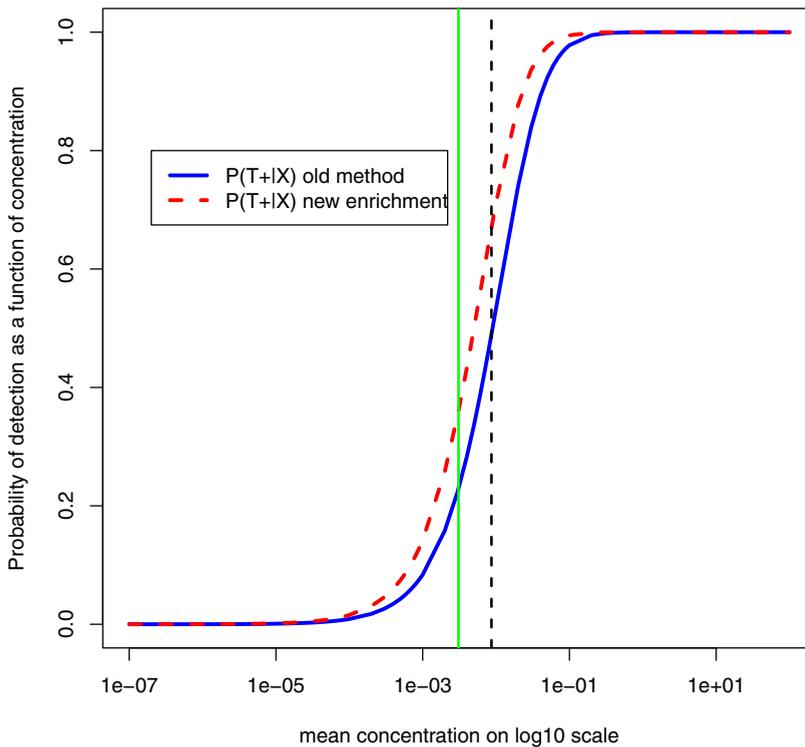


Fig. 3. Summary of the test sensitivity models. The discrepancies between the two models predominantly occur at concentrations much lower than 1 cfu per g. The *LOD* is given by the solid vertical line. The dashed vertical line indicates the concentration at which the maximum difference in sensitivity between the old and new enrichment methods occurs.

The vertical line in Figure 3 represent the mean concentration at the theoretical *LOD* of a single organism. If this is interpreted to mean that the test could truly identify 1 organism in a 325 g sample, than there will be one or more organisms contained in the sample with probability $P(T+) = 1 - e^{-1} = 0.63$. A test with less than perfect sensitivity will have a lower probability of detection. The probabilities of detection at this concentration for both tests are substantially

less than this theoretical upper bound, with the probabilities of a positive test being 0.21 and 0.40.

The level of the target organism at which the maximum difference between the two tests occurs can be estimated from the two models and is presented graphically in Figure 4. This value could then be used to determine both the appropriate concentration for testing and the number of samples required to achieve a test with a specified power. From a practical standpoint this may not be a reasonable approach, given that for this application the concentrations are very low and it becomes very difficult to accurately dilute concentrations at these low levels.

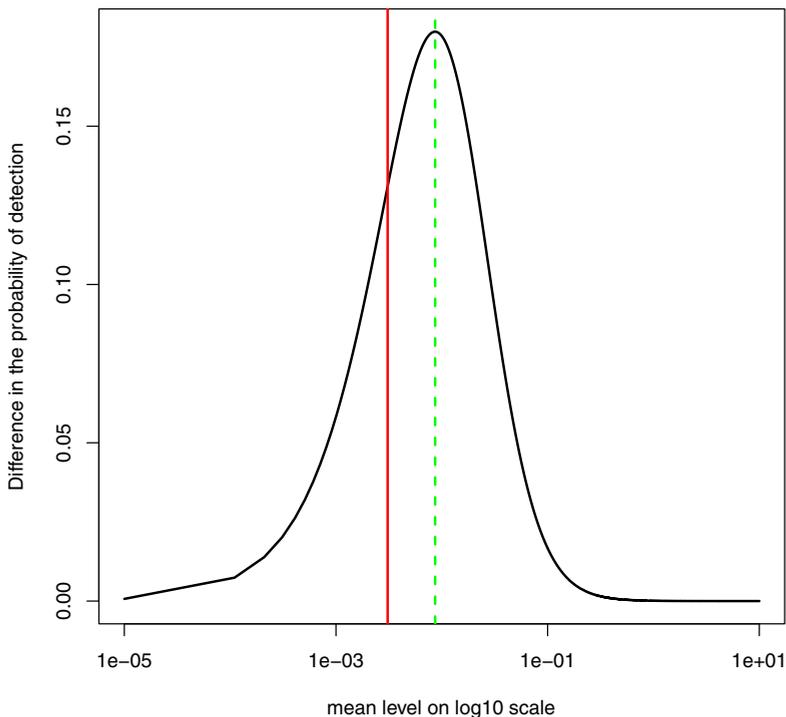


Fig. 4. Results of the model can be used to determine the average concentration at which the maximum discrepancies occurs. This concentration can be used in a validation test to determine if differences in the two enrichment media exist. The *LOD* is given by the solid vertical line. The dashed vertical line indicates the concentration at which the maximum difference in sensitivity between the old and new enrichment methods occurs.

The solid vertical line on Figure 4 represents the difference in the probability of detection at the mean level of 1 organisms per 325 g, which is the *LOD* for the current sampling program. The maximum difference between the two test sensitivity models occurs for a concentration of roughly 1 organism per 114 g (dashed vertical line), at which point the probability of a positive test result is 0.45 and 0.67 for the old and new enrichment techniques, respectively.

In retrospect it seems fairly obvious that equivalency testing for enrichment methodologies should be performed at levels near the *LOD*. The reasoning for this claim is twofold:

1. Qualitative tests, such as PCR methods, are so sensitive that only a small number of cells are required for detection.
2. The purpose of enrichment techniques is to selectively instigate reproduction of the target organism. Given the exponential growth in the number of organisms once reproduction has begun, a positive test result is expected.

10. Conclusions

There are surprising number of applications where this simplified framework can be used to integrate surveillance data with data describing basic demographics of the population (e.g., prevalence and/or levels of contamination) and counts derived from a surveillance system. The advantage of using Bayesian methods with simplified risk assessment models is their ability to combine the available data, objectively calibrate the model to match the surveillance data, and to estimate latent variables in the model for which there are few data.

11. Acknowledgement

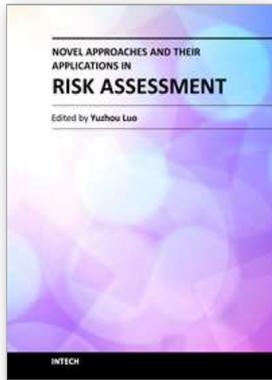
This research utilized the Colorado State University ISTE C Cray HPC System supported by NSF Grant CNS-0923386.

12. References

- Albert, I., Grenier, E., Denis, J. & Rousseau, J. (2008). Quantitative risk assessment from farm to fork and beyond: A global bayesian approach concerning food-borne diseases, *Risk Analysis* 28: 557–571.
- Allos, B., Moore, M., Griffin, P. & Tauxe, R. (2004). Surveillance for sporadic foodborne disease in the 21st century: The foodnet perspective, *Clinical Infectious Diseases* 38: 115–120.
- Bartholomew, M., Vose, D., Tollefson, L. & Travis, C. (2005). A linear model for managing the risk of antimicrobial resistance originating in food animals, *Risk Analysis* 25(1): 99–108.
- de Jong B & K., E. (2006). The comparative burden of salmonellosis in the european unionmember states, associated and candidate countries, *BMC Public Health* 6(4). URL: www.biomedcentral.com/1471-2458/6/4
- Ebel, E., Williams, M. & Schlosser, W. (2012). Parametric distributions of under-diagnosis parameters used to estimate annual burden of illness for five foodborne pathogens, *Journal of Food Protection* 74: (in press).
- FSIS (1998). Microbiology laboratory guidebook. 3rd ed, *Technical report*, Food Safety and Inspection Service, US Department of Agriculture. http://www.fsis.usda.gov/Science/Microbiological_Lab_Guidebook/index.asp; accessed October 23, 2011.
- Furumoto, W. & Mickey, R. (1967). A mathematical model for the infectivity-dilution curve of tobacco mosaic virus: Theoretical considerations, *Virology* 32: 216–223.
- Givens, G. H. (1993). *A Bayesian framework and importance of sampling methods for synthesizing multiple sources of evidence and uncertainty linked by a complex process model*, PhD thesis, University of Washington, Seattle, Washington.

- Haas, C., Rose, J. & Gerba, C. (1999). *Quantitative Microbial Risk Assessment*, John Wiley & Sons, New York, New York.
- Hald, T., Vose, D., Wegener, H. & Koupeev, T. (2004). A bayesian approach to quantify the contribution of animal-food sources to human salmonellosis, *Risk Analysis* 24: 255–269.
- Harrigan, W. (1998). *Laboratory Methods in Food Microbiology*, Academic Press, San Diego, California.
- Helsel, D. (2005). *Nondetects and Data Analysis: Statistics for Censored Environmental Data*, John Wiley & Sons., Hoboken, New Jersey.
- Herikstad, H., Motarjemi, Y. & Tauxe, R. (2002). Salmonella surveillance: A global survey of public health serotyping, *Epidemiology and Infection* 129(1): 1–8.
- Knaus, J., Porzelius, C., Binder, H. & Schwarzer, G. (2009). Easier parallel computing in r with snowfall and scluster, *The R Journal* 1: 54–59.
- Limpert, E., Stahel, W. & Abbt, M. (2001). Log-normal distributions across the sciences: Keys and clues, *Bioscience* 51: 341–352.
- Lunn, D., Spiegelhalter, D., Thomas, A. & Best, N. (2009). The bugs project: Evolution, critique and future directions (with discussion), *Statistics in Medicine* 28: 3049–3082.
- Lunn, D., Thomas, A., Best, N. & Spiegelhalter, D. (2000). Winbugs - a bayesian modelling framework: Concepts, structure, and extensibility, *Statistics and Computing* 10(4): 325–337.
- Oliver, J. (2005). The viable but nonculturable state in bacteria, *Journal of Microbiology* 43: 93–100.
- Oliver, J., Dagher, M. & Linden (2005). Induction of *escherichia coli* and *salmonella typhimurium* into the viable but nonculturable state following chlorination of wastewater, *Journal of Water Health* 3: 249–257.
- Parsonsa, D., Ortona, T., D’Souza, J., Mooreb, A., Jonesc, R. & Dodd, C. (2005). A comparison of three modelling approaches for quantitative risk assessment using the case study of *salmonella* spp. in poultry meat, *International Journal of Food Microbiology* 98: 35–51.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL: <http://www.R-project.org/>
- Raftery, A., Givens, G. & Zeh, J. (1995). Inference from a deterministic population dynamics model for bowhead whales, *Journal of the American Statistical Society* 90(430): 402–416.
- Rossini, A., Tierney, L. & Li, N. (2007). Simple parallel statistical computing in r, *Journal of Computational and Graphical Statistics* 16(2): 399–420.
- Rubin, D. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The sir algorithm, *Journal of the American Statistical Society* 82(398): 543–546.
- Scallan, E., Hoekstra, R., Angulo, F., Tauxe, R., Widdowson, M., Roy, S., Jones, J. & Griffin, P. (2011). Foodborne illness acquired in the united states major pathogens, *Emerging Infectious Diseases* 17(1): 7–15.
- Vose, D. (2008). *Risk Analysis: A Quantitative Guide*. 3rd edition, John Wiley & Sons, West Sussex, England.

- Williams, M. & Ebel, E. (2012). Estimating changes in public health following implementation of hazard analysis and critical control point in the united states broiler slaughter industry, *Foodborne Pathogens and Disease* 9: 59-67.
- Williams, M., Ebel, E. & Hoeting, J. (2011). Bayesian analysis for food-safety risk assessment: Evaluation of dose-response functions within winbugs, *Journal of Statistical Software* 43: 1-14.
- Williams, M., Ebel, E. & Vose, D. (2011a). Framework for microbial food-safety risk assessments amenable to bayesian modeling, *Risk Analysis* 31(4): 548-565.
- Williams, M., Ebel, E. & Vose, D. (2011b). Methodology for determining the appropriateness of a linear dose-response function, *Risk Analysis* 31(3): 345-350.



Novel Approaches and Their Applications in Risk Assessment

Edited by Dr. Yuzhou Luo

ISBN 978-953-51-0519-0

Hard cover, 344 pages

Publisher InTech

Published online 20, April, 2012

Published in print edition April, 2012

Risk assessment is a critical component in the evaluation and protection of natural or anthropogenic systems. Conventionally, risk assessment is involved with some essential steps such as the identification of problem, risk evaluation, and assessment review. Other novel approaches are also discussed in the book chapters. This book is compiled to communicate the latest information on risk assessment approaches and their effectiveness. Presented materials cover subjects from environmental quality to human health protection.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Michael S. Williams, Eric D. Ebel, Jennifer A. Hoeting and James L. Withee (2012). A Bayesian Approach for Calibrating Risk Assessment Models, *Novel Approaches and Their Applications in Risk Assessment*, Dr. Yuzhou Luo (Ed.), ISBN: 978-953-51-0519-0, InTech, Available from: <http://www.intechopen.com/books/novel-approaches-and-their-applications-in-risk-assessment/simplified-framework-for-microbial-food-safety-risk-assessment-methods-and-applications>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.