The Annals of Applied Statistics
2015, Vol. 9, No. 3, 1415–1432
DOI: 10.1214/15-AOAS845

A BAYESIAN APPROACH TO THE EVALUATION OF RISK-BASED MICROBIOLOGICAL CRITERIA FOR *CAMPYLOBACTER* IN BROILER MEAT

BY JUKKA RANTA*, ROLAND LINDQVIST†,‡, INGRID HANSSON§, PIRKKO TUOMINEN* AND MAARTEN NAUTA¶

Finnish Food Safety Authority Evira, National Food Agency†, Swedish University of Agricultural Sciences‡ National Veterinary Institute§ and Technical University of Denmark¶*

Shifting from traditional hazard-based food safety management toward risk-based management requires statistical methods for evaluating intermediate targets in food production, such as microbiological criteria (MC), in terms of their effects on human risk of illness. A fully risk-based evaluation of MC involves several uncertainties that are related to both the underlying Quantitative Microbiological Risk Assessment (QMRA) model and the production-specific sample data on the prevalence and concentrations of microbes in production batches. We used Bayesian modeling for statistical inference and evidence synthesis of two sample data sets. Thus, parameter uncertainty was represented by a joint posterior distribution, which we then used to predict the risk and to evaluate the criteria for acceptance of production batches. We also applied the Bayesian model to compare alternative criteria, accounting for the statistical uncertainty of parameters, conditional on the data sets. Comparison of the posterior mean relative risk, $E(RR|\text{data}) = E(P(\text{illness}|\text{criterion is met})/P(\text{illness})|\text{data})$, and relative posterior risk, $RPR = P(\text{illness}|\text{data, criterion is met})/P(\text{illness}|\text{data})$, showed very similar results, but computing is more efficient for RPR. Based on the sample data, together with the QMRA model, one could achieve a relative risk of 0.4 by insisting that the default criterion be fulfilled for acceptance of each batch.

1. Introduction. Campylobacteriosis is the most commonly reported bacterial enteric disease in humans in many industrial countries [EFSA (2013)]. One risk factor for human campylobacteriosis is handling and consuming contaminated poultry meat [Kapperud et al. (2003); Wingstrand et al. (2006)]. During slaughter, broiler carcasses can become contaminated with *Campylobacter*, and this contamination in the slaughter batch can originate from the intestinal contents or from the environment [Rosenquist et al. (2003, 2006); Lindqvist and Lindblad (2008); Nauta et al. (2009)].

Received March 2014; revised May 2015.
Key words and phrases. Bayesian modeling, hierarchical models, evidence synthesis, uncertainty, OpenBUGS, 2D Monte Carlo, quantitative microbiological risk assessment, food safety, *Campylobacter*.

Quantitative Microbiological Risk Assessments (QMRA) of food-borne pathogens in a production chain aim to provide numerical estimates of consumer risk. They involve modeling several steps in the production chain or a complete farm-to-fork chain starting from the primary production models and ending with dose-response models [Nauta et al. (2009)]. At best, the available data can only partially cover the steps. A common approach uses detailed simulation models that aim to provide a realistic representation of the mechanistic nature of various processes known to influence the survival of pathogens and their transmission, leading to possible consumer exposure. Other, parsimonious statistical models aim to bridge the data gaps by using the observed data to provide estimates of the overall effects without making too many assumptions. Detailed models inevitably involve many parameters. Because of limited data, however, the resulting estimates can have considerable uncertainty, which is often suppressed, and can lead to predictions whose uncertainties are more assumption-driven than data-driven. Although assumptions can provide insight for possible or hypothetical what-if scenarios, data can be more fully exploited through formal statistical inference, which in turn can provide estimates with uncertainty bounds based on the empirical evidence.

Since the Bayesian interpretation of probability is necessarily conditional and depends on the available evidence (prior + empirical data), it provides a logical way to assess multidimensional parameter uncertainty that can be explicitly updated by new data [Gelman and Hill (2007); Spiegelhalter and Best (2003)], specifically in the context of microbiological risk assessments [Albert et al. (2008)] addressing bacterial growth [Spor et al. (2010)] or management interventions for better food safety [Ranta and Maijala (2002); Ranta et al. (2010, 2013)]. Even so, some parameters may be so inherent in the problem that they need to be included regardless of whether sufficient, or any, data exist. Hence, they become part of the model uncertainty, which can be considered one of the several levels of uncertainty [Spiegelhalter and Riesch (2011)].

We present a Bayesian method for evaluating and comparing the effects of microbiological criteria (MC) in broiler production on consumer risk. Microbiological criteria have been recognized as practical measures for defining the level of acceptability in food product testing for decades, with the earliest versions already available in the 1960s [Codex Alimentarius (1997); NRC (1985)]. Our evaluation of MC is based on uncertainty analysis concerning model parameters, for which national sample data are available on broiler carcasses. We analyze two types of such data combined in tandem with the subsequent QMRA model of consumer risk. The computations were implemented using freely available OpenBUGS software in the R environment [Lunn et al. (2013)], (http://www.openbugs.net/), and the model code is available in the supplementary material (Section A.5) [Ranta et al. (2015)].

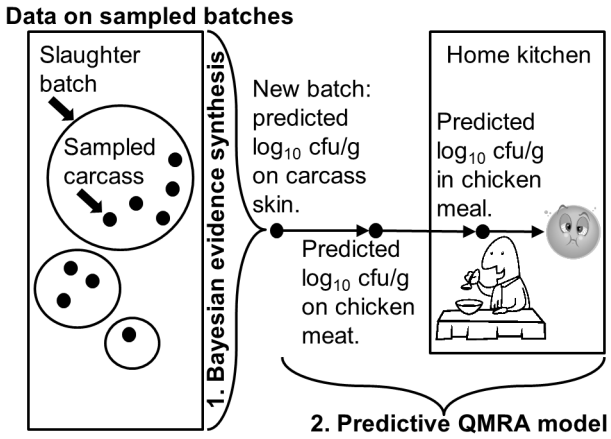

FIG. 1. *Simplified diagram of the food chain from slaughter batches to consumption, covering the steps modeled. Data from slaughter batches are used for Bayesian evidence synthesis* (1). *The resulting posterior distribution of underlying parameters is then used as an input distribution in the existing QMRA model* (2) *for predicting the number of colony forming units of bacteria per gram* (*cfu/g*) *of meat and the consequent risk of illness. Another predictive QMRA model could be used instead of the one adopted here.*

2. Methods and data.

2.1. *General model structure.* Figure 1 shows the simplified food chain from slaughter batches to consumption, covering the main steps modeled. In the first part, we conduct Bayesian evidence synthesis of two data sets on sampled broiler carcasses after slaughter. Batch structures and microbial levels have also been studied using Bayesian models in Commeau et al. (2012). In the sample data sets here, a batch originates from one broiler flock. Generally, batches in the model could represent a single broiler flock, part of a flock or a combination of flocks. Usually a batch contains more than 10,000 carcasses. If contamination occurs, prevalence within a batch is usually high. For contaminated carcasses, the concentration level varies from carcass to carcass, and the mean concentration varies between batches. Therefore, the hierarchical model involves parameters for both prevalence and concentration. Generally, we use q for prevalence of contaminated batches and p_j for within-batch prevalence in batch j. The concentration model is defined for contaminated carcasses only, so the average concentration (as $\log_{10}$) of all contaminated carcasses is μ, and the within-batch mean is μ_j. Concentrations of individual carcasses ($i = 1, 2, \ldots$) are modeled as $\mathrm{N}(y_{ij}|\mu_j, \sigma_w^{-2})$, where σ_w^2 is the within-batch variance component and $\sigma_w^{-2} = \tau_w$ is the corresponding precision parameter. Likewise, batch means are modeled as $\mathrm{N}(\mu_j|\mu, \sigma_b^{-2})$, where σ_b^2 is the between-batch variance component. The data for this model would ideally consist of measurements from several individual carcasses from a sample of batches. Such data are not always available, and we show an example of Bayesian

analysis with two qualitatively very different data sets. The underlying parameters are needed for predicting contamination of *a random new carcass from a random new batch*. A log-normal distribution is commonly used for microbial concentrations, and Q–Q-plots were used for checking approximate normality. Q–Q-plots of equally many random points were generated from the standard normal distribution to see that the Q–Q-plot of data falls reasonably within sampling error.

In the second part, an existing QMRA model describes the subsequent processing chain from carcasses to fresh meat, meal preparation, consumption and the probability of illness. This subsequent QMRA model contains a sequence of conditional distributions, which are taken as given in Nauta, Sanaa and Havelaar (2012); see the supplementary material Section A.2 [Ranta et al. (2015)]. We treat this QMRA model as a template for computing the risk, and the essential link to the first part is that the parameters describing carcass contamination are input parameters in the second part. These parameters are specific to each country, with uncertainties depending on national carcass sample data.

Finally, Bayesian posterior predictive distributions for a random serving from a random batch were used to study the effect of various microbiological criteria. The criteria define critical levels of contamination per batch so that batches can be rejected or accepted based on sample results. As a default scenario, the criterion is defined as "$n = 5, c = 1, m = 1000$," which means that at most one (c) sample out of five (n) is allowed to have $\log_{10}$ cfu/g > 3 ($m > 1000$ colony forming units per gram). Knowing whether a batch was accepted provides additional evidence, which has an effect on the posterior distribution of parameters for that batch and, consequently, on the predictions for servings stemming from the batch. Notation is given in Table 1.

2.2. *Two types of carcass sample data*. It is usually not possible to devise a sampling plan beforehand to serve the ideal data needs of a risk assessment model. Two Swedish data sets represent the types of historical data that could be commonly available. The first set is a one-year baseline study, conducted by Lindblad et al. (2006) between September 2002 and August 2003, that reports data on the prevalence and levels of thermophilic *Campylobacter* species in Swedish broiler chickens. Batches were sampled from ten slaughterhouses that represent 99.9% of the yearly production, and the number of samples per slaughterhouse was proportional to the annual production. One chilled carcass per batch was analyzed, and *Campylobacter* was quantified by direct plating in 88 out of 617 carcasses. These data have been used in another quantitative risk assessment by Lindqvist and Lindblad (2008). Bacteria concentrations as $\log_{10}$ values on positive whole carcasses represent bacteria concentration per carcass skin. According to Nauta, Sanaa and Havelaar (2012), in order to transform the values to $\log_{10}$ cfu per gram of skin, it was assumed that skin weight is approximately 100 g, so $\log_{10}(100)$ was subtracted from each measurement. The first data set, data1, also denoted by

TABLE 1
List of notation

j'	batch index in Lindblad et al. "1/batch" data
j''	batch index in Hansson et al. "$N_{j''}$/batch$^+$" data
j	generic batch index for prediction
J'	# positive carcasses in Lindblad et al. data.
N'	# sampled carcasses in Lindblad et al. data (= # sampled batches)
$x_{j''}$	# positive carcasses in j''th batch in Hansson et al. data
$N_{j''}$	# sampled carcasses in j''th batch in Hansson et al. data
q	prevalence of contaminated batches
$p_{j''}$	within-batch prevalence in batch j''
α	parameter for distribution of within-batch prevalence
μ	mean $\log_{10}$ cfu/g of all contaminated carcasses
μ_j	mean $\log_{10}$ cfu/g of contaminated carcasses in batch j; j' or j'' or generic
y_{ij}	$\log_{10}$ cfu/g of contaminated carcass i in batch j; j' or j'' or generic
I_j	true contamination status for a generic batch
σ_b^2	between-batch variance of μ_j's; either j' or j'' or generic
σ_w^2	within-batch variance of y_{ij}'s; either j' or j'' or generic
τ_b	between-batch precision σ_b^{-2}
τ_w	within-batch precision σ_w^{-2}
y_c	predicted $\log_{10}$ cfu/g for a contaminated carcass to be used for a serving
w	weight (g) of a broiler serving
n_c	bacteria count in a raw broiler serving of weight w
r	transfer probability for a bacteria cell from raw broiler meat to salad
d	bacteria count, the dose, in final serving
$P_0(\text{ill}\|d)$	probability of illness (dose response)
θ_j	batch-specific parameters (I_j, p_j, μ_j) in predictions
θ_s	serving-specific parameters (d, n_c, r, w, y_c) in predictions
L	# Monte Carlo draws of batch-specific parameters per MCMC iteration
M	# Monte Carlo draws of serving-specific pars. per batch per MCMC iteration
MC:$n/c/m$	microbiological criterion: n = sample size, c = max # positives exceeding m cfu/g

"1/batch" data, hence represents one sample for each of the $j' = 1, \ldots, N' = 617$ batches.

The second data set, presented by Hansson et al. (2010), describes a collection of carcass samples taken between July and October 2006, from 20 batches delivered by producers with a history in the Swedish *Campylobacter* surveillance program of often delivering *Campylobacter*-positive broilers. All batches were positive. The sample size per batch varied from 5 to 25, and the percentage of positives per sample varied from 85% to 100% with a mean of 98%. Sample means and sample standard deviations for positive carcasses were reported per batch. In these data, the measurements represent $\log_{10}$ cfu per ml of rinse water when 400 ml of water was used. According to Nauta, Sanaa and Havelaar (2012), these were transformed into values per carcass by adding $\log_{10}(400)$ to the original values, then

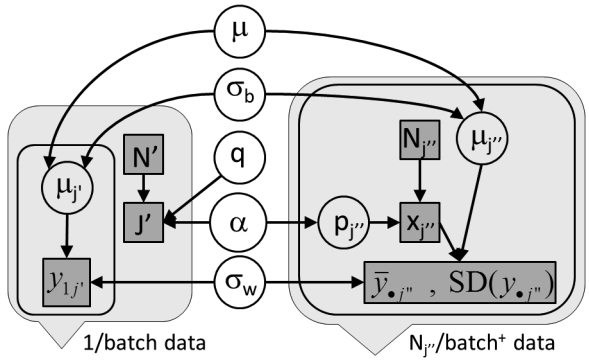

FIG. 2. *Directed acyclic graph of the final evidence synthesis model combining two qualitatively different data sets with common parameters in the middle. Incoming arrows denote conditional distributions as explained in the text.*

subtracting $\log_{10}(100)$, with the result that $\log_{10}(4)$ was added to each measurement. Transformation to cfu/g is necessary for compatibility with dose-response models. The second data set, data2, also denoted by "$N_{j''}$/batch^{+}" data, hence represents $N_1, \ldots, N_{J''}$ samples from $J'' = 20$ positive batches.

These two data sets provide complementary evidence. From the first sample we obtain some information about total variance and overall batch prevalence but nothing about within-batch prevalence. The second data set provides information on within-batch parameters for positive batches, but nothing on the overall batch prevalence. Therefore, evidence synthesis is needed.

2.3. *Evidence synthesis from carcass sample data.* A common challenge of QMRAs arises from the limited amount of data available. Typically, only one data set of the types presented here may be available. These limitations have a direct influence on the uncertainties, which can be quantified and presented as a posterior distribution of the parameters. Below, we present results based on both of the sample data sets taken separately and combined to illustrate this point. The full evidence synthesis model is shown as a directed acyclic graph (DAG) in Figure 2.

2.3.1. *Modeling with baseline data*: *One sample per batch* (*"*1*/batch" data*). The first data set provides a representative baseline sample for one year. To compute the posterior distribution of model parameters from these data alone, one has to decide what to assume concerning within-batch prevalence, for which there was no information at all. An expert elicitation might provide a prior distribution, but for the example, we use the background information that contaminated batches have a high within-batch prevalence, and we simply set $p_{j'} = 1$ for all such batches, $j' = 1, \ldots, J'$. From Bayes's theorem, the posterior distribution is

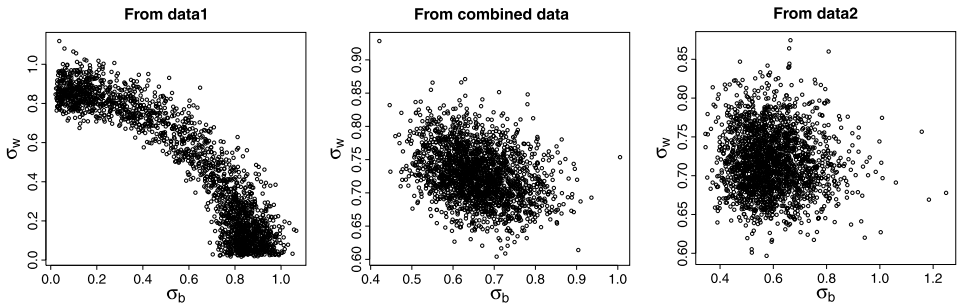

FIG. 3. *Marginal posterior distributions of* (σ_b, σ_w) *based on each data set alone* (1*/batch left*, $N_{j''}/\text{batch}^+$ *right*) *and the two data sets combined* (*middle*).

then

$$\begin{aligned}\pi(\mu, \tau_b, \tau_w, q, \mu_1, \ldots, \mu_{J'} | J', N', y_{11}, \ldots, y_{1J'}) \\ \propto \text{Binomial}(J' | N', q) \prod_{j'=1}^{J'} \text{N}(y_{1j'} | \mu_{j'}, \tau_w) \\ \times \text{N}(\mu_{j'} | \mu, \tau_b) \pi(\mu) \pi(\tau_b) \pi(\tau_w) \pi(q),\end{aligned} \tag{1}$$

where $N' = 617$, $J' = 88$, and $y_{1j'}$ are the adjusted $\log_{10}$ concentrations. Conventionally, the variance parameters were replaced with the precision parameters $\tau_b = \sigma_b^{-2}$, $\tau_w = \sigma_w^{-2}$. The posterior distribution can now be computed using the prior distributions $q \sim \text{U}(0, 1)$, $\mu \sim \text{N}(0, 10^{-4})$, $\tau_b \sim \text{Gamma}(0.001, 0.001)$ and $\tau_w \sim \text{Gamma}(0.001, 0.001)$. The Gamma priors resemble the uninformative priors $\pi(\tau_b) \propto \tau_b^{-1}$, $\pi(\tau_w) \propto \tau_w^{-1}$. The Gamma prior for between-batch precision can be problematic in some situations and should not be an automatic choice. Sensitivity to priors is discussed below. The normal prior for μ is practically flat, to reflect the lack of prior knowledge. The marginal posterior distribution of (σ_b, σ_w), based on these data, is shown at the left in Figure 3. Since "1/batch" data provide evidence on total variance, we can infer what combinations of values for variance parameters are more probable than others. If these were the only available data, the consequent uncertainty would be described by this joint distribution, which would be further reflected in the predictive distributions of the log-cfu concentrations.

2.3.2. *Modeling based on more than one sample per batch, for positive batches only* ("$N_{j''}/\text{batch}^+$" *data*). With only "$N_{j''}/\text{batch}^+$" data on positive batches, there is no information about batch prevalence q. However, it is possible to estimate within-batch prevalence $p_{j''}$ for each batch, and also to construct a hierarchical model for $p_{j''}$ with hyperparameter α to describe variation in within-batch prevalences. Knowing that prevalence is usually high, we restrict ourselves to distributions peaked near 1 by choosing prior $p_{j''} \sim \text{Beta}(\alpha, 2)$. The reported sample

means and sample standard deviations of the $\log_{10}$ concentrations summarize the observed $x_{j''}$ positive carcasses per batch, so the posterior becomes

$$\pi\big(\mu, \mu_1, \ldots, \mu_{J''}, \tau_b, \tau_w, p_1, \ldots, p_{J''}, \alpha | \{x_{j''}, N_{j''}, \bar{y}_{\cdot j''}, \mathrm{SD}(y_{\cdot j''})\}_{j''=1,\ldots,J''}\big)$$

$$(2) \quad \propto \prod_{j''=1}^{J''} \mathrm{Binomial}(x_{j''}|N_{j''}, p_{j''})\pi\big(\bar{y}_{\cdot j''}, \mathrm{SD}(y_{\cdot j''})|\mu_{j''}, \tau_w, x_{j''}\big)\mathrm{N}(\mu_{j''}|\mu, \tau_b)$$
$$\times \mathrm{Beta}(p_{j''}|\alpha, 2)\pi(\mu)\pi(\tau_b)\pi(\tau_w)\pi(\alpha).$$

Implementing the posterior is less straightforward because of the term $\pi(\bar{y}_{\cdot j''}, \mathrm{SD}(y_{\cdot j''})|\mu_{j''}, \tau_w, x_{j''})$; see the supplementary material, Section A.1 [Ranta et al. (2015)]. The priors were the same as before, complemented by the flat prior $\pi(\alpha) = \mathrm{U}(0, 10^4)$. The marginal posterior distribution of (σ_b, σ_w) is shown at the right in Figure 3. Compared with "1/batch" data, the variance components are now better identified. However, to fully use the two data sets jointly, a Bayesian evidence synthesis is applied below.

2.3.3. *Modeling based on combined data and results*. The posterior distribution was constructed by combining the likelihood functions from the two data sets while keeping the same priors. However, the common parameters made it possible to include "1/batch" data without the restrictive assumption of 100% within-batch prevalence. The corresponding factor of the likelihood was then written as $\mathrm{Binomial}(J'|N', q\alpha/(\alpha+2))$, since $E(p_{j'}|\alpha) = \alpha/(\alpha+2)$. The marginal posterior distribution for (σ_b, σ_w) based on the combined data is shown in the middle panel of Figure 3. Both data sets were crucial for estimating the full set of parameters in Table 2.

2.3.4. *Sensitivity to priors*. With "$N_{j''}$/batch$^+$" data, the results are robust for the reasonably uninformative prior choices because the data are informative enough for the parameters. With "1/batch" data, the only sensitive choice is the prior for the variance components because the data are informative for the total variance. Uniform priors for the standard deviations σ_b and σ_w were tested, and they led to quite similar overall conclusions and point estimates. However, the bimodality of marginal distributions was more pronounced with the Gamma priors. If these data were the only data, then the uniform priors could be preferred for robustness. In the evidence synthesis of the two data sets, the choice of priors is less critical because the combined data are fairly informative for all parameters. Because all of the results and predictions were ultimately based on the evidence synthesis, the default priors above were considered sufficient. With more seriously limited data, the priors could have more effect. Simple uninformative, or improper, standard priors might not work as such then, a known pitfall for hierarchical models. The choice of prior becomes critical for the between-batch variance σ_b^2 if the number of batches is small and/or σ_b^2 is nearly zero [Gelman (2006)].

TABLE 2
Summary of the parameter estimates from evidence synthesis based on two data sets

Parameter	Mean	95% credible interval	
"1/batch" data:			
μ	2.1	1.9	2.3
q	0.14	0.12	0.17
σ_b	0.58	0.046	0.96
σ_w	0.48	0.034	0.94
$\phi = \sigma_w^2/(\sigma_w^2 + \sigma_b^2)$	0.44	0.0015	1.0
Combined data:			
μ	2.4	2.2	2.5
q	0.15	0.12	0.18
α	85	38	177
σ_b	0.66	0.52	0.82
σ_w	0.74	0.68	0.80
$\phi = \sigma_w^2/(\sigma_w^2 + \sigma_b^2)$	0.55	0.43	0.68
"$N_{j''}$/batch$^+$" data:			
μ	2.9	2.6	3.2
α	85	39	172
σ_b	0.60	0.42	0.86
σ_w	0.74	0.68	0.81
$\phi = \sigma_w^2/(\sigma_w^2 + \sigma_b^2)$	0.61	0.42	0.77

2.4. *Consumer risk and microbiological criteria* (*MC*).

2.4.1. *Batch-specific inference, given the MC status of the batch.* Next we focus on predicting risk resulting from servings from a generic new batch j. The predictions depend on batch-specific parameters: the batch contamination status I_j (binary, "0/1"), within-batch prevalence p_j, and batch mean $\log_{10}$ concentration μ_j in the contaminated carcasses. These parameters have conditional distributions, given the previous parameters: $I_j|q \sim \text{Bern}(q)$ and $p_j|\alpha \sim \text{Beta}(\alpha, 2)$ and $\mu_j|\mu, \tau_b \sim \text{N}(\mu, \tau_b)$. Therefore, for servings from a given batch the disease probability is $P(\text{ill}|\theta_j, \sigma_w) = I_j p_j P_0(\text{ill}|\mu_j, \sigma_w)$, conditional on batch parameters $\theta_j = (I_j, p_j, \mu_j)$ and σ_w, with P_0 computed from the given QMRA model (see the supplementary material, Section A.2 [Ranta et al. (2015)]).

Alongside the risk associated with a batch, as expressed by the batch-level parameters, we are also interested in the probability that the batch complies with a particular microbiological criterion [Nauta, Sanaa and Havelaar (2012)]. We have interpreted a stated criterion, defined by the triplet $n/c/m$, as a condition for accepting a batch. Hence, only batches where at most c out of n sampled carcasses exceed the contamination level of m cfu/g are accepted for consumption.

By taking the Bayesian approach, we treat the MC status as an observation that provides additional evidence for a batch. This knowledge subsequently updates

the posterior distribution of the parameters concerning such a batch. The posterior distribution of the risk is then computed conditionally based on one of the following: (1) the criterion is met (batch accepted), (2) the criterion is not met (batch rejected), and (3) the criterion is not applied or, equivalently, criterion status is not known. Based on the two data sets, batches are accepted with a high probability: $P(\text{MC is met}|\text{data1, data2}) = 0.95$. Knowing that the batch was accepted leads to lower risk estimates for that batch. Without knowing the MC status, the batch is contaminated with probability $P(I_j = 1|\text{data1, data2}) = 15\%$, but when the batch status is given, the probability becomes either $P(I_j = 1|\text{data1, data2, MC met}) = 10\%$ or $P(I_j = 1|\text{data1, data2, MC not met}) = 1$.

Also, the observed batch status affects the probability of concentrations on contaminated carcasses in the batch: $E(\mu_j|\text{data1, data2}) = 2.37$, but $E(\mu_j|\text{data1, data2, MC not met}) = 2.96$. For the case "MC not met," the batch is contaminated with certainty (no false positives allowed). When the MC status is unknown or the MC is met, it is possible that the batch is completely free of contaminated carcasses. For the case "MC met," it may be of interest to compute the posterior mean of the mean concentration μ_j depending also on this hidden variable: $E(\mu_j|\text{data1, data2, MC met}, I_j = 1) = 2.05$. Hence, for a compliant *but contaminated* batch, the mean concentration of contaminated carcasses is probably lower than for a similar batch with unknown MC status.

The posterior distribution of the batch parameters $\theta_j = (I_j, p_j, \mu_j)$, conditional on given underlying parameters $q, \mu, \sigma_w, \sigma_b, \alpha$ and the batch status "MC met" is $\pi(\theta_j|\text{MC met}, q, \mu, \sigma_w, \sigma_b, \alpha)$. By Bayes's theorem, it is proportional to

$$P(\text{MC met}|I_j, p_j, \mu_j, \sigma_w)\,\text{Bern}(I_j|q)\,\text{Beta}(p_j|\alpha, 2)\text{N}\big(\mu_j|\mu, \sigma_b^{-2}\big). \tag{3}$$

Because the underlying parameters are unknown, and because we had the evidence from the two data sets, the marginal posterior distribution of the batch parameters is $\pi(\theta_j|\text{MC met, data1, data2})$, which is proportional to

$$\begin{aligned}\int P(\text{MC met}|I_j, p_j, \mu_j, \sigma_w)\,\text{Bern}(I_j|q)\,\text{Beta}(p_j|\alpha, 2)\text{N}\big(\mu_j|\mu, \sigma_b^{-2}\big)\\ \times \pi(q, \mu, \sigma_w, \sigma_b, \alpha|\text{data1, data2})\,d(q, \mu, \sigma_w, \sigma_b, \alpha),\end{aligned} \tag{4}$$

where the integral is taken over the underlying set of parameters $(q, \mu, \sigma_w, \sigma_b, \alpha)$, representing, for example, a country. Because observing MC status changes the probability for the hidden batch status I_j and the log-cfu distribution, the added knowledge can also be expected to affect the batch-specific consumer risk estimates; see Figure 4. The posterior risk, given the MC status and the data, is a single number,

$$\begin{aligned}&P(\text{ill}|\text{MC met, data1, data2})\\ &\quad = \int_{\Theta_j} P(\text{ill}|\theta_j)\pi(\theta_j|\text{MC met, data1, data2})\,d\theta_j,\end{aligned} \tag{5}$$

resulting from integrating over all of the uncertain parameters, both aleatoric and epistemic.

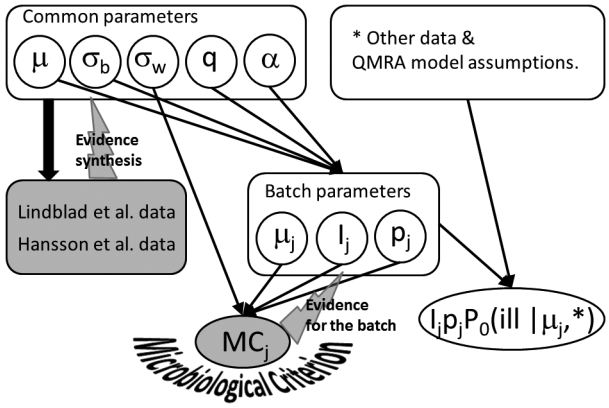

FIG. 4. *Directed acyclic graph of the Bayesian model combining the two carcass sample data sets for common parameters and predicting the batch parameters conditionally, now also based on the status of the batch* (MC_j = *accepted/rejected*). *The batch-specific illness probability* (*via QMRA model*) *then depends both on the carcass sample data and on the batch-specific status, for a generic new batch* j *to be predicted.*

3. Comparisons of microbiological criteria using measures of relative risk.

3.1. *Measures of relative risk.* Microbiological criteria were further studied by changing the values for n, c and m in the criterion, as shown in Table 3. In comparisons, it is of interest to know the risk level of the accepted batches relative to a situation where no criterion would be applied. For this purpose, we defined the relative posterior risk (*RPR*) as the ratio of the two batch-specific posterior probabilities:

$$RPR = \frac{P(\text{ill}|\text{MC met})}{P(\text{ill}|\text{MC not applied})}. \tag{6}$$

TABLE 3

Posterior means of $RR(q, \mu, \sigma_w, \sigma_b, \alpha)$, *with various MC* $(n/c/m)$. *Default in bold.* $RR(q, \mu, \sigma_w, \sigma_b, \alpha)$ *is risk conditional on acceptance of the batch* (*MC is met*) *divided by the risk conditional on "MC not applied." Corresponding means of rejection percentages are shown in subscripts. The same results were obtained for RPR*

	$m = 1000$		$m = 100$	
	$n = 5$	$n = 10$	$n = 5$	$n = 10$
$c = 0$	$0.20_{9\%}$	$0.10_{11\%}$	$0.01_{14\%}$	$0.00_{14\%}$
$\mathbf{c = 1}$	$\mathbf{0.42_{5\%}}$	$0.22_{8\%}$	$0.06_{12\%}$	$0.01_{14\%}$
$c = 2$	$0.61_{3\%}$	$0.34_{6\%}$	$0.14_{10\%}$	$0.03_{13\%}$
$c = 3$	$0.78_{1\%}$	$0.45_{5\%}$	$0.30_{7\%}$	$0.05_{12\%}$
$c = 4$	$0.92_{0\%}$	$0.56_{3\%}$	$0.58_{4\%}$	$0.09_{11\%}$

The two probabilities to be compared are single numbers obtained by integrating over the parameter uncertainties. Alternatively, one could study a parametric expression for relative risk,

$$RR(q, \mu, \sigma_w, \sigma_b, \alpha) = \frac{P(\text{ill}|\text{MC met}, q, \mu, \sigma_w, \sigma_b, \alpha)}{P(\text{ill}|q, \mu, \sigma_w, \sigma_b, \alpha)}, \tag{7}$$

as a function of the underlying (country-level) parameters $q, \mu, \sigma_w, \sigma_b, \alpha$. To account for uncertain parameters, we would compute the posterior mean $E(RR(q, \mu, \sigma_w, \sigma_b, \alpha)|\text{data})$. In practice, this computation requires 2D Monte Carlo: when the underlying parameters $q, \mu, \sigma_w, \sigma_b, \alpha$ are sampled from their posterior distribution using MCMC methods, within each iteration step the batch parameters are Monte-Carlo integrated depending on the current values of the underlying parameters.

The parametric approach to relative risk is much more computer-intensive than *RPR*. Nauta, Sanaa and Havelaar (2012) introduced a related measure, minimum relative residual risk *MRRR*. In implementing it, they assumed 100% within-batch prevalence ($p = 1$ for all batches) for contaminated batches. In our notation, we obtain the expression

$$\begin{aligned} &MRRR(q, \mu, \sigma_w, \sigma_b, p) \\ &\quad = \frac{q \int_{-\infty}^{\infty} p P_0(\text{ill}|\mu_j, \sigma_w) P(\text{MC met}|p, \mu_j, \sigma_w) \pi(\mu_j|\mu, \sigma_b)\, d\mu_j}{P(\text{ill}|q, \mu, \sigma_w, \sigma_b, p)} \\ &\quad = \frac{\int_{-\infty}^{\infty} P(\text{ill}, \text{MC met}|q, \mu_j, \sigma_w, p) \pi(\mu_j|\mu, \sigma_b)\, d\mu_j}{P(\text{ill}|q, \mu, \sigma_w, \sigma_b, p)} \\ &\quad = \frac{P(\text{ill}, \text{MC met}|q, \mu, \sigma_w, \sigma_b, p)}{P(\text{ill}|q, \mu, \sigma_w, \sigma_b, p)}, \end{aligned} \tag{8}$$

where $P(\text{ill}, \text{MC met}|q, \mu, \sigma_w, \sigma_b, p)$ is, in fact, the following total probability of illness:

$$\begin{aligned} &P(\text{ill}|q, \mu, \sigma_w, \sigma_b, p, \text{intervention}) \\ &\quad = \underbrace{P(\text{ill}|\text{MC met}, q, \mu, \sigma_w, \sigma_b, p) P(\text{MC met}|q, \mu, \sigma_w, \sigma_b, p)}_{= P(\text{ill}, \text{MC met}|q, \mu, \sigma_w, \sigma_b, p)} \\ &\qquad + \underbrace{P(\text{ill}|\text{MC not met}, q, \mu, \sigma_w, \sigma_b, p)}_{:= 0, \text{due to intervention}} \\ &\qquad \times P(\text{MC not met}|q, \mu, \sigma_w, \sigma_b, p). \end{aligned} \tag{9}$$

This calculation assumes that contaminated batches are used, but only after treatment that eliminates contamination. *MRRR* evaluates the quotient between the *total probability* of illness with such intervention and the total probability of illness without an intervention.

As a point of comparison, *RR* evaluates the *conditional probability* of illness for batches where MC was met, divided by the probability of illness for batches where MC was not applied. Therefore, if the same underlying parameter values ("$\cdot$") are used to evaluate the expressions, $MRRR(\cdot) = RR(\cdot) \times P(\text{MC met}|\cdot)$, so $MRRR \leq RR$. In our example, these are nearly equal because $P(\text{MC met}|\cdot) \approx 1$ with the Swedish data. In earlier implementations of *MRRR*, the parameters $p, q, \mu, \sigma_w, \sigma_b$ either are fixed values (e.g., $p = 1$) or else result from the assigned independent uncertainty distributions for each parameter, but with *RR* and *RPR*, the parameters are drawn from their joint posterior distribution.

3.2. *Evaluating relative risks based on the posterior distribution.* To calculate the probability $P(\text{ill}|q, \mu, \sigma_w, \sigma_b, \alpha)$ in the denominator in equation (7) for *RR*, we can use the following integral:

$$\int_{\Theta_j} P(\text{ill}|\theta_j)\pi(\theta_j|q, \mu, \sigma_w, \sigma_b, \alpha)\, d\theta_j, \tag{10}$$

because illness is conditionally independent of $q, \mu, \sigma_w, \sigma_b, \alpha$, given the batch parameters $\theta_j = (I_j, p_j, \mu_j)$. The illness probability involves integrating the serving-specific parameters θ_s [which include number of bacteria from the broiler y_c, n_c, serving size w, cross-contamination (transfer) probability r in the salad making, and dose d]. The whole expression can be approximated (see the supplementary material, Section A.3 [Ranta et al. (2015)]) as

$$\approx q\frac{\alpha}{\alpha+2}\frac{1}{L}\sum_{l=1}^{L}\frac{1}{M}\sum_{m=1}^{M}P_0\big(\text{ill}|\theta_s^{(m,l)}\big), \tag{11}$$

where $\theta_s^{(m,l)}$ are Monte Carlo draws for the serving-specific parameters within batches, sampled with the current values of $q, \mu, \sigma_w, \sigma_b, \alpha$ at each MCMC iteration step, so that $\theta_j^{(l)}$ is sampled first, then $\theta_s^{(m,l)}$ depending on each $\theta_j^{(l)}$. L batches and M servings within each of the L batches are simulated.

Next, to calculate the probability, $P(\text{ill}|\text{MC met}, q, \mu, \sigma_w, \sigma_b, \alpha)$, in the numerator in equation (7) for the *RR*, we can write it as

$$\frac{P(\text{ill, MC met}|q, \mu, \sigma_w, \sigma_b, \alpha)}{P(\text{MC met}|q, \mu, \sigma_w, \sigma_b, \alpha)}. \tag{12}$$

The denominator in equation (12), $P(\text{MC met}|q, \mu, \sigma_w, \sigma_b, \alpha)$, can be approximated (see the supplementary material, Section A.3 [Ranta et al. (2015)]) as

$$\approx q\frac{1}{L}\sum_{l=1}^{L}1_{\{\text{MC met}\}}\big(p_j^{(l)}, \mu_j^{(l)}, \sigma_w\big) + (1-q), \tag{13}$$

where the batch parameters are L Monte Carlo draws from $\pi(\theta_j|q, \mu, \sigma_w, \sigma_b, \alpha)$ and $1_{\{\text{MC met}\}}$ is the indicator variable for whether the batch complies, so the aver-

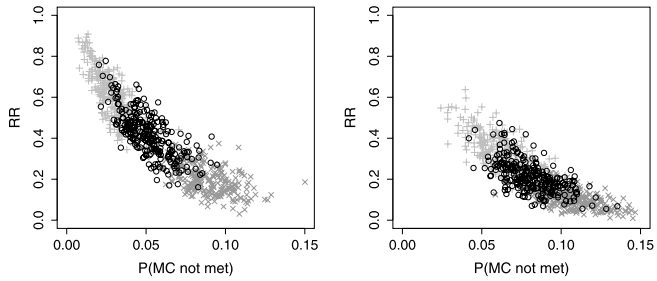

FIG. 5. *Thinned samples from the joint posterior distributions of* $(P(\text{MC not met}), RR)$, *both of which depend on the same unknown parameters*, $(q, \mu, \sigma_w, \sigma_b, \alpha)$, *for which the posterior distributions were computed. A single dot represents one MCMC draw of* $q, \mu, \sigma_w, \sigma_b, \alpha$ *used in evaluating* $P(\text{MC not met})$ *and RR. Three MC with* $m = 1000$ *cfu. Left*: $\mathbf{n} = \mathbf{5}$, $c = 0$ (“$\times$”), $c = 1$ (“$\circ$”), $c = 2$ (“$+$”). *Right*: $\mathbf{n} = \mathbf{10}$, $c = 0$ (“$\times$”), $c = 1$ (“$\circ$”), $c = 2$ (“$+$”). *Posterior means were* $(0.09, 0.20)$, $(0.05, 0.42)$, $(0.03, 0.61)$ *and* $(0.11, 0.10)$, $(0.08, 0.22)$, *and* $(0.06, 0.34)$, *respectively. MCMC run with* 16,000 *iterations*, *with* 40 *batches and* 10 *servings per batch at each iteration*.

age for the Monte Carlo sample is an approximation of the probability. The numerator in equation (12), $P(\text{ill, MC met}|q, \mu, \sigma_w, \sigma_b, \alpha)$, can be approximated (see the supplementary material, Section A.3 [Ranta et al. (2015)]) as

$$\approx q \frac{1}{L} \sum_{l=1}^{L} p_j^{(l)} \left[\frac{1}{M} \sum_{m=1}^{M} P_0\big(\text{ill}|\theta_s^{(m,l)}\big) \right] 1_{\{\text{MC met}\}}\big(p_j^{(l)}, \mu_j^{(l)}, \sigma_w\big). \tag{14}$$

With all of these Monte Carlo integrations, sampling batch parameters and serving parameters within batches, we compute the approximation of $RR(q, \mu, \sigma_w, \sigma_b, \alpha)$ within each step of the MCMC simulation, which, in turn, draws samples for the underlying parameters $(q, \mu, \sigma_w, \sigma_b, \alpha)$ from their posterior distribution.

It is more efficient to compute the ratio of two posterior probabilities, $RPR = P(\text{ill}|\text{MC met})/P(\text{ill}|\text{MC not applied})$, than the posterior distribution of RR or $MRRR$, which requires 2D Monte Carlo. However, an advantage of 2D Monte Carlo is that we then obtain credible intervals, for example, of $RR(q, \mu, \sigma_w, \sigma_b, \alpha)$, which describe the uncertainty in the underlying parameters. With the default MC $(n = 5, c = 1, m = 1000)$, the posterior mean was $E(RR(q, \mu, \sigma_w, \sigma_b, \alpha)|\text{data1, data2}) \approx 0.42$ and $RPR \approx 0.40$; see the supplementary material, Section A.4 [Ranta et al. (2015)]. In our example application, $E(MRRR|\text{data1, data2})$ was also quite similar (≈ 0.39). Ideally, the percentage of rejected batches (with “MC not met”) in the total production and the relative risk (RR) should both be low. The 2D-uncertainty plot for these is shown in Figure 5. The final result for a particular criterion, which accounts for all uncertainties, can be obtained by taking the overall posterior means.

4. Discussion. The uncertainties of the risk estimates emerge roughly for two qualitatively different reasons: (1) existing but partial or limited data from the specific production system and (2) external assumptions. The latter cannot easily be avoided when microbiological risk assessments aim to cover production chains and processes ranging from food production to consumption. Here we focused on uncertainties that can be quantified more explicitly based on production sample data. This analysis was illustrated with two data sets that contained partial but complementary evidence.

The posterior distribution of the core parameters was used to predict the consequent risk for consumers, so the uncertainties were propagated into the final risk estimates. However, this assessment is contingent upon the often unquantifiable uncertainties concerning the QMRA model for the remaining food pathway. Our approach can be combined with any available QMRA model. Parallel to the posterior predictive consumer risk, we also predicted the outcome of a batch-specific microbiological criterion, as defined by the triplet $n/c/m$. The parametric risk, which depends on the acceptance of batches, as determined by the MC, divided by the parametric risk without any MC at all, was described by the relative risk $RR(q, \mu, \sigma_w, \sigma_b, \alpha)$. When comparing MC in Figure 5, it is evident how much *RR* can be reduced by increasing the sample size n or by decreasing the number of positives c allowed to exceed the concentration m among them. At the same time, the expected percentage of noncompliant batches increases, which will increase costs if noncompliant batches are rejected or otherwise specially treated. It is then possible to optimize the criterion to achieve a higher risk reduction with tolerable costs. In this way, the chosen criterion would also be risk-based [Nauta, Sanaa and Havelaar (2012)]. By comparison, some MC are nearly equivalent. For example, both "$n = 5, c = 2, m = 1000$" and "$n = 10, c = 4, m = 1000$" have $RR \approx 0.6$ and an expected rejection percentage of 3%. Of course, the latter has a higher sampling cost because of double the number of samples. Further, *RPR* and $E(RR|\text{data1}, \text{data2})$ only describe relative effects. If the absolute risk level is already low, statistically significant reductions in relative risks might not be epidemiologically significant. The burden of disease, in number of cases, is more difficult to judge than *RR*'s, because of uncertainties along the carcass-to-serving path; see Ternhag et al. (2005).

Computationally, *RR* depends on the unknown parameters and involves 2D Monte Carlo. The simulation of batches and of servings within batches needs some optimization and can be slow to run. Instead, relative posterior risk *RPR* is much faster to compute and gives practically the same result. The same posterior illness probability can be obtained either by computing $E_d(P_0(\text{ill}|d))$ by Monte Carlo at each MCMC iteration and then taking the average over iterations or simply by computing $P_0(\text{ill}|d)$ with all the current parameters at each

MCMC iteration and then averaging over iterations. Eventually, *all* unknown parameters will become integrated when computing the final posterior probability. The order of integration does not matter, and it is more efficient not to run the 2D Monte Carlo within the MCMC. Code for the models with sample data is freely available, so that risk assessors can compute results that explicitly depend on their own data. Since the model code contains two qualitatively different examples, with individual and summary data, it could easily accommodate typical situations.

Acknowledgments. We thank the Nordic Working Group for Microbiology & Animal Health and Welfare (NMDD) www.norden.org/en/ for funding, and the editors for very careful and detailed comments that greatly improved the manuscript.

SUPPLEMENTARY MATERIAL

Appendix: A Bayesian approach to the evaluation of risk-based microbiological criteria for *Campylobacter* in broiler meat (DOI: 10.1214/15-AOAS845SUPP; .pdf). More details of computations and the BUGS codes are described in the supplementary materials.

REFERENCES

ALBERT, I., GRENIER, E., DENIS, J.-B. and 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.

CODEX ALIMENTARIUS COMMISSION COMMISSION, C. A. (1997). Principles and guidelines for the establishment and application of microbiological criteria related to foods. CAC/GL 21-1997.

COMMEAU, N., CORNU, M., ALBERT, I., DENIS, J.-B. and PARENT, E. (2012). Hierarchical Bayesian models to assess between- and within-batch variability of pathogen contamination in food. *Risk Anal*. **32** 395–415.

EUROPEAN FOOD SAFETY AUTHORITY, EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL (2013). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2011. *EFSA Journal* **10** 1–442.

GELMAN, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal*. **1** 515–533 (electronic). MR2221284

GELMAN, A. and HILL, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge Univ. Press, New York.

HANSSON, I., PUDAS, N., HARBOM, B. and OLSSON-ENGVALL, E. (2010). Within-flock variations of *Campylobacter* loads in caeca and on carcasses from broilers. *International Journal of Food Microbiology* **141** 51–55.

KAPPERUD, G., ESPELAND, G., WAHL, E., WALDE, A., HERIKSTAD, H., GUSTAVSEN, S., TVEIT, I., NATÅS, O., BEVANGER, L. and DIGRANES, A. (2003). Factors associated with increased and decreased risk of *Campylobacter* infection: A prospective case-control study in Norway. *Am. J. Epidemiol.* **158** 234–242.

LINDBLAD, M., LINDMARK, H., LAMBERTZ, S. T. and LINDQVIST, R. (2006). Microbiological baseline study of broiler chickens at Swedish slaughterhouses. *Journal of Food Protection* **69** 2875–2882.

LINDQVIST, R. and LINDBLAD, M. (2008). Quantitative risk assessment of thermophilic *Campylobacter spp.* and cross-contamination during handling of raw broiler chickens evaluating strategies at the producer level to reduce human campylobacteriosis in Sweden. *Int. J. Food Microbiol.* **121** 41–52.

LUNN, D., JACKSON, C., BEST, N., THOMAS, A. and SPIEGELHALTER, D. (2013). *The BUGS Book, a Practical Introduction to Bayesian Analysis*. CRC Press, Boca Raton, FL.

NATIONAL RESEARCH COUNCIL (1985). An evaluation of the role of microbiological criteria for foods and food ingredients. National Research Council, The National Academies Press, Washington, DC.

NAUTA, M. J., SANAA, M. and HAVELAAR, A. H. (2012). Risk based microbiological criteria for *Campylobacter* in broiler meat in the European Union. *Int. J. Food Microbiol.* **158** 209–217.

NAUTA, N. J., HILL, A., ROSENQUIST, H., BRYNESTAD, S., FETSCH, A., VAN DER LOGT, P., FAZIL, A., CHRISTENSEN, B., KATSMA, E., BORCK, B. and HAVELAAR, A. H. (2009). A comparison of risk assessments on *Campylobacter* in broiler meat. *International Journal of Food Microbiology* **129** 107–123.

RANTA, J. and MAIJALA, R. (2002). A probabilistic transmission model of *Salmonella* in the primary broiler production chain. *Risk Anal.* **22** 47–58.

RANTA, J., SIEKKINEN, K.-M., NUOTIO, L., LAUKKANEN, R., HELLSTRÖM, S., KORKEALA, H. and MAIJALA, R. (2010). Causal hidden variable model of pathogenic contamination from pig to pork. *Stat. Model.* **10** 69–87. MR2758057

RANTA, J., MIKKELÄ, A., TUOMINEN, P. and WAHLSTRÖM, H. (2013). Bayesian risk assessment for *Salmonella* in egg laying flocks under zero apparent prevalence and dynamic test sensitivity. *J. SFdS* **154** 8–30. MR3147063

RANTA, J., LINDQVIST, R., HANSSON, I., TUOMINEN, P. and NAUTA, M. (2015). Supplement to "A Bayesian approach to the evaluation of risk-based microbiological criteria for *Campylobacter* in broiler meat." DOI:10.1214/15-AOAS845SUPP.

ROSENQUIST, H., NIELSEN, N. L., SOMMER, H. M., NØRRUNG, B. and CHRISTENSEN, B. B. (2003). Quantitative risk assessment of human campylobacteriosis associated with thermophilic *Campylobacter* species in chickens. *International Journal of Food Microbiology* **83** 87–103.

ROSENQVIST, H., SOMMER, H. M., NIELSEN, N. L. and CHRISTENSEN, B. B. (2006). The effect of slaughter operations on the contamination of chicken carcasses with thermotolerant *Campylobacter*. *International Journal of Food Microbiology* **108** 226–232.

SPIEGELHALTER, D. J. and BEST, N. G. (2003). Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat. Med.* **22** 3687–3709.

SPIEGELHALTER, D. J. and RIESCH, H. (2011). Don't know, can't know: Embracing deeper uncertainties when analysing risks. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **369** 4730–4750. MR2853273

SPOR, A., DILLMANN, C., WANG, S., DE VIENNE, D., SICARD, D. and PARENT, E. (2010). Hierarchical Bayesian modelling for *Saccharomyces cerevisiae* population dynamics. *Int. J. Food Microbiol.* **142** 25–35.

TERNHAG, A., TÖRNER, A., SVENSSON, Å., GIESECKE, J. and EKDAHL, K. (2005). Mortality following *Campylobacter* infection: A registry-based linkage study. *BMC Infectious Diseases* **5** 70.

WINGSTRAND, A., NEIMANN, J., ENGBERG, J., NIELSEN, E. M., GERNER-SMIDT, P., WEGENER, H. C. and MØLBAK, K. (2006). Fresh chicken as main risk factor for campylobacteriosis, Denmark. *Emerging Infect. Dis.* **12** 280–285.

J. RANTA
RISK ASSESSMENT RESEARCH UNIT
FINNISH FOOD SAFETY AUTHORITY EVIRA
MUSTIALANKATU 3
FIN-00790 HELSINKI
FINLAND
E-MAIL: jukka.ranta@evira.fi

R. LINDQVIST
NATIONAL FOOD AGENCY
P.O. BOX 622
SE-75126 UPPSALA
SWEDEN
AND
SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES SLU
P.O. BOX 7070
SE-75007 UPPSALA
SWEDEN
E-MAIL: roland.lindqvist@slv.se

I. HANSSON
NATIONAL VETERINARY INSTITUTE SVA
SE-75189 UPPSALA
SWEDEN
E-MAIL: ingrid.hansson@sva.se

P. TUOMINEN
RISK ASSESSMENT RESEARCH UNIT
FINNISH FOOD SAFETY AUTHORITY EVIRA
MUSTIALANKATU 3
FIN-00790 HELSINKI
FINLAND
E-MAIL: pirkko.tuominen@evira.fi

M. NAUTA
NATIONAL FOOD INSTITUTE
TECHNICAL UNIVERSITY OF DENMARK DTU
MØRKHØJ BYGADE 19
DK-2860 SØBORG
DENMARK
E-MAIL: maana@food.dtu.dk