Journal de la Société Française de Statistique
Vol. 154 No. 3 (2013)

Bayesian risk assessment for Salmonella in egg laying flocks under zero apparent prevalence and dynamic test sensitivity

Titre: L'évaluation bayésienne des risques de salmonellose, pour des élevages de poules pondeuses de prévalence apparente nulle et un test de sensibilité évoluant autour du temps

Jukka Ranta[1], Antti Mikkelä[1], Pirkko Tuominen[1] and Helene Wahlström[2]

Abstract: A continuous time two-state hidden Markov process model was used to describe prevalence of salmonella infected flocks over laying phase in egg production. The infection status of a flock was treated as a binary hidden variable that can be detected as salmonella positive only by imperfect microbiological testing. Sensitivity of the test depends on the sampling type and analysis method used, but also on the unknown phase of epidemic among the hens within the flock. In a data set obtained from a national control programme under very low prevalence, all tests at all ages may show negative results. However, some temporally varying uncertainty remains about the unknown true prevalence, due to temporal changes in overall test sensitivity. By defining the sensitivity as a function of duration of within flock epidemic, Bayesian modeling was developed for quantitative risk assessment. Using minimal assumptions derived from expert knowledge or plausible scenarios, the effect of dynamically changing test sensitivity was accounted for by integration over the unknown time of infection. The sensitivity model was combined with the hidden Markov process model, conditional to temporal sequence of test results. Computations were performed using OpenBUGS.

Résumé : Un modèle à processus de Markov caché à deux états en temps continu a servi à décrire la prévalence de cheptels infectés par les salmonelles durant la phase de ponte dans le cadre de la production d'œufs. L'état infectieux d'un cheptel a été traité comme une variable cachée binaire susceptible d'être détectée comme étant positive aux salmonelles uniquement par des tests microbiologiques imparfaits. La sensibilité du test dépend du type d'échantillonnage et de la méthode d'analyse employés mais aussi de la phase inconnue de l'épidémie parmi les poules du cheptel. Dans un jeu de données issu d'un programme de contrôle national sous prévalence très basse, il est possible que tous les tests à tous les âges génèrent des résultats négatifs. Cependant, une certaine incertitude temporellement variable demeure en regard de la prévalence réelle inconnue, du fait des évolutions temporelles de la sensibilité d'ensemble des tests. En définissant la sensibilité comme une fonction de la durée de l'épidémie au sein du cheptel, un modèle Bayésien a été développé pour l'évaluation quantitative des risques. En employant des hypothèses minimales dérivées de connaissances expertes ou de scénarios plausibles, les effets de la sensibilité des tests changeante au cours du temps ont été pris en compte par intégration sur la durée inconnue de l'infection. Le modèle de sensibilité a été combiné avec le modèle de processus de Markov caché, conditionnellement á la séquence temporelle des résultats de tests. Les calculs ont été effectués avec OpenBUGS.

Keywords: hidden Markov process, sensitivity, detection, identifiability, uncertainty, Bayesian Risk Assessment, QMRA, OpenBUGS, laying hens, salmonella
Mots-clés : processus de Markov caché, sensibilité, détection, identifiabilité, incertitude, évaluation bayésienne des risques, QMRA, OpenBUGS, poule pondeuse, salmonella
AMS 2000 subject classifications: 62F15, 93B35, 60J28, 68T37

[1] Risk Assessment Unit, Finnish Food Safety Authority, Evira, Finland.
E-mail: jukka.ranta@evira.fi
[2] Swedish Zoonosis Center. The National Veterinary Institute, SVA, Sweden.

 ISSN: 2102-6238

1. Introduction

1.1. Risk assessment application

Salmonella control programmes aim to control risks over entire food production chains, to comply with the microbial criteria set and eventually to meet the Appropriate Level Of Protection, ALOP. On one hand, they are intended to act both directly and indirectly as interventions against the risk. On the other hand, they provide data from microbial testing results which can be informative about the current salmonella prevalence in food production. A positive test result means salmonella is detected, while a negative result means no detection. Under low prevalence situation the control samples are often all negative, showing no observable variation of results. However, the underlying true prevalence may not be zero, because all detection methods are imperfect. There are reported examples where *Salmonella* could not be detected in pooled faeces samples and mixed dust samples used in control programmes, nor in on-farm cloacal swabs in randomly selected (vaccinated) flocks. Yet, in the particular example, 26% of the flocks were found positive both in cloaca swabs and in the caeca, suggesting that the proportion of *Salmonella* infected laying flocks is underestimated based on the official monitoring programme (Van Hoorebeke S., Van Immerseel F., De Vylder J., Ducatelle R., Haesebrouck F., Pasmans F., de Kruif A., Dewulf J., 2009). This could be due to intermittently shedding hens and/or low within herd prevalence. Consequently, all this has an impact on the uncertainty of risk assessment and it can have implications on planning of control programmes. The detection methods may be improved in various ways, e.g. by developing more sensitive tests or by allocating the sampling differently. For egg production, this has been studied in simulation models (Klinkenberg D., Thomas E., Calvo Artavia F.F., Bouma A., 2011), based on parameter estimates from experimental settings (Thomas M.E., Klinkenberg D., Ejeta G., Van Knapen F., Bergwerff A.A., Stegeman J.A., Bouma A., 2009). Another experimental study was reported in De Vylder J., Dewulf J., Van Hoorebeke S., Pasmans F., Haesebrouck F., Ducatelle R., Van Immerseel F. (2011). However, the estimates are specific to the experiment and different experiments would be needed to reflect different actual conditions e.g. due to different housing systems. Ideally, relevant data should be collected from the real life production chain in question. Unfortunately, research driven sampling in real ongoing production systems can be very difficult to carry out in practice. Even if possible, it is obviously difficult to estimate transmission parameters from real flocks when the apparent prevalence is zero. Rare epidemics would provide valuable information, but often risk managers have already intervened before sampling for research purposes could be planned and executed. All these difficulties add up to the overall uncertainty in quantitative modeling when trying to produce estimates of true prevalence for real flock populations of laying hens, and consequent risk estimates for the egg production.

Results from surveillance programmes can be informative for estimation of some quantities along the production chain whereas for the estimation of others there may not be data at all. Auxiliary information about the structure of the production system may help to construct probabilistic dependency structures, and expert knowledge can be used to derive plausible estimates of some unknown parameters within the full model. Most probabilistic Quantitative Microbial Risk Assessment (QMRA) approaches employ 'direct' Monte Carlo simulation based on assigned distributions and parameter values, each based on a separate analysis that does not utilize other

 ISSN: 2102-6238

parts of evidence. Bayesian modeling offers a theory with well structured framework for evidence synthesis from diverse sources of data frequently occurring in quantitative food safety risk assessments (Albert I., Grenier E., Denis J.-B., Rousseau J., 2008; Ranta J., Siekkinen K.M., Nuotio L., Laukkanen R., Hellström S., Korkeala H., Maijala R., 2010; Albert I., Espié E., de Valk H., Denis J.-B., 2011). These models are always hierarchical and computationally more demanding than direct Monte Carlo. Monte Carlo models and Bayesian belief networks in QMRA were discussed in Smid J.H., Verloo D., Barker G.C., Havelaar A.H. (2010).

1.2. Model setting

Bayesian modeling was developed for the QMRA of salmonella in laying hens with the aim of introducing as few parameters as possible to reduce the need for overly many arbitrary assumptions and to facilitate more explicit estimation from the actual data presented from a control programme. In other words, the idea is to make inference from concrete data representing a particular production chain and to *encapsulate* the existing information rather than to *inflate* it by unknown parameters. Therefore, respecting the limited data and the unavoidable uncertainty, parsimonious parametrization was preferred over overly detailed computer simulations of e.g. within flock epidemics or contamination of individual eggs from individual hens, etc. Even if technically possible *in silico*, the inferential basis of detailed mechanistic simulations would be shallow or nonexistent, considering the actual and limited data on the observable process in question. The lack of data on more detailed processes poses limitations for formal analysis which cannot be overcome by increasing the complexity of the modeling. The uncertainties in a more detailed and complex model would be "*inevitably contingent on yet more assumptions that may turn out to be misguided*" as stated in Spiegelhalter D.J., Riesch H. (2011). The uncertainties in complex modeling, and the balance between assumptions and empirical data are increasingly discussed in many areas of science. In a NordForsk report 'Hot Topic - Cold Comfort' (Hernes G., 2012) the dilemma was summarized as: "*Some have argued that the models are too simplistic to mirror the real world. Others have argued the opposite - that they are too complex and nonlinear so that conclusions depend more on the tangled feedbacks in the models themselves rather than on the observational material about the interconnections in the real world.*"

Suggesting a compromise for the dilemma, Bayesian approach combines prior assumptions with data in a consistent framework. In probabilistic Bayesian inference with a parsimonious QMRA model, it is hoped that the results become more evidence driven than assumption driven compared to a more complex model, unless overly influential prior information is deliberately fed in where it is judged to be necessary. Also, the advantage of parsimonious Bayesian modeling is that results are reasonably easy to recalculate for each new data set to enable risk assessments on a regular basis, reflecting actual data obtained. This is called 'updating of probabilities', or 'probabilistic learning' based on new evidence. Nevertheless, even with parsimonious models in QMRA, some parameters remain more uncertain than others, and some of them could not be eliminated from the model without losing essential structures of the original problem. For those parameters, important expert knowledge can be elicited to formulate informative priors, or bounds for scenarios, if judged to be reliable and essential where data are lacking. Bayesian methods in some public policy and government settings have been discussed generally in Fienberg S.E.

 ISSN: 2102-6238

(2011). For sound representation of the uncertainty of all unknown parameters, it is important that the uncertainties are incorporated avoiding too naïve ad hoc point estimates drawn separately. The multidimensional uncertainty can be managed by using conditional distributions in a hierarchical structure modularly contingent on the body of evidence. In this paper we present one example where a model of test sensitivity was combined with the hidden Markov process model of the infection. This model is based on a generalization of an earlier Bayesian discrete time modeling for broiler production (Ranta J., Maijala R., 2002; Maijala R., Ranta J., Seuna E., Pelkonen S., Johansson T., 2005). An earlier simpler version of the continuous time model was applied in a quantitative risk assessment of laying hens by EFSA, European Food Safety Authority (2010). The complete assessment can be found in another EFSA report, EFSA Panel on Biological Hazards (BIOHAZ) (2010).

2. Materials

Comparable data on salmonella in egg production was based on Swedish and Finnish Salmonella Control Programmes, SSCP & FSCP, representing two similar production systems under similar control programmes. The number of flocks is of the same order, and it is rare to find salmonella positive samples. From this background information, we specified the number of flocks and the number of tests and their timing for the model. For the computational example we choose a (not uncommon) scenario where all the tests prove to be negative over a time period corresponding approximately to annual production. Such 'all negative results' provide an interesting example for modeling, when the underlying prevalence is estimated under dynamically changing overall sensitivity of the testing scheme. Expert opinion is needed to construct prior distributions or default values for the parameters determining the overall sensitivity.

There is no genuinely flock specific national (Finnish or Swedish) database in the sense that every testing result would be attached with a unique flock ID, enabling records of exact flock specific event histories to be exploited. The number of flock units is known, and it is assumed that one flock unit is occupied by a single flock at a time. The flock unit is the unit that is sampled and usually this is also the epidemiological unit where the within flock epidemic may take place. A holding may have one or more flock units. The national control programme specifies the ages at which tests are done and all flocks are tested according to this scheme. The egg laying period varies but can be from 18 weeks of age up to 70-75 weeks. For the model, each flock is assumed to be tested four times, at weeks 24,37,52 and 67 of age. The laying phase is assumed to start at week 21 and to end at week 70 of age. These values closely reflect the current practice, but can be easily changed as needed. A positive test result typically leads to destruction of the flock which is assumed to take place without delay.

Based on the sampling scheme and the number of flock units and the approximate life span, it can be approximately calculated how many tests and at what age they are performed during a calendar year. As a result, the model is presented with a specific number of flocks, $N = 900$, for which we assume all four testing results to be known at the given ages. Although not exact, and ignoring that samples from some flocks would fall into the previous or the next calendar year, this

 ISSN: 2102-6238

may be taken to represent data corresponding to annual production.

3. Modeling infection process of flocks

In the following, we define that a flock is infected by salmonella whenever salmonella still survives in the flock house or its hens which may be excreting salmonella at various rates. The model for infection and recovery of flocks is based on a well known continuous time Markov process with two states. These are usually denoted as infected and uninfected (susceptible) in infectious disease modeling, hence so called SIS-models. The process has two parameters: the infection intensity $\lambda \in \Re^+$ and recovery intensity $\mu \in \Re^+$. Given that a flock is free of salmonella, the waiting time of a new infection is then exponentially distributed with expected time $1/\lambda$. Similarly, if a flock is infected the expected waiting time for recovery is $1/\mu$. Additionally, test sensitivity p is defined as an uncertain model parameter, which will be later defined as a function of age. According to a two-state Markov process (Karlin S., Taylor H.M., 1975), the probability of being infected at time (age of the flock) t depends on the starting state, I_0, i.e. whether the flock was already infected in the beginning or not. This provides the following transition probabilities

$$\begin{array}{ll} \text{transition: } 0 \to 1 & p_{01} = P(I_t = 1 \mid I_0 = 0, \lambda, \mu) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}\exp(-(\lambda+\mu)t) \\ \text{transition: } 1 \to 1 & p_{11} = P(I_t = 1 \mid I_0 = 1, \lambda, \mu) = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\exp(-(\lambda+\mu)t). \end{array} \quad (1)$$

These probabilities represent the 'undisturbed' process of infections and recoveries of a flock up to arbitrary time (age) t. The remaining transition probabilities are determined as $p_{00} = 1 - p_{01}$ and $p_{10} = 1 - p_{11}$. If the state of the process can be observed without error, such set of observations (panel data) provide information about transition probabilities over specific time intervals and the likelihood function has a closed form solution (Cook R.J., 1999). Flocks may also be infected from the beginning of their laying period, and this is described by probability ν. Hence, the total probability of a flock being infected at time t is of the form:

$$P(I_t = 1 \mid \nu, \lambda, \mu) = (1-\nu)P(I_t = 1 \mid I_0 = 0, \lambda, \mu) + \nu P(I_t = 1 \mid I_0 = 1, \lambda, \mu), \quad (2)$$

assuming no interventions occur and that there are no observed testing results from the flock. Although the process mathematically allows infinite number of infections and recoveries in the long run, both intensities can be so low that practically the flock remains in its initial state over the limited lifetime, or there may be just one event, not many. In fact, we can expect most of the flocks to be uninfected and to remain so. Although infection is a rare event, a flock is likely to be persistently infected once it happens. In principle, also recovery is possible although its expected waiting time can exceed the lifetime of a flock.

Production of eggs is roughly proportional to the number of days spent in laying phase. Hence, the proportion of egg production under nonzero risk is proportional to the number of days the flock was infected while laying. This proportion of total production at risk (from a flock) is described by calculating the expected infection prevalence over its laying phase:

$$E(Q) = \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} P(I_t = 1 \mid \nu, \lambda, \mu)dt. \quad (3)$$

 ISSN: 2102-6238

where T_1 is the beginning of the laying phase, and T_2 is the end of it. The above expression is conditional to the given parameter values only, not assuming any observations yet. Assuming the parameters are the same for all flocks, so that there are no differences in rates of infection and recovery, $E(Q)$ can be interpreted to represent the proportion of total egg production at risk. Alternatively, if these would depend on flock specific variables such as flock size, then we might define a stratification of flock population, with distinct parameters in each stratum. However, a meaningful application would then require observations to be reported accordingly for each flock type. For simplicity of exposition, we do not assume any background variables of the flocks, such as flock size or housing type although the model could be extended in that direction. However, the results will usually be conditional to national data which implicitly implies stratification by countries. Therefore, $E(Q)$ can be used for comparisons between production systems in different countries. Once the uncertainties of the parameters are described as posterior distributions derived from production chain specific observations, we also obtain a posterior distribution of $E(Q)$. We will consider the role of observations next.

3.1. Modeling observed testing results and intervention

According to the Salmonella Control Programmes, all detected positive flocks are destroyed. The testing times during laying period are denoted t_k, where $k = 1,2,\ldots,K$ is the index of these times and: $T_1 = t_0 < t_1 < \ldots < t_K < t_{K+1} = T_2$. Hence, the probability of surviving up to a given testing time is needed for the likelihood function based on test results. The expression for $E(Q)$ should also account for this.

Since a detected positive flock is assumed to be destroyed, the observations consist of a series of negative results, possibly ending with a positive test result. In our example scenario, all results were negative. The probability of a test result depends on the hidden true infection status at that time, and the test sensitivity. In turn, the probability of the true infection status depends on the whole history of past observations. The probability can be written using a recursive formula given by Nagelkerke N.J., Chunge R.N., Kinoti S.N. (1990). Also, EM-algorithm could be exploited for likelihood inference as shown by Bureau A., Shiboski S., Hughes J.P. (2003). More general methods for Bayesian inference are discussed in Douc R., Garivier A., Moulines E., Olsson J. (2011). For a Bayesian approach, we use the recursive solution to evaluate the likelihood function as follows. Define $\rho_{t_k} = P(I_{t_k} = 1 \mid D_{t_1},\ldots,D_{t_k})$ for observed (binary) testing results D_t for a flock, then trivially $P(I_{t_k} = 1 \mid D_{t_1},\ldots,D_{t_{k-1}},D_{t_k} = 1) = 1$ but if $D_{t_k} = 0$, then

$$\rho_{t_k} = \begin{cases} \frac{(1-p)[(1-p_{01}-p_{10})\rho_{t_{k-1}}+p_{01}]}{1-p[(1-p_{01}-p_{10})\rho_{t_{k-1}}+p_{01}]} & k > 1 \\ \frac{(1-p)P(I_{t_1}=1)}{(1-p)P(I_{t_1}=1)+P(I_{t_1}=0)} & k = 1 \end{cases},$$

where the probabilities $P(I_{t_1} = 1)$ at the first testing time are given as in equation (2), and p_{ij} are defined as transition probabilities between two testing times, as given in equation (1). Let us denote the observation history $\{D_{t_1} = 0,\ldots,D_{t_k} = 0\}$ up to some testing time t_k as $H^0_{t_k}$, and the history at a time t_0 before the first observation as $H^0_{t_0} = \emptyset$. Now, the conditional probability of a positive test result at time t_k, given the past history of negative results, can be written as

$$P(D_{t_k} = 1 \mid H^0_{t_{k-1}}) = pP(I_{t_k} = 1 \mid H^0_{t_{k-1}}) = p(p_{01}(1-\rho_{t_{k-1}}) + p_{11}\rho_{t_{k-1}}). \quad (4)$$

 ISSN: 2102-6238

Note that the test sensitivity p could be replaced by a time dependent sensitivity p_{t_k}, to be modeled too. This detection probability provides the binomial likelihood, given that the number of tested flocks and test results at each testing time are specified. The proportion of total production at risk, the expected prevalence over laying phase, given the observations and assuming removal of detected positive flocks may be defined as

$$\begin{array}{lll} E(Q \mid H^0_{t_K}) = & & \\ \frac{1}{t_1-t_0}\int_{t_0}^{t_1} P(I_t = 1 \mid H^0_{t_0})dt & \times & P(I_{t_1} = 1 \mid H^0_{t_0})p \\ +\sum_{k=2}^{K}\left[\left[\frac{1}{t_k-t_0}\sum_{i=1}^{k}\int_{t_{i-1}}^{t_i} P(I_t = 1 \mid H^0_{t_{i-1}})dt\right] \right. & \times & \left.\left[\prod_{i=1}^{k-1}(1-P(I_{t_i} = 1 \mid H^0_{t_{i-1}})p)\right] P(I_{t_k} = 1 \mid H^0_{t_{k-1}})p\right] \\ \frac{1}{t_{K+1}-t_0}\sum_{i=1}^{K+1}\int_{t_{i-1}}^{t_i} P(I_t = 1 \mid H^0_{t_{i-1}})dt & \times & \left[\prod_{i=1}^{K}(1-P(I_{t_i} = 1 \mid H^0_{t_{i-1}})p)\right]. \end{array} \tag{5}$$

The expression makes exhaustive stratification of possible events into 'detection at 1st test', 'detection at 2nd test', etc., up to 'detection at the last test', and finally 'no detection'. The probabilities of these events are given by survival probabilities (of a flock) over previous tests multiplied by detection probability at a given testing time, respectively, and the survival probability over all tests for 'no detection'. The survival probabilities account for the effect of 'pruning' due to elimination of possible detected positive flocks. Each probability depends on the observed past testing results. Within this stratification of possible events, the integrals are calculated of $P(I_t = 1|H^0_{t-1})$ analogous to equation (3). It can be noted that if the sensitivity of testing (p) becomes zero, the equation yields the sum of $K+1$ integrals covering the entire interval $[T_1, T_2]$ (whole laying period), similar to $E(Q)$ in the simple case which did not account for observed testing results. $E(Q \mid H^0_{t_K})$ is simplified because it only describes the predicted proportion of production time under infection, i.e. proportion of time when the eggs are produced under risk. It does not take into account variation in the rate of laying contaminated eggs which may depend on finer state of the epidemic cycle within the flock. It also assumes that detected positive flocks are eliminated without delay. Note also that the expression is only predictive for this type of production process and the 'events' described here are not used for Bayesian inference of model parameters. $E(Q)$ is not likelihood function and the posterior inference relies on priors and the binomial likelihood based on detection probabilities given in equation (4). This also means that it is not necessary to evaluate $E(Q \mid H^0_{t_K})$ during MCMC in OpenBUGS. It may be more convenient to calculate it afterwards e.g. in R, from the already drawn MCMC sample of parameters.

3.2. Modeling overall test sensitivity

In the above formulations, test sensitivity was assumed constant over time (age). The detection sensitivity due to way of sampling (sock samples, pooled samples or other) and the associated laboratory methods may be assumed reasonably constant (p) and even fairly high in ideal situation. However, the overall sensitivity of detecting an infected flock is a more complicated function, mainly of the within flock infection process, but also of other possible factors. For comparison, in other applications, e.g. paratuberculosis in cattle (van Schaik G., Schukken Y.H., Crainiceany C., Muskens J., VanLeeuwen J.A., 2003), the sensitivity of testing has been reported to depend significantly on the clinical state of infection, thus compromising the overall sensitivity when

 ISSN: 2102-6238

the infection is in the early stage. Likewise, when detecting salmonella in cattle herds (Ranta J., Tuominen P., Maijala R., 2005), the herd level sensitivity can depend on a combination of events. For example, whether the infection is clinical or subclinical, within herd prevalence, possible pooling of a varying number of individual samples, in addition to the analytical sensitivity of the test. Bayesian estimation of flock-level sensitivity for laying-hen houses was described by Mahé A., Bougeard S., Huneau-Salaün A., Le Bouquin S., Petetin I., Rouxel S., Lalande F., Beloeil P.A., Rose N. (2008) and sensitivity of pooled vs individual samples have been compared (Arnold M.E., Carrique-Mas J.J., McLaren I., Davies R.H., 2011). Related examples are found in human epidemiology, where the diagnosis probabilities can be specific to disease stage and calendar time, e.g. in HIV diagnosis (Sweeting M.J., De Angelis D., Aalen O.O., 2005). At the risk of introducing unidentifiable parameters, modeling of all these effects that influence the overall sensitivity can be feasible, provided that sufficient information exist about each contributing factor. However, in the absence of gold standard, working only with sensitivity (Se), specificity (Sp) and prevalence θ_i for $i = 1,\ldots,n$ populations, vaguely informative ('objective') priors may still suffice, assuming that testing results are conditionally independent given the parameters and true infection status, and that properties of testing are constant while prevalence is variable between populations (Bonde M., Toft N., Thomsen P.T., Sørensen J.T., 2010; Branscum A.J., Gardner I.A., Johnson W.O., 2005).

In its elementary form, in a population with prevalence θ, the detection model with M individual samples tested, with X detected positive, is simply

$$X \sim \text{Bin}(M, \theta \times \text{Se})$$

when assuming Sp to be 100%. Using the observed data X and M, the goal would be to compute posterior distribution

$$\pi(\theta, \text{Se} \mid X, M) \propto \binom{M}{X} (\theta \times \text{Se})^X (1 - \theta \times \text{Se})^{M-X} \pi(\theta)\pi(\text{Se})$$

where the prior $\pi(\theta)$ is typically vaguely informative, whereas $\pi(\text{Se})$ aims to be informative about the sensitivity. If both are vague, the resulting distribution would show strong posterior dependency between θ and Se, due to non-identifiability. Also, we could not infer much about prevalence if we believe the sensitivity to be very low. Hence, it is crucial to the problem to specify what can be assumed of the sensitivity. In our application, we elaborate this by modeling an age dependent sensitivity, p_t.

3.3. Modeling age dependent sensitivity

3.3.1. Parameterizations of simple functions

After the onset of infection in a laying flock, there can be a short period during which the new infection might not be detectable due to low within flock prevalence. The sampling methods are designed for reliable detection when the within flock prevalence is at least 5 %. Based on expert

 ISSN: 2102-6238

opinions and other studies Thomas M.E., Klinkenberg D., Ejeta G., Van Knapen F., Bergwerff A.A., Stegeman J.A., Bouma A. (2009), this level can be reached in 2-4 weeks time. After this, the detection probability can peak, possibly to decline later again. The possible decline may be due to hens gradually recovering or acquiring immunity to salmonella, and particularly if shedding of salmonella is gradually decreasing or intermittent. The process is not too well known and only some expert opinions could be drawn, which are still difficult to elicit exactly in the form of a mathematical function. For an expert, it may also be difficult to separate the loss of sensitivity on one hand from the actual recovery from the infection on the other hand, since these are not directly observable events in practice. The observations inevitably depend on the testing method, giving the apparent status of the flock instead of the true status. Therefore, we choose to sketch simple scenarios for overall sensitivity, and then compute results under each scenario. Microbial detection problems have been discussed and further elaborated in an EFSA report on salmonella in laying hens EFSA Panel on Biological Hazards (BIOHAZ) (2010). In a very simplified model, detection probability would jump to p after some time (d_1) from the start of infection and drop to zero after some more time (d_2) from the onset of infection. Hence, expert knowledge would be required to set plausible values for d_1, d_2 and p. To specify the sensitivity at a given testing time, we consider a possible presence of infection at that time and how long it might have lasted. This depends on its starting time. Let τ_0 denote the unknown starting time of such *last* (yet undetected) infection before being in infected state, $I_t = 1$, at testing time t at which it may become detected. The simplest model of sensitivity is then:

$$P(D_t = 1 \mid \tau_0, I_t = 1) = p(d) = \begin{cases} 0 & \text{, if } d < d_1 \\ p & \text{, if } d_1 < d < d_2 \\ 0 & \text{, if } d > d_2 \end{cases} \tag{6}$$

where $d = t - \tau_0$. Taking into account the uncertainty of when the infection started τ_0, $(0 < \tau_0 < t)$ we calculate the integral

$$p_t = P(D_t = 1 \mid I_t = 1) = \int_0^t P(D_t = 1 \mid \tau_0, I_t = 1)\pi(\tau_0 \mid I_t = 1)d\tau_0$$

where we choose (see appendix)

$$\pi(\tau_0 \mid I_t = 1) = \frac{e^{-(t-\tau_0)\mu}}{(1 - e^{-\mu t})/\mu}.$$

The integration results to:

$$p_t = P(D_t = 1 \mid I_t = 1) = \frac{pe^{-\mu t}}{1 - e^{-\mu t}} \left(e^{\mu \max\{t-d_1,0\}} - e^{\mu \max\{t-d_2,0\}} \right) \tag{7}$$

which is a continuous smooth function although the simplistic $p(d)$ was not. Finally, taking into account the possibility (ν) of infection starting at time zero, we would obtain

$$p_t = P(D_t = 1 \mid I_t = 1) = (1 - \nu)\frac{pe^{-\mu t}}{1 - e^{-\mu t}} \left(e^{\mu \max\{t-d_1,0\}} - e^{\mu \max\{t-d_2,0\}} \right) + \nu p \mathbf{1}_{\{d_1 < t < d_2\}}(t)$$

 ISSN: 2102-6238

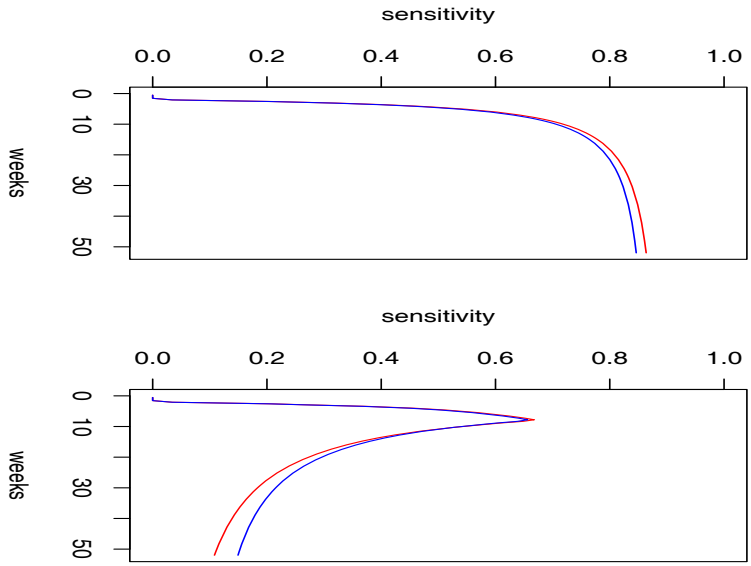

FIGURE 1. *Detection probability for an infected flock at different ages, assuming fixed values* $\nu = 0$, $p = 0.9$, $d_1 = 2$, *and* $\mu = 0.1$ *(red) and* $\mu = 1$ *(blue), and* $d_2 = \infty$ *(left),* $d_2 = 8$ *(right). At each time* t, *uncertain start of infection* τ_0 *is accounted for by integration over possible values* $\tau_0 \in [0,t]$.

This overall sensitivity is thus a function of age t, and it can be plotted, see Figure (1), with chosen values for p, μ, d_1, d_2, ν. The function is always zero when $t < d_1$ because then the duration of infection would have been too short for detection. After $t = d_1$ the detection probability rises until time $t = d_2$ after which it decreases due to increasing possibility that the duration may have lasted longer than d_2 in which event the detection probability would be zero again. If the initial infection probability ν is high, there is more pronounced stepwise change in the detection probability. In the extreme case where $\nu = 1$, the detection probability is p for $d_1 \leq t < d_2$ and zero elsewhere. In our application, initial infection probability is very low, so that practically the sensitivity can be calculated from the equation (7). This corresponds to the shapes of the example functions in Figure (1).

The above simple step function in equation (6) easily leads to cumbersome numerical evaluations in more complicated settings. A computationally more feasible alternative could be to define sensitivity as a smooth function of duration of infection. This can also be more appropriate assumption in the application. Three different simple functions could be proposed to represent different assumptions. One assumption is that the overall sensitivity will be monotonically increasing to its maximum after which it similarly decreases to zero, with the shape of a Gaussian function.

$$p(d) = pe^{-0.5(d-a)^2/\sigma^2} \quad \text{for } d = t - \tau_0 > 0.$$

This has a peak value of p at $d = a$, the duration at which maximum sensitivity is reached. Parameter σ determines the width of time window around optimal duration, within which detection

 ISSN: 2102-6238

is reasonably possible, $\approx a \pm 1.96\sigma$. If parameter σ is small, detection is possible only very near the optimal duration. With large σ, the detection is not drastically decreased even when testing much earlier or later relative to the optimal duration. Therefore, an expert opinion would be needed to specify a and σ. This model expresses the simple idea of a peaking sensitivity, assuming it to be symmetric in time with respect to the optimal duration. This may not be epidemiologically ideal assumption, though. An alternative assumption is that the sensitivity will only increase as a duration of infection. Hence the function would be:

$$p(d) = p(1 - e^{-ad}) \text{ for } d = t - \tau_0 > 0.$$

This requires only one parameter, which can be assigned by expert opinion by setting the duration d^* at which e.g. 95% of the maximum sensitivity is reached, solving $a = -\log(0.05)/d^*$. An extension to this assumption is that the sensitivity would ultimately decrease after sufficient duration. The function would then be:

$$p(d) = p(1 - e^{-ad})e^{-a\max(d-d_2,0)} \text{ for } d = t - \tau_0 > 0 \text{ and } a = -\log(0.05)/d_1.$$

This assumes that the slope of the decreasing part of the function is similar as in the increasing part. Hence, two parameters would be needed as expert opinion, d_1 and d_2. For example, with $d_1 = 4$ and $d_2 = 8$ we would assume that 95% of the maximum overall sensitivity is reached in 4 weeks after start of infection, and that the decline of sensitivity will begin after 8 weeks since start of infection.

With the Gaussian function, the sensitivity at the first testing time would be calculated as:

$$p_t = P(D_t = 1 \mid I_t = 1) = \int_0^t p e^{-0.5(t-\tau_0-a)^2/\sigma^2} \frac{e^{-(t-\tau_0)\mu}}{(1-e^{-\mu t})/\mu} d\tau_0$$

$$= \frac{p\mu}{1-e^{-\mu t}} e^{-0.5(a^2-(a-\mu\sigma^2)^2)/\sigma^2} \sqrt{2\pi}\sigma[\Phi((t-a+\mu\sigma^2)/\sigma) - \Phi((-a+\mu\sigma^2)/\sigma)]$$

Here, Φ is the cumulative distribution function of standard normal density. Similarly, we could solve the integration based on the other two functions. As a result, we obtain the sensitivity as a function of age only, and we have accounted for the uncertain time of infection, τ_0. However, things become more complicated after knowing that negative test results have been obtained at some earlier time points because these provide evidence for τ_0.

3.3.2. Accounting for previous negative tests

By time t, we also know the number of negative test results so far, and their timing as our observations. This provides evidence about when the infection probably has started, given that it has started. Consequently, this has an effect on what the sensitivity at time t probably is. The more we have negative test results in the history, the less plausible it becomes that the infection started before all these tests - it probably would have been caught then.

 ISSN: 2102-6238

When information from the previous negative tests is included, the solution of the conditional density $\pi(\tau_0 \mid I_t = 1)$ involves integration over a piecewise function. For a time point t before the first testing time t_1, $t \in [0, t_1]$, we obtain

$$\pi(\tau_0 \mid I_t = 1) = \frac{e^{-(t-\tau_0)\mu}}{(1-e^{-\mu t})/\mu}$$

and p_t is solved as before. For a time point between the first and the second testing time, $t \in [t_1, t_2]$, assuming the increase-decrease function, we obtain

$$\pi(\tau_0 \mid I_t = 1) \propto \begin{cases} e^{-(t-\tau_0)\mu}(1-p(1-e^{-a(t_1-\tau_0)})e^{-a\max(t_1-\tau_0-d_2,0)}) & \text{if } \tau_0 \in [0,t_1] \\ e^{-(t-\tau_0)\mu} & \text{if } \tau_0 \in [t_1,t] \end{cases}$$

where $a = -\log(0.05)/d_1$. Likewise, for a time point between the second and third testing time, $t \in [t_2, t_3]$, with the increase-decrease function, the density of τ_0 would be

$$\pi(\tau_0 \mid I_t = 1) \propto$$
$$\begin{cases} e^{-(t-\tau_0)\mu}(1-p(1-e^{-a(t_1-\tau_0)})e^{-a\max(t_1-\tau_0-d_2,0)})(1-p(1-e^{-a(t_2-\tau_0)})e^{-a\max(t_2-\tau_0-d_2,0)}) & \text{if } \tau_0 \in [0,t_1] \\ e^{-(t-\tau_0)\mu}(1-p(1-e^{-a(t_2-\tau_0)})e^{-a\max(t_2-\tau_0-d_2,0)}) & \text{if } \tau_0 \in [t_1,t_2] \\ e^{-(t-\tau_0)\mu} & \text{if } \tau_0 \in [t_2,t] \end{cases}$$

In any case, after computing the normalized density $\pi(\tau_0 \mid I_t = 1)$, a second integration is needed to obtain

$$p_t = \int_0^t p(1-e^{-a(t-\tau_0)})e^{-a\max(t-\tau_0-d_2,0)}\pi(\tau_0 \mid I_t = 1)d\tau_0.$$

For the computational task, there are basically two numerical possibilities: either to simulate the unknown τ_0 or to solve the integral by numerical integration routines. Notice that the conditional distribution of τ_0 was defined to depend only on μ and the negative test results in the past. The distribution is not a marginal distribution from the full posterior density which would involve all remaining parameters too. This is to separate the sensitivity model from the infection process model which leaves only one common parameter, μ, in both modules. We will discuss this in the section on posterior computations.

4. Computations

4.1. Prior distributions and assumptions from expert opinion

Expert opinion is needed mainly for the following parameters:

- p maximum sensitivity of testing in a infected flock
- d_1 duration of infection when the maximum test sensitivity is reached
- d_2 duration of infection when the test sensitivity starts declining, $d_2 \geq d_1$

These will determine the overall test sensitivity and how it changes over the age of flock. Informative prior for recovery rate μ is also needed because real life observations from test results are not

 ISSN: 2102-6238

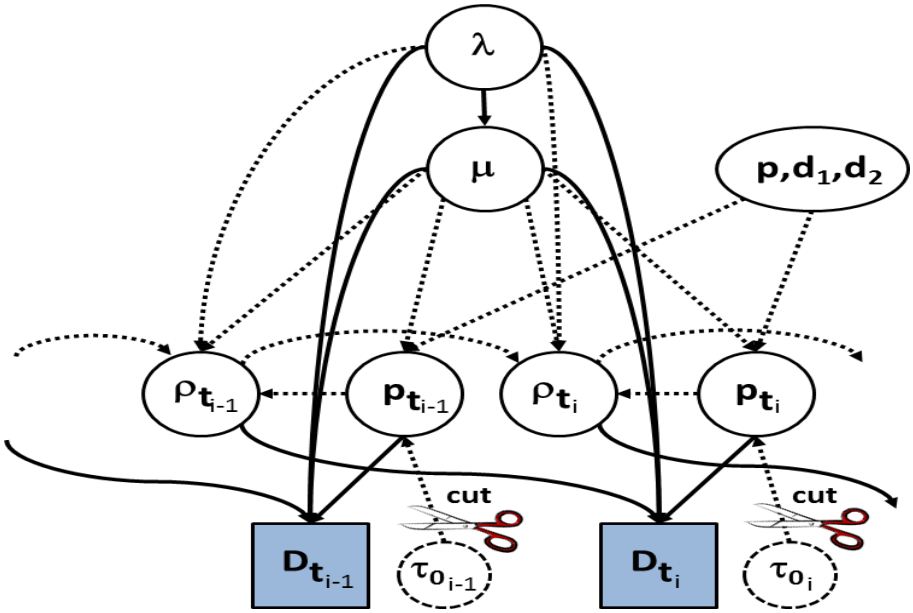

FIGURE 2. *DAG of the detection model for two consecutive testing times. Primitive model parameters together with probability of hidden infection* ρ_t*, detection sensitivity* p_t *and observed testing results* D_t *are shown. Dotted arrows denote deterministic dependency, solid arrows stochastic dependency.* τ_0 *is explicitly included only in the model where evidence feedback from the rest of the model is cut.*

available after a positive test due to elimination of detected positive flocks. Hence, the data rarely, if ever, contain information about how long time a natural recovery would take, but experts might provide some knowledge about the persistence of infection (Gradel K.O., Andersen J., Madsen M., 2002), or experimental results might be used. Generally, salmonella contamination in a flock, in hens and/or environment, can be persistent. It can survive long times e.g. in dust (Carrique-Mas J.J., Davies R.H., 2008). A quick recovery and clearance seems therefore very unlikely. Indeed, this is the reason for depopulating detected infected flocks. Without extensive investigations, pending on still limited data, it could be assumed that the expected time for recovery is longer than the expected time for infection. Therefore, the conditional prior would be $\mu \mid \lambda \sim \mathrm{U}(0,\lambda)$. Parameter p, the maximum sensitivity, can be given either as a fixed value or as a prior distribution. This should describe the plausible sensitivity under an ideal situation for detection. An assumption could well be that in ideal conditions, the sensitivity is 95% since the sampling is designed to detect an infected flock in such conditions. This has been phrased in alternative ways (EFSA Panel on Biological Hazards (BIOHAZ), 2010) which are thought to correspond to sampling 300 individual faeces, which would give $1-(1-1\%)^{300} > 0.95$, assuming 1% design prevalence. For example, 60 faecal droppings cultured as one pool would give $1-(1-5\%)^{60} > 0.95$, assuming pooling does not affect laboratory sensitivity. Parameter d_1 can be described by an informative prior $52(d_1-1) \sim \mathrm{Gamma}(6.8,(6.8-1)/2)$ with a minimum of 1 week, mode at 3 weeks, and upper 95th percentile at 5 weeks. This is roughly based on other modeling results in Thomas M.E., Klinkenberg D., Ejeta G., Van Knapen F., Bergwerff A.A., Stegeman J.A., Bouma A. (2009). Parameter d_2 is more difficult because it is hard to judge how long exactly it takes before the

 ISSN: 2102-6238

sensitivity begins to decline. We cannot use very vague priors because there is no information about this parameter in the observed data. Without assigning very arbitrarily informative prior distributions it may be best to assume alternative fixed values describing the 'worst case' and 'best case'. In the worst case, sensitivity starts declining after only three weeks after the maximum, $d_2 = d_1 + 3/52$, which could make detection difficult. In the best case, there is no decline, i.e. $d_2 = \infty$. For the remaining parameters, priors were chosen as vaguely informative. Knowing that the infections are rare to begin with, Jeffreys' prior was adopted for $\nu \sim \text{Beta}(0.5, 0.5)$ (Miconnet N., Cornu M., Beaufort A., Rosso L., Denis J.-B., 2005). For the infection intensity the prior was $\lambda \sim \text{U}(0, 10^3)$ with an arbitrarily large upper bound. A flat prior $\pi(\log(\lambda)) \propto 1$, i.e. $\pi(\lambda) \propto \lambda^{-1}$, was also tried but this lead to extremely small values of λ in the posterior sample, $< 10^{-22}$, and numerical errors in evaluations. This is due to the combined effect of a prior density increasing towards infinity at zero and the data showing no observed infections.

The effects of various prior assumptions are shown in Figure (3). Since we can expect μ to be very small, the values of d_1 and d_2 are most influential on p_t, which can be seen as a function $p_t(\mu, d_1, d_2)$ after averaging over τ_0.

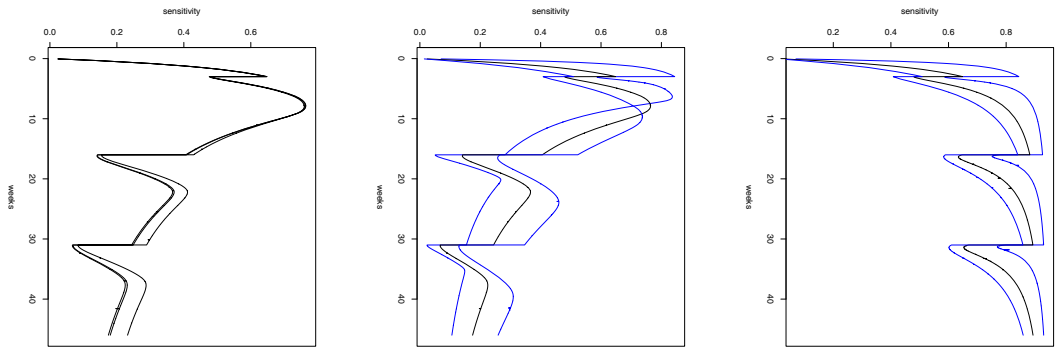

FIGURE 3. *Overall sensitivity. Left:* $d_1 = 3/52, d_2 = 6/52, p = 0.95, \mu \in \{0.00001, 0.1, 1\}$. *Center:* $d_1 \in \{1/52, 3/52, 5/52\}, d_2 = d_1 + 3/52, p = 0.95, \mu = 0.00001$. *Right:* $d_1 \in \{1/52, 3/52, 5/52\}, d_2 = \infty, p = 0.95, \mu = 0.00001$. *Testing times were at weeks 3, 16, 31 and 46 after start of laying. Figure produced by R-function, using the procedure* `integrate()`.

4.2. *Posterior computations: sensitivity model within the larger model*

For a given set of parameters, the sensitivity p_t needs to be computed as described above in the corresponding submodel, requiring integration over the unknown τ_0 which determines p_t, together with p, d_1 and d_2. Firstly, it should be noted that the proposed numerical integration involves a conditional density $\pi(\tau_0 \mid I_t = 1, H_{t^-}^0, \mu, p, d_1, d_2)$ which is not the same as the marginal posterior density of τ_0 from the full model. The density of τ_0 represents the information given in the *submodel* which is to be used nested within the larger model as a 'fixed' prior distribution of p_t, apart from the one common parameter, μ. Secondly, by doing this integration we effectively obtain the expected sensitivity, i.e. averaged over all possible values of τ_0 instead of a sensitivity

 ISSN: 2102-6238

that is conditional to a specific τ_0. Hence, we obtain marginalized model for p_t from the more detailed model by eliminating parameter τ_0. The marginalization corresponds to viewing any sensitivity as a conditional probability, where the conditions include true infection status $I_t = 1$ among other specified conditions. In principle, knowing *all* conditions in detail would lead to a sensitivity (detection probability) of either zero or one. In this case, the essential condition is τ_0, the onset time of the infection. Since τ_0 is an unknown parameter which nevertheless defines p_t in our model, we actually solve $E_{\tau_0}(p_t(\tau_0,\mu,d_1,d_2))$ with respect to the specific density of τ_0. However, we do not want integration over the full posterior density of τ_0 because that would constitute posterior dependency between τ_0 and the unknown true prevalence $P(I_t = 1)$, and consequently between τ_0 and λ and ν. This can be experimented in MCMC simulations where τ_0 is explicitly sampled among all other parameters from full posterior. The convergence becomes poor and results useless, likewise in the introductory example with flat priors on both Se and prevalence. Interestingly, when the MCMC simulations are done in BUGS using `cut`-function to cut evidence feedback from the remaining model to τ_0, (apart from μ), the results become similar to the above model. The cutting effectively removes likelihood contributions stemming from the child nodes when sampling the parameter in question. In particular, when simulating each τ_0 separately from its specific density $\texttt{cut}(\tau_{0_{t_1}}),\ldots,\texttt{cut}(\tau_{0_{t_4}})$ we obtain distributions of p_t which range from zero to p (95%), depending on the corresponding value of τ_0 at each iteration. The simulated average of p_t then corresponds to the $E_{\tau_0}(p_t)$ above. This exemplifies the issue of non-identifiability occurring if both sensitivity and true prevalence are unknown parameters in the same full model. To avoid the problem, we need conditionally independent information to model p_t, which was achieved by eliminating τ_0 by integration over the specific density in the nested submodel.

Computational burden of integration is that, firstly, we need to solve the normalizing constant of the density of τ_0 by numerical integration. Secondly, p_t is then obtained by yet another numerical integration, divided by this normalizing constant. Integration over defined densities can be done by using the procedure `integral()` provided in OpenBUGS. This procedure was only used for computing p_t as a function of parameters at each iteration step within the MCMC, while the parameters were simulated in OpenBUGS. In the results, we monitor the MCMC sample of the true infection prevalence $P(I_{t_k} \mid H^0_{t_{k-1}})$ and the overall sensitivity p_t at the given testing times. In order to choose a sufficient level of error tolerance for the numerical integration, bivariate plots were used, Figure (4). With too large error tolerance the mapping between p_t and d_1 becomes ragged and piecewise continuous, which should not be the case. The posterior distribution of all unknown parameters of interest is of the form:

$$P(\lambda,\mu,\nu,d_1 \mid \{D\},\{N\}) \propto \prod_{k=1}^{K} \text{Binomial}(D_{t_k} = 0 \mid p_{t_k} P(I_{t_k} = 1 \mid H^0_{t_{k-1}}), N_{t_k}) \times \text{Prior}(\lambda,\mu,\nu,d_1),$$

where $t_1,\ldots,t_K$ represent all testing times $(K = 4)$ during the laying period and N_{t_k} are the corresponding numbers of flocks under testing. For Finnish flocks, $N_{t_k} = 900$. Notice that the probability of infection $P(I_{t_k} \mid H^0_{t_{k-1}})$ is nevertheless a function of model parameters according to the recursive equation, and the posterior distribution of these parameters is computed conditional to whole data $D_{t_1},\ldots,D_{t_K}$, depicted in the Directed Acyclic Graph (DAG) in Figure (2). In this way, the data are informative not only on the current time point, but also retrospectively when we assess the probabilities after finally having all data. Alternatively, one might compute strictly stepwise

 ISSN: 2102-6238

by using only the data that have been revealed before and up to a 'current' time point. Likewise, $E(Q \mid H^0_{t_K})$ is a function of model parameters, but since it is not part of likelihood function or prior, one can evaluate it also afterwards from the obtained MCMC sample of parameters. This was done by using an R-function to evaluate the corresponding equation (5).

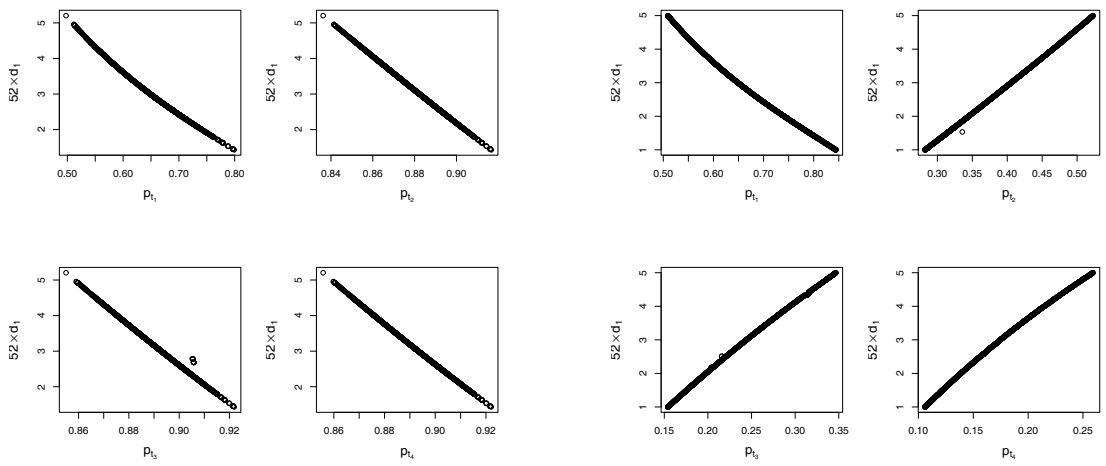

FIGURE 4. *Each value of d_1 corresponds to a nearly unique value of $p_{t_i}(\mu, d_1, d_2)$ $(i = 1,2,3,4)$ because μ is very small and d_2 is either ∞ or $d_1 + 3/52$. Clear outliers from the curve then indicate possible numerical errors due to the tolerance level set in numerical integration over τ_0. With a small tolerance level errors can be minimized, although computation time will increase. Left: $d_2 = \infty$. Right: $d_2 = d_1 + 3/52$. 50,000 iterations.*

5. Results

The estimates of true prevalence (in percentages %) at the four testing times are listed in Table (1) for two sensitivity models. Marginal posterior distributions of model parameters are shown in Figure (5). For comparison, some results from simulation of τ_0 and using `cut`-function are also shown. The first model assumes that sensitivity is an increasing function of duration of infection, whereas the second model assumes that it will also decrease after the infection has lasted over three weeks. The latter is a somewhat pessimistic assumption because it leads to very quickly decreasing sensitivity. These two alternative models could nevertheless be thought to provide the worst and the best case estimates of true prevalence. Although the dynamically changing sensitivity leads to different detection probabilities over time, the upper bound of posterior 95% intervals are in the range of 0.15% to 1.1% at all testing times, given that the test results were all negative in 900 flocks. There is logical connection between the estimates and the sensitivity function: when the sensitivity is clearly increasing, the estimates become smaller, i.e. the negative result is more likely to be truly negative. Likewise, when the sensitivity is decreasing, the estimates become higher, i.e. there is more uncertainty about whether the negative result indicates a truly negative status or not. From the posterior distribution of the sensitivity it may be concluded that although the ideal sensitivity was set to 95%, this is most likely not reached over the laying period. In the worst case, the sensitivity could be expected to be as low as 18%. On the other hand, the

 ISSN: 2102-6238

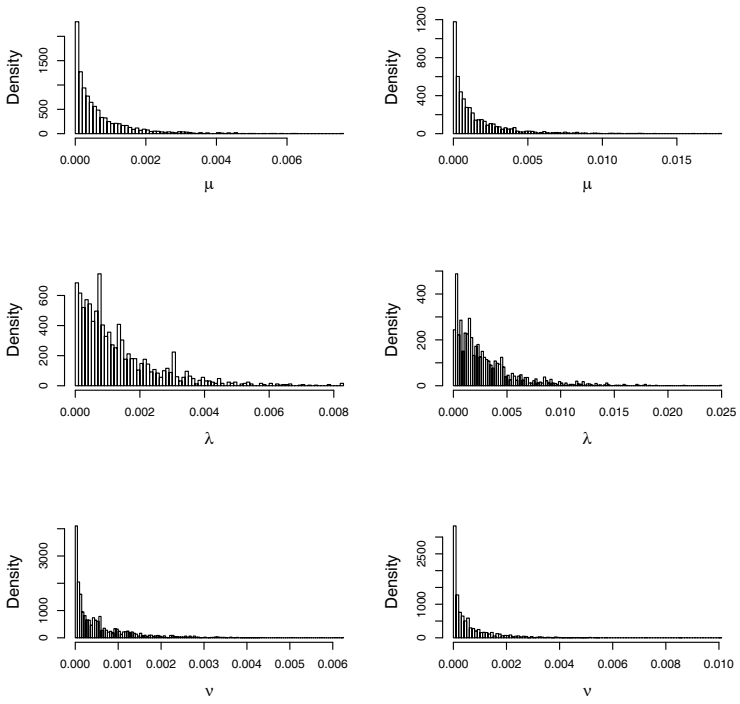

FIGURE 5. *Marginal posterior distributions of* μ, λ *and* ν. *Left:* $d_2 = \infty$. *Right:* $d_2 = d_1 + 3/52$.

estimates of true prevalence based on all data remained low even in the worst case.

Based on the MCMC sample of parameters, posterior distribution of the 'proportion of production under infection of flock', $E(Q \mid H^0_{t_K})$, was computed for the two scenarios, Figure (6). In the best case, the posterior 95% CI was [0.00%, 0.21%], but in the worst case it was [0.01%,1.08%]. The underlying assumption in both is then that any possible detected positive flocks would be removed from production at the time of their detection. This describes only the proportion of production that could be produced while the flocks are infected and still not removed. The resulting egg prevalence is expected to be much smaller because only a proportion of eggs produced by an infected flock are actually contaminated. Reported percentages of *Salmonella* Enteritidis positive eggs for such flocks have been below 1.8% (eggshell) and below 0.4% (egg content), Dewaele I., Van Meirhaeghe H., Rasschaert G., Vanrobaeys M., De Graef E., Herman L., Ducatelle R., Heyndrickx M., De Reu K. (2012). Also, egg contents have been reported to be clean in randomly selected battery cages (n=50) even at a *Salmonella* Enteritidis positive farm, with 92% of fecal samples positive, and 34% of eggshells positive (García C., Soriano J.M., Benítez V., Catalá-Gregori P., 2011). Any presence of salmonella in a flock may cause contamination of eggs, either eggshells or contents. For now, salmonella contamination of eggshell and egg content were not further described, although this would be the natural extension in a farm-to-fork continuum. Model codes are available from authors by request.

 ISSN: 2102-6238

TABLE 1. *Posterior estimates (%), conditional to full data, of the underlying true flock prevalence* $P(I_{t_k} = 1 \mid H^0_{t_k})$ *and sensitivity* p_{t_k} *at the testing times, and of* $E(Q \mid H^0_{t_K})$. *50,000 iterations (R/OpenBUGS).*

	mean	sd	2.5%	median	97.5%
Assumed sensitivity function: $p(d) = p(1-e^{-ad})e^{-a\max(d-d_2,0)}$ with $d_2 = \infty$					
Numerical integration of $E_{\tau_0}(p_{t_1}(\tau_0,\mu,d_1,d_2)),\ldots,E_{\tau_0}(p_{t_4}(\tau_0,\mu,d_1,d_2))$					
$P(I_{t_1} = 1 \mid H^0_{t_0})$	0.06	0.07	0.00	0.04	0.28
$P(I_{t_2} = 1 \mid H^0_{t_1})$	0.06	0.04	0.01	0.04	0.17
$P(I_{t_3} = 1 \mid H^0_{t_2})$	0.04	0.04	0.00	0.03	0.15
$P(I_{t_4} = 1 \mid H^0_{t_3})$	0.04	0.04	0.00	0.03	0.15
p_{t_1}	64	5.6	54	64	75
p_{t_2}	88	1.4	85	88	91
p_{t_3}	89	1.2	87	89	91
p_{t_4}	89	1.2	87	89	91
$E(Q \mid H^0_{t_K})$	0.06	0.06	0.00	0.04	0.21
Simulating $\texttt{cut}(\tau_{0_{t_1}}),\ldots,\texttt{cut}(\tau_{0_{t_4}})$					
$P(I_{t_1} = 1 \mid H^0_{t_0})$	0.06	0.08	0.00	0.04	0.28
$P(I_{t_2} = 1 \mid H^0_{t_1})$	0.06	0.05	0.00	0.04	0.19
$P(I_{t_3} = 1 \mid H^0_{t_2})$	0.04	0.04	0.00	0.03	0.16
$P(I_{t_4} = 1 \mid H^0_{t_3})$	0.04	0.04	0.00	0.03	0.16
p_{t_1}	62	25	6.2	70	92
p_{t_2}	87	17	25	95	95
p_{t_3}	88	16	29	95	95
p_{t_4}	89	16	29	95	95
$E(Q \mid H^0_{t_K})$	0.06	0.06	0.00	0.05	0.21
Assumed sensitivity function: $p(d) = p(1-e^{-ad})e^{-a\max(d-d_2,0)}$ with $d_2 = d_1 + 3/52$					
	mean	sd	2.5%	median	97.5%
Numerical integration of $E_{\tau_0}(p_{t_1}(\tau_0,\mu,d_1,d_2)),\ldots,E_{\tau_0}(p_{t_4}(\tau_0,\mu,d_1,d_2))$					
$P(I_{t_1} = 1 \mid H^0_{t_0})$	0.08	0.09	0.00	0.05	0.34
$P(I_{t_2} = 1 \mid H^0_{t_1})$	0.10	0.09	0.01	0.08	0.35
$P(I_{t_3} = 1 \mid H^0_{t_2})$	0.15	0.14	0.01	0.11	0.54
$P(I_{t_4} = 1 \mid H^0_{t_3})$	0.20	0.20	0.01	0.14	0.75
p_{t_1}	64	6.3	51	64	75
p_{t_2}	42	4.8	34	41	52
p_{t_3}	26	4.0	19	25	34
p_{t_4}	18	3.2	13	18	26
$E(Q \mid H^0_{t_K})$	0.28	0.28	0.01	0.19	1.08
Simulating $\texttt{cut}(\tau_{0_{t_1}}),\ldots,\texttt{cut}(\tau_{0_{t_4}})$					
$P(I_{t_1} = 1 \mid H^0_{t_0})$	0.10	0.12	0.01	0.06	0.42
$P(I_{t_2} = 1 \mid H^0_{t_1})$	0.13	0.13	0.01	0.10	0.48
$P(I_{t_3} = 1 \mid H^0_{t_2})$	0.19	0.21	0.01	0.13	0.76
$P(I_{t_4} = 1 \mid H^0_{t_3})$	0.25	0.29	0.01	0.16	1.08
p_{t_1}	63	25	6.2	71	92
p_{t_2}	42	39	0.0	31	95
p_{t_3}	26	37	0.0	0.6	95
p_{t_4}	19	33	0.0	0.0	94
$E(Q \mid H^0_{t_K})$	0.33	0.39	0.01	0.20	1.39

 ISSN: 2102-6238

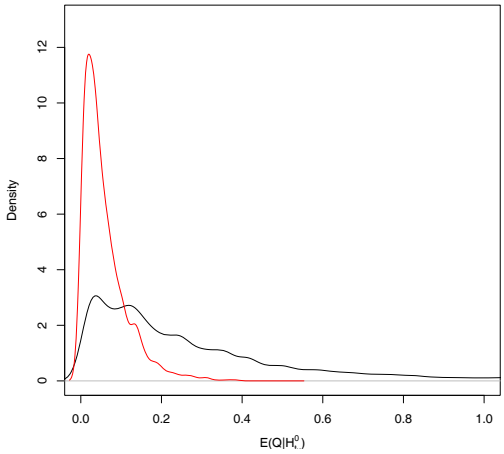

FIGURE 6. *Posterior distribution of* $E(Q \mid H_{t_K}^0)$*, the expected proportion of total production under infection of the flocks. Black density: assuming sensitivity would start decreasing after* $3+52*d_1$ *weeks of infection. 95% interval [0.01%,1.08%]. Red density: assuming duration of infection does not decrease sensitivity. 95% interval [0.00%, 0.21%]. 50,000 iterations.*

6. Discussion

The estimation of true flock prevalence under dynamic test sensitivity can potentially reveal age-dependent prevalence patterns in egg production chain under typical salmonella control programme. However, the probabilities of infection remain low, given that all test results were negative as they typically are in e.g. Swedish and Finnish production chains. With a parsimonious model, estimates of the true prevalence could be obtained by inserting the actual data and computing the posterior estimates based on that empirical evidence. Also, the model might be further used to study the optimal timing and number of the flock tests. The model utilizes only a small number of parameters, with minimal prior assumptions, aiming at a compact implementation in R/OpenBUGS. With robust but minimal external assumptions, data based results can be obtained individually for each production chain data, instead of forcing the same assumed default values and, consequently, similar simulation results for all populations. Hence, the probabilistic risk assessment can be made more data based process where probabilities are updated by new evidence. Sensitivity of the results to the most uncertain assumptions still needs to be explored. Since there are only a few of them in a parsimonious model, extensive cross tabulations of large number of assumptions is avoided. In this analysis, the most uncertain parameter concerns the duration of infection that would be required before the detection sensitivity would start declining. This can have considerable effects on results. Plausible parameter value for this remains difficult to obtain as expert opinion and it is impossible to estimate it from currently available data. Therefore, two scenarios were computed based on the best case and worst case assumptions.

Specification of the overall sensitivity of the surveillance method implemented in the control programme is a prerequisite for assessing the underlying true prevalence (hidden infection status of laying flocks). Albeit the optimal sensitivity may be assumed fairly high and well known in

 ISSN: 2102-6238

ideal conditions, the actual sensitivity over time is more complicated and depends on the within flock epidemic process. However, detailed within flock data are seldom, if ever, available and directly linked to the same test positive flocks in national data. Making fixed assumptions about the evolution of the within flock process, based on completely external sources of information would easily drive the outcome of the estimation, overrunning the actual data obtained from testing results. The national testing results currently only report whether the flock as a whole was detected positive or not. When the test sensitivity can be modeled, the underlying true flock status can be estimated from the hidden Markov process model, using the reported observations. An assumption was made that the sensitivity can be roughly determined from the duration of infection, which depends on the unknown starting time τ_0 of the infection. The observed testing results then provide also information about when the infection probably started, given that the flock would be infected at time t, $\tau_0 < t$. Hence, both the hidden infection status, and the overall sensitivity conditionally depend partially on the same parameters and the posterior distribution of parameters depends on the entire history of testing results. The probability distributions of parameters cannot be assessed completely separately and independently of each other. Therefore, the combined uncertainties and probabilistic inference of all model parameters are not sufficiently described by mechanistic models with 'forward simulation' or expert opinions, drawing independently from external sources of information. In our example, the sensitivity parameter was modeled nested within the larger model, to supply partially independent information for its submodel, to be used as a fixed prior in the larger model. Related examples of `cut`-function in OpenBUGS are e.g. from pharmacokinetic modeling (Lunn D., Best N., Spiegelhalter D., Graham G., Neuenschwander B., 2009), but also joint models may be used when information feedback in full model is considered advantageous and if poor identifiability of parameters is not a problem.

Every model contains some core assumptions which need to be acknowledged. The hidden Markov process here makes the assumption of time-homogeneous intensities. In other words, intensities λ and μ are assumed to be constant in time. If these were assumed to be some functions of time, describing e.g. varying infection pressure over the age of flocks, we would need strong additional assumptions about the nature of these functions. This would further complicate the analysis and effectively override the information in the available data. It can be hard to obtain meaningful results from the data unless one assumes either that the intensities are constant over time together with a sensitivity that is a function of time, or vise versa. More informative data might be obtained for further modeling concerning within flock prevalence at the times of default samples. Also, comparison of more effective microbial detection methods with the current method could be used to address the remaining uncertainty about possible percentage of false negatives and their significance for the consequent consumer risk. Nevertheless, extensions of the current example model would hardly be feasible without significant additional data.

7. Appendix A

Sensitivity of the testing method at time t depends on the status of the within flock epidemic process. This depends, at least, on the duration of the epidemic, determined by the time of onset τ_0. Therefore, we aim to quantify the sensitivity as a function of duration $d = t - \tau_0$, and finally, to account for the uncertain time of onset. First, the conditional density function of the onset time

 ISSN: 2102-6238

τ_0 of the still ongoing infection at t is derived as

$$\pi(\tau_0 \mid I_t) \propto \pi(I_t = 1 \mid \tau_0)\pi(\tau_0)$$

where the prior $\pi(\tau_0)$ is chosen as uniform. More exactly, the prior should be the density of the last onset time of the still ongoing infection, based on the two-state Markov process, conditional to its observed history. But, apparently, this does not lead to an easily computable function. In our case, both λ and μ are rather small because infections are rare, and recovery slow. The waiting times exceed the life time of a flock. Therefore, we can expect that there is only one infection time, if any. Conditional to $I_t = 1$ and the constant intensity λ, the density of the event time is U$(0,t)$, so that this can be very reasonable approximation. Given the onset time τ_0 of the ongoing infection, the probability to be still infected at time t is the 'survival probability'

$$\pi(I_t = 1 \mid \tau_0) = 1 - F(t - \tau_0 \mid \mu) = e^{-(t-\tau_0)\mu}.$$

Therefore, by normalizing we obtain:

$$\pi(\tau_0 \mid I_t = 1) = \frac{e^{-(t-\tau_0)\mu}}{\int_0^t e^{-(t-\tau_0)\mu}\mathbf{d}\tau_0} = \frac{e^{-(t-\tau_0)\mu}}{(1-e^{-\mu t})/\mu}.$$

Conditionally on $I_t = 1$, it is always more probable that the ongoing infection started recently than a longer time ago, but the steepness of the density depends on the recovery rate μ. For example, if μ is large, it is even less likely that the infection started long time ago, because the flock would have recovered already - if μ indeed was large.

8. Appendix B

TABLE 2. *List of notations. Time unit is one year, hence e.g. '3 weeks' is transformed to 3/52. Time zero is defined as the beginning of laying period.*

λ	intensity (rate) of new infections of flocks.
μ	intensity (rate) of recovery of infection of flocks.
ν	probability of being infected already in the beginning of laying period.
I_t	binary indicator of infection of a flock at age t.
D_t	number (or indicator for single flock) of positive testing results at testing age t.
N_t	number of flocks tested at age t.
τ_0	starting time of infection, given that it has started.
d	duration of infection.
p	maximum overall sensitivity of detecting an infected flock.
d_1	duration of infection at which the maximum sensitivity is reached.
d_2	duration of infection at which the sensitivity starts declining.
p_t	overall test sensitivity of detecting a flock at age t.
t_1, t_2, t_3, t_4	testing ages of the flocks.
$E(Q)$	expected infection prevalence over laying period under no intervention (percentage of infected flocks in a large population of laying flocks).
$E(Q \mid H^0_{t_K})$	as above, accounting for the effect of eliminating all detected positives at the time of detection, i.e. survival of non-detects only.
$\pi(\cdot)$	probability density of "$\cdot$".

 ISSN: 2102-6238

References

Albert I., Espié E., de Valk H., Denis J.-B. (2011). A bayesian evidence synthesis for estimating campylobacteriosis prevalence. *Risk Analysis*, 31(7):1141–1155.

Albert I., Grenier E., Denis J.-B., Rousseau J. (2008). Quantitative risk assessment from farm to fork and beyond: A global bayesian approach concerning food-borne diseases. *Risk Analysis*, 28(2):557–571.

Arnold M.E., Carrique-Mas J.J., McLaren I., Davies R.H. (2011). A comparison of pooled and individual bird sampling for detection of *salmonella* in commercial egg laying flocks. *Preventive Veterinary Medicine*, 99:176–184.

Bonde M., Toft N., Thomsen P.T., Sørensen J.T. (2010). Evaluation of sensitivity and specificity of routine meat inspection of danish slaughter pigs using latent class analysis. *Preventive Veterinary Medicine*, 94:165–169.

Branscum A.J., Gardner I.A., Johnson W.O. (2005). Estimation of diagnostic-test sensitivity and specificity through bayesian modeling. *Preventive Veterinary Medicine*, 68:145–163.

Bureau A., Shiboski S., Hughes J.P. (2003). Applications of continuous time hidden markov models to the study of misclassified disease outcomes. *Statistics in Medicine*, 22:441–462.

Carrique-Mas J.J., Davies R.H. (Jan-Mar, 2008). Salmonella enteritidis in commercial layer flocks in europe: Legislative background, onfarm sampling and main challenges. *Brazilian Journal of Poultry Science*, 10(1):1–9.

Cook R.J. (1999). A mixed model for two-state markov processes under panel observation. *Biometrics*, 55:915–920.

De Vylder J., Dewulf J., Van Hoorebeke S., Pasmans F., Haesebrouck F., Ducatelle R., Van Immerseel F. (2011). Horizontal transmission of salmonella enteritidis in groups of experimentally infected laying hens housed in different housing systems. *Poultry Science*, 90:1391–1396.

Dewaele I., Van Meirhaeghe H., Rasschaert G., Vanrobaeys M., De Graef E., Herman L., Ducatelle R., Heyndrickx M., De Reu K. (2012). Persistent *salmonella* enteritidis environmental contamination on layer farms in the context of an implemented national control program with obligatory vaccination. *Poultry Science*, 91:282–291.

Douc R., Garivier A., Moulines E., Olsson J. (2011). Sequential monte carlo smoothing for general state space hidden markov models. *The Annals of Applied Probability*, 21(6):2109–2145.

EFSA Panel on Biological Hazards (BIOHAZ) (2010). Scientific opinion on a quantitative estimate of the public health impact of setting a new target for the reduction of salmonella in laying hens. *EFSA Journal*, 8(4):1546. Available online: www.efsa.europa.eu. [86 pp.]. doi:10.2903/j.efsa.2010.1546.

European Food Safety Authority (2010). Quantitative risk assessment of salmonella enteritidis in shell eggs in europe. *EFSA Journal*, 8(4):1588. Available online: www.efsa.europa.eu. [43 pp.]. doi:10.2903/j.efsa.2010.1588.

Fienberg S.E. (2011). Bayesian models and methods in public policy and government settings. *Statistical Science*, 26(2):212–226.

García C., Soriano J.M., Benítez V., Catalá-Gregori P. (2011). Assessment of *salmonella* spp. in feces, cloacal swabs, and eggs (eggshell and content separately) from a laying hen farm. *Poultry Science*, 90:1581–1585.

Gradel K.O., Andersen J., Madsen M. (2002). Comparisons of sampling procedures and time of sampling for the detection of *salmonella* in danish infected chicken flocks raised in floor systems. *Acta vet. scand.*, 43:21–30.

Hernes G. (2012). *Hot Topic - Cold Comfort. Climate Change and Attitude Change*. NordForsk.

Karlin S., Taylor H.M. (1975). *A First Course in Stochastic Processes*. Academic Press, San Diego, 2nd edition.

Klinkenberg D., Thomas E., Calvo Artavia F.F., Bouma A. (2011). *Salmonella* enteritidis surveillance by egg immunology: Impact of the sampling scheme on the release of contaminated table eggs. *Risk Analysis*, 31(8):1260–1270.

Lunn D., Best N., Spiegelhalter D., Graham G., Neuenschwander B. (2009). Combining MCMC with 'sequential' PKPD modelling. *J Pharmacokinet Pharmacodyn*, 36:19–38.

Mahé A., Bougeard S., Huneau-Salaün A., Le Bouquin S., Petetin I., Rouxel S., Lalande F., Beloeil P.A., Rose N. (2008). Bayesian estimation of flock-level sensitivity of detection of *salmonella* spp., enteritidis and typhimurium according to the sampling procedure in french laying-hen houses. *Preventive Veterinary Medicine*, 84:11–26.

Maijala R., Ranta J., Seuna E., Pelkonen S., Johansson T. (2005). A quantitative risk assessment of the public health impact of the finnish salmonella control program for broilers. *Int. J. Food Microbiology*, 102(1):21–35.

Miconnet N., Cornu M., Beaufort A., Rosso L., Denis J.-B. (2005). Uncertainty distribution associated with estimating a proportion in microbial risk assessment. *Risk Analysis*, 25(1):39–48.

Nagelkerke N.J., Chunge R.N., Kinoti S.N. (1990). Estimation of parasitic infection dynamics when detectability is imperfect. *Statistics in Medicine*, 9:1211–1219.

Ranta J., Maijala R. (2002). A probabilistic transmission model of salmonella in the primary broiler production chain. *Risk Analysis*, 22(1):47–58.

 ISSN: 2102-6238

Ranta J., Siekkinen K.M., Nuotio L., Laukkanen R., Hellström S., Korkeala H., Maijala R. (2010). Causal hidden variable model of pathogenic contamination from pig to pork. *Statistical Modelling*, 10(1):69–87.

Ranta J., Tuominen P., Maijala R. (2005). Estimation of true salmonella prevalence jointly in cattle herd and animal populations using bayesian hierarchical modeling. *Risk Analysis*, 25(1):23–37.

Smid J.H., Verloo D., Barker G.C., Havelaar A.H. (2010). Strengths and weaknesses of monte carlo simulation models and bayesian belief networks in microbial risk assessment. *International Journal of Food Microbiology*, 139:S57–S63.

Spiegelhalter D.J., Riesch H. (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. *Phil. Trans. R. Soc. A.*, 369:4730–4750.

Sweeting M.J., De Angelis D., Aalen O.O. (2005). Bayesian back-calculation using a multi-state model with application to HIV. *Statistics in Medicine*, 24:3991–4007.

Thomas M.E., Klinkenberg D., Ejeta G., Van Knapen F., Bergwerff A.A., Stegeman J.A., Bouma A. (2009). Quantification of horizontal transmission of *salmonella enterica* serovar enteritidis bacteria in pair-housed groups of laying hens. *Applied and Environmental Microbiology*, 75(19):6361–6366.

Van Hoorebeke S., Van Immerseel F., De Vylder J., Ducatelle R., Haesebrouck F., Pasmans F., de Kruif A., Dewulf J. (2009). Faecal sampling underestimates the actual prevalence of *salmonella* in laying hen flocks. *Zoonoses and Public Health*, 56:471–476.

van Schaik G., Schukken Y.H., Crainiceany C., Muskens J., VanLeeuwen J.A. (2003). Prevalence estimates for paratuberculosis adjusted for test variability using bayesian analysis. *Preventive Veterinary Medicine*, 60:281–295.

 ISSN: 2102-6238