Establishment of Risk based microbiological criteria in the Nordic countries: A case study on Campylobacter in broiler meat.

Nauta, Maarten

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Establishment of Risk based microbiological criteria in the Nordic countries: A case study on Campylobacter in broiler meat.

NMDD project 2011-2012,
15 February 2013

Maarten Nauta,
National Food Institute,
Department Epidemiology and Microbial Genomics,
DTU,
Denmark
maana@food.dtu.dk

on behalf of the project participants, with thanks for providing national data and helpful discussions.

Roland Lindqvist, SLV, Sweden
Franklin Georgsson, Matis, Iceland
Helga Hogåsen, NVI, Norway
Sebastian Hielm and Pirkko Tuominen, MMM and EVIRA, Finland
Hanne Rosenquist, Jens Kirk Andersen, DTU, Denmark.

Appendix 4 on uncertainty analysis using a Bayesian approach is written by Jukka Ranta and Pirkko Tuominen, EVIRA, Finland

Summary

Microbiological criteria (MCs) offer a practical tool for food safety control and they are currently under discussion internationally. To meet the present scientific standards, there is an increasing demand for so- called “risk based” microbiological criteria that are based on risk assessment. In this project we studied the potentials for setting risk based microbiological criteria on *Campylobacter* in chicken meat by studying the potential impact that specific microbiological criteria would have in different Nordic countries. This is done on the basis of different data sets that have been collected in these countries in the past, and for the 2008 EU baseline survey data. The approach used is similar to that applied for the EFSA opinion of *Campylobacter* control (EFSA 2011, Nauta, Sanaa and Havelaar 2012), but in this study additional data sets are analysed. Next, as an alternative approach for setting risk based microbiological criteria, the “case-by-case” risk assessment methodology is used (Christensen et al 2013) and its impact is analysed on the basis of the same data sets. In both approaches the same risk assessment model for Campylobacter in broiler meat is used. The difference between the approaches is that for the microbiological criterion the rule for compliance is based on the traditional definition, defined by the number of samples containing more than a critical concentration of bacteria (Van Schothorst et al 2009), whereas for the “case-by-case” it is based on a critical risk estimate.

The study confirms that the risk of campylobacteriosis from broiler meat produced in the Nordic countries (and especially Norway and Finland) is low compared to most other European countries. When using different data sets from the same country, the results differ between them, but the ranking of countries is unaltered. It is for example found that microbiological criterion based on n= 5 samples, with a threshold concentration of m=1000 cfu/g that may be exceeded in c=1 out of 5 samples, gives between 0 and 10% non-complying batches of poultry meat in the Nordic countries. The risk reduction obtained by implementation of this MC varies greatly, and is, in general, larger when more non-complying batches are no longer accepted on the fresh meat market. Detailed results per country can be obtained from the report.

The analyses in the report focus on the evaluation of one specific microbiological criterion, which was previously selected as an example scenario by EFSA (2011), and one comparable “case-by-case” criterion. Results suggest that the efficiencies, in terms of potential risk reduction versus the percentage of non-complying batches, are similar for both methods. However, when studying the uncertainties, the uncertainty attending the “case by case” approach seems to be a little smaller. This preliminary result suggests that the “case by case” approach may be a more reliable method. One way to study this further will be to proceed with Bayesian data analysis as presented in this report.

We have shown that risk based microbiological criteria can be established, given the availability of a risk assessment model that links the bacterial concentration measured at the point where the criterion is set to the public health risk. The methods described in this

report can offer a risk management tool where the choice for the optimum criterion can be based on combination of the potential risk reduction and the percentage of non-complying batches that, in some way, require sanctioning. It is, however, unsure to what extent the anticipated effects on public health risk will be achieved. Not only is every microbiological risk assessment attended by considerable uncertainty, also the effects of setting criteria in terms of decreased concentrations of bacterial pathogens on the meat as an effect of (potential) sanctioning, are difficult to predict.
A software tool will be developed to facilitate the use of the methods described in this report by interested parties, such as industrial and governmental food safety managers in the international community.

Content

1. Introduction

In recent years there is an increased interest in establishing risk-based Microbiological Criteria (MCs) as a tool in food safety control. MCs can offer a practical tool, providing well defined critical limits in terms of acceptable prevalence and concentrations of microorganisms, as well as the analytical method and sampling scheme to be used. When the MCs are "risk-based" they are set on the basis of their expected impact on public health risks.

In this project we aim to develop tools for setting risk based Microbiological Criteria (MCs), focusing on *Campylobacter* in poultry meat, and apply them to the Nordic countries. We apply a method for evaluating MCs that has previously been used in the EFSA opinion on *Campylobacter* control (EFSA 2011, Nauta, Sanaa and Havelaar, 2012) on risk based microbiological criteria for *Campylobacter*. Several Nordic data sets are studied and compared. A new element in this project is that we compare the effect of setting an MC with the "case by case" risk assessment approach applied in Denmark (Christensen et al., 2013). In both approaches food lots are tested by taking a number of samples, in which the *Campylobacters* are enumerated. The results are evaluated against a well-defined criterion that can be used as a decision rule for compliance of the food lot. This criterion is established on the basis of quantitative microbiological risk assessment. We explore the impact of MCs in the Nordic countries based on available data. These data are

1. The EU baseline survey data 2008, for Denmark, Sweden, Norway and Finland. (EFSA 2010).
2. The Swedish data sets published by Lindblad et al (2006), and Hansson et al (2010).
3 The Icelandic data set published by Stern et al (2007).
4. The Danish retail monitoring data set (2001-2010).
5. The Finish data from a professional kitchen survey (2001)
6. The Norwegian data from the Kalkunkylling project (2006-2007).

2. Method

2.1. General approach

The modelling concept is shown in Figure 1. Given a certain concentration of *Campylobacter* on a food product (typically chicken meat), (1) the health risk attending a food product can be assessed by applying quantitative microbiological risk assessment and (2) the outcome of a microbiological test protocol can be evaluated by a model simulating the testing procedure. This test protocol is typically linked to a criterion that defines whether the tested food lot is complying with this criterion or not. If we do both (1) and (2) simultaneously, we can study the relation between the initial concentration, the test protocol and the risk. This relation informs us on the impact that the application of a test protocol and criterion, and the attending corrective action, can have on public health.

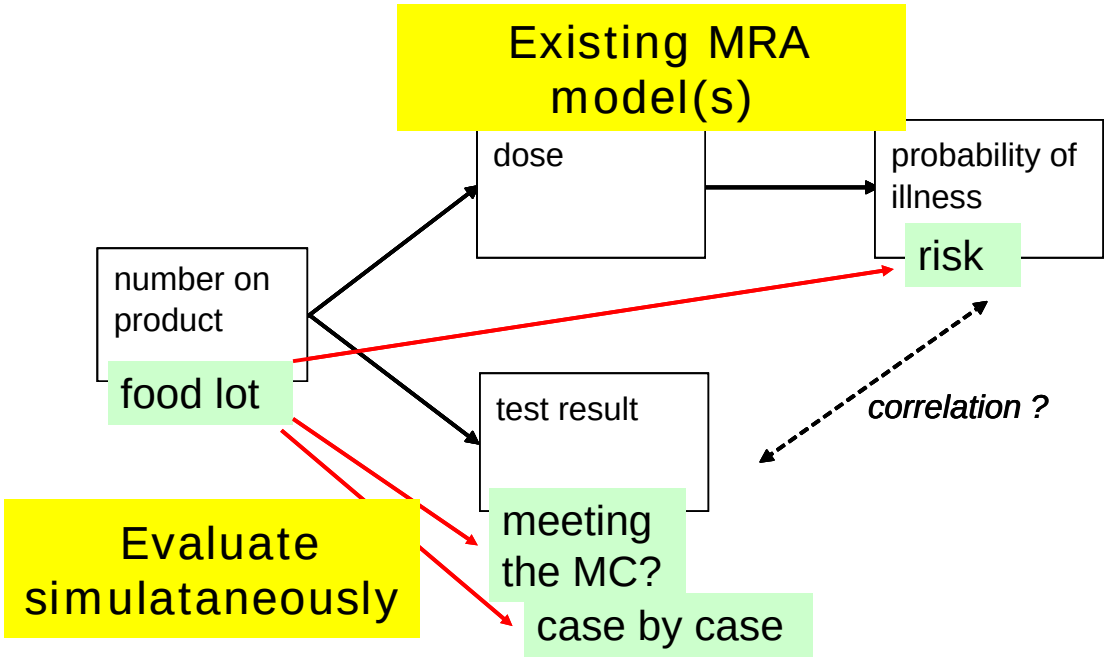

Figure 1. The conceptual approach: For different food lots we evaluate the attending risk of campylobacteriosis (by QMRA) and the impact of a test protocol, where samples are evaluated against a microbiological criterion (MC) or with the "case by case" (CC) risk assessment methodology

In this project we apply three models

1) A QMRA model to assess the risk of a sample, a food lot and a set of food lots
2) A model for the probability of compliance of a food lot given a predefined Microbiological criterion (MC)
3) A model for the probability of compliance of a food lot using the Danish case by case risk assessment approach, and the attending criterion (Case by case criterion, CC).

These models are linked to evaluate for different sets of food lots

1) the risk attending this food lot
2) the percentage of food lots that is not complying with a MC or CC.
3) the residual risk of the food lots that do comply with the MC or CC.

In this approach non-complying food lots (2) are considered as a cost and the decreased residual public health risk (3) is regarded as a benefit.

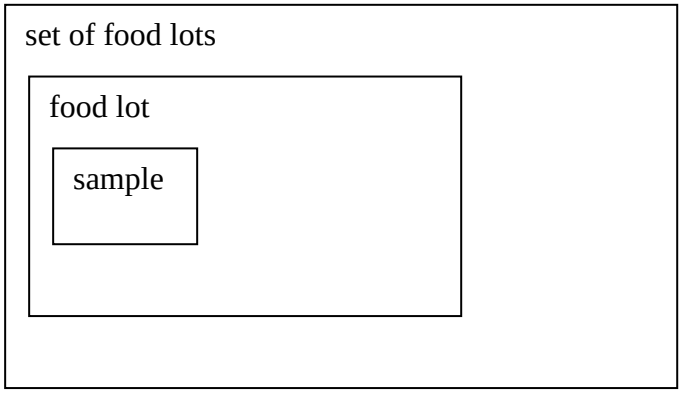

Figure 2. The modelling approach distinguishes between samples, food lots and sets of food lots.

For a clear understanding it is important to differentiate between samples, food lots and sets of food lots (see Figure 2). Food lots (or batches[1]) of food are tested for compliance, by taking a number of samples (*n*) from this food lot. Each sample is analysed, resulting in an estimate for the concentration of *Campylobacter* in that sample. In the end we want to evaluate a set of food lots, that for example represents all meat produced in a country, but can also represent the daily production of a company, or for example the fresh chicken legs produced in a region during the summer months.

2.2. The QMRA model.

The QMRA model is described in Appendix 1 (with text copied from Nauta, Sanaa and Havelaar, 2012). It allows an assessment of the probability of illness per serving derived

[1] Terms ”food lot” and ”batch” are used interchangably in this report

from a sample, a food lot or a set of food lots given the concentration, or distribution of concentrations of *Campylobacter* on the meat. This probability of illness is the risk estimate.

There are some important assumptions made to establish this model, that relate to

- the transition factor τ that expresses the difference between the observed concentration (e.g. on a skin sample or from a carcass wash) and the concentration per g of meat.
- the consumer phase model (Nauta et al 2008, Nauta and Christensen 2011): the relation between the concentration on the meat and the ingested dose is assumed to be the same, independent of meat product and consumer population.
- the dose response model: the one most frequently used in QMRA is assumed to be appropriate.

In general we want to evaluate this risk on a relative scale (that is compared to some "baseline" risk) because the absolute risk estimate may be biased, and is attended by a large uncertainty. This uncertainty will be smaller for a relative risk. However, in this report most calculations are done with the absolute risk estimate. This is discussed in section 5.4.

2.3. The microbiological criterion (MC)

Given a definition of the sample size *n*, and the values of the maximum concentration in cfu/g *m* and the maximum number of sample that may exceed *m*, *c*, we can estimate the probability of compliance, given the concentration(s) in the meat. Here we use the percentage of non-complying batches, referred to as BNMC.

The mathematics of evaluation of an MC is explained in Appendix 2 (with text from Nauta, Sanaa and Havelaar, 2012).

If we can distinguish between complying and non-complying food lots, we can also estimate the effect on risk. The minimum residual risk (MRR) is the risk of the complying food lots, which can be interpreted as the risk of the meat that is available for consumption in the extreme scenario where all food lots are tested and the meat of non-complying food lots is not available for consumption (or all that meat is first treated so that all *Campylobacters* on that meat are inactivated, for example by heating). The minimum relative residual risk (MRRR) is the relative measure that indicates which percentage of risk remains from the complying batches.

2.4. The Case by case criterion (CC)

The Danish case by case risk assessment method can be evaluated in the same way as microbiological criteria. Both are based on a sample size *n*. In the case by case method the risk is assessed for each individual sample, and the mean risk of *n* samples is calculated as

$$Risk_{n\ samples} = \frac{\sum_{i=1..n} Q_{ill}(C_{obs,i} - \tau)}{n}$$

With $C_{obs,\ I}$ the observed concentration in sample *i*.
Compliance can be defined by getting a value of this $Risk_{n\ samples}$ below a critical value. Alternatively, this mean risk is evaluated against a "baseline risk estimate", and the relative risk is calculated as

$$RR_b = \frac{Risk_{n\ samples}}{Risk_{baseline}}$$

Using this relative risk estimate can be more informative for the risk manager, and will also be less uncertain, as some of the model uncertainties are cancelled out in the quotient, as discussed in section 2.3.
In Denmark the criticality of the relative risk of a batch is determined on a case by case basis.
Here we assume a fixed critical value for the mean risk of the *n* samples, $Risk_{n\ samples}$, which later may be translated in a critical relative risk (see section 5.5). It is important that the values for $Risk_{n\ samples}$ should be considered on a relative scale (the risk is 2, 3 etc. times larger or smaller), not as absolute values.

2.5. Input requirements

To evaluate the impact of MC and CC, we need to derive distributions of concentrations in sets of food lots. There are several parameters of interest:

- The food lot prevalence indicates the percentage of food lots that is contaminated (where the other food lots are *Campylobacter* free).
- The within food lot prevalence indicates the percentage of food products within the food lot that is contaminated (with at least one cfu). This prevalence may vary between food lots.
- The mean concentration, which is actually the mean of all mean concentrations in contaminated products in contaminated food lots. (We only include contaminated products in this mean, because we usually work with log concentrations and log(0) does not exist)
- The standard deviation in means between food lots, because the mean concentrations will differ between food lots.
- The standard deviation within food lots. This will usually vary between food lots, but for simplification it is commonly assumed that the within food lot sd is identical for all food lots.

Below, our approach is to assume that the logs of concentrations are normally distributed within food lots, and that the mean log concentrations are normally distributed between

food lots as well. Normal distributions are fitted to the data by using a Maximum Likelihood Estimation (MLE) method that includes censored data (that is it includes the fact that the enumeration has a detection limit, so some (low) concentrations will not be observed) (Lorimer and Kiermeyer 2007). This MLE can be used to estimate the mean and standard deviation of concentrations, given an (observed) prevalence (a 2 parameter MLE), and it can also be used to estimate both the prevalence and mean and standard deviation of concentrations (a 3 parameter MLE).
In some cases (data from Hansson et al (2010) and Stern et al (2007)) we can use the empirical data, without fitting a distribution. That allows us to evaluate the impact of fitting distributions through the data sets (see 4.3.3.).
When interpreting the empirical data in terms of model input, one should realize that the measured concentration is usually expressed as log cfu/ml of rinsing fluid, log cfu/g skin or otherwise. The model requires log cfu/g of (fresh retail) meat as input. As in EFSA (2011), Nauta and Christensen (2011) and Nauta, Sanaa and Havelaar (2012), we assume that the relation between the observed concentration and the concentration per gram of meat is

$$\log C_{meat} = \log C_{obs} - \tau.$$

This τ is unknown. In this study we assume $\tau = 1$, as has been done to relate the skin samples in the EFSA baseline survey (cfu/g skin) to meat (cfu/g meat) (EFSA 2011).
In the EU baseline survey, which is the basis for the exploration of the effect of MC setting described in the EFSA opinion on Campylobacter control (EFSA 2011), one skin sample is analysed per food lot. This means that the within flock standard deviation of log concentrations cannot be differentiated from the between flock standard deviation. Therefore we define parameter φ so that

$$var_{between} = (1-\varphi)\, var_{obs} \quad \text{and} \quad var_{within} = \varphi\, var_{obs}$$

In EFSA (2011), it is assumed that $\varphi = 30\%$.

2.6. Model implementation

For this project, the model is implemented as a Monte Carlo simulation model in @Risk. This allows an evaluation of the MCs and the CC in the same runs of the model, and therefore a direct comparison of the two approaches.
In the Monte Carlo simulations, *n* samples are taken from a set of different contaminated batches, with contamination described by the distributions of concentration, hence depending on the mean concentration on the meat μ_{obs}, the total variance var_{obs}, and the parameters τ and φ. The between batch prevalences (p_{MS}) is applied to correct for the fact that only contaminated batches are simulated.

The *n* samples are evaluated for:
- (MC) the number that is not complying (so has a concentration >*m*). If in total more than *c* samples are not complying, the batch is not complying. The risk of non-complying batches is stored, and the mean risk of these batches is evaluated against the mean risk of all batches.
- (CC) the risk attending each sample. The mean risk of all samples is calculated, and compared to a threshold risk. (Preferably the mean risk is expressed as a relative risk, compared to a baseline risk, and evaluated against a threshold relative risk.)

2.7. Overall evaluation

For this study, project participants from different Nordic countries provided national data, as described in chapter 3. These data were evaluated by the QMRA model, which gave a risk estimate for the risk attending the meat that is represented by the data sets. Next, the same data sets were evaluated for the impact of
1) setting a microbiological criterion with n=5, m = 1000, c=1. (the default MC in EFSA, 2011)
2) performing a case by case risk assessment with n=5 and $Risk_{n\ samples}$ = 1%, as this appeared to yield comparable results for the percentage on non-complying batches.

2.8. Uncertainty analyses

Quantitative microbiological risk assessments are commonly attended by substantial uncertainties, among others due to limited knowledge about the impact of consumer food handling on exposure, and about the dose response relation. As a consequence it is difficult to quantify the uncertainty about risk estimates. This uncertainty is nonetheless of importance for the risk manager, as it describes to what extent the consequences of control measures may be different than anticipated.
Nauta, Sanaa and Havelaar (2012) performed a sensitivity analysis for the MC approach, and find that BNMC and MRRR are sensitive to the values the unknown parameters τ and φ, and identify the need for improving the estimates of these parameters. However, the impact on the total model uncertainty is small, compared to the impact of the parameters describing the concentration distribution on the meat samples and the parameters defining the MC.
In this project additional uncertainty analyses are performed by two different methods: Monte Carlo analysis (4.3.1. and Appendix 3) and Bayesian Analysis (4.3.2 and Appendix 4). Also, the impact of using fitting distributions through the data, instead of using the raw data, is studied for the two data sets where raw data per flock was available (4.3.2. using data from Hansson et al 2010; Stern et al 2007).

3. Data from Nordic countries

3.1. Sweden

3.1.1. EU baseline data

A normal distribution (mean and standard deviation (sd)) is fitted through the EU baseline data (log cfu/g skin) using (2 par) MLE for censored data; the prevalence estimate is obtained from the enrichment data. The estimates are: prevalence (prev) 13.4 %, mean (mu) 1.42 cfu/g skin, standard deviation (sd) 1.01 (EFSA, 2011).

3.1.2. SLV data Lindblad et al 2006.

The following information on this data set has been provided:

Name data set	**SLV_2005_Baseline**
sampling point	Processing after chilling
who did the sampling and why	SLV, national survey
type of data	P/A, quantitative
method of microbial analysis	NMKL no. 119 (qual), Direct plating of 1 ml, or 0.1 ml rinse fluid on mCCDD (quant.)
sample type	Whole carcass rinse
sampling scheme, amount of data	representative 1-year, 617 samples
considered representative for	National baseline
year(s) of sampling	September 2002 - August 2003
raw data available, not just means (and st. devs)	Yes
additional comments	http://www.ncbi.nlm.nih.gov/pubmed/17186653

The abstract of Lindblad et al 2006 states:
"*Campylobacter, predominantly Campylobacter jejuni, were detected on 15% (by enrichment) or 14% (by direct plating) of the carcasses. With one exception, all samples from late December through April were Campylobacter negative. The 10th and 90th percentiles of Campylobacter numbers per carcasses were 3.0 and 5.0 log CFU, respectively, and the maximum was 7.1 log CFU.*"
The data set consists of 88 quantified concentrations per carcass, the lower limit is 2.6 log cfu/carcass (1 in 1 ml from 400 ml). Enrichment has been done in 10 ml.
The observed data can be used, as well as a normal fit through the data, using a (3 par) MLE to assess prevalence, mean and sd of the normal distributed logs. The data set holds 88 positive samples with enumeration among 617, i.e. an observed prevalence of 14.3%. If, in the censored data analysis, the threshold for the detection limit is set at 2.3 (this

equals a count of 0.5 cfu, as 2.6 is 1 counted cfu, the estimated prev = 0.147, mu =4.06, sd =0.91, see the graph.

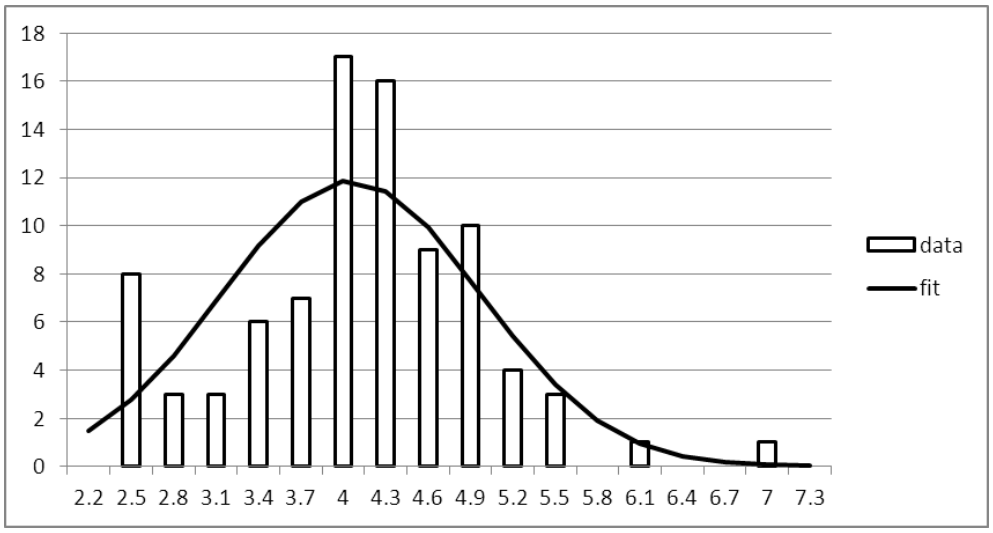

Figure 3. The Lindblad et al. data set and the fitted distribution.

The log cfu/carcass is transposed into log cfu/g meat to fit it in the method and it is assumed that the MC/CC is evaluated by the same microbiological method as the samples in this study. Transposing the data requires an assumption on the mass of skin (in g) on a carcass. The 100 g skin per chicken from the Danish QMRA (Christensen et al 2001) is adopted here.
So to obtain cfu/ g of skin we have to subtract 2 logs, so the mean gets 2.06.
There is no information on the within flock variance, so φ is put at 30% (EFSA 2011).
τ is assumed to be 1 (EFSA 2011).

3.1.3. Hansson et al (2010) data

Hansson et al 2010 published a data set with carcasses from 20 flocks, as shown in the table below. These are not representative for the whole year. The data are expressed as log cfu/ml carcass rinse (using the same method as Lindblad et al 2006, but expressed per ml). To transform the data into "per carcass" log(400) is added, because 400 ml rinsing fluid was used. To transform it into "per g of skin" log (100) gas to be subtracted again (see 3.1.2), so in total we have to add log(4) to the published data.
Interestingly, these data include within flock prevalence. The mean mean of logs is 2.92 log cfu/g. The mean standard deviation is 0.54. The standard deviation of the standard deviations is 0.16. The standard deviations of the means is 0.60.
A distribution that fits through the standard deviations is a Gamma (11.6, 0.046) (obtained using @Risk distribution fitting).
The total sd is the square root of the sum of variances: $\sqrt{(0.6*0.6 +0.54*0.54)} = 0.81$; so $\varphi = 0.445$.

Table: Results of Hansson et al 2010)

mean conc. per ml rinse	**mean per g. skin**	**st. dev.**	**within flock prev.**
2.31	2.91	0.61	0.96
1.96	2.56	0.51	1.00
1.38	1.98	0.60	0.91
2.98	3.58	0.48	1.00
2.87	3.47	0.71	1.00
2.76	3.36	0.39	1.00
3.02	3.62	0.58	1.00
2.69	3.29	0.40	1.00
3.15	3.75	0.49	1.00
2.63	3.23	0.37	1.00
2.74	3.34	0.37	1.00
2.32	2.93	0.26	1.00
2.62	3.23	0.49	1.00
2.62	3.22	0.35	1.00
1.35	1.95	0.81	0.85
1.21	1.81	0.80	0.86
2.19	2.79	0.48	1.00
1.39	1.99	0.75	1.00
2.13	2.74	0.69	0.95
2.11	2.71	0.61	1.00
means	**2.92**	**0.54**	**0.98**

When we apply this data set we can use both the fitted distribution through the data, as the original data set. This is interesting because the means and standard deviations are correlated (see figure 4).
The between flock prevalence is not known in this data set, from the other Swedish data we can assume it is 14%.

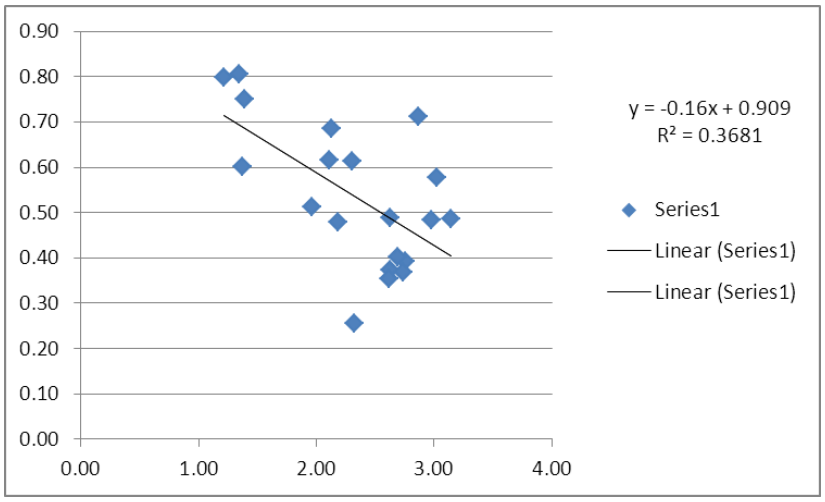

Figure 4. The mean concentrations (x axis) on a log scale, vs. the within food lot standard deviation of the Hansson et al 2010 data. An interesting finding is that a higher concentration associates with a lower standard deviation.

3.1.4. Summary

For Sweden four different input distributions are available.

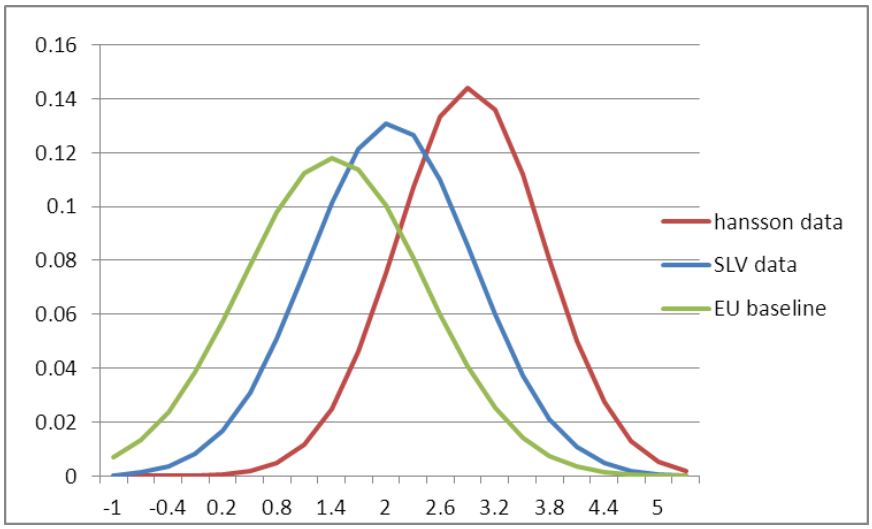

Figure 5. The three distributions obtained from Swedish data. When considering the differences one should realize that the data are obtained by different methods, and the fact that the Hansson data set is not representative for the whole year

	prev bet (%)	prev within	mean (log cfu/g skin)	sd total	τ	φ
EU baseline	13.4	1	1.48	1.01	*1*	*30%*
SLV	14.7	1	2.06*	0.91	*1*	*30%*
Hansson 1°	*14*	0.98	2.92*	0.81	*1*	44.5%
Hansson 2°	*14*	0.98	Sampled		*1*	

* from carcass rinse data, transformed data.
° not representative for the year
uncertain values in italics

3.2. Denmark

3.2.1. EU baseline data

A normal distribution (mean and sd) is fitted through the EU baseline data (log cfu/g skin) using (2 par) MLE for censored data; prevalence estimate is obtained from the enrichment data. The estimates are: prevalence 31.1 %, mean 2.1 cfu/g skin, sd 1.28 (EFSA, 2011).

3.2.2. Retail data

The following information on this data set has been provided:

Name data set	**Surveillance at retail**
sampling point	Retail and wholesale
who did the sampling and why	Regional laboratories for the Danish Veterinary and Food Administration – National surveillance
type of data	semi quantitative
method of microbial analysis	NMKL 119
sample type	min. 15 g sample
sampling scheme, amount of data	Samples supposed to be taken evenly over the year (not always applied). Between ~400 and ~2400 samples per year
considered representative for	Whole country, annually
year(s) of sampling	2000→
raw data available, not just means (and st. devs)	Yes
additional comments	chilled and frozen meat; sampling of whole carcasses until 2006, from 2007 whole carcasses and parts

There is a large amount of data available. These are retail data for many years (2001 – 2010), differentiated by origin (imported Danish), fresh or frozen meat, and per season. (Boysen et al, 2013). These data are used for the Case by Case methodology (Christensen et al 2013).
The overall mean over the observations for three years (2005-2007), not weighted for production etc., yields 5736 samples, and estimates of prevalence 69.94%, mean - 0.68 and sd 2.11. Here the semi quantitative data (log cfu/g) are fitted to a normal distribution using (3 par) MLE for censored data. Note that
1) when data are weighted over the season, this gives a different (lower) result for the risk estimates
2) the prevalence estimate is very high, but this is coincides with a low mean concentration, as a result of the method for fitting a distribution

	prev bet (%)	prev within	mean (log cfu/g skin)	sd total	τ	φ
EU baseline	31.1	1	2.1	1.28	*1*	*30%*
retail data	69.9	*1*	-0.68	2.11	*1*	*30%*

3.3. Finland

3.3.1. EU baseline data

A normal distribution (mean and sd) is fitted through the EU baseline data (log cfu/g skin) using (2 par) MLE for censored data; prevalence estimate is obtained from the enrichment data.
The estimates are: prevalence 5.7 %, mean 0.32 cfu/g skin, sd 1.59

3.3.2. Professional kitchen survey 2001.

Thermotolerant *Campylobacters* were analysed from raw poultry meat in 6 municipalities. The samples were taken during the months known to be campy-rich in Finnish poultry (June, August, September and October). 244 samples were taken from 199 professional kitchens: - 217 were domestic, 13 from Thailand, 3 from Brazil. 2 from Denmark, 1 from Hungary

n	Finnish			Outlandish		
	n	Pos	%	*n*	Pos	%
244	217	44	20.3	27	3[x]	11.1

[x]2 French and 1 Thai sample (frozen)

results include MPNs for 50 positive samples (three of them not broiler).
A censored data (2 par) MLE assuming a normal distribution of the log through the broiler data (n=47), gives a mean -0.67 and sd 0.92 for log cfu/g, and prevalence 47/244 = 19.2% . It is assumed that this is the same type of data as in the other studies.
This gives two data sets for Finland.

	prev bet (%)	prev within	mean (log cfu/g skin)	sd total	τ	φ
EU baseline	5.7	1	0.32	1.59	*1*	*30%*
prof. kitchens	19.2	*1*	-0.67	0.92	*1*	*30%*

3.4. Norway

3.4.1. EU baseline data

A normal distribution (mean and sd) is fitted through the EU baseline data (log cfu/g skin) using (2 par) MLE for censored data; prevalence estimate is obtained from the enrichment data. The estimates are: prevalence 5.1 %, mean 0.95 cfu/g skin, sd 0.51.

3.4.2. Kalkunkylling project

The following information on this data set has been provided:

Name data set	**KalkunKylling20062007**
sampling point	Producer
who did the sampling and why	Research project, products ready for sale sent by producers
type of data	Quantitative and presence/absence
method of microbial analysis	NMKL no.119, 2007
sample type	CF. under

sampling scheme, amount of data	Random samples collected weekly, according to estimated consumption rate 223 breast 18 whole chicken 64 meat cuts 29 minced meat with given date and origin (region)
considered representative for	Norwegian consumption, seasonal and regional variation
year(s) of sampling	2006 (second half of November and December) and 2007 (until second half of November)
raw data available, not just means (and st. devs)	Yes, in Excel
additional comments	

The data set contains meat samples, in total 334 observations, in 29 of those *Campylobacter* was detected, but in only 4 of them to countable levels (more than 100 cfu/g). This implies a prevalence of 8.7%. Unfortunately it is not possible to use the raw data for exploration of the impact of MC and CC, because the within food lot variability is not known, and the majority of data is not quantitative. Also, it is pretty hard to fit a distribution through the data. Assuming the lower limit for detection in the presence/absence test is 1 cfu/10 g, the best fit would be mean 1.26, sd 0.63. This is based on too little data, but we can explore it to see the impact.

	prev bet (%)	prev within	mean (log cfu/g skin)	sd total	τ	φ
EU baseline	5.1	1	0.95	0.519	*1*	*30%*
kalkunkylling	8.7	*1*	*1.26*	*0.63*	*1*	*30%*

3.5. Iceland

3.5.1. Stern et al. 2007 data.

The following information on this data set has been provided:

Name data set	**Campylobacter in fresh broilers**
sampling point	Processing - Retail packs at processing
who did the sampling and why	Research project
type of data	Quantitative
method of microbial analysis	Direct plating – rinse and weep samples
sample type	Whole carcass rinse and weep from packs
sampling scheme, amount of data	32 flocks, 20 samples per flock
considered representative for	Whole country
year(s) of sampling	2003, 2004
raw data available, not just means (and st. devs)	Yes

These data contain weep fluids and carcass rinses taken from 32 batches (mainly summer, July 2002 – September 2004.) The data represent fresh broilers that tested negative when faecal samples from herds were tested 3-5 days before slaughter (after about 32 days of rearing) but became colonized from that time until the day of slaughter as proofed by ceca samples collected at the slaughterhouse.

Carcass rinse is performed in 100 ml fluid **after** removal of the weep, which contains quite a lot of *Campylobacter*. (The rinse data are used here, but this is an underestimation of the total carcass load, which should be a measure that sums the weep and rinse.)

In Iceland 2004 three types of broiler meat can be considered:

(1) Contaminated fresh meat, from flocks tested negative 3-5 days before slaughter, but positive at slaughter. This is the type of meat in the data set.

(2) Frozen meat, from flocks tested positively at the farm 3-5 days before slaughter. This meat contains *Campylobacter*, and data are available in two data sets that have not been fully analysed. A quick analysis shows this meat holds about 0.5 log lower concentrations than the fresh contaminated meat. This is less than expected, given the effect of freezing on concentrations, which again may be explained by higher caecal concentrations / higher within flock prevalences in flocks that get positive 3-5 days before slaughter.

(3) Meat from flocks tested negative at the farm and during slaughter. This may be assumed to be not contaminated.

Unfortunately it is not clear which percentage of flocks falls in categories (1) and (2) (F Georgsson, pers. comm).

A major assumption would be that the rinse data reflect the current contamination. This is a common assumption, but in this case it is explicit that part is taken away as weep. The volume taken away is usually less than 10 ml (F. Georgsson, pers. comm).

The results published by Stern et al. (2007) for average log cfu/100 ml rinse are

mean	sd	n	within flock prev
3.28	0.56	4	0.2
5.05	0.48	20	1
4.34	0.39	20	1
3.56	0.82	12	0.6
3.49	0.87	10	0.5
4.06	0.55	17	0.85
4.74	0.8	20	1
4.96	1.01	18	0.9
3.63	0.59	14	0.7
3.89	0.81	9	0.45
4.49	0.58	20	1
5.55	0.66	20	1
3.21	0.24	6	0.3
4.9	0.69	20	1
3.62	0.59	7	0.35
4.1	0.75	7	0.35
4.62	0.8	20	1
5.24	0.52	20	1
3.28	0.58	20	1
3.47	0.75	16	0.8
5.37	0.91	20	1
3.19	0.83	19	0.95
3.79	0.56	20	1
4.76	0.79	20	1
5.94	0.53	20	1
4.17	0.71	20	1
4.88	0.53	20	1
2.95	0.49	12	0.6
4.98	0.59	20	1
5.27	0.34	20	1
3.47	0.74	17	0.85
3.06	0.46	17	0.85

Expressed as cfu/g of skin (100 g skin per carcass, 100 ml rinse per carcass), this gives: The mean is 4.23 log cfu/carcass, or 2.23 log cfu/g skin; the standard deviation of the means is 0.84; the mean standard deviation is 0.64; the standard deviation of the standard deviations is 0.17. A distribution that fits through the standard deviations is a Gamma (12.4, 0.052) (using @Risk)). The total sd is the square root of the sum of variances: $\sqrt{(0.84*0.84 +0.64*0.64)} = 1.05$, $\varphi = 0.371$.
The mean within flock prevalence is 82%.
The between flock prevalence is unknown, in the calculations it is assumed to be 10%.

Like with the Hansson et al 2010 data we can sample from the fitted distributions and from the values in the table directly.

	prev bet (%)	prev within	mean (log cfu/g skin)	sd total	τ	φ
Stern 1[°]	*?*	0.82	2.23[*]	1.05	*1*	37.1%
Stern 2[°]	*?*	0.82	sampled		*1*	

For Iceland there are no EU baseline survey data. Data are representative if flocks detected positive at the farm are scheduled and are not consumed as fresh meat.

4. Results

4.1. The risk estimates for different data sets

The QMRA model (section 2.2) is used to assess the risks of the meat from the different studies mentioned in chapter 3. The results are shown in Figure 6. There are clear differences in the risks assessed from different data sets within and between countries.

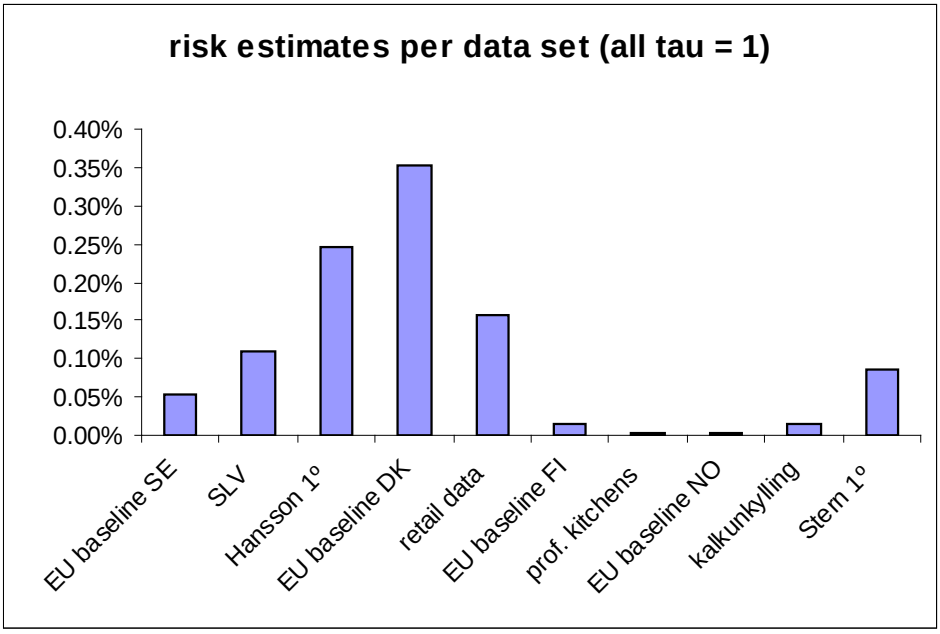

Figure 6. The risk estimates obtained from the different data sets.

It may be of interest to explore the association between the risk estimates and the flock prevalence and the mean concentrations in the contaminated meat. Figure 7 shows that neither of these surrogate risk measures are strong indicators for the assessed risk. A similar result was obtained by Nauta, Sanaa and Havelaar (2012).

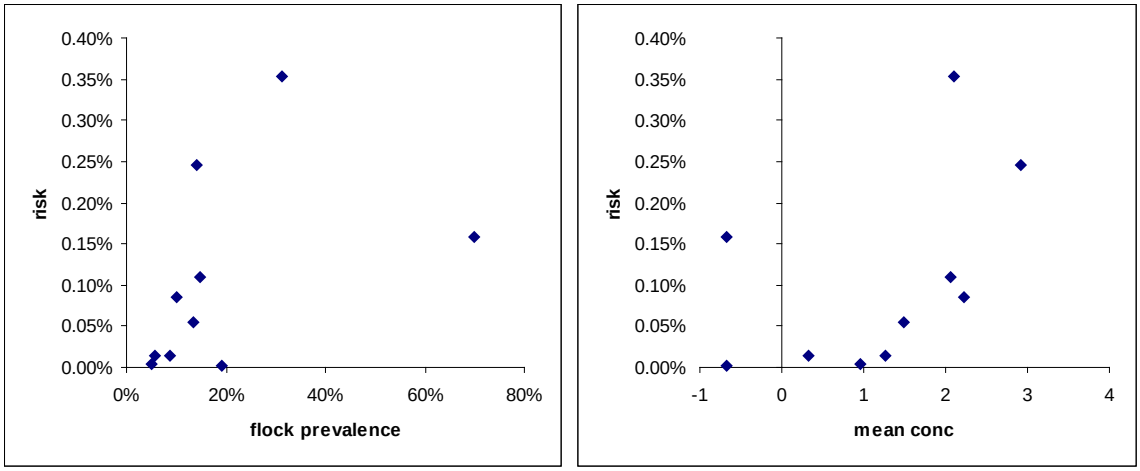

Figure 7 The association between flock prevalence and mean concentration and the assessed risk. Each dot represents one data set,

4.2. Scenario studies for Microbiological criteria and case by case

4.2.1. "Default" scenarios

Evaluation of the MCs and case by case risk assessment as explained in chapter 2, with parameters as given in section 2.7, gave the results presented in the table below. Figure 8 presents the relation between the assessed risk and the percentage of non-complying food lots[2] for the different data sets, and shows that more food lots are non-complying when the risk attending a set of food lots is higher. Figure 9 shows the relation between the percentage of non-complying batches and the MRRR value, which can be used to evaluate the efficiency of the MC or CC: A lower relative residual risk is usually associated with a higher percentage of non-complying batches. (Note that this graph will look different for different MCs (other values of *n, m* and *c*) and for different CC (other values of *n* and $Risk_{n\ sample}$. (Nauta, Sanaa and Havelaar, 2012).) Figure 10 shows the relation between the percentage of non-complying food lots and the (absolute) minimum residual risk, MRR. This gives a different picture, showing that when the MRRR is relatively high, the MRR is usually low, and that the MRR may still be high, even if the percentage of non complying food lots is high and the MRRR is low.

Table. Results of the analysis for MC (with *n*=5, *m*=1000, *c*=1) and CC (with *n*=5 and $Risk_{n\ samples}$ = 1%,) applied to the different data sets.

[2] The abbreviation BNMC stands for the percentage of non-compliying batches for the Microbiological Criterion, BNCC is the same for the "case by case" risk assessment approach. BN.C is used when either of them is addressed.

data	BNMC	MRRR MC	BNCC	MRRR CC
EU baseline SE	1.23%	62.51%	1.41%	58.79%
SLV	3.03%	48.14%	3.52%	42.64%
Hansson 1º	8.71%	19.17%	9.40%	15.13%
EU baseline DK	9.86%	28.50%	10.79%	24.75%
DK retail data	3.50%	49.97%	4.38%	44.22%
EU baseline FI	0.34%	55.90%	0.39%	51.72%
prof. kitchens	0.00%	100.00%	0.00%	100.00%
EU baseline NO	0.00%	100.00%	0.00%	100.00%
kalkunkylling	0.03%	97.34%	0.05%	95.76%
Stern 1º	2.74%	50.79%	2.86%	48.26%

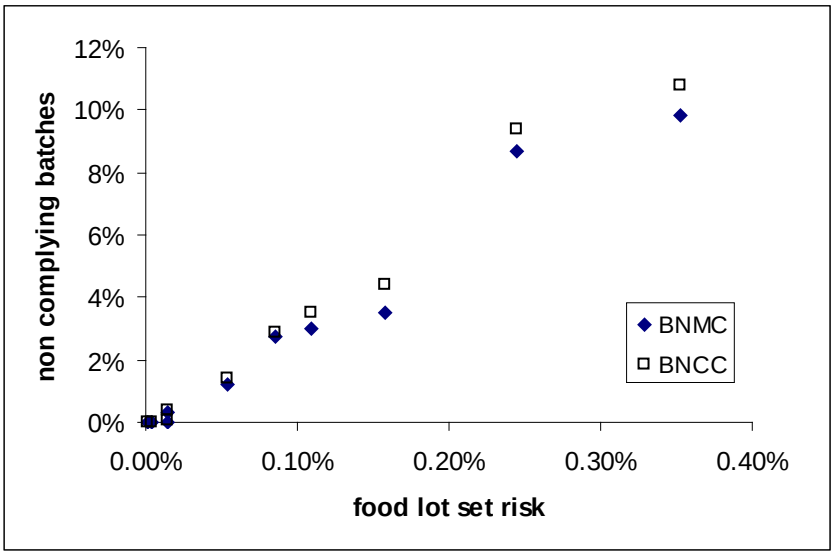

Figure 8 The association between the percentage of non-complying food lots, using a microbiological criterion (diamonds) or the case by case methodology (squares), and the risk estimates of the sets of food lots (the different data sets).

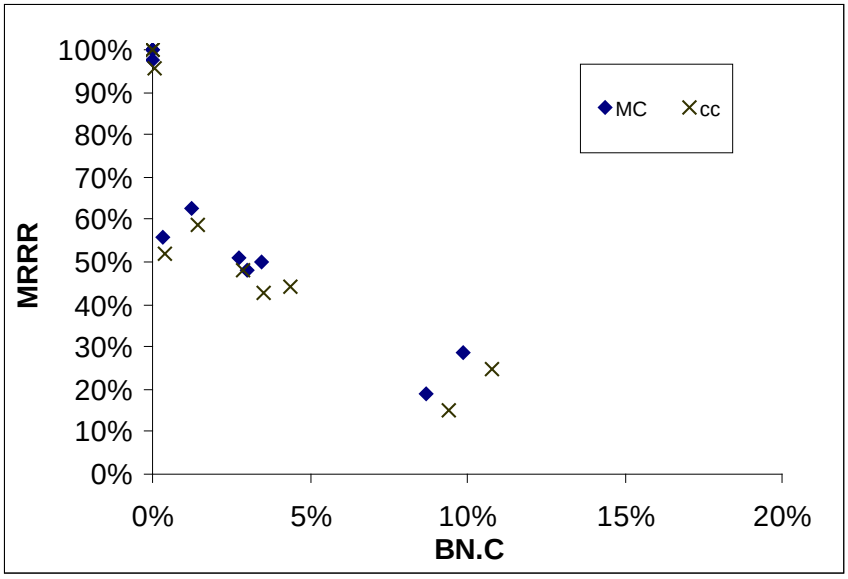

Figure 9. The minimum relative residual risks vs. the percentages of non complying food lots for MC and CC. For the data, see the table above.

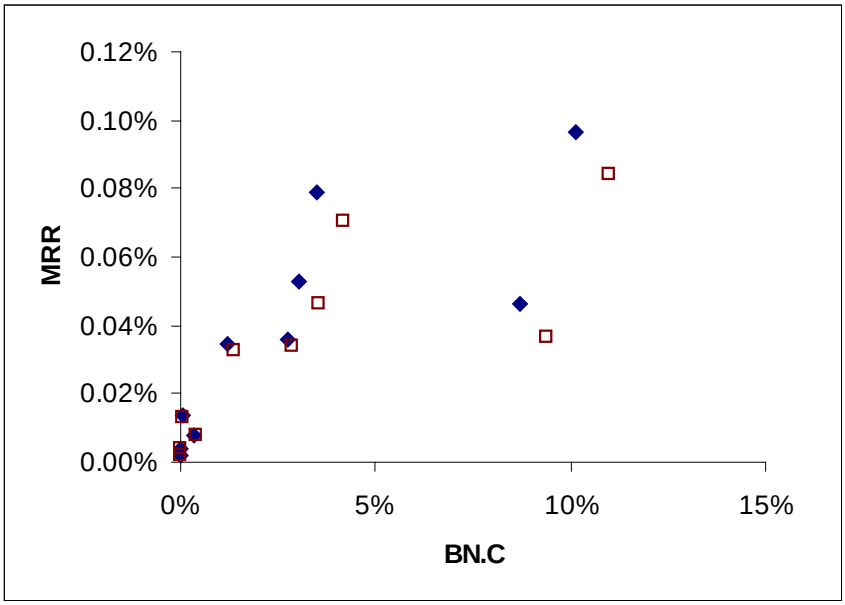

Figure 10. The minimum residual risk vs. the percentage of non-complying batches for MC and CC.

4.2.2. Alternative application of the model: searching an appropriate criterion

With the models described above, and the available concentration data, one can search an appropriate Microbiological Criterion, or "case by case criterion". For example, if the objective is to achieve a MRRR of 50%, it is of interest to find the most suitable combination of *n*, *m* and *c* values, or *n* and $Risk_{n\ sample}$ values. When doing this one should realize that

- there are various sets of parameter combinations that will lead to the desired MRRR value
- so far the only available method to find those is by "trial and error" using the simulation model developed for this project

To illustrate a search for an appropriate criterion, we explored how an MRRR of 50% can be achieved for the Swedish data collected for the EU baseline survey, as described in 3.1.1.

	prev bet (%)	prev within	mean (log cfu/g skin)	sd total	τ	φ
SE EU baseline	13.4	1	1.48	1.01	*1*	*30%*

After some preliminary exploration, ten scenarios were defined, with different values for the MC and CC parameters. Results of 30000 iterations are shown in Figure 11 and the table below.

It can be observed that, for the MC, many combinations of *n*, *m* and *c* lead to values of MRRR close to 50%. Higher values of *n* allow more variation in options for *c*, which allows a better fine-tuning of the MC. Assuming a limited precision of the enumeration methods of bacteria, values of *m* should not be chosen too precise. For example: The two scenario's with *n*=1 (*c*=0 and *c*=1) give values for MRRR that differ substantially. It is not possible to define a scenario for *n*=1 where they are closer together. With n=10, many more parameter combinations (*c*=0 ... 10) can be chosen, and therefore it is easier to find a value of MRRR closer to 50%.

As noted before, an increasing MRRR usually leads to a decreasing BNMC.

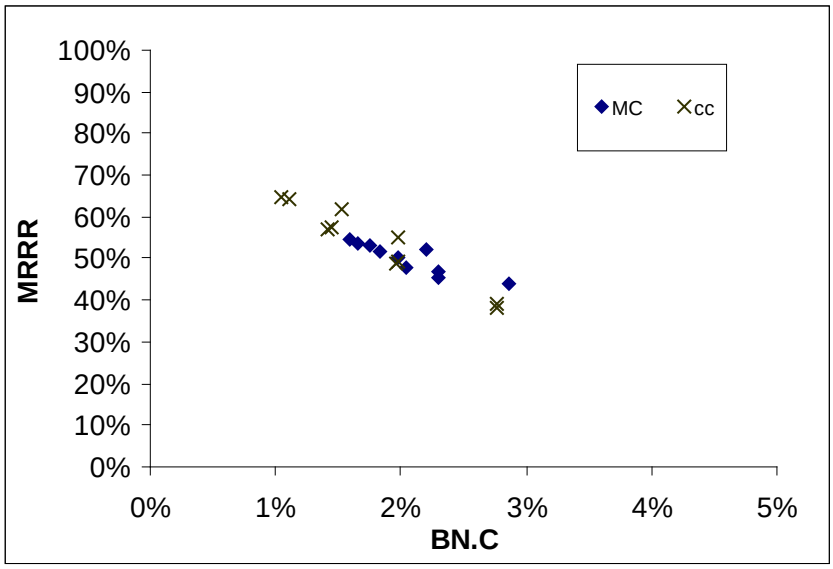

Figure 11. Results for ten different MCs / CCs applied to the Swedish EU baseline data, explore to find those that result in an MRRR = 50%. See Table below.

n	*m*	*c*	$Risk_{n\ sample}$	BNMC	MRRR MC	BNCC	MRRR CC
10	1000	1	0.006	1.8%	51.5%	2.8%	38.0%
10	500	3	0.008	1.6%	53.7%	2.0%	48.6%
10	200	5	0.010	2.1%	47.4%	1.4%	57.7%
10	200	6	0.012	1.6%	54.6%	1.0%	65.0%
5	1000	0	0.006	2.3%	46.7%	2.8%	38.9%
5	400	1	0.008	2.3%	45.4%	2.0%	49.2%
5	500	1	0.010	2.0%	50.0%	1.4%	58.0%
5	600	1	0.012	1.7%	53.6%	1.1%	64.8%
1	300	0	0.008	2.2%	52.5%	2.0%	55.1%
1	200	0	0.010	2.8%	44.6%	1.5%	61.9%

For the case by case approach, the results depend on *n* and $Risk_{n\ sample}$. Theoretically, for each *n*, there will exist a $Risk_{n\ sample}$ that yields an MRRR = 50%. When searching for the desired parameter value combination to define the Microbiological Criterion and/or "case by case criterion" one should, once again, realize that

- the MRRR only translates to a risk reduction if all batches are tested, and all non complying batches are withdrawn, or diverted away to undergo complete inactivation of all Campylobacter.
- the predictions are attended by various important uncertainties.

Therefore, definition of the criterion based on such a detailed modelling exercise only may not be appropriate.

Note that the same approach can be used if one aims at finding an MC/CC that yields a specific BN.C (BNMC or BNCC), for example 5%. Again, note that with a higher number of samples *n*, more precise MCs can be formulated:

n	*m*	*c*	$Risk_{n\ sample}$	BNMC	MRRR MC	BNCC	MRRR CC
10	100	3	0.001	4.8%	21.3%	8.9%	4.5%
10	100	2	0.002	5.8%	15.8%	6.7%	11.3%
10	200	1	0.003	5.1%	19.8%	5.1%	18.9%
10	400	0	0.004	5.1%	20.3%	4.1%	26.0%
10	500	0	0.005	4.6%	24.0%	3.3%	32.4%
5	300	0	0.002	4.7%	24.2%	6.5%	12.9%
5	200	0	0.003	5.6%	18.2%	5.0%	20.5%
3	100	0	0.0025	6.4%	14.6%	5.6%	18.1%
2	100	0	0.0025	5.6%	20.0%	5.4%	20.5%
1	100	0	0.0025	4.1%	32.0%	4.9%	26.1%

4.2.3. Residual risk if non-complying batches undergo non-lethal treatment

The assumption in the analyses is now that all non- complying batches undergo treatment that is lethal for all *Campylobacter*s. Alternatively, products may get a treatment that gives a reduction in the concentration, like freezing.
In that case the percentage of non complying batches will be the same, but the residual risk, MRRR, is larger. The table below gives the results for a hypothetical 1 log reduction.

	BNMC	MRRR MC	BNCC	MRRR CC
EU baseline SE	1.1%	75.5%	1.4%	70.1%
SLV	2.9%	64.5%	3.7%	58.6%
Hansson 1º	8.4%	46.8%	9.6%	41.7%
EU baseline DK	9.8%	55.0%	11.1%	51.6%
retail data	3.4%	72.3%	4.1%	67.8%
EU baseline FI	0.3%	72.4%	0.4%	69.2%
prof. kitchens	0.0%	100.0%	0.0%	100.0%
EU baseline NO	0.0%	100.0%	0.0%	100.0%
kalkunkylling	0.0%	98.4%	0.1%	96.7%
Stern 1º	2.6%	62.3%	2.9%	59.4%

4.3. Uncertainty analyses

4.3.1. Uncertainty analyses by Monte Carlo analysis

Some uncertainty analyses have been performed by Monte Carlo analyses as described in Appendix 3. The models were run with varying values for the prevalence, mean concentration, standard deviation and the factor ϕ, expressing the uncertainty about the microbiological data. The uncertainty of the risk models was studied by running the model for different consumer phase models. Next, the uncertainty analyses were performed for different data sets, and different microbiological criteria (that is varying sample sizes and critical limits).

The results showed that , in general, the uncertainty attending the "case by case" (CC) approach is smaller than that of the "Microbiological criteria" (MC) approach. As expected, uncertainty increased if the data set was smaller. As previously found, the choice of the CPM has larger impact on the absolute risk estimate than on the relative risk, but the effect of the CPM on the residual risk estimate, MRRR, is still considerable, especially for the MC approach. However, for microbiological criteria the impact of the choice of the CPM seems smaller than the impact of the microbiological data sets used in this study.

A smaller sample size *n* seems to have little impact on the performance of the model as read from a BNCC-MRRR plot. Still, with *n*=1, the agreement between the risk estimate for the batch, based on the (single) sample(s), and the risk estimate for the whole batch, is weaker than with *n*=5. This implies that with low sample size *n*, errors in identifying "high risk" batches are more likely.

The uncertainty analysis suggests that the "case by case" approach is less sensitive for the uncertainty attending the data and the risk model than the "risk based microbiological criterion" approach. The latter approach, however, has the advantage that the predicted percentage of non-complying batches does not depend on the risk model, so the uncertainty attending this model has no impact on this result.

It should be stressed that the uncertainty about the dose response model is not yet studied here, among others because it is quite complicated to quantify this uncertainty.

The method for uncertainty analysis applied here can be used to assess the uncertainty of the performance of microbiological criteria or the "case by case" approach for any data set is considered for application, and it can inform the user about the suitability of the data set.

4.3.2. Uncertainty analysis: a Bayesian approach

In this approach, model parameters are estimated jointly by their posterior probability distribution. Probability is thus used and interpreted as a degree of uncertainty which depends on the stated evidence (=data) in each case. The Swedish data sets provided a

pilot example in which this could be explored. These data were from two published sources, each insufficient alone for estimating the complete set of parameters. However, these data sets could be combined for the final analysis showing an example of Bayesian evidence synthesis. Posterior distributions of model parameters were also computed separately from the two data sets, which shows how the parameter distributions behave in each case. In the Bayesian approach, uncertainty depends explicitly on the stated evidence, and therefore changes when the evidence (data) is changed. With growing evidence, the distributions get narrower and the uncertainty decreases. With insufficient evidence they are wider. Even then, the data may impose partial evidence about a set of parameters, so that e.g. the sum of two parameters can be constrained by that. The advantage is that all parameters are considered jointly, and their uncertainty distribution produced probabilistically coherently, to describe what can be inferred from the explicit data. In this way, estimating each parameter in isolation of others by ad-hoc methods is avoided, and their combined uncertainty becomes described. When similarly structured data sets are available from different countries, it is then possible to produce posterior distributions for each country individually, based on the evidence from each without borrowing parameter estimates from other countries or remote conditions. The approach in short: ‘uncertainty of parameters is computed, given the known data, not the distribution of observables (or other quantities), given the unknown parameters’
Based on the posterior distribution of parameters, further predictions were made down to the probability of illness (dose-response). These further predictions take as an ‘input distribution’ the posterior distribution of the above mentioned parameters. However, the predictions describe a food preparation process and a consumption pattern based on assumed scenarios. These reflect mainly the scenario of Danish consumption amounts of young adults, the cross contamination experiment with salad making, and finally the assumed dose response model itself. The final risk estimate largely depends on these assumptions (among others), but the uncertainty distributions can be reflected through them. Finally, it is possible to study batch level parameters in the light of the batch specific result of the Microbial Criterion ('MC unknown' vs 'compliance with the MC' vs 'non- compliance with the MC'). This added evidence will again change the posterior distribution of the parameters, which in turn is reflected on risk estimates concerning such batch. Accomplishing the Bayesian approach required re-defining and re-programming essential model structure in OpenBUGS, and analyzing the example data sets. First results from this are reported in the appendix and further studies are planned to be published.

4.3.3. Using raw data vs. aggregated data

As described in section 2.5, all data sets are parameterised in the same way, by assuming the log concentrations are normally distributed, characterized by a mean, a between flock variance and a within flock variance. One important assumption here may be that the

within flock variance is the same in each flock. However, two data sets provide data from a set of flocks, and can be used in another approach where the empirical data are used to assess the risk, and to simulate the MC or Case by Case risk assessment. In the Monte Carlo simulation method a random flock is sampled first, and then *n* meat samples are taken from the normal distribution describing the concentrations in that flock. The objective is to sort out whether this will give the same MRRR / BNMC values.

4.3.3.1. Hansson data set

The original data set (where we sample from the flocks in the table in 3.1.3.) was compared with the aggregated data, where we fit a normal distribution and use a constant φ.
The risk estimates for both methods are almost identical (1.77% vs. 1.79 %).
Results for MRRR and BNMC/ BNCC are given in the tables below and in Figures 12 and 13. The two approaches are comparable, although the percentage of non-complying batches (BN.C) is a bit higher for the raw data set.

				raw data		aggregate data	
n	*m*	*c*	τ	BNMC	MRRR	BNMC	MRRR
10	1000	0	1	12.4%	3.7%	11.1%	6.1%
10	1000	1	1	11.1%	8.0%	9.7%	10.8%
10	1000	2	1	9.9%	13.0%	8.7%	15.6%
5	1000	0	1	11.4%	7.6%	10.0%	10.4%
5	1000	1	1	9.2%	17.3%	8.0%	19.3%
5	1000	2	1	7.3%	28.9%	6.3%	29.5%
1	500	0	1	9.3%	19.7%	8.3%	20.8%
1	1000	0	1	7.0%	35.4%	6.3%	33.5%
1	5000	0	1	2.1%	79.7%	2.3%	68.1%
5	1000	1	0.5	9.2%	19.0%	8.0%	22.2%

			raw data		aggregate data	
	threshold					
n	risk	τ	BNCC	MRRR	BNCC	MRRR
10	0.005	1	12.3%	3.8%	12.2%	2.8%
10	0.010	1	10.5%	9.7%	9.3%	12.4%
10	0.015	1	7.9%	23.9%	6.8%	25.9%
5	0.005	1	12.2%	4.1%	12.1%	3.1%
5	0.010	1	10.3%	11.4%	9.2%	13.4%
5	0.015	1	8.0%	24.1%	6.7%	26.8%
1	0.005	1	11.3%	8.7%	11.0%	8.9%
1	0.010	1	9.3%	19.7%	8.3%	20.9%
1	0.015	1	6.9%	36.3%	6.2%	34.0%
5	0.010	0.5	12.1%	5.0%	12.1%	3.9%

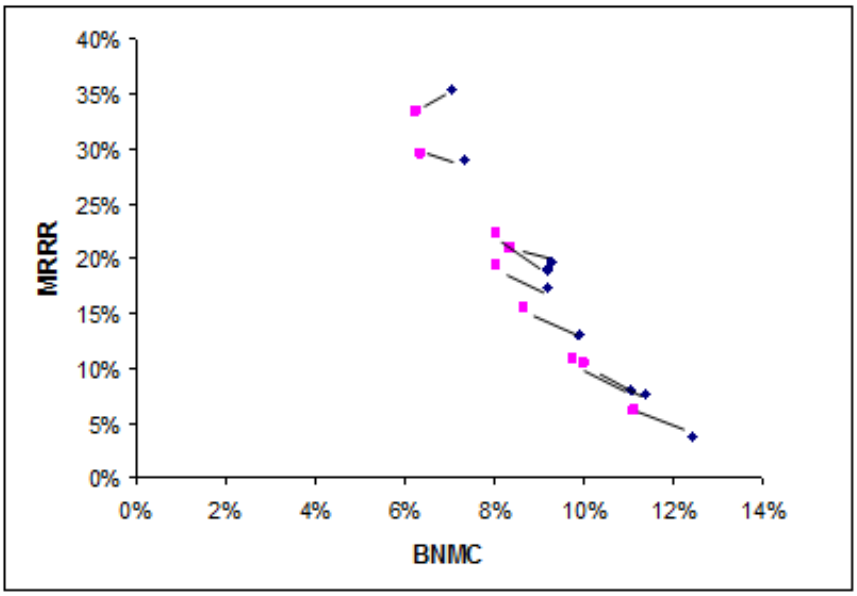

Figure 12. Results for BNMC and MRRR for the Hansson et al data set, using the raw data (blue diamonds) and aggregate data (pink squares). The lines connect the point with similar criterion.

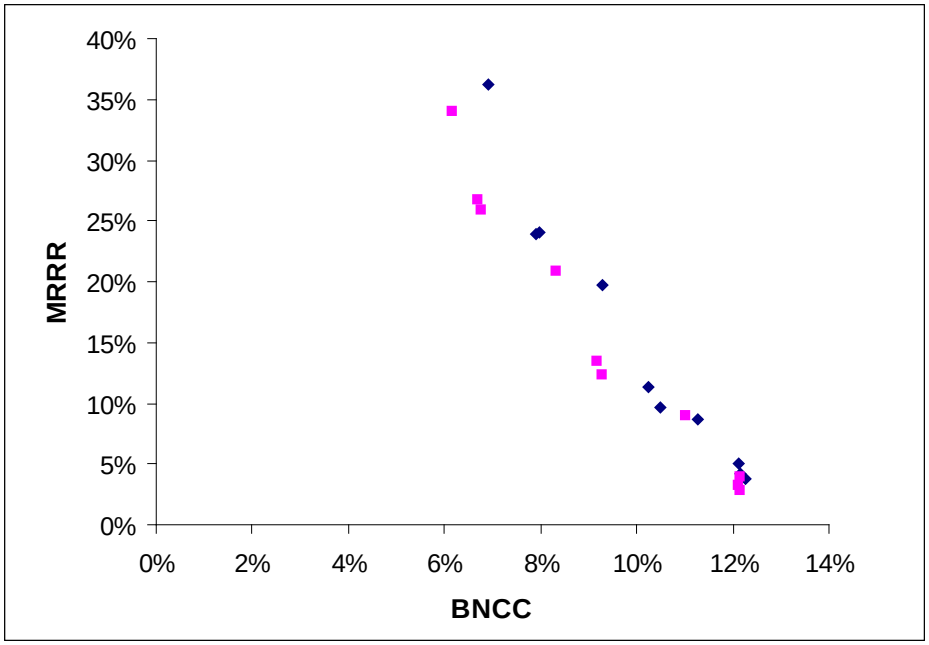

Figure 13. Results for BNCC and MRRR for the Hansson et al data set, using the raw data (blue diamonds) and aggregate data (pink squares).

4.3.3.2. Stern et al data set.

The same comparison is performed for the Icelandic data set (3.5.1.), where the unknown between flock prevalence is assumed to be 14%.
This gives similar results, although the difference in results between the two data sets is larger: the aggregate data give MRRR values that are considerably lower. Risks are similar again (1.04 and 1.06% for the raw data set and aggregate resp.).

				raw data		aggregate data	
n	*m*	*c*	τ	BNMC	MRRR	BNMC	MRRR
10	1000	0	1	8.0%	17.3%	7.3%	16.5%
10	1000	1	1	5.7%	31.2%	6.1%	23.3%
10	1000	2	1	4.3%	43.7%	5.3%	30.3%
5	1000	0	1	6.4%	28.0%	6.4%	22.5%
5	1000	1	1	3.7%	52.1%	4.7%	38.6%
5	1000	2	1	2.1%	73.9%	3.2%	58.6%
1	500	0	1	3.9%	59.4%	4.6%	45.3%
1	1000	0	1	2.7%	71.9%	3.4%	60.8%
1	5000	0	1	1.0%	98.1%	1.2%	96.1%
5	1000	1	0.5	3.7%	57.0%	4.7%	43.0%

			raw data		aggregate data	
	threshold					
n	risk	τ	BNCC	MRRR	BNCC	MRRR
10	0.005	1	7.5%	18.7%	7.1%	16.2%
10	0.010	1	4.1%	44.6%	5.6%	26.6%
10	0.015	1	2.4%	65.9%	4.1%	45.0%
5	0.005	1	7.2%	22.1%	7.1%	16.9%
5	0.010	1	4.1%	46.9%	5.4%	30.8%
5	0.015	1	2.5%	67.3%	3.9%	48.7%
1	0.005	1	5.8%	43.8%	6.5%	27.6%
1	0.010	1	3.8%	60.0%	4.6%	45.3%
1	0.015	1	2.7%	72.4%	3.4%	61.1%
5	0.010	0.5	7.0%	26.9%	7.1%	20.9%

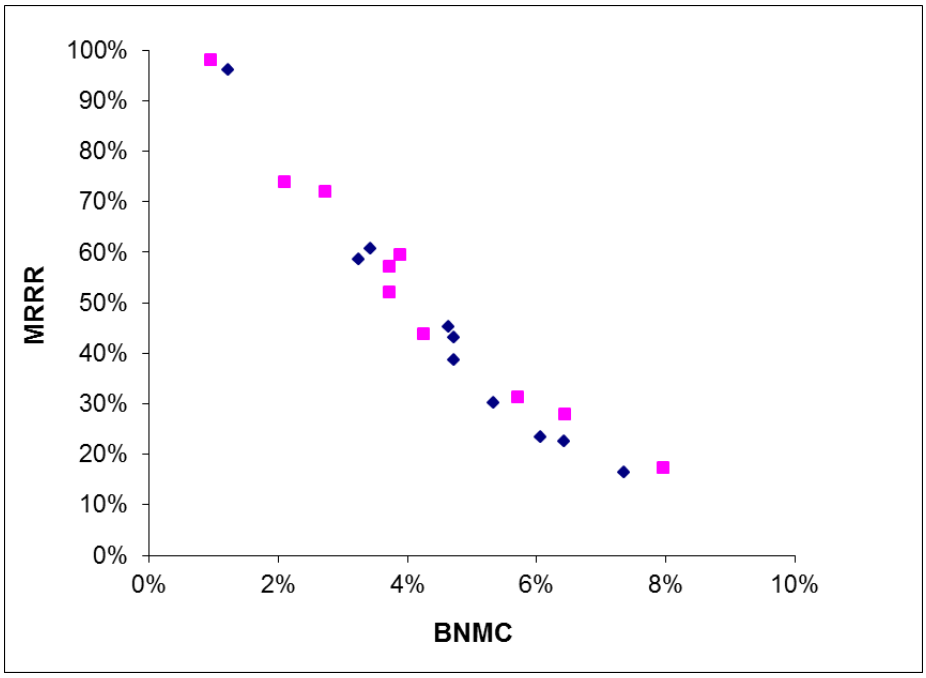

Figure 14. Results for BNMC and MRRR for the Stern et al data set, using the raw data (blue diamonds) and aggregate data (pink squares).

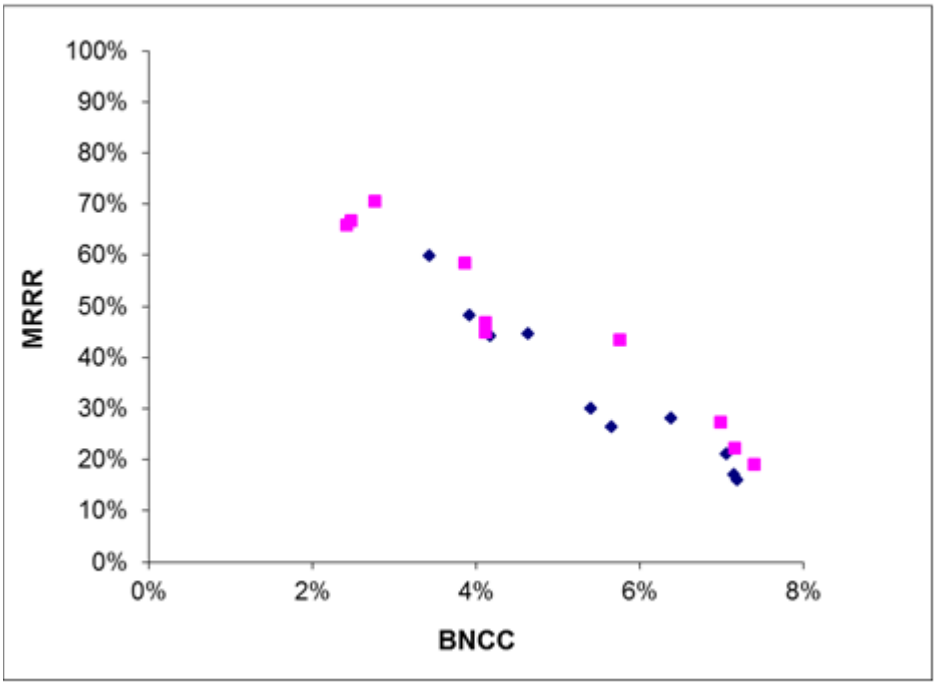

Figure 15. Results for BNCC and MRRR for the Stern et al data set, using the raw data (blue diamonds) and aggregate data (pink squares).

5. Discussion

5.1. Risk estimates in different Nordic countries

Public health risk estimates have been obtained for different Nordic countries, using the same QMRA model for all of them. It shows that these risk estimates are different for different Nordic countries and for different data sets within the countries. The differences between the Nordic countries show the same tendency as observed on the basis of the EU baseline survey (EFSA 2011), with very low risk estimates for Finland and Norway, and a relatively high risk estimate for Denmark (which, however, is still low compared to many other EU countries).

The variation in risk estimates for different data sets can be explained in many ways. Apart from real differences in prevalence and concentrations of *Campylobacter* on the meat in different countries and during different studies, methods of microbial analysis and the types of sample (skin, carcass rinse, etc.) are different, the sample set need not be representative for the whole year and may also be biased if for example only flocks positive at the farm are sampled. Also, the fitting of normal distributions to the data may not always be appropriate, for example if there are too little quantitative data available. Finally, one should realize that the risk estimates obtained from this sort of QMRA are usually imprecise, due to the uncertainty about the dose response relation, the uncertainty about the impact of consumer behaviour, and the translation of concentration data on skins and carcasses into concentrations on the meat. It is therefore best to only consider the risk estimates relative to each other, and not as absolute values.

5.2. The impact of MC and CC

The impact of setting microbiological criteria (MC) and the case by case risk assessment methodology (CC) is compared for one specific MC and CC, with sample size n=5. It shows that, in this case, MC and CC perform similarly. For the chosen criteria, the minimum (relative) residual risk is a bit lower and the percentage of non complying food lots is a bit higher for the CC, but this can change for different criteria. The results do not suggest that one method is performing better (more efficient in reducing the risk) than the other. However, the uncertainty analysis using Monte Carlo analysis (Appendix 3) suggests that the uncertainty attending the case by case risk assessment methodology is probably smaller.

The results also show that if the current risk is higher, the percentage of non-complying batches will also be higher, and the benefit in terms of low MRRR is higher as well. However, the absolute minimum residual risk is, in general, still higher for those sets of food lots where the risk was initially highest.

An important assumption behind the applied method is that all food lots are tested, all non complying batches are identified and corrective action is taken for all of them. This is

probably not a realistic scenario with the application of culturing techniques. If only X% of the food lots is tested, both the percentage of non-complying batches (BN.C) and the value for 1- MRRR will be reduced by X%, so the direct impact of the MC or CC will be smaller. However, the MC or CC may have an additional effect, when producer make extra efforts to reduce the probability that their food lots are non-complying. For example, Christensen et al. (2013) show that the case by case risk assessment approach may be effective in this way. It is, however, difficult to quantify this additional effect.

5.3. Strengths and weaknesses of the method

The method described in this report, and the analyses of the data, show how criteria can be derived to secure microbiological food safety. By using quantitative microbiological risk assessment, the impact of these criteria can be evaluated in terms of reduction of public health risk. The comparison of the impact of different criteria on both the expected risk reduction and the percentage of non-complying batches, offers a nice toll for risk managers to compare different options. The method does not require definition of Food Safety Objectives or Performance Objectives, and is therefore simpler than methods to establish risk based microbiological criteria that do require such definitions (Van Schothorst et al. 2009). The advantage above criteria that are not risk-based (i.e. those that are not derived based on a formal risk assessment) is that the impact can not only be evaluated against the expected percentage of non-complying batches (BN.C), but also in terms of expected risk reduction.

When comparing the MC and CC approaches, the advantage of the MC approach is that it applies the terminology and methods applied for microbiological criteria (Van Schothorst et al 2009) and is therefore probably easiest to communicate. The advantage of the CC approach is that it requires less parameters to define the criterion, and that the attending uncertainty is probably smaller. The disadvantage of the CC approach may be that it is harder to define the critical (relative) risk, and/or the baseline.

The methods apply risk assessments, and therefore the uncertainty attending the estimated effects on public health risk is substantial. The risk assessment is based on some assumptions that need further research (like the transition factor τ). For some people, numerical estimates may suggest an unjustified precision. However, the alternative for using risk assessments (based on the present scientific knowledge) as the basis for assessing the impact of criteria, is the use of surrogate measures for risk. As shown in section 4.1 such surrogates need not be well associated with risk estimates.

5.4. Relative risk estimates: definition of the baseline risk

To perform the case by case risk assessment in a similar way as it is currently done in Denmark, a baseline has to be defined (see 2.4). This baseline represents the current mean risk and reflects the current situation. In Denmark the baseline has for example

been based on retail data collected in a monitoring programme 2005-2009. The relative risk RR_b then gives the risk of the analysed food lot compared to the mean risk over those years. Risk managers can decide whether this RR_b is too high or not.
This baseline has not been defined in the present study. It may be identical for all Nordic countries, or it may be country specific, and that choice (which is a risk managers decision) has not been made. The advantage of this approach is that the risk manager need not decide on the acceptability of food lots on the basis of a highly uncertain risk estimate, but on basis of a less uncertain relative risk, that evaluates the risk against a current risk (for example: the risk from this food lot would be ten times larger than the average yearly risk from the meat consumed in the country).

5.5 Future perspectives

This report describes the potentials for setting risk based criteria for *Campylobacter* in broiler meat, to show the principle of setting such risk based criteria. *Campylobacter* in broiler meat was chosen for its significance for public health, and because a risk assessment model is available. In the future, the same approach can be applied to other pathogens and food products, for example for *Salmonella* on pork.
An interesting finding of the analyses described in this report is that sampling plans with a low sample size (like n=1) are almost equally efficient as those with a larger sample size. This is somewhat counterintuitive, and should be explored in more detail. An obvious approach would be to do more detailed performance analyses of the different sampling plans that include an analysis of the percentages of false negatives and false positives in terms of compliance, as a function of the sample size.
The analyses performed in this study, supported by the experience with the case by case risk assessment in Denmark (Christensen et al 2013), show that a future implementation of a microbiological criterion, or a case by case risk assessment, may result in a considerable reduction in the human incidence of campylobacteriosis obtained from the consumption of broiler meat. A more stringent criterion will result in a larger risk reduction, but at the expense of a larger amount of non complying batches of broiler meat. The optimum balance between risk reduction and costs of non-complying batches is one that has to be taken by the food safety risk managers.
If it would be decided to implement a criterion as described in this report, it would be important to study the effect on the basis of data on prevalence and concentrations of *Campylobacter* on the meat, as well as human incidence of campylobacteriosis among consumers.
Application of the methods described in this report may be complex for those with limited experience in mathematical modelling and risk assessment. Therefore, a freely available software tool will be developed, that allows the users to assess the effect of setting microbiological criteria, and/or application of the case by case risk assessment approach for *Campylobacter*, based on the data sets considered appropriate by the user.

This has been made possible by a new research grant, kindly provided by NMDD in 2013.

Acknowledgements

This study is performed in the project “Establishment of Risk Based Microbiological Criteria in the Nordic Countries”, subsidized by Nordisk Ministerråd, Nordisk Arbetsgrupp för Mikrobiologi, djurhälsa och djurvälfärd (NMDD), including a special additional grant for uncertainty analysis.

References

Anonymous. 1997. Recommended international code of practice – General principles of food hygiene. CAC/GL 21-1997: Rome: Codex Alimentarius Joint Office.

Black RE, Levine MM, Clements ML, Hughes TP and Blaser MJ, 1988. Experimental *Campylobacter jejuni* infection in humans. J Infect Dis, 157, 472-9.

Boysen, L., Nauta, M., Duarte, A.S.R, Rosenquist, H. 2013. Human risk from thermotolerant Campylobacter on broiler meat in Denmark International Journal of Food Microbiology. in press.

Christensen, B., Sommer, H., Nielsen, N., Rosenquist, H., 2001. Risk assessment of Campylobacter jejuni in chicken products. 138 pp. The Danish Veterinary and Food Administration, Copenhagen, Denmark.

Christensen, B.B., Rosenquist, H., Sommer, H.M., Nielsen, N.L., Fagt, S., Andersen, N.L.,Norrung, B., 2005. A model of hygiene practices and consumption patterns in the consumer phase. Risk Analysis 25, 49–60.

Christensen, B., Nauta, M., Korsgaard, H., Sørensen, A.I., Rosenquist, H., Boysen, L., Perge, A., Nørrung, B. 2013 Case-by-case risk assessment of broiler meat batches: An effective control strategy for Campylobacter**,** Food Control, 31:485-490.

EFSA, 2010. Analysis of the baseline survey on the prevalence of Campylobacter in broiler batches and of Campylobacter and Salmonella on broiler carcasses, in the EU, 2008; part B: Analysis of factors associated with Campylobacter colonisation of broiler batches and with Campylobacter contamination of broiler carcasses; and investigation of the culture method diagnostic characteristics used to analyse broiler carcass samples. The EFSA Journal 8 (8), 1522.

EFSA, 2011. Scientific Opinion on *Campylobacter* in broiler meat production: control options and performance objectives and/or targets at different stages of the food chain EFSA Journal 2011; 9(4):2105.

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.

Landis, J.R. and Koch, G.G. 1977. The measurement of observer agreement for categorical data. Biometrics 33 (1): 159-174.

Lindblad, M. Lindmark, H. Thisted Lambertz, S. and Lindqvist, R., 2006 Microbiological Baseline Study of Broiler Chickens at Swedish Slaughterhouses. Journal of Food Protection 69:2875-2882.

Lorimer, M.F. and Kiermeier, A., 2007. Analysing microbiological data: Tobit or not Tobit? International Journal of Food Microbiology 116, 313–318.

Nauta, M.J., Jacobs-Reitsma, W.F., and Havelaar, A.H., 2007. A risk assessment model for campylobacter in broiler meat. Risk Analysis 27: 845-61.

Nauta, M.J., Fischer, A.R.H., Van Asselt, E.D., De Jong, A.E.I, Frewer, L.J. and De

Jonge, R., 2008.Food Safety in the Domestic Environment: the effect of consumer risk information on human disease risks. Risk Analysis, 28, 179-192.

Nauta, M.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.

Nauta, M. and Christensen, B. 2011 The impact of consumer phase models in microbial risk analysis. Risk Analysis 31:255-265.

Nauta, M.J., Sanaa, M. and Havelaar, A.H. 2012 Risk based microbiological criteria for *Campylobacter* in broiler meat in the European Union, . International Journal of Food Microbiology 158:209-217.

Stern, N. J., Georgsson, F., Lowman, R., Bissailon, B.-R., Reiersen, J., Callicot, K.A. et al. 2007, Frequency and Enumeration of Campylobacter species from Processed broiler carcasses by weep and Rinse samples. 2007. Poultry Science 86: 394-399.

Teunis, P.F. and Havelaar, A.H., 2000. The Beta Poisson dose-response model is not a single-hit model. Risk Analysis 20, 513-520.

Van Schothorst, M., Zwietering, M.H., Ross, T., Buchanan, R.L., Cole, M.B., International Commission on Microbiological Specifications for Foods (ICMSF), 2009. Relating microbiological criteria to food safety objectives and performance objectives. Food Control 20, 967–979.

Abbreviations / model parameters

MC	Microbiological criterion
CC	Case by case risk assessment, or case by case risk assessment-based criterion
BNMC	Percentage of non-complying batches using the MC approach
BNCC	Percentage of non-complying batches using the MC approach
MRRR	Minimum Relative Residual Risk
MRR	Minimum Residual Risk
BN.C	Either BNMC or BNCC.
n	Number of samples taken per food lot
m	Threshold for bacterial concentration in MC
c	Number of samples, from n, that is allowed to have a concentration larger than m, for compliance in the MC
QMRA	Quantitative Microbiological Risk Assessment
τ	Transition factor: the difference between the log concentrations measured on the meat (or skin) during sampling, and the log concentration per g of meat at retail.
φ	The fraction of the observed variance in concentrations attributable to within flock variance.
prev	Prevalence of contaminated food lots in a set of food lots.
mu	Mean concentration (log cfu/g) in a contaminated food lot
sd	Standard deviation

Appendix 1. Baseline risk assessment: QMRA model

To assess the risk of illness associated with a chicken meat product with a concentration of *Campylobacter* C_{meat}, (in cfu per gram of meat) we apply a consumer phase model (CPM) referred to as the "Nauta CPM" by Nauta and Christensen (2010). In this model, the distribution of weights of consumed portions of chicken meat w_C is derived from Danish data for young adult males (Christensen et al., 2001; Christensen et al., 2005)

$w_C \sim$ Lognormal(189,127) (in gram) (eq. 1)

(where ~ means "is a sample from"; truncated at 1000 g)

To obtain a discrete number of bacteria in a raw portion of meat ($N_{portion}$), assuming a homogeneous random distribution in the meat, the load per raw portion is given by

$N_{portion} \sim$ Poisson($C_{meat} \times w_C$) (eq. 2)

For the transfer rates from meat at retail to portion of ready to eat meal a discrete distribution is used with an identical likelihood of each value. It is based on an observational study for volunteers preparing a salad containing (cooked) chicken breast (Nauta et al., 2008)

$-\log(p_{tr}) \sim$ [2.24, 2.36, 2.37, 2.58, 2.82, 2.86, 3.16, 3.17, 3.47, 3.52, 3.57, 3.83, 3.83, 3.84, 3.87, 3.89, 3.89, 3.90, 3.94, 4.03, 4.09, 4.42, 4.53, 4.54, 4.54, 4.62, 4.62, 4.68, 4.73, 4.76, 4.84, 4.92, 4.93, 4.95, 4.97, 5.20, 5.25, 5.27, 5.39, 5.47, 5.60, 5.83, 5.89, 5.95, 5.96, 6.02, 6.23, 6.38, 6.96, 7.37, 7.90, 8.20, 9.00, 9.00, 9.00] (eq. 3)

The ingested doses are then sampled from a binomial distribution

$d \sim$ Binomial[$N_{portion}$, p_{tr}] (eq. 4)

Using the widely applied dose response model for *Campylobacter* based on the data of Black et al (1988) (Nauta et al., 2009; Teunis and Havelaar, 2000) and a fixed probability of illness given infection of 0.33 (Nauta et al., 2007), the probability of illness with (discrete) dose d is given by the Beta-binomial model:

$$P_{ill}(d) = 0.33 \times \left(1 - \frac{\Gamma(\alpha+\beta)\Gamma(\beta+d)}{\Gamma(\beta)\Gamma(\alpha+\beta+d)}\right) \quad \text{(eq. 5)}$$

with $\alpha = 0.145$ and $\beta = 7.59$ (Teunis and Havelaar, 2000).

The model (eq. 1 – eq. 5) is implemented by first running eqs. (1) to (4) with 91 x 50,000 iterations in a Monte Carlo simulation model implemented in @Risk 5.5. (Palisade, Newfield, USA), for a range of 91 values of C_{meat} = {-2, -1.9, ... , 7}. Each iteration provides a value of the dose *d* (eq. 4), so that 50000 iterations allow us to approximate the distribution of doses as a function of the concentration in the meat, $P_{EX}(d| C_{meat})$, representing the probability of ingesting dose *d* when the concentration on the meat product at retail is $C_{meat.}$

With

$$Q_{ill}(C_{meat}) = \sum_{d=1}^{\infty}\left(P_{ill}(d) \times P_{EX}(d \mid C_{meat})\right) \quad \text{(eq. 6)}$$

using 91 values of C_{meat}, we can derive a vector of 91 values for the population risk as a function of the concentration on the meat, $Q_{\text{ill}}(C_{\text{meat}})$.

The input of this model is the concentration per gram of chicken meat C_{meat} in log cfu/g meat. However, the data available from the EU baseline survey are given as the observed concentrations per gram of skin (C_{skin}). As *Campylobacters* are generally believed to reside on the skin or on the meat and not within the meat, this C_{skin} is likely to be larger than C_{meat}. The uncertain difference between the observed C_{skin} and the model input C_{meat} is included in the model as an unknown constant transition factor τ, such that for the concentrations in any batch *b*

$$\log C_{\text{meat, b}} = \log C_{\text{skin, b}} - \tau \quad \text{(eq.7)}$$

Based on discussion with several European experts, we assumed as a default that $\tau = 1$. This implies that the concentration per gram of meat is one log unit less than the concentration per gram of skin.

Given a batch of chicken meat products with a distribution of concentrations, such that $g(C_{skin,\,b}\,;\,\mu_b)$[3] represents the probability density for the concentration on the skin $C_{skin,\,b}$ with mean μ_b in that batch, the population risk for that batch can be calculated as

$$Risk_b = \int_0^{\infty} Q_{ill}(C_{skin,b} - \tau)\, g(C_{skin,b}; \mu_b)\, dC_{skin,b} \quad \text{(eq.8)}$$

A within batch prevalence less than 100%, can easily be included in (eq. 8) by multiplying the integral with this within batch prevalence.

The mean risk for a set of batches, for example all those produced in one of the EU Member States, is a function of the prevalence of contaminated batches in the MS, p_{MS}, and the between batch distribution of the within batch distributions, which is characterized by the variation in the mean per batch, μ_b. With $f(\mu_b)$ representing the normal probability density of μ_b, the mean risk in a MS is

$$Risk_{MS} = p_{MS} \int_{\mu_b} f(\mu_b) Risk_b(\mu_b)\, d\mu_b \quad \text{(eq. 9)}$$

This integral is approximated by the summation in a Microsoft Excel spreadsheet splitting up the normal distribution in 61 bins with equal width.

[3] The distribution of concentrations on the skin is conditional on a set of parameters; only μ_b is given here because it is the only one that is varying between batches.

Appendix 2. Evaluating the Microbiological Criterion

The components of a microbiological criterion include a precise definition of the microorganisms and food products of concern, the sampling protocol and the corresponding microbiological analysis of the samples (Anonymous, 1997). Next, per food lot it defines

- the sample size *n*,
- the critical concentration *m* (maximum concentration)
- the critical number of samples *c*, that may yield a value larger than *m*.

In our study, we calculate the probability of compliance with the MC for all batches produced in each MS, and simultaneously assess the probability of illness for batches complying and batches not complying with the MC.

First, for each batch, we calculate the probability of non-compliance with the MC, $P_{\text{Batch,nc}}$, by evaluating how often more than *c* out of *n* samples will be found with a concentration larger than *m*. Assuming a perfect test and given the normal distribution of log concentrations in a batch with mean concentration μ_b and standard deviation σ_b, the probability that the concentration $C_{\text{skin,b}}$ is larger than *m*, $P_{>m}$, is calculated using the normal distribution cumulative density function $\Phi(.)$, so

$$P_{>\text{m}} = 1\text{-}\Phi\left(m, \mu_b, \sigma_b\right) \qquad \text{(eq. 11)}$$

If the within batch prevalence would be smaller than 100%, it can be included in (eq. 11) by multiplying the right hand side of the equation by this within batch prevalence, as for eq. (8).

Next, for each batch, the probability of getting more than *c* out of *n* samples not complying with the MC follows from the probability function of the Binomial (n, $P_{>m}$) distribution, and can be calculated using its cumulative probability function BINOM(.)

$$P_{Batch,\,nc} = P(Binomial(n, P_{>m}) > c) \; = 1 - \text{BINOM}(c, n, P_{>\text{m}}) \qquad \text{(eq. 12)}$$

For a set of batches, for example all those produced in one of the EU member states, with $f(\mu_b)$ representing the normal probability density function of μ_b

$$P_{MS,nc} = p_{MS} \int_{\mu_b} f(\mu_b) P_{Batch,\,nc}\, d\mu_b \qquad \text{(eq. 13)}$$

Hence, the probability obtained is a function of the distribution of concentrations found in the EU baseline survey and the values of *n*, *c* and *m*. It can be calculated by integration and is approximated in a Microsoft Excel spreadsheet splitting up the normal distributions in 61 bins with equal width. This probability is the expected percentage of batches that is not complying with the MC. $P_{\text{MS,nc}}$ is an important model result, referred to as Batches Not complying with the MC or **BNMC** in a MS.

Next, we assess the risk using the model described in 2.2. With this model, the probability of illness from a random serving is a function of the distribution of skin concentrations $C_{skin,b}$, as calculated with eq. 9. This mean risk can be assessed for every feasible set of batches, that is for each MS, for the whole of Europe, and for batches that do or do not comply with the MC.
For each MS, the maximum risk reduction that can be achieved by setting a MC, is the risk reduction obtained by testing all batches and subsequent rejection of all batches that do not comply with the MC. The risk associated with batches complying with the MC is

$$Risk_{MS,c} = p_{MS} \int_{-\infty}^{+\infty} f(\mu_b)(1 - P_{Batch,nc}) Risk_b(\mu_b) d\mu_b \qquad \text{(eq. 14)}$$

The quotient of the risk of all batches complying with the MC in a MS and all batches in a MS is referred to as the Minimum Relative Residual Risk (**MRRR**).

$$MRRR = \frac{Risk_{MS,c}}{Risk_{MS}} \qquad \text{(eq. 15)}$$

An assumption behind this is that complying batches pose no longer a risk, and are replaced by zero risk batches, they are not replaced by food lots from the population of non complying batches. This is not stated correctly in the paper Nauta et al 2012. If we do assume the rejected batches are replaced by those from the complying batches, we get

$$MRRR^{*} = \frac{Risk_{MS,c}}{Risk_{MS}(1 - BNMC)} \qquad \text{(eq 15a)}$$

For consistency we do not apply this here.

Appendix 3. Uncertainty analyses for Risk based Microbiological Criteria: A Monte Carlo approach

Introduction and Methods

Like all models, the presented models are built on a set of simplifying assumptions, and restricted data sets. The practical value of the results, and the uncertainty attending the model outputs, will depend on the impact of these assumptions and data. Here we analyse the impact of some of the sources of uncertainty by using Monte Carlo simulations.

Model outline and sources of uncertainty

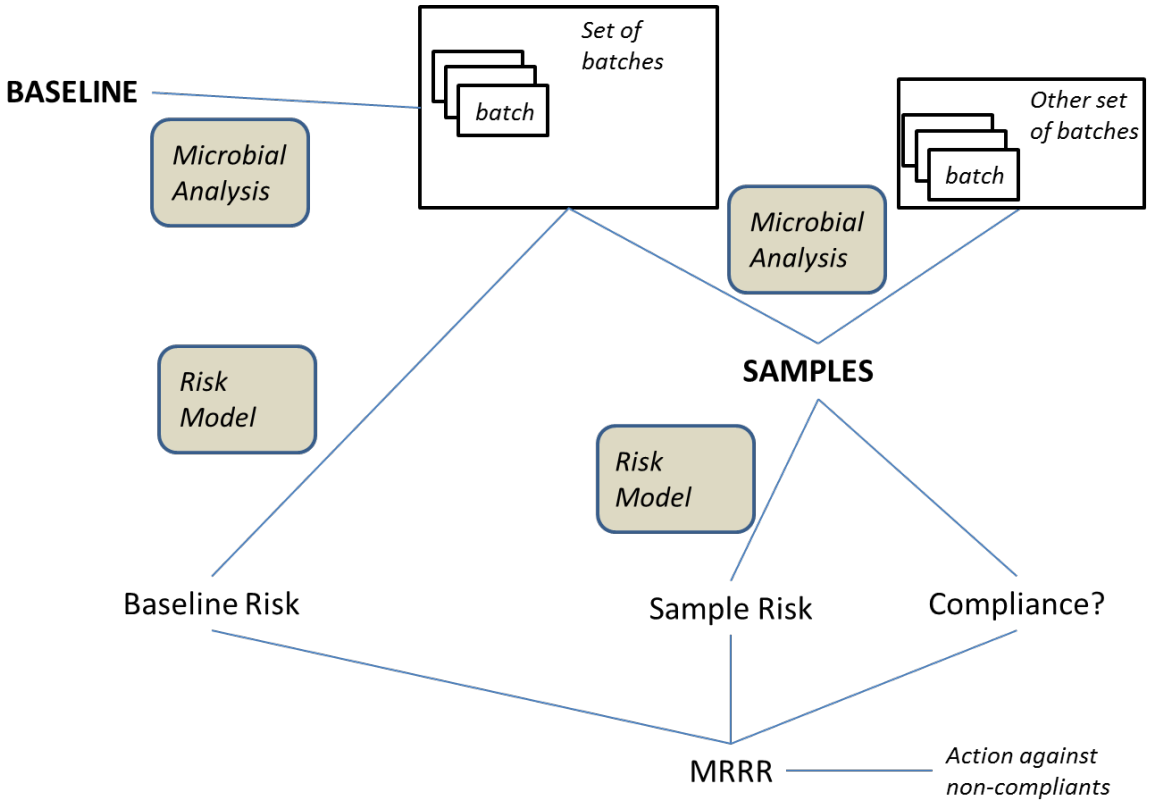

Figure U1. Schematic overview of the analysis of Microbiological Criteria. The Baseline and the Samples (the sampling results) depend on the Microbial analysis methods. The risk estimates for the baseline and the samples depend on the Risk Model used. For the case by case (CC) risk assessment, the compliance is also tested with the Risk Model, but in the "Microbiological Criteria" (MC) method it is not.

The elements of the model used for evaluating risk based Microbiological Criteria and the ”Case by Case” method are shown in figure U1, and described below. The sources of uncertainty are given in italics.

1. For the baseline data, k_{bas} batches are tested, usually with n_{bas}= 1 sample per batch. In the analyses done for the NMDD project these are the only data used (apart from some data used for the risk model).
2. The (semi-) quantitative data from the baseline are used to estimate the prevalence of contaminated batches, p_{bas} (= p_{MS} in Nauta et al 2012).
3. The true within batch prevalence is assumed to be 100%. This seems to be a fair assumption for *Campylobacter* in broiler meat. In some samples *Campylobacter* may not be found due to a low concentration, though. In that case the observed within batch prevalence may be < 100%.
4. It is assumed that the measured log concentrations of the contaminated samples are normally distributed. The baseline data are fitted to this distribution[4], which gives a mean μ and standard deviation σ_{tot}. Then, the mean per batch, $\mu_b \sim N(\mu, \sigma_{bet})$ and the concentration $\log(C) \sim N(\mu_b, \sigma_{wit})$ within batch. So we assume that the within batch variance is identical in all batches in a set of batches, and that a fraction φ of the variance is attributable to within batch variance, such that the within batch variance $\sigma_{wit}^2 = \varphi\ \sigma_{tot}^2$ and the between batch variance equals $\sigma_{bet}^2 = (1-\varphi)\ \sigma_{tot}^2$.
5. *With the translation from data to distribution one source for uncertainty is the microbial analysis: measurement error (including imperfect recovery) and stochasticity (randomness in plate counts, and the probability of a heterogeneous distribution of cells in the samples).*
6. *Next, the assumption about the “true distribution” (e.g. log normality, fixed within flock standard deviation) is a simplification. The parameter φ is not well known. These are additional sources of uncertainty*
7. When testing batches for compliance, n samples are taken per batch, from k_{test} batches. Here either the same microbial analysis is performed as for the baseline, or another one. If the same method of microbial analysis is used, potential biases in the method will cancel out; this need not be the case if different methods are used. *Uncertainties mentioned under 5) are relevant here as well. Another source of uncertainty comes from stochasticity in taking random samples, which is a binomial process.*
8. *To establish the baseline risk, a risk assessment model is used that covers the food chain from the sampling point to the estimate of the risk per serving. This risk model contains many assumptions and simplifications, and in general it is hard to quantify the uncertainty.*
9. *The risk assessment model applied for Campylobacter includes*

[4] The method of fitting is not prescribed here. MLE is an option

- *A transition from concentration on the sample (e.g. skin) to concentration per g consumed meat, at the point of sampling[5], the factor τ.*
- *A consumer phase model that describes exposure as a function of the concentration on the meat at the point of sampling, including assumptions on the method of preparation and the effect thereof, and the portion size. This is a source of uncertainty, which includes both the uncertainty attending the data behind the model, representativity of the data, and potential model misspecification.*
- *The dose response model. This is based on one data set, and not perfect.*

10. The same risk model is used to assess the risk of a sample. This may be a sample from the set of batches behind the baseline, but it can also be from an independent batch.
11. The compliance of the batch (BN.C) is determined on the basis of the n samples. This is a predefined rule, either based on the numbers c and m (MC approach) or on the basis of the risk estimate (CbC approach).
12. The minimum relative residual risk is defined on the basis of the sample risks, the baseline risks and the compliance. Here it is important to be aware of the assumption about the action taken to batches that do not comply.

The challenge is to describe the uncertainty about the MRRR and BN.C. There may be a correlation between the two, and this should be considered.

Uncertainties

Roughly, uncertainties may be a consequence of (1) the uncertainty about the data, that are obtained from a microbial analysis (grey box in figure U1) or (2) the risk model (another grey box in figure U1).

Microbial analysis

With the microbial analysis we obtain plate count data that inform us on the concentrations and the prevalence in the set of batches.
If you have a sample with a number of bacteria on it, the count from the sample should reflect what really is in the sample. But there is uncertainty because

- An individual viable bacterial cell is not necessarily a colony forming unit (cfu). What exactly is a cfu may depend on the growth medium used. This gives some uncertainty that we will ignore.
- The sample is assumed to be representative for the sampled product, so a homogeneous distribution of cells over the product is assumed. This gives some uncertainty that we will ignore.
- When doing a dilution series and plate counts, we assume good homogenisation of the samples, and a Poisson process. In principle the uncertainty about the count can

[5] The meat is not consumed at sampling. So it is the concentration per g of meat that will be consumed after preparation

be quantified under this assumption. However, very often we do not know the original plate counts, only the result of the interpretation as prescribed in the protocol for microbial analysis (e.g. ISO 10272-2 used for Campylobacter quantification as used in the EU baseline study). Therefore it is hard to quantify the uncertainty about the counts.

- We can ignore all before mentioned uncertainty because it cannot be quantified, and assume that the concentration estimates from the plate counts from the samples do reflect the true numbers on sampled the food products. Then there is uncertainty about the prevalence p_{bas}, mean μ and variance σ_{tot}^2 from the fact that you have a limited data set. These three parameters can be estimated in different ways.

  - Directly from the data, taking the observed prevalence $p_{bas,\ obs} = N\ (C>0)/k_{bas}$, with N(C>=0) the positive samples from enumeration and plate counts, as in the EU baseline; the mean log C (cfu/g) found, m, and the variance therein, s^2.
  - Doing a maximum likelihood fit, like e.g. done in Nauta et al 2012, and in the NMDD report. The advantage of this method is that you can accommodate your estimate for censored data instead of assuming a value for C for the results above or below the limit of detection. Disadvantage is that it is a bit more complicated, and builds on an assumption on the limit of detection, LoD. Also, you can decide to include your estimate of the prevalence in it or not. Another disadvantage is that it is more complex to express your uncertainty about your estimate.

- The microbial analysis is performed twice, once for the baseline and once when sampling for the criterion (see Figure 1). I am not sure what the impact of this will be

  - If the same microbial analysis is performed twice, the bias in the method will be partially cancelled out: If your count is lower than what is really in the product, it will be lower both in the baseline and in the sample, so you compare apples with apples. But
    - If there is a lack of precision, variability between counts of the same product, this will result in a difference between “baseline sample” and “sample sample”;
    - The risk model may be based on yet another microbial analysis. So there may be a bias there, as well as a lack of precision that has an impact.

  These issues demand a substantial amount of research, which does not fit in the current project. Therefore it is not included here.

  - If two different methods for microbial analysis are used, the biases may be different, and they need not cancel out. This will increase the uncertainty, but it will be hard to quantify this.

It can be concluded that there are quite a number of uncertainties attending the microbial analysis, which cannot be easily quantified. Some can be, if the raw count data are available, some will demand more fundamental research.

Method used:

We choose to assess the uncertainty from the microbial analysis by describing the statistical uncertainty attending the data, and to study its impact on MRRR and BN.C.

1) For estimating the parameters describing the contamination of the set of batches in the baseline (Going from BASELINE to Set of batches):
   a. $P_{bas} \sim$ Beta(N(C>0)+1,k_{bas}-N(C>0)+1)
   b. $\mu \sim m + Tdist(N(C>0))\cdot s/\sqrt{(N(C>0))}$
   c. $\sigma_{tot}^2 \sim s^2(N(C>0)-1)/ChiSq(N(C>0)-1)$

   Using the Beta distribution (Beta), the Student T distribution (Tdist) and the Chi Squared distribution (ChiSq).

   We also need an estimate for the uncertainty about φ, and there we choose a BetaPert[6]:

   d. $\varphi \sim$ BetaPert(0.1, 0.3, 0.5), an arbitrary choice, based on the assumption that the between batch variation will be larger than the within batch variation
2) For the sampling as a basis of the criterion:

   Apart from the stochasticity as already implemented in the model (the random sampling of samples from the products) there is a Poisson process of counting. Assuming plate counts with a dilution d (so that there is a "limit of detection" L =1/d), the plate counts gives

   $C_{plate} \sim$ Poisson(C/L)*L cfu/g

   with C the sampled true concentration

   Detection limit was set at 100 cfu (so L=100, the dilutions were such that if one cfu was counted, this implies 100 cfu/g.)

Risk Model

In general, there is quite some uncertainty attending the risk estimates from "farm to fork" risk assessment models, and this uncertainty is hard to quantify.

Here the impact of the choice of the CPM is studied, analogous to the approach of Nauta and Christensen (2011). Three CPMs are compared: The FAO, Nauta and van Asselt CPM, which are considerably different (Nauta and Christensen 2011), see Figure U2.

[6] the impact of variation in the within flock standard deviation was explored. It seems to have little impact and therefore it was decided not to study it any further.

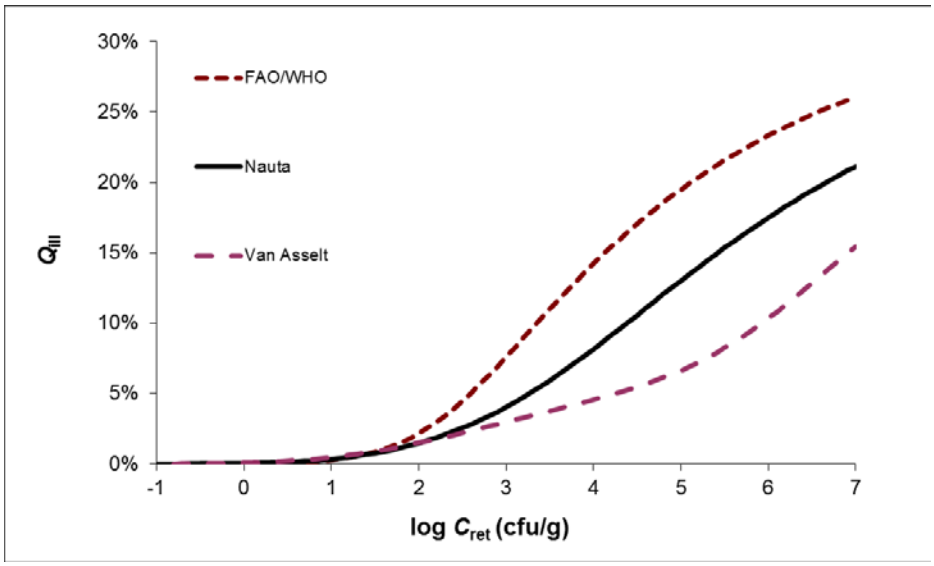

Figure U2. Three consumer phase models, predicting the probability of illness (Q_{ill}) per serving associated with the consumption of chicken meat with the indicated concentration on the meat at retail (See Nauta and Christensen 2011).

Simulations

Monte Carlo simulations were performed to investigate the impact of the uncertainty described above on the MRRR and BN.C results.

As a baseline MC / CC scenario we choose n=5, m=1000, c=1 and a critical relative risk RR_{crit}=1.5. The data set primarily studied was the DK EU baseline data set, with 123 observed concentrations from 396 samples, comparative analyses were done with the Swedish SLV data set (Lindblad et al. 2006) with 88 observed concentrations from 617 samples and the Finnish EU baseline data set with 21 observed concentrations from 411 samples. Alternative MC/CC were studied as well.

The uncertainty in the data sets was implemented by using 200 sets of samples for m, s, p and ϕ, and running the model with 5000 iterations (expressing the variability in sampling) for each of these 200 sets, to get 200 estimates of the mean risk, mean MRRR and mean BN.C, reflecting the uncertainty.

Additional simulations were done to explore the impact of changes in τ (default value τ=1) and the impact of leaving out the uncertainty in ϕ.

For the CC, an additional analysis was performed to study the benefit of increasing n, by considering how well the individual risk estimates of each batch ($R_{b,\,est}$, based on n

samples) agree with the "true" risks, that is the risk estimates based on the (known) whole distribution within the batches (R_b). This is done by comparing the relative risk estimate based on the samples, $RR_{b,\,est}$, with the relative risk estimate based on the "true distribution", RR_b [7]. If the sampling indicates $RR_{est} > RR_{crit}$, but in reality $RR < RR_{crit}$, a batch is incorrectly judged as non-complying. This correctness and incorrectness can be scored for all sampled batches in the 5000 iterations, which yields a 2x2 table,

	$RR_{b,\,est} > RR_{crit}$	$RR_{b,\,est} > RR_{crit}$
$RR_b > RR_{crit}$	a	b
$RR_b < RR_{crit}$	c	d

from which the Odds Ratio (OR) and the kappa value can be calculated.

OR = ad/bc

kappa = $\frac{a+d-(a+b)(a+c)-(b+d)(c+d)}{1-(a+b)(a+c)-(b+d)(c+d)}$ (Landis and Koch, 1977)

The interpretation of the quantitative results of these tests in the current context is not obvious, so here they are only regarded in a relative way.

Results

Baseline: DK EU baseline data, Nauta CPM, *n*=5, *m*=1000, *c*=1, RR_{crit} =1.5

The results of the baseline model, using the Nauta CPM, are given in table U1 and figureU3. Note that the results may differ from the results in the report, as we didn't apply MLE fitting here.
The table shows that the uncertainty about the risk estimates is large, with a 95% confidence interval from 0.24% -0.44%. The mean risk estimate based on the samples (Risk est), the mean risk of the sampled batches (Risk) and the risks of the sets of batches Risk set (which is uncertain because the μ, σ and ϕ are uncertain) are very similar for individual simulations, which is expected. Also, the uncertainty about the BN.C and MRRR results is large, especially for the MC approach: the 95% confidence interval for BNMC ranges from 6.7% to 12.7%, MRRR from 26.5% to 46.9%, the graph shows that there is a correlation in the results, as the dots appear in a band. For the CC approach the uncertainty is smaller: the 95% confidence interval for BNCC ranges from 5.6% to 7.7%, MRRR from 40.6% to 52.9%, the graph shows that the dots appear in a circular cloud,

[7] $RR_{b,\,est} = Risk_{b,\,est}/Risk_{set}$; $RR_b = Risk_b/Risk_{set}$ for each batch, with $Risk_{set}$ the risk of the set of batches (with mean μ and standard deviation σ_{tot}).

suggesting little correlation between BNCC and MRRR. The figure shows that the CC results are clustered closer together than the MC results, and that the uncertainty in the results is smaller.

Table U1. Results of the uncertainty analysis for the microbiological data from the DK baseline study, using n=5, m=1000, c=1, RR_{crit}=1.5.
The presented values are the mean μ and between and within food lot standard deviation of concentrations σ_{bet} and σ_{wit}, and the between flock prevalence p; The mean risk R_b of the 5000 sampled food lots, the estimated mean risk R_{est} of the 5000 sampled food lots based on n samples, and the calculated risk of the set of food lots as a whole, R_{set}; BNMC and MRRR for the MC approach, with the mean risk estimates of the non complying (nc) and complying food lots(c); BNCC and MRRR for the CC approach, with the mean risk estimates of the non complying (nc) and complying food lots (c). The rows give the results for the mean of the 200 simulations, the 2.5, 50 and 97.5 percentile of the 200 simulations, and the result for a simulation of 5000 iterations using the most likely (ML) values of μ, σ_{bet}, σ_{wit} and p.

	μ	σ bet	σ wit	p	Risk	Risk est	Risk set	BNMC	risk nc	risk c	MRRR mc	BNCC	risk nc	risk c	MRRR cc
mean	2.21	0.93	0.59	0.31	0.33%	0.33%	0.33%	9.8%	2.3%	0.12%	35.2%	6.7%	2.8%	0.16%	46.8%
2.5%	1.98	0.79	0.42	0.26	0.24%	0.23%	0.24%	6.7%	1.9%	0.09%	26.5%	5.6%	2.0%	0.11%	40.6%
50%	2.21	0.93	0.59	0.31	0.33%	0.32%	0.33%	9.8%	2.3%	0.11%	34.9%	6.7%	2.8%	0.15%	46.7%
97.5%	2.41	1.07	0.77	0.36	0.44%	0.43%	0.44%	12.7%	2.7%	0.14%	46.9%	7.7%	3.5%	0.22%	52.9%
ML	2.21	0.92	0.60	0.31	0.33%	0.32%	0.33%	9.9%	2.3%	0.12%	35.1%	6.7%	2.8%	0.16%	47.2%

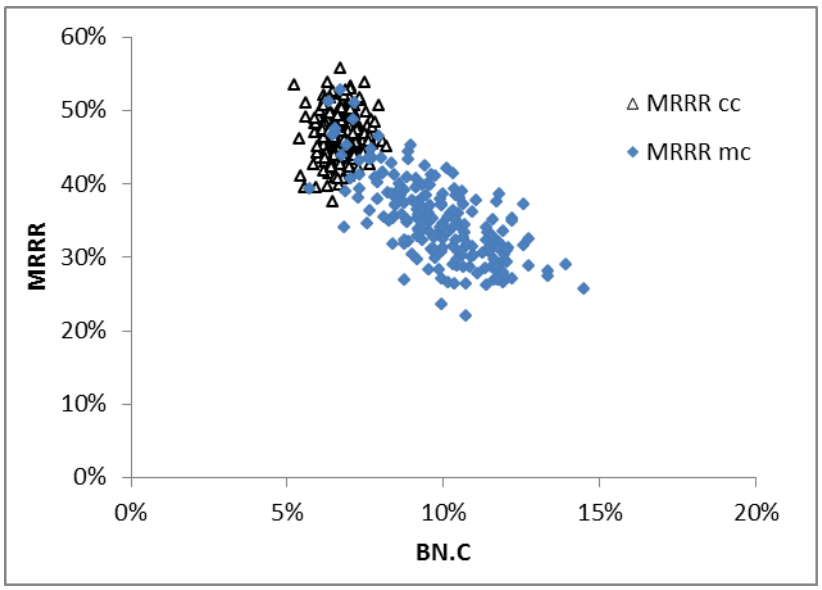

Figure U3: Uncertainty in the MRRR and BN.C results from 200 simulations of the baseline model. Each dot represents the result of one random simulation.

For the CC approach the OR and kappa values are calculated to check the agreement between the relative risk estimates based on the samples $RR_{b,\,est}$ and the "true" relative risks of the batches RR_b.
It shows that:

	OR RR cc	a	b	c	d	KAPPA
mean	121.8	0.75	0.04	0.03	0.18	0.79
2.5%	58.5					0.73
50%	109.0					0.79
97.5%	242.7					0.86

There is a strong agreement between test and "true values", but it is interesting that the kappa value is sensitive for n: If n=1 it drops to 0.6:

	OR RR cc	a	b	C	d	KAPPA
mean	25.77	0.72	0.07	0.07	0.15	0.60
2.5%	13.38					0.51
50%	23.25					0.60
97.5%	49.84					0.71

Changing n, c and RR_{crit}.

Results for some other scenarios are given in table U2 and figure U4.

Table U2. Mean results for scenarios with different values of c and n (for the MC) and RR_{crit} and n (for the CC), as well as the OR and kappa value for the latter. Based on 200 x 5000 iterations of the baseline model (so all other assumptions and parameter values are the same as in the previous paragraph). Risk estimates do not depend on these parameters, and are therefore not given again.

	BNMC	risk nc	risk c	MRRR mc		
c=1	9.8%	2.3%	0.1%	35.2%		
c=2	6.7%	2.8%	0.2%	47.6%		
c=0	14.7%	1.9%	0.1%	21.2%		
n=1, c=1	7.6%	2.4%	0.2%	48.9%		
	BNCC	risk nc	risk c	MRRR cc	OR	KAPPA
RR_{crit}=1.5	6.7%	2.8%	0.2%	46.8%	122	0.79
RR_{crit}=2	4.5%	3.4%	0.2%	58.7%	158	0.80
n=1, RR_{crit}=1.5	6.7%	2.5%	0.2%	52.5%	25.8	0.60

Figure U4 confirms the smaller uncertainty of CC compared to MC. It also shows that within each of these approaches, the clouds of points are about equal in size, which means that the uncertainties are comparable. For different values of c the clouds are partially overlapping. Also for *n*=1 and *n*=5 with R=1.5 there is a large overlap in the clouds, but it is also clear that only MRRR is affected by the change in *n*, not so much BNCC.

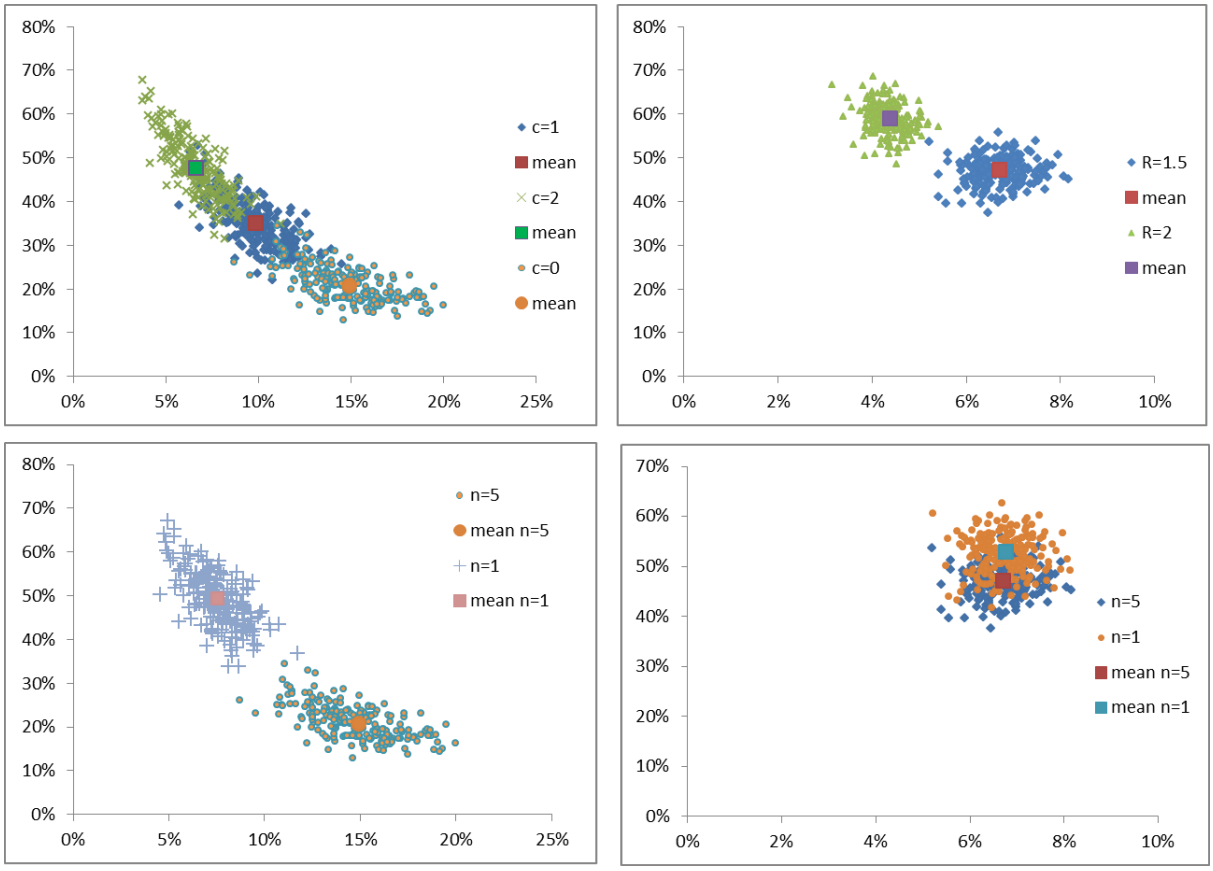

Figure U 4. Plots of MRRR (y axis) and BN.C (x-axis) for the DK baseline data and the Nauta CPM. Graphs on the left show MC results, on the right CC results. Upper graphs are default for n=5 (upper) with m=1000 and varying values of c (upper left) and varying value of RR_{crit} (given as R=1.5 and R = 2; upper right). Lower left: c=0 and m=1000, varying n. Lower right: RR_{crit} = 1.5, varying n. All graphs show the results of 200 x 5000 iterations.

Different data sets

The analyses performed for the DK EU baseline data set, were repeated for the data sets of Lindblad et al. 2006 (Sweden, SLV data) and the EU baseline data for Finland. Results are shown in Table U3 and Figure U6. Note that the results may differ from the results in the report, as we didn't apply MLE fitting here.

Table U3. Results for data sets from three different countries (data from Table U1 are presented again for comparison).

	μ	σ bet	σ wit	p	Risk	BNMC	risk nc	risk c	MRRR mc	BNCC	risk nc	risk c	MRRR cc	OR RR cc	Kappa
Lindblad et al 2006 / SLV / Sweden															
mean	2.11	0.72	0.47	0.14	0.11%	3.1%	1.8%	0.05%	50.6%	3.0%	1.9%	0.06%	50.2%	119	0.79
2.5%	1.92	0.60	0.34	0.12	0.08%	1.9%	1.5%	0.04%	37.3%	2.5%	1.4%	0.04%	43.4%	56	0.72
50%	2.11	0.71	0.46	0.14	0.11%	3.1%	1.8%	0.05%	50.6%	3.0%	1.9%	0.05%	50.2%	112	0.79
97.5%	2.30	0.89	0.60	0.17	0.15%	4.9%	2.1%	0.07%	64.2%	3.7%	2.4%	0.08%	56.6%	232	0.85
ML	2.11	0.71	0.47	0.14	0.11%	3.0%	1.8%	0.05%	50.9%	3.0%	1.8%	0.05%	50.4%	107	0.79
Finland EU baseline															
mean	1.14	0.66	0.42	0.05	0.01%	0.1%	1.3%	0.01%	90.3%	1.0%	0.5%	0.00%	48.3%	50	0.68
2.5%	0.83	0.49	0.27	0.03	0.00%	0.0%	1.0%	0.00%	67.8%	0.6%	0.2%	0.00%	38.4%	21	0.55
50%	1.13	0.64	0.41	0.05	0.01%	0.1%	1.3%	0.01%	91.9%	1.0%	0.5%	0.00%	48.3%	46	0.69
97.5%	1.46	0.90	0.62	0.08	0.02%	0.4%	1.6%	0.01%	99.6%	1.4%	1.0%	0.01%	59.3%	101	0.77
ML	1.11	0.63	0.41	0.05	0.01%	0.0%	1.2%	0.01%	94.6%	0.9%	0.4%	0.00%	49.2%	47	0.68
Denmark EU baseline															
mean	2.21	0.93	0.59	0.31	0.33%	9.8%	2.3%	0.12%	35.2%	6.7%	2.8%	0.16%	46.8%	122	0.79
2.5%	1.98	0.79	0.42	0.26	0.24%	6.7%	1.9%	0.09%	26.5%	5.6%	2.0%	0.11%	40.6%	59	0.73
50%	2.21	0.93	0.59	0.31	0.33%	9.8%	2.3%	0.11%	34.9%	6.7%	2.8%	0.15%	46.7%	109	0.79
97.5%	2.41	1.07	0.77	0.36	0.44%	12.7%	2.7%	0.14%	46.9%	7.7%	3.5%	0.22%	52.9%	243	0.86
ML	2.21	0.92	0.60	0.31	0.33%	9.9%	2.3%	0.12%	35.1%	6.7%	2.8%	0.16%	47.2%	122	0.79

The data sets differ in size, prevalence estimate, and mean and standard deviation of concentrations. As found in the main report, The risk estimate is lowest for the Finnish data, followed by the Swedish and Danish data. The main difference between Sweden and Denmark is the lower prevalence in Sweden, Finland has both a lower prevalence and a lower mean, and the results are based on less data, so the uncertainty attending the estimates is larger.

As a consequence, the results for the Swedish data look very similar to the Danish data (figure U4 and U5), but the results for the Finnish data look different. For the uncertainty analysis it is particularly interesting to see that the clouds in Figure U6 are largely overlapping. This implies that there is not much difference between the different alternatives for *c*, RR_{crit}, and *n*. The most important factor is probably the limited size of the data set.

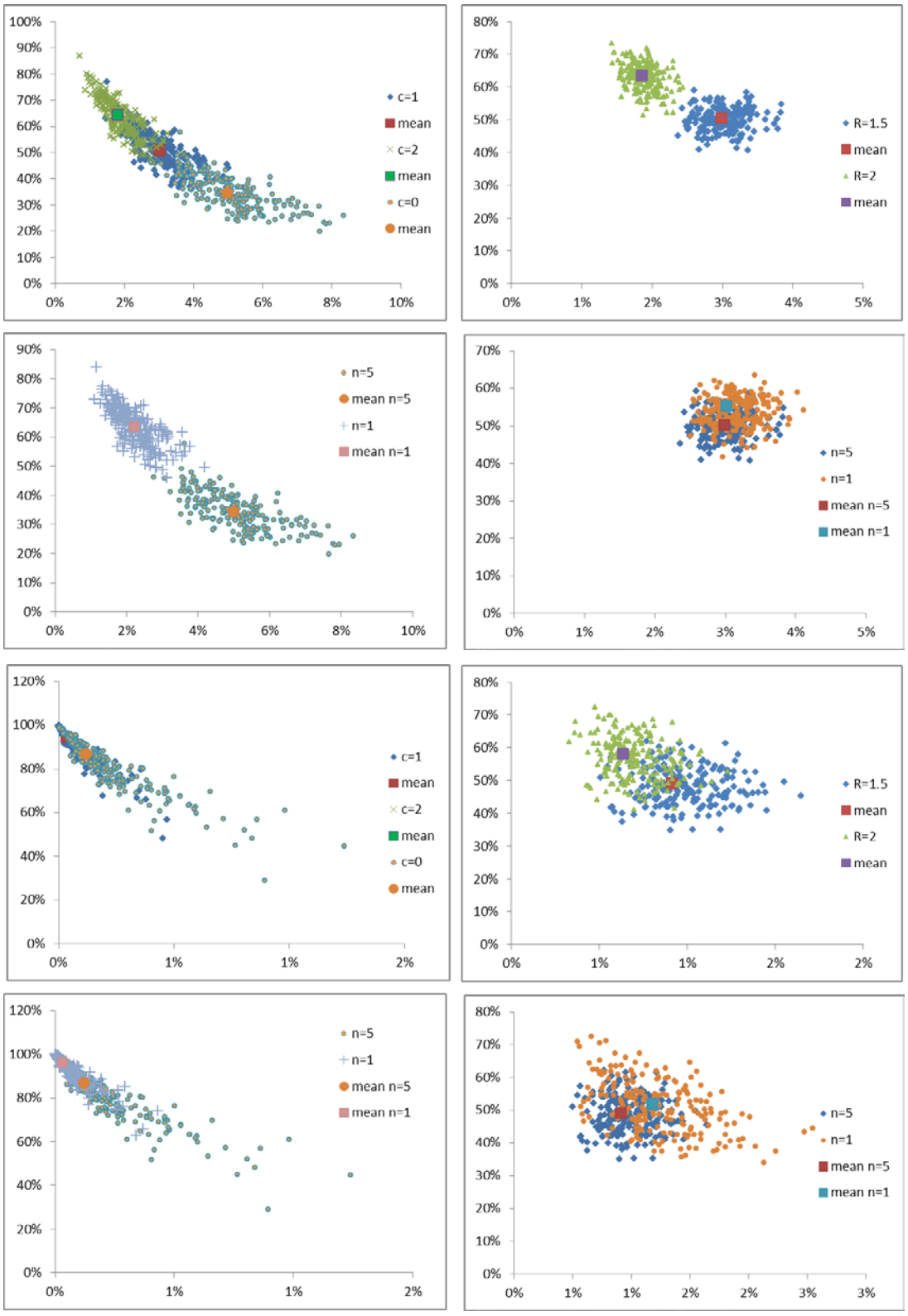

Figure U5. Results for the Swedish SLV (Lindblad et al. 2006) data (four upper graphs)and the Finnish EU baseline data (four lower graphs). For explanation see Figure U4.

Different Risk Models

The model for the DK EU baseline data is rum with two alternative CPMs: The FAO CPM and the Van Asselt et al CPM. Results are shown in Table U4 and Figures U6, U7 and U8.

Table U4. Results using two alternative Consumer Phase models

	μ	σ bet	σ wit	p	Risk	BNMC	risk nc	risk c	MRRR mc	BNCC	risk nc	risk c	MRRR cc	OR RR cc	Kappa
FAO															
Mean	2.21	0.93	0.59	0.31	0.53%	10.0%	3.9%	0.14%	28.3%	6.9%	4.8%	0.21%	39.5%	123	0.79
2.5	1.98	0.79	0.42	0.26	0.35%	6.8%	3.2%	0.11%	20.3%	5.8%	3.4%	0.13%	33.0%	53	0.72
50	2.21	0.93	0.59	0.31	0.52%	10.0%	3.9%	0.14%	27.6%	6.9%	4.8%	0.20%	39.2%	112	0.80
97.5	2.41	1.07	0.77	0.36	0.70%	12.8%	4.6%	0.18%	38.7%	7.9%	6.0%	0.30%	46.6%	236	0.86
ML	2.21	0.92	0.60	0.31	0.52%	10.1%	3.9%	0.15%	27.9%	6.9%	4.7%	0.21%	39.6%	104	0.79
Van Asselt															
Mean	2.21	0.93	0.59	0.31	0.30%	9.9%	1.9%	0.13%	42.7%	7.0%	2.1%	0.16%	53.6%	124	0.80
2.5	1.98	0.79	0.42	0.26	0.23%	6.8%	1.6%	0.10%	34.0%	6.0%	1.7%	0.12%	47.2%	58	0.73
50	2.21	0.93	0.59	0.31	0.30%	9.9%	1.9%	0.13%	42.7%	7.0%	2.1%	0.16%	53.4%	110	0.80
97.5	2.41	1.07	0.77	0.36	0.37%	12.9%	2.1%	0.15%	53.7%	8.1%	2.5%	0.21%	59.7%	244	0.86
ML	2.21	0.92	0.60	0.31	0.30%	10.0%	1.8%	0.13%	42.4%	7.0%	2.1%	0.16%	53.8%	107	0.79

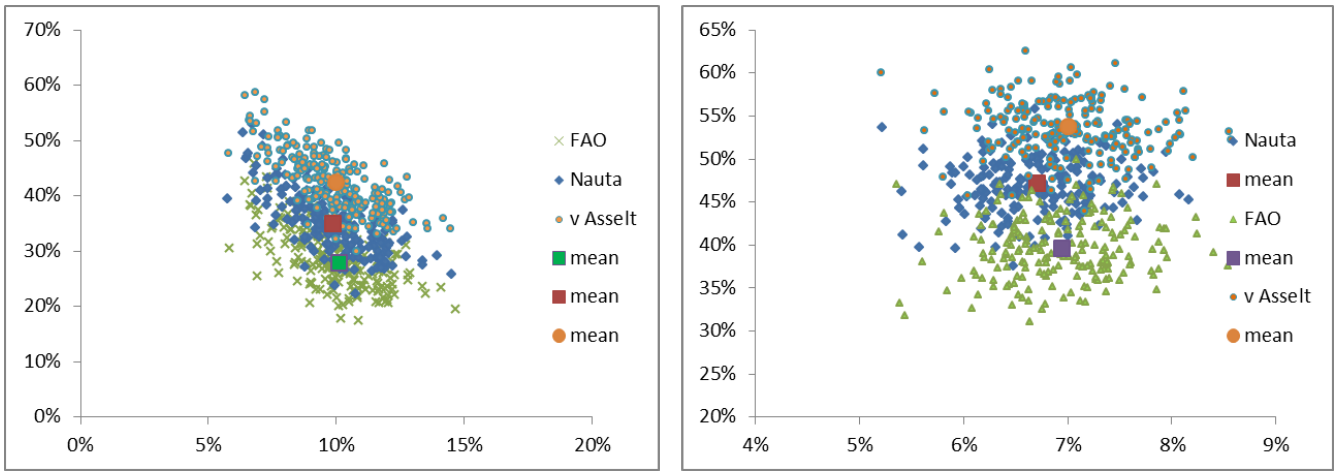

Figure U6 Plots of BNMC and MRRR (left) and BNCC and MRRR (right) for the three different CPMs (FAO, Nauta and Van Asselt).

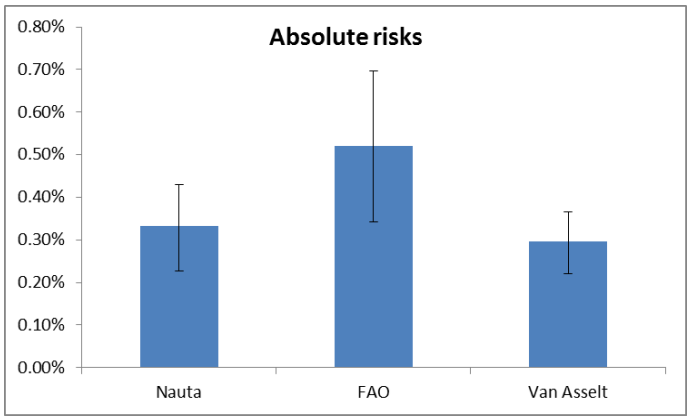

Figure U7 Risk estimates ($Risk_{set}$) for the three CPMs with the 95% confidence interval, as given in tables U4 and U1.

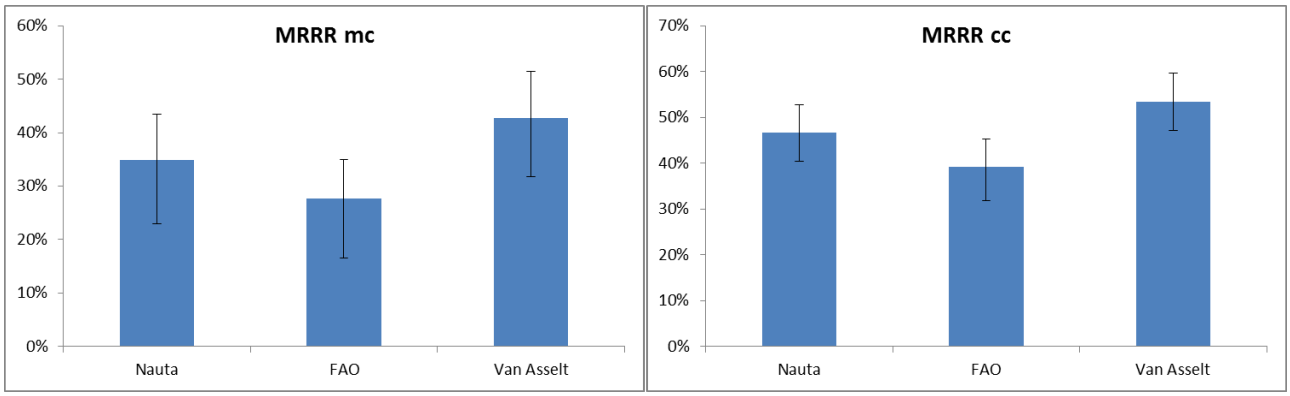

Figure U8. Values for the MRRR for the default MC/CC scenarios for the three CPMs, with 95% confidence intervals.

As expected from the graph (Figure U2), the risk estimate is highest for the FAO model, followed by the Nauta model and the Van Asselt model. The order is reversed when we look at the values for the MRRR. The graphs (figure U6) show that the clouds are overlapping, which indicates that the differences between the three CPMs are small, certainly for the MC approach. Interestingly, the BNMC value does not depend on the CPM, as the risk model is not applied to calculate BNMC. It is however used in the CC approach, and it can be seen that here, too, there is very little difference between the three CPMs.

The impact of τ and φ

The sensitivity analysis of τ and φ in the MC approach model has been described by Nauta et al. 2012.
Here we show the impact of τ (tau; figure U9) and φ (phi; figure U10) graphically.

The impact of the unknown parameter τ is quite large, certainly for the estimate of MRRR, although the clouds are partially overlapping. Omitting the uncertainty in φ decreases the uncertainty in MRRR, but not so much the uncertainty in BNMC or BNCC.

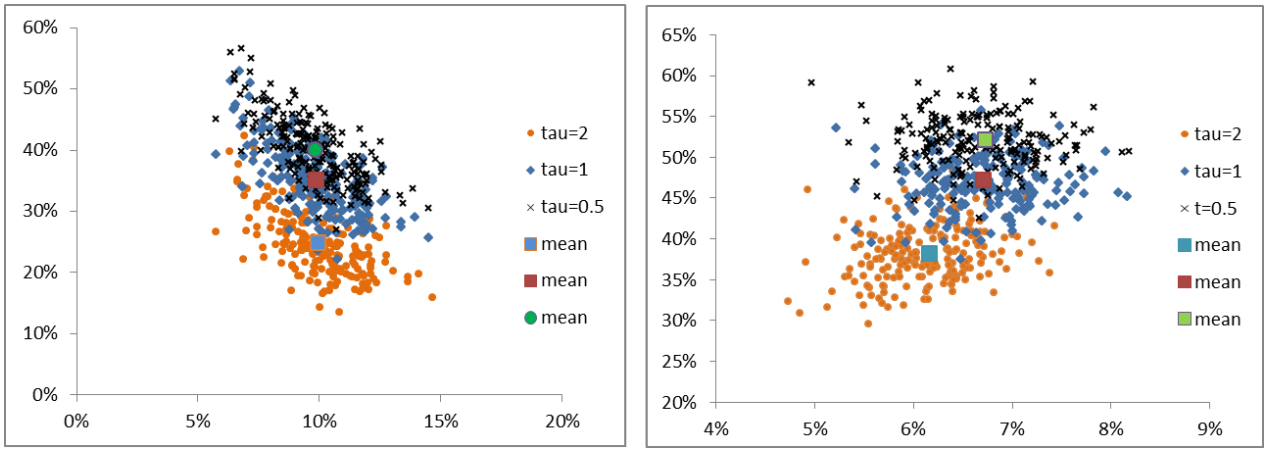

Figure U9. Plots of BNMC and MRRR (left) and BNCC and MRRR (right) for three different values of τ.

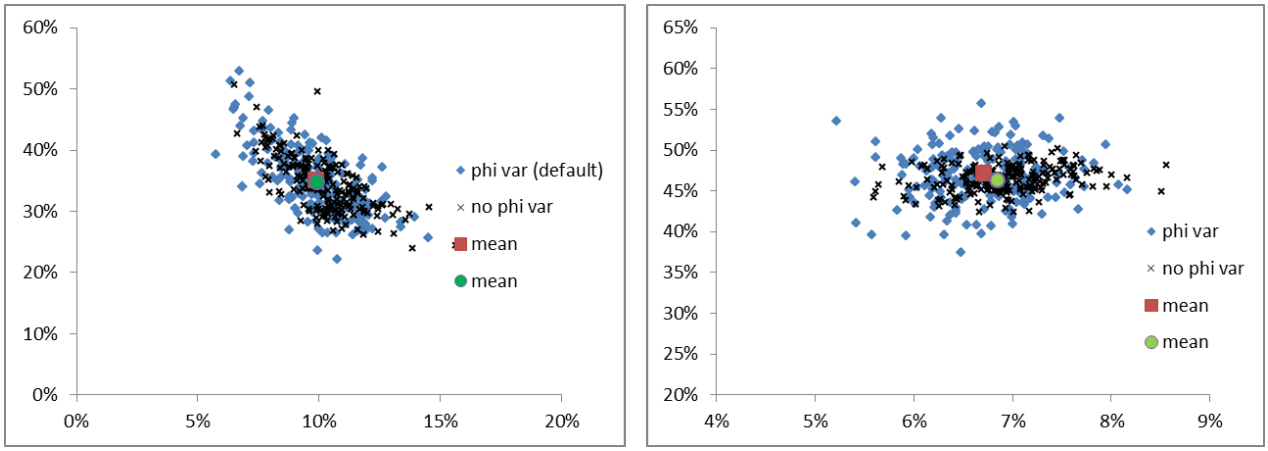

Figure U10. Plots of BNMC and MRRR (left) and BNCC and MRRR (right) for the baseline model (phi var) and an alternative where the value of φ is fixed to 0.3.

Discussion

Uncertainty analyses have been performed by means of Monte Carlo analyses, to explore the impact of uncertainty in the data sets applied, and uncertainty in the risk model used.

A baseline analysis has been done for the Danish EU baseline data set, with 123 samples with a quantified concentration, among 329 samples, using the Nauta *et al*. consumer phase model, one particular MC (n=5, m=1000, c=1) and one CC scenario (n=5, RR_{crit}=1.5). It showed that the uncertainty in the results is substantial, in terms of risk estimates, Minimum Relative Residual Risk and percentage of non-complying batches

(BNMC and BNCC). However, the uncertainty is smaller when the case by case approach is applied. This may be a consequence of the fact that the relative risk estimate and the criterion for compliance are based on the same model, which may cancel out some of the uncertainties.

This general result appears also in the other analyses. In these it is found that

- effect of *n*:

The sample size of the test is compared for *n*=5 and *n*=1. (Larger sample size showed only minor differences with *n*=5). It shows that the "efficiency" of the MC/CC, as expressed in the BN.C-MRRR plot is hardly influenced by the sample size, certainly compared with the overall uncertainty (the clouds are largely overlapping). This means that the risk reduction that can be achieved at the expense of a certain percentage of non-complying batches is almost the same for any value of *n*, so small sample sizes may be sufficient to apply the method. However, for the CC approach, the calculation of the OR and the kappa value show that, with small *n*, it is more often the case that the wrong batch is identified as non- complying. Apparently this has little impact on the results for all batches, but for individual batches, and their producers, this may not be acceptable.

- effect of *c*

As previously found, for the MC approach, the acceptable number of samples > *m* has an impact on the balance between BNMC and MRRR. The uncertainty ranges for adjacent *c*-values are overlapping, which indicates that the precision in which the effect of the choice of *c* on the balance between BNMC and MRRR can be predicted, is limited.

- effect of RR_{crit}

The values of RR_{crit} used in these analyses are quite low, which was necessary to get results for the CC approach that are comparable with the MC approach. As the clouds are not overlapping, it shows that the results are significantly different with RR_{crit}=1.5 and RR_{crit}=2.

- effect of the data set used

In terms of the observed uncertainties, the results for the Swedish and Danish data sets are similar. For the Finnish data set, with less concentration data and a lower mean concentration, the uncertainties seem to be larger, and the discriminatory power between the different parameter values for the MC/CC approach and the sample size is smaller. This confirms that it is relevant to have a sufficiently large data set.

- effect of the CPM

The choice of the CPM has an effect on the risk estimate and the value of MRRR. This implies that the choice of the consumer phase model, which is an important source of uncertainty in the risk assessment model, is important when the effect of a MC/CC on the risk is assessed. This effect seems to be larger for the MC than the CC approach. The clouds in the graphs are largely overlapping, which suggests that, in the light of the total

uncertainty, the effect of the choice of the CPM is limited. Interestingly, unlike BNCC, BNMC does NOT depend on the CPM as there is no risk estimate involved.

- impact of "unknown" parameters τ and φ

The values of the parameters τ and φ are based on expert judgement, and it is difficult to assess their uncertainty. Uncertainty of φ is included in the analyses, and when it is omitted, the uncertainty about BN.C seems to remain similar, but the uncertainty about MRRR decreases. This implies that a more precise estimate of φ can reduce the uncertainty about MRRR.
For τ the results of Nauta et al (2012) are confirmed. It has considerable impact on, mainly, the MRRR value, but also on BNCC. The uncertainty clouds of $\tau = 1$ and $\tau = 2$ are hardly overlapping, which suggests and confirms that it is important to obtain a good estimate for this parameter.

It should be stressed that the uncertainty about the dose response model is not yet studied here, among others because it is quite complicated to quantify this uncertainty. One possibility would be to include the uncertainty in the estimates of the dose response parameters α and β, as derived from the Black et al (1988) data set, into the analysis. This would however not solve the problem that this data set may not be representative for all population groups and all *Campylobacter* strains.
Future research may focus on performance analysis of the impact of the sample size *n* in both the "case by case" and the MC approach. This can help to make the approaches comparable to the FSO based MCs. Also, this will increase the insight in the finding that the performance of $n=1$ is much comparable to $n=5$, but the kappa values differ.
Next, in the current analyses we now first define a set of batches, determine a baseline from that, and then consider how often batches from the same set of batches are complying or not. One can also consider what happens with batches from other sets of batches: what happens for example if batches from DK are compared to the Finnish baseline? This comes closer to what the current Danish case by case is actually doing.

With the methods described in this appendix the uncertainty in the estimated effect of a microbiological criterion, or application of the "case by case" approach, is possible for any data set that is considered for usage. This can be supportive for the decision on the implementation of either of these.

Appendix 4: Bayesian analysis on Swedish data - an example

J Ranta, P Tuominen

February 13, 2013

1 Summary of two SWE data sets providing complementary evidence

The first data set is described in Lindblad et al paper [1], providing measured campylobacter concentrations in carcasses. It is assumed that each carcass is from a different batch (flock). It is also assumed that the data represent the whole production so that there is no selection bias. There were **529 negative results** and **88 positives**. The concentrations were measured from the 88 positives. (From these, 2 $\log_{10}$ need to be subtracted to get cfu/g, according to expert assumption. See section 'Data from Nordic countries' in the report).

```
2.60,2.60,2.60,2.60,2.60,2.60,2.60,2.60,2.90,2.90,2.90,
3.08,3.08,3.20,3.26,3.34,3.40,3.41,3.51,3.51,3.56,3.56,
3.60,3.64,3.64,3.64,3.68,3.88,3.88,3.92,3.92,3.94,3.94,
3.94,3.94,3.96,3.97,4.01,4.02,4.02,4.06,4.07,4.09,4.10,
4.16,4.16,4.18,4.19,4.23,4.25,4.26,4.26,4.29,4.30,4.31,
4.32,4.36,4.38,4.41,4.41,4.50,4.60,4.60,4.64,4.65,4.68,
4.70,4.71,4.71,4.76,4.77,4.82,4.86,4.90,4.96,4.98,4.99,
4.99,5.02,5.06,5.14,5.26,5.32,5.40,5.42,5.42,6.17,7.15
```

The second data set is described in Hansson et al paper [2]. These provide data from positive flocks (batches) only, shown below. (We need to add $\log_{10}(4)$ to get cfu/g according to expert assumption. See section 'Data from Nordic countries' in the report). However, the data represent only mean values and standard deviations per flock. Also the sample size per flock (N) and number of positives per sample (X) are known.

means	2.31	1.96	1.38	2.98	2.87	2.76	3.02	2.69	3.15	2.63	2.74	2.32	2.62	2.62	1.35	1.21	2.19	1.39	2.13	2.11
SDs	0.61	0.51	0.60	0.48	0.71	0.39	0.58	0.40	0.49	0.37	0.37	0.26	0.49	0.35	0.81	0.80	0.48	0.75	0.69	0.61
Ns	25	10	23	16	12	13	5	10	20	11	20	15	17	17	20	21	23	20	20	11
Xs	24	10	21	16	12	13	5	10	20	11	20	15	17	17	17	18	23	20	19	11

Note: a 'flock' is here effectively assumed to be processed as one 'batch' which will be one 'lot'.

2 Modeling for predicting cfu-values in carcasses, with uncertainties

2.1 General model structure

The general structure involves a hierarchy in three levels: (1) measurements from individual carcasses, belonging to (2) flocks, belonging to (3) a group of flocks produced. The measurements consist of bacteria log-concentrations for positive carcasses and the number of negative (zero concentration) carcasses. The log-concentration y_{ij} in a positive carcass i in flock j may be described by normal distribution, conditionally on the flock average concentration (among positive carcasses), $\mathrm{N}(\mu_j, \sigma_w^2)$. This distribution describes variability, σ_w^2 , of the positive concentrations within a flock. Variability of the averages μ_j may be described by $\mathrm{N}(\mu, \sigma_b^2)$, where μ describes the overall mean concentration of all the positive carcasses produced. The negative carcasses (i.e. zero concentrations) in a positive flock can be described by binomial model with a parameter describing within flock prevalence. Then, the number of positives x_j among all N_j sampled in a positive flock j is described by $\mathrm{Bin}(N_j, p_j)$. In a general model, also

the between flock variability of prevalences p_j could be modeled. For simplicity, we could assume the expected within flock prevalence (in positive flocks) to be the same p in all such flocks. Finally, the flock prevalence q is the proportion of positive flocks among all flocks in production. This is to account for those flocks that are completely negative, i.e. within prevalence is zero.

The average concentration γ in whole production is thus a weighted average of the zero concentrations in negative flocks, and the zero and positive concentrations in positive flocks: $\gamma = 0 \times (1-q) + 0 \times q(1-p) + \mu qp = \mu qp$. The uncertainty distribution of this would be of interest in describing the average concentration of an arbitrary carcass in production. Furthermore, an individual measurement would then be described by a mixture distribution taking values of zero with probability $(1-q) + q(1-p) = 1 - qp$ and positive values with probability qp from $\mathrm{N}(\mu, \sigma_b^2 + \sigma_w^2)$. Of course, all this is conditional to assuming these parameters are known.

The general model requires data to represent several positive concentrations from the flocks where positives have been identified, to be able to estimate within and between flock variability of the positive concentrations. For example, it is not enough to have only one positive concentration per flock, nor to have only the flock specific average concentrations. Therefore, depending on the data, the model may be over-parameterized which means that we would not be able to estimate well individual parameters. Instead, some combinations of parameter values would still explain the data, so that some parameter combinations could be found more probable than others. This will be seen later when analyzing data that are insufficient for the full model. Below, we show example with two data sets, that are both insufficient in different ways, but that are complementary to each other. Therefore, the combination of the two data sets is necessary to provide reasonable estimates. In contrast, analyzing the two data sets separately shows the resulting uncertainty from an over-parameterized model. Then, assuming a fixed value or assuming an ad hoc distribution would not solve the problem. However, with modestly over parameterized models, there can still be some constraining information in the data that could be used. As result, we could obtain a posterior distribution for the parameters, showing what combinations of parameter values are most plausible based on the available evidence.

2.2 Lindblad et al data

Starting with the Lindblad et al data, the model structure distinguishes positive and (assumingly) negative flocks (batches) so that we may define a national average concentration μ in *positive* carcasses in all positive flocks, and a flock specific average concentration μ_j of *positive* carcasses for each positive flock. Hence, the concentration model concerns positive carcasses only. According to the hierarchical structure there are two variance components describing the between-flock variation σ_b^2 and the within-flock variation σ_w^2. The conditional distribution of positive measurements from carcasses (in a flock) could then be defined as

$$y_{ij} \sim \mathrm{N}(\mu_j, \sigma_w^2)$$

Note that there are no data about within flock prevalence in Lindblad et al., because the data only reports one positive measurement per flock. Concerning both positives and zeros, we (at first) assume that all measurements from a 'positive flock' would be positive, and all from a 'negative flock' would be negative. The assumption can be relaxed later, but only by combining additional data about within batch prevalence which was not available in this data set. Even with this assumption, the model remains over parameterized, concerning the two variance components. Nevertheless, with the assumption, the positive and zero measurements in a flock j are described by a mixture distribution

$$\pi(y_{ij}) = q\mathrm{N}(y_{ij} \mid \mu_j, \sigma_w^2) + (1-q)1_{\{y_{ij}=0\}}(y_{ij})$$

where q is the prevalence of positive flocks and μ_j is the average log-prevalence for that flock. Effectively, for the whole production, the overall average concentration would be simplified to $\gamma = q\mu$. This model has four basic parameters $(q, \mu, \sigma_b^2, \sigma_w^2)$ together with $\mu_1, \ldots, \mu_x$ (where x is the number of positive carcasses, i.e. positive flocks). For a standard Bayesian approach with no additional prior knowledge, we need to define fairly vaguely informative prior distributions e.g. as follows

$$\begin{array}{lcl} q & \sim & \mathrm{U}(0,1) \\ \mu & \sim & \mathrm{N}(0,10^4) \\ \mu_j & \sim & \mathrm{N}(\mu,\sigma_b^2) \\ \sigma_b & \sim & \mathrm{U}(0,1000) \\ \sigma_w & \sim & \mathrm{U}(0,1000) \end{array}$$

2.2.1 Choice of priors

The above chosen priors are practically as uninformative as possible, to let the data drive the results. Alternatively, if we had relevant prior information this could be incorporated by constructing an informative prior. To do that properly would require either careful expert elicitation, meta-analysis based on literature, or other data sources which could be combined here explicitly by extending the model. (Actually: the latter will be done when we combine both the Lindblad et al data and the Hansson et al data in later sections). Mathematically, prior that are intended to be uninformative could also be constructed in slightly different ways. For example, the flat normal density prior $\mathrm{N}(0,10^4)$ for μ could also be uniform density over a wide range. In general, if all priors are sufficiently wide uniform densities (and if the posterior exists), the posterior density corresponds to likelihood function and the posterior mode to maximum likelihood estimate. In the current example, there is no unique maximum likelihood estimate of variance components because the data are not sufficient to single out unique best estimate. Instead, there are combinations of values of the two parameters that are equally good in terms of likelihood function. (Compare: if you were told only the sum of two dice, you could only infer about the combination of values of the two dice). Since the posterior distribution is a joint distribution of all the unknown parameters, it can be used to summarize the combined evidence of data and priors, also in the case of practically uninformative priors. However, it is worth doing some sensitivity analysis of the priors. A commonly applied prior for variance parameter is defined for the precision (inverse of variance) as Gamma-distribution with small parameters to be practically uninformative. The effect of this is discussed in results section.

2.2.2 Results

The posterior distribution, based on Lindblad et al data, would then be based on the priors and the normal distribution of observed positive concentrations, and the binomial distribution of the number of positive flocks (assuming within flock prevalence can be either 0% or 100%):

$$\pi(\mu,\sigma_b,\sigma_w,q,\mu_1,\ldots,\mu_x \mid x,N,y_1,\ldots,y_x) \propto$$

$$P_{\mathrm{Binom}}(x \mid N,q)\prod_{j=1}^{x}\pi(y_j \mid \mu_j,\sigma_w^2)\pi(\mu_j \mid \mu,\sigma_b^2)\pi(\mu)\pi(\sigma_b)\pi(\sigma_w)\pi(q)$$

Since the model is over-parameterized when only one measurement (carcass) per flock is known, we do not obtain good point estimates. Nevertheless, the joint uncertainty can be seen from the posterior distribution showing that the variance parameters are not independent. Their joint distribution shows what we can learn about their uncertainty from these data. This joint distribution (of all parameters) can be further reflected in simulating predictions from this model. For a sensitivity analysis of the priors, the wide normal density prior for μ was changed to a wide uniform density, but the results remained practically the same. Moreover, for variance parameters σ_b^2 and σ_w^2, alternative Gamma(0.001,0.001)-priors were studied so that these were assigned for the precisions $\tau_b = \sigma_b^{-2}$ and $\tau_w = \sigma_w^{-2}$. Effectively, the means, modes and marginal posterior probability intervals (of 95%) of both σ_w and σ_b remained essentially the same for conclusions. However, the marginal posteriors for σ_b and σ_w became slightly different than with the uniform prior, taking a more pronounced bimodal shape. This occurs because the two parameters are unidentifiable, i.e. the data only informs about their combination, not their individual values, and the possible individual values could just as well (as judged by the data) be infinitely small (for one parameter) and fairly large (for the other parameter), and vice versa. They cannot be infinitely large because the observed data has finite variation. In such situation, the seemingly uninformative Gamma-prior can put inappropriately too much weight on very small and large values, even

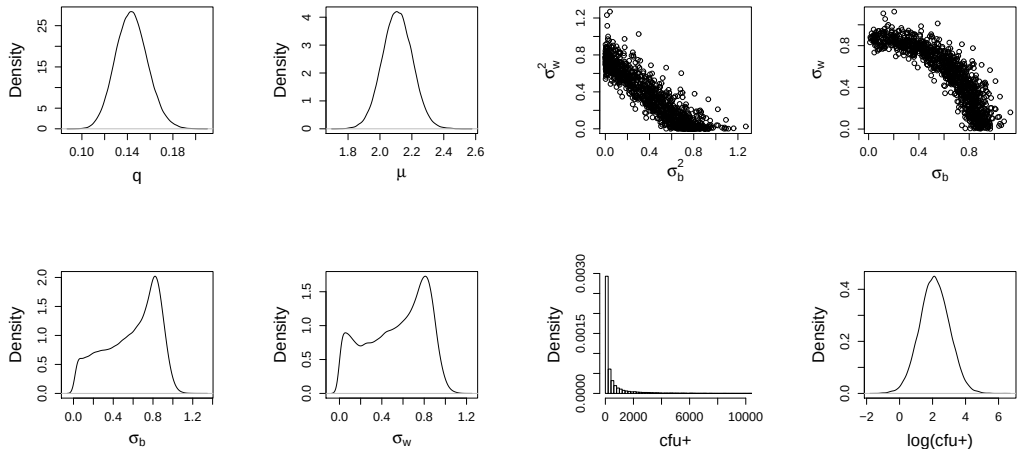

Figure 1: Results from Lindblad et al data: Marginal posterior distributions of model parameters (left, 4 frames). Joint posterior of variance component parameters (right, up 2 frames) and posterior predictive distribution of cfu+ and $\log_{10}$(cfu+) in any (positive) carcass **in positive flocks** (right, down 2 frames).

though this does not drastically affect the estimated mean or mode. After all, it is not a reasonable prior to think that extremely large or small values are plausible. With that respect, the original uniform prior on σ_b and σ_w seems more reasonable, giving uniformly weight on all possible values over a sufficiently wide range. For the essential conclusions, both priors led to similar results. They represent qualitatively different prior densities that are both reasonably uninformative in slightly different ways, for positive valued parameters. Other priors for variance parameters would be clearly more informative and that would have potentially very large effect on the results. Use of informative priors would require explicit analysis of that background information.

Results from Lindblad et al data alone:

The uncertainty of the flock prevalence q can be seen in its marginal posterior distribution and this is probably not too biased if the data were representative collection from all flocks, assuming that each flock is correctly identified as negative or positive. This is now based on only one measurement per flock, assuming the within flock prevalence is either 0% or 100%. We could not estimate within flock prevalence because there was only one measured concentration per flock. Therefore, estimation of average concentration in whole production is not feasible without these strong assumptions of within flock prevalence. This obviously causes bias, unless better assumption is available. Total variability of concentrations is the sum of within flock variability and between flock variability, and the uncertainty of the two can be described by a 2D-plot of their joint posterior distribution. This shows which combinations of values are more probable than others, Figure (1). Posterior distribution of $\phi = \sigma_w^2/(\sigma_w^2 + \sigma_b^2)$ is in Figure (2).

2.3 Hansson et al data

Next, we analyze Hansson et al data. These data represent a selected sample from positive flocks only. In this case, there is no information about overall flock prevalence. Instead, there is evidence about within flock prevalence and concentrations in positive flocks. However, the data represent averages and SDs per flock, together with sample size and positives per sample. The carcass specific measurements are not reported. In this case, we can even model within flock prevalence hierarchically, with the prior knowledge that it is likely to be large. Reflecting this, a prior could be $p_i \sim \text{Beta}(\alpha, 2)$ with hyper parameter $\alpha \sim \text{U}(0, 1000)$, giving the expected within flock prevalence $\alpha/(\alpha+2)$. The posterior distribution in this case becomes

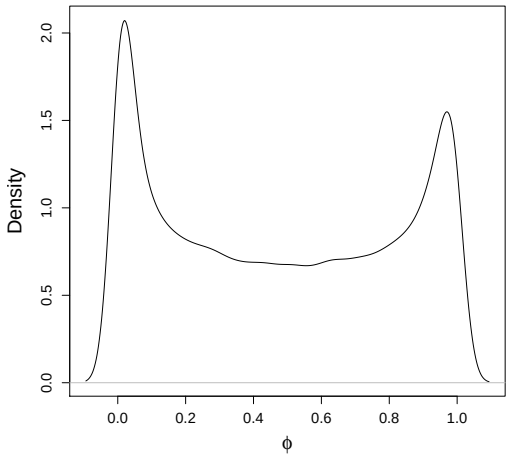

Figure 2: Results from Lindblad et al data: Posterior distribution of $\phi = \sigma_w^2/(\sigma_w^2 + \sigma_b^2)$. (Kernel smoothing applied to MCMC sample).

$$\pi(\mu, \mu_1, \ldots, \mu_n, \sigma_b, \sigma_{w_1}, \ldots, \sigma_{w_n}, p_1, \ldots, p_n, \alpha \mid X_1, \ldots, X_n, N_1, \ldots, N_n, \bar{y}_1, \ldots, \bar{y}_n, \mathrm{SD}(y)_1, \ldots, \mathrm{SD}(y)_n) \propto$$

$$\prod_{i=1}^{n} P_{\mathrm{Binom}}(X_i \mid N_i, p_i) \underbrace{\pi(\bar{y}_i \mid \mu_i, \sigma_{w_i}^2, N_i)}_{\mathrm{N}(\mu_i, \sigma_{w_i}^2/N_i)} \pi(\mu_i \mid \mu, \sigma_b^2) \pi(p_i \mid \alpha) \pi(\tau_{w_i} \mid \mathrm{SD}(y)_i, N_i) \pi(\mu) \pi(\sigma_b) \pi(\alpha)$$

In each group i, we first solve posterior distribution independently for each σ_{w_i} known as $\tau_{w_i} = 1/\sigma_{w_i} \sim \Gamma((N_i - 1)/2, (N_i - 1)\mathrm{SD}(y)_i^2/2)$. This is based on the standard improper prior $\pi(\mu_i, \log(\sigma_{w_i})) \propto 1$ which leads to the known solution of marginal posterior of τ_{w_i} which does not depend on μ_i, but only on the sample size and sample standard deviation. The resulting distribution is taken as a prior for τ_{w_i}. Since the posterior of σ_{w_i} is obtained in each group, we pool these by calculating their average: $\bar{\sigma}_w = \frac{1}{20} \sum_{i=1}^{20} \sigma_{w_i}$ which therefore also has a distribution implied by the distributions of each σ_{w_i}. Alternatively, we might consider more elaborated hierarchical modeling for σ_{w_i} with hyper parameters.

2.3.1 Choice of priors

In this case, the data contains more information about both variance parameters, so that they are even less influenced by priors than with the Lindblad et al data. Therefore, the practically uninformative priors can be used here too. Prevalence of positive flocks q is not a parameter in this model because there is no information at all about it in this data set. (The data represents only positive flocks).

2.3.2 Results

Results from Hansson et al data alone:

In this case, we obtain better estimates of within and between flock variance for positive concentrations. Their joint distribution shows better identifiability than previously. Also, within flock prevalence can be estimated for all flocks. See Figures (3), (4) But we obviously have no estimate of overall flock prevalence q, because the data only represents a sample of positive flocks. Hence, we could not provide estimates for overall prevalence in all flocks, based on these data alone.

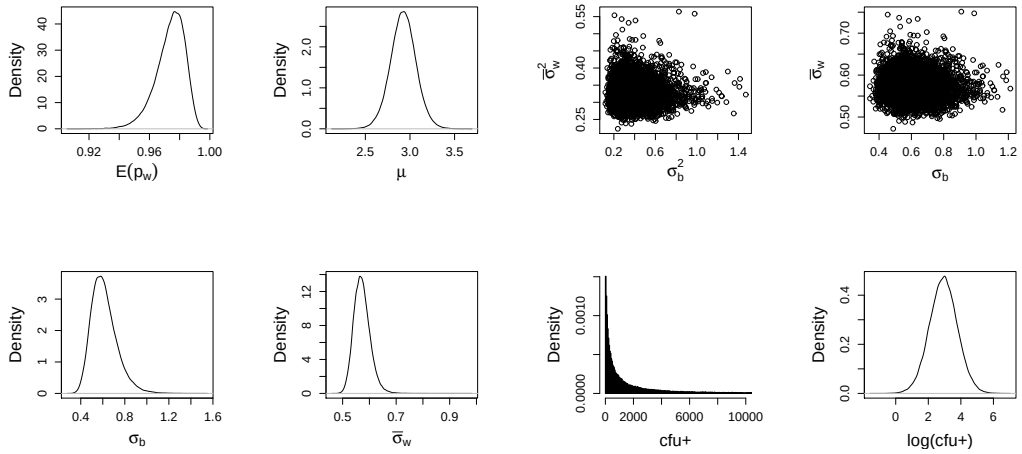

Figure 3: Results from Hansson et al data: Marginal posterior distributions of model parameters (left, 4 frames). Joint posterior of variance component parameters (right, up 2 frames) and posterior predictive distribution of cfu+ and $\log_{10}$(cfu+) in any carcass **in positive flocks** (right, down 2 frames).

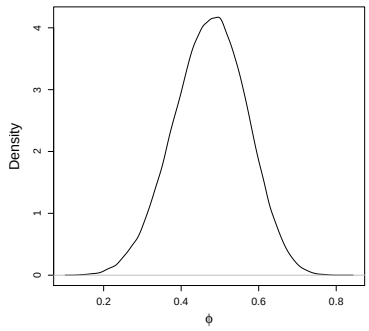

Figure 4: Results from Hansson et al data: Posterior distribution of $\phi = \bar{\sigma}_w^2/(\bar{\sigma}_w^2+\sigma_b^2)$. (Kernel smoothing applied to MCMC sample).

2.4 Combining the two data sets

Since both of these data were limited alone, but because they provide complementary information, it makes sense to combine them in a jointly defined model. The models for each are exactly as above, only put together, except that we can now try to take into account the within batch prevalence in the first data too, using the information drawn from the second. The corrected probability to sample a positive carcass in the first study would be $qE(p_w) = q\alpha/(\alpha+2)$. Also, the prediction for a random carcass in the whole production can be modeled by first sampling the carcass status with probability $qE(p_w)$, and then, conditionally on positive status, the log-cfu-value from $\mathrm{N}(\mu, \sigma_b^2 + \bar{\sigma}_w^2)$.

It remains to be considered how these data are truly comparable and what is the adequate way to transform measurements to a common scale, because the transformation of original reported log-cfu values to assumed log-cfu/g itself changes the data very much in terms of absolute cfu-levels. This can have effects on predictions with the dose-response model eventually. Comparability relies on the assumption of similar detection methods in all samples and that the sampling is not purposively or otherwise selective.

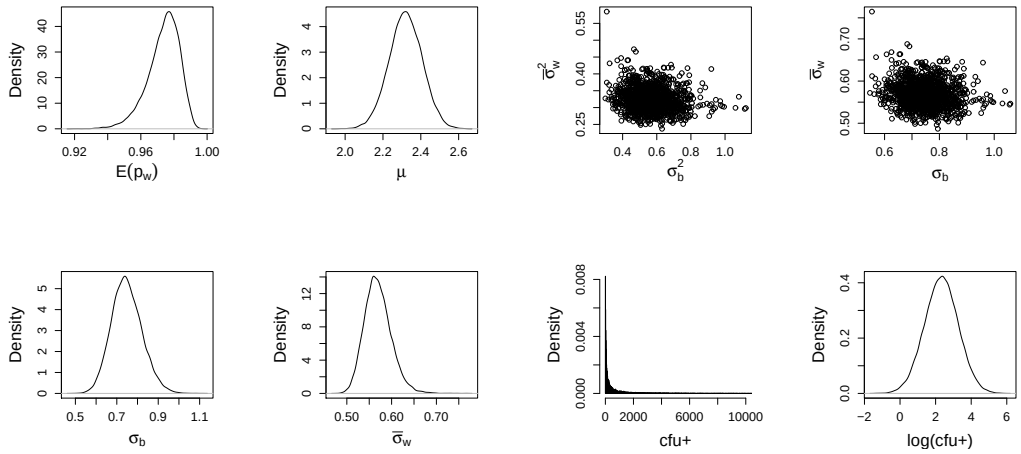

Figure 5: SWE data: Marginal posterior distributions of model parameters (left, 4 frames). Joint posterior of variance component parameters (right, up 2 frames) and posterior predictive distribution of cfu+ and $\log_{10}$(cfu+) in any carcass **in positive flocks** (right, down 2 frames).

After obtaining the joint posterior distribution for model parameters, Figures (5), (6), (7), we can reflect this uncertainty in further predictions concerning the risk estimates.

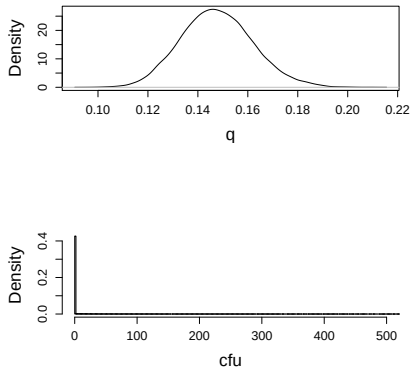

Figure 6: SWE data: Posterior distribution of flock prevalence q (mean: 0.15, 95% CI: [0.12,0.18]) and posterior predictive distribution of cfu in an individual sample. (Percentiles: 50%: 0, 95%: 543, 99%: 5222).

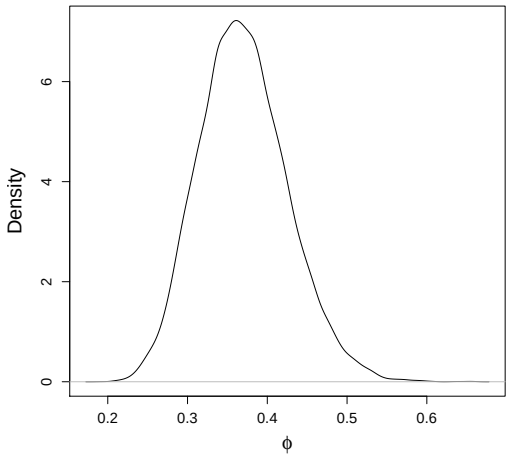

Figure 7: SWE data: Posterior distribution of $\phi = \bar{\sigma}_w^2/(\bar{\sigma}_w^2 + \sigma_b^2)$. (Kernel smoothing applied to MCMC sample).

3 BUGS codes

3.1 Model code for Lindblad et al

```
 model{
# Lindblad data model:
# Assuming within batch prevalence is either 0 or 1.
for(i in 1:NBpos){  # number of batches
for(j in 1:SB[i]){ # number of samples in batch
# concentration data in pos samples:
# conversion of cfu, subtract 2:
logcfu2[i,j] <- (logcfu[i,j]-2)
logcfu2[i,j] ~ dnorm(mub[i],tau_w)
}
mub[i] ~ dnorm(mu,tau_b)
}
# pos/neg-data:
NBpos ~  dbin(q,N)
N <- NBpos+NBneg

# prediction:
cfupred <- cfupredpos*pos + (1-pos)*0
pos ~  dbern(q)
logcfupredpos ~ dnorm(mu,tau_tot)
cfupredpos <- pow(10,logcfupredpos)

# priors:
q ~ dunif(0,1)  # flock prevalence
mu ~ dnorm(0,0.001) # overall mean log-cfu of pos concentrations
tau_tot <- 1/(var_b+var_w) # total precision = 1/(total variance)
var_b <- sigma_b*sigma_b; tau_b <- 1/var_b
var_w <- sigma_w*sigma_w; tau_w <- 1/var_w
sigma_b ~ dunif(0,100) # between batch SD (pos batches)
sigma_w ~ dunif(0,100) # within batch SD (pos batches)'
```

```
phi <- var_w/(var_w+var_b)
}
```

3.2 Model code for Hansson et al

```
model{
# Hansson data model:
for(i in 1:Nflocks){
# sample means reported as data in each batch:
# conversion of cfu, add log10(4):
mlogcfu2[i] <- (mlogcfu[i]+0.60206)
mlogcfu2[i] ~ dnorm(mub[i],tau_sample[i])
tau_sample[i] <- 1/(var_w[i]/n.carcass[i])
mub[i] ~ dnorm(mu,tau_b)
# distribution of tau_w[i] based on reported sample SDs only,
# assuming no information about mub[i]:
tau_w[i] ~ dgamma(alpha[i],beta[i])
sigma_w[i] <- 1/sqrt(tau_w[i]); var_w[i] <- 1/tau_w[i]
alpha[i] <- (n.carcass[i]-1)/2
beta[i] <- (n.carcass[i]-1)*pow(sdlogcfu[i],2)/2
pos.carcass[i] ~ dbin(pwithin[i],n.carcass[i])
}
msigma_w <- mean(sigma_w[])

# prediction (for pos carcasses in pos batches only):
logcfupredpos ~ dnorm(mu,tau_tot)
cfupredpos <- pow(10,logcfupredpos)

# priors:
mu ~ dnorm(0,0.001)
tau_tot <- 1/(var_b+msigma_w*msigma_w)
var_b <- sigma_b*sigma_b; tau_b <- 1/var_b
sigma_b ~ dunif(0,100) # between batch SD (pos batches)
phi <- msigma_w*msigma_w/(var_b+msigma_w*msigma_w)
for(i in 1:Nflocks){pwithin[i] ~ dbeta(apw,2)}
apw ~  dunif(1,10000)
mpw <- apw/(apw+2)
}
```

3.3 Model code for combined Lindblad & Hansson

```
model{
# Lindblad data:
for(i in 1:NBpos){  # number of batches
for(j in 1:SB[i]){ # number of samples in batch
# concentration data in pos samples:
# conversion of cfu, subtract 2:
logcfu2[i,j] <- (logcfu[i,j]-2)
logcfu2[i,j] ~ dnorm(Lmub[i],mtau_w)
}
Lmub[i] ~ dnorm(mu,tau_b)
}
# pos/neg-data:
# probability to sample a positive carcass from a positive flock
NBpos ~  dbin(prpos,N);  prpos <- q*mpw
N <- NBpos+NBneg
##########################
# Hansson data:
for(i in 1:Nflocks){
# sample means reported as data in each batch:
# conversion of cfu, add log10(4):
mlogcfu2[i] <- (mlogcfu[i]+0.60206)
mlogcfu2[i] ~ dnorm(Hmub[i],tau_sample[i])
tau_sample[i] <- 1/(var_w[i]/n.carcass[i])
Hmub[i] ~ dnorm(mu,tau_b)
# sample SDs reported as data:
tau_w[i] ~ dgamma(alpha[i],beta[i])
sigma_w[i] <- 1/sqrt(tau_w[i]); var_w[i] <- 1/tau_w[i]
alpha[i] <- (n.carcass[i]-1)/2
beta[i] <- (n.carcass[i]-1)*pow(sdlogcfu[i],2)/2
pos.carcass[i] ~ dbin(pwithin[i],n.carcass[i])
}
msigma_w <- mean(sigma_w[])
mtau_w <- 1/(msigma_w*msigma_w)
#########################
# prediction:
cfupred <- cfupredpos*pos + (1-pos)*0
pos ~  dbern(prpos)
logcfupredpos ~ dnorm(mu,tau_tot)
cfupredpos <- pow(10,logcfupredpos)

# priors:
q ~ dunif(0,1)
mu ~ dnorm(0,0.001)
tau_tot <- 1/(var_b+msigma_w*msigma_w)
var_b <- sigma_b*sigma_b; tau_b <- 1/var_b
sigma_b ~ dunif(0,100) # between batch SD (pos batches)
phi <- msigma_w*msigma_w/(var_b+msigma_w*msigma_w)
for(i in 1:Nflocks){pwithin[i] ~ dbeta(apw,2)}
apw ~  dunif(1,10000)
mpw <- apw/(apw+2)
}
```

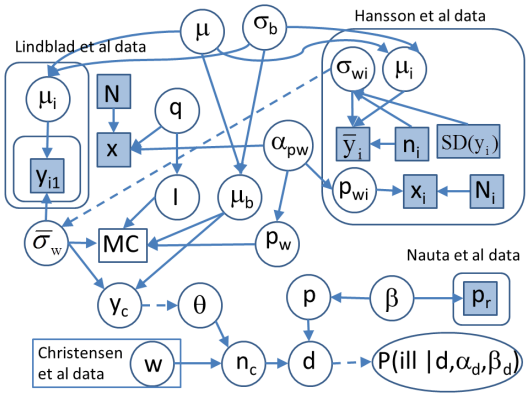

Figure 8: Directed Acyclic Graph (DAG) of the whole model.

4 Modeling risk of cross contamination to a salad

With the posterior distribution of model parameters based on the two data sets above, further predictions can be made about the risk of servings produced. The predictions are based on a sequence of probability distributions so that each of them is conditional to the output of the previous distribution. This makes a hierarchical model, which eventually depends on the top level parameters which can be drawn from the posterior distribution. Hence, the posterior distribution (joint distribution of all parameters) provides a complete description of uncertainty, conditional to the stated data sets and the postulated model. These data form the evidence base for the analysis. The sequence of predictive distributions is described as follows (adapting from the model structure of Nauta et al [3]).

We follow the path of a positive carcass, by starting with the distributions

$$\mu_b \sim \mathrm{N}(\mu, \sigma_b^2)$$
$$y_c \sim \mathrm{N}(\mu_b, \bar{\sigma}_w^2)$$

Which gives the log-cfu for a positive carcass. For the resulting meat, the same assumption is made that the log-cfu per gram of meat is $y_m = y_c - 1$. The cfu per gram of positive meat is then $\theta = 10^{y_m}$. Distribution of serving size w is taken to be log-normal with mean 189 g, and variance 127. (Danish young adult males, see Appendix 1). Then, it may be expected that the number of bacteria in a serving would be $w\theta$, and the number of bacteria in a predicted serving as

$$n_c \sim \mathrm{Poisson}(w\theta) = \mathrm{Poisson}(\lambda)$$

Instead of the risk of the meat meal, we predict the risk of the salad that may get cross contamination during preparation. The fraction of bacteria that may be cross contaminated has been measured in a salad making study, from which we have the results. Instead of using the 55 measurements as *de facto* we can make a small model to describe probable transmission percentages. Each of the measured percentages p_r in the study can be described by a beta-density, $p_r \sim \mathrm{Beta}(2, \beta)$, with prior $\beta \sim \mathrm{U}(1, 10^4)$. An uninformative prior is sufficient here. Then, posterior predictive distribution is computed for these data, which gives $\pi(p \mid p_1, \ldots, p_{55})$. This can be used for drawing values for p. Next, the number of bacteria in the salad is modeled as

$$d \sim \mathrm{Bin}(n_c, p)$$

Note that d is the dose in those servings that have been prepared when in contact with a positive broiler. The dose in any serving would be obtained from a mixture distribution describing both negative broilers and positive broilers so that the above distribution applies only to the latter. Finally, the dose response, given d, is assumed to be given with fixed parameters, $\alpha_d = 0.145, \beta_d = 7.59$, so the probability of illness from eating a salad resulting from a preparation involving a positive broiler is

$$P(\text{ill} \mid d) = 0.33(1 - \frac{\Gamma(\alpha_d + \beta_d)\Gamma(\beta_d + d)}{\Gamma(\beta_d)\Gamma(\alpha_d + \beta_d + d)})$$

The overall probability of illness, conditional to d, would be this multiplied by flock prevalence q and within flock prevalence p_w.

$$P(\text{ill} \mid d, q, p_w) = P(\text{from pos broiler})P(\text{ill} \mid d) = qp_wP(\text{ill} \mid d)$$

This is the probability for a single consumer, given that the dose was d in case there was a positive broiler (and given q, p_w). It can also be interpreted as the proportion of ill people in a population if all ingested dose d whenever the serving was resulting from a contact with a positive broiler. In other words, the probability of illness depends on a set of uncertain parameters, all described by conditional probability densities constructed as a hierarchy. Parameters q and p_w we can simulate from the posterior density described earlier. Parameter d, the dose, depends conditionally on some parameters (n_c, p), which further depend on other parameters explained above, essentially starting from the underlying parameters μ_b, μ, σ_w and σ_b. Also these are parameters which can be drawn from the posterior distribution described earlier.

The probability of illness resulting from some unknown dose d (i.e. unconditional to a specific d), in a serving resulting from a contact with a positive broiler is obtained by integrating over all possible values of d, which further depends on other unknown parameters. This 'average probability' over possible values of d would still be conditional on the other parameters. All these unknown parameters are to be integrated over their distribution describing their uncertainty. In Bayesian analysis, probability is a measure of uncertainty. Therefore, uncertainties may be described either as epistemic or aleatoric, but nevertheless they are descriptions of our uncertainty, given what we assume or know. The goal is not to provide a *prior probability distribution* as a final result (based on assumed underlying parameters) but to provide a *posterior distribution* of the parameters, conditional to data. Predictions of outcome variables then depend on this parameter distribution. Therefore, the goal is to provide $P(\text{ill} \mid \text{data})$ instead of prior distribution $P(\text{ill} \mid \text{unknown parameters})$. Results from this integration are explained below.

The conditional distribution of dose d expresses our uncertainty about what the dose in a serving could be, and this describes our aleatoric uncertainty based on the assumption that the doses may vary from serving to serving, according to the conditional probability $\pi(d \mid n_c, p)$ which depends on the parameters describing that food preparation with n_c cells of bacteria in the raw broiler, and expected proportion of bacteria transferred, p, for that food preparation. In turn, uncertainty of these would be described by the conditional distribution of cell count n_c expressing the aleatoric uncertainty, conditional to serving size w and concentration in raw meat θ of the broiler, etc. Taking into account all uncertain parameters requires integration over the joint posterior distribution of all the parameters. This eventually gives the result, $P(\text{ill} \mid \text{data})$, the conditional probability of illness, given the data (a single number). This can be interpreted as the probability for an individual as well as the expected proportion of individuals who become sick, based on the data set (=evidence) and the assumed model.

If we denote by Ω the set of all such uncertain parameters, the prior probability of the illness (in case of contact with a positive broiler) involves integration of the form

$$P(\text{ill} \mid \text{from pos broiler}, \mathbf{prior}) =$$

$$\int_\Omega \underbrace{P(\text{ill} \mid d)}_{\text{dose resp model}} \pi(d \mid n_c, p)\pi(n_c \mid w, \theta)\pi(\theta \mid \mu_b, \bar{\sigma}_w^2)\pi(\mu_b \mid \mu, \sigma_b^2)\pi(p, w, \mu, \sigma_b^2, \bar{\sigma}_w^2)\, d\Omega$$

The whole integration depends on the prior distribution $\pi(p, w, \mu, \sigma_b^2, \bar{\sigma}_w^2)$ of the underlaying parameters. In this integration, the dose response probability that is conditional to dose d, becomes integrated over all possible values of d, according to the conditional probability of d which further depends on other parameters that are also integrated over their conditional distributions, down to the underlying parameters which have prior distribution. It is possible to plot the distribution of the conditional dose

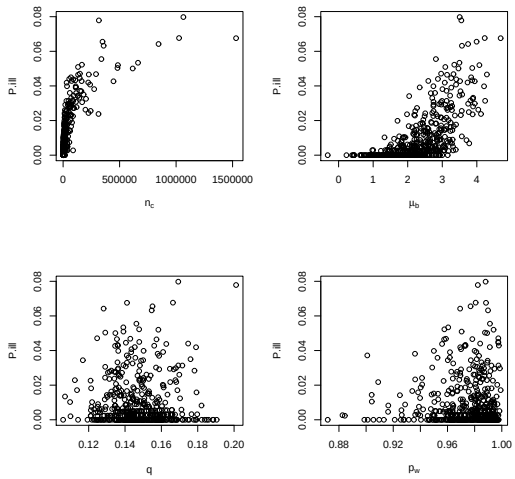

Figure 9: Marginal posterior distributions of $(n_c,\ \text{P.ill}(d,q,p_w))$, $(\mu_b,\ \text{P.ill}(d,q,p_w))$, $(q,\ \text{P.ill}(d,q,p_w))$ and $(p_w,\ \text{P.ill}(d,q,p_w))$ where 'P.ill' in the vertical axis is $\text{P.ill}(d,q,p_w)$ and denotes the conditional probability, $qp_wP(\text{ill} \mid d)$, i.e. given q, p_w, d.

response probability (which has distribution when d has distribution), or to compute the result of the whole integration which gives a single number. What is the distribution of the underlaying parameters and what is it based on, is a key question of scientific inference. The Bayesian solution to this was given in the preceding sections (1)-(3). The prior becomes replaced by the posterior distribution $\pi(p, w, \mu, \sigma_b^2, \bar{\sigma}_w^2 \mid \text{data})$. To get the posterior probability of illness, the integration is:

$$P(\text{ill} \mid \text{from pos broiler}, \mathbf{data}) =$$

$$\int_\Omega P(\text{ill} \mid d)\pi(d \mid n_c, p)\pi(n_c \mid w, \theta)\pi(\theta \mid \mu_b, \bar{\sigma}_w^2)\pi(\mu_b \mid \mu, \sigma_b^2)\pi(p, w, \mu, \sigma_b^2, \bar{\sigma}_w^2 \mid \mathbf{data})\, d\Omega$$

This is the posterior probability concerning a serving resulting from a contact with a positive broiler.

To get the overall posterior probability for any serving, we can start from the conditional probability

$$P(\text{ill} \mid d, q, p_w) = qp_wP(\text{ill} \mid d)$$

which can also be interpreted as the proportion of ill in the population resulting from any servings, given they would ingest dose d whenever the serving was related to positive broiler, and given the probability of a positive broiler is qp_w. This is to be integrated over the unknowns d, q, p_w, with respect to the posterior distribution, so that what is left is

$$P(\text{ill} \mid \text{data}) =$$

$$\int_\Omega qp_wP(\text{ill} \mid d)\pi(d \mid n_c, p)\pi(n_c \mid w, \theta)\pi(\theta \mid \mu_b, \bar{\sigma}_w^2)\pi(\mu_b \mid \mu, \sigma_b^2)\pi(q, p_w, p, w, \mu, \sigma_b^2, \bar{\sigma}_w^2 \mid \mathbf{data})\, d\Omega$$

which is the probability (a single number) for a single individual, but can also be interpreted as the expected proportion of ill in the population, under the same kind of food preparation and food source.

All these integrations are done using MCMC in OpenBUGS, to simulate Monte Carlo samples of all the unknown parameters from their joint posterior distribution. For this, it is sufficient to define all the conditional distributions, and finally the priors of the underlying basic parameters.

Note that whenever we obtain new data (with similar structure), the effect of such new data is directly reflected in the posterior distribution of the parameters, and predictions based on them. Also, the

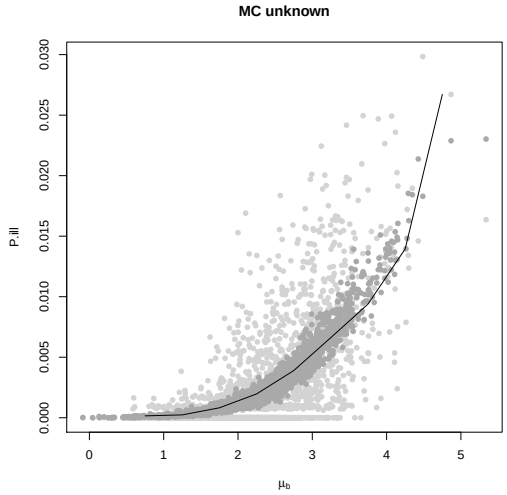

Figure 10: The marginal posterior distribution of $(\mu_b, \text{P.ill}(d, q, p_w))$ (light gray) and $(\mu_b, \overline{\text{P.ill}}_{50}(q, p_w))$ (dark gray), assuming no knowledge of Microbial Criterion status (MC). Here $\text{P.ill}(d, q, p_w)$ means the conditional probability $qp_wP(\text{ill} \mid d)$ and $\overline{\text{P.ill}}_{50}(q, p_w)$ means the 'average' probability of illness in 50 predicted servings each with a different d (average over d), that are each simulated from the common parameters $\mu_b, \sigma_w, E(w), V(w), \beta$. The line is computed from the averages of the $\text{P.ill}(d, q, p_w)$. The distribution $\pi(\mu_b, \overline{\text{P.ill}}_n(q, p_w) \mid \text{data})$ approaches the line, when $n \to \infty$.

Swedish data set could be replaced by similarly structured data from another country, which provides an assessment of that country, based on the specific evidence from that country. In this way, "pan-European" parameter assumptions could be replaced by evidence based uncertainty distributions for each country, based on their data. Also, the uncertainty distribution (of parameters as a whole) is probabilistically coherent statistical inference based on the evidence in the given data sets (as a whole) and we do not need ad hoc choices of fixed parameters for them, or separate reasoning for each parameter estimate that would be unrelated to the reasoning of others, when making inference from a given data set.

4.1 Results

The risk, measured as (posterior) probability of illness, $P(\text{ill} \mid \text{data})$, based on the previous modeling was estimated as 0.00303. This is a single number, and corresponds to the posterior mean of the expression for conditional probability $qp_wP(\text{ill} \mid d)$ over 50,000 MCMC iterations. Considering this conditional probability, a 95% credible interval was [0,0.016], describing the uncertainty due to unknown parameters q, p_w, d. The result should be judged against the number of disease cases that would result from consumption of a large number of servings. This could (probably) amount to an overestimate when compared to official statistics of human infections, although underreporting should also be accounted for. An estimate of campylobacteriosis (reported cases) incidence has been about 78-96 per 100,000 in Sweden, of which about 55% were said to be imported [4]. It is not straightforward to make comparison between epidemiological incidence estimates and the absolute values of illness probability since the model provides illness probability per consumed serving and these servings are assumed similar to the ones in the cross contamination study (salads with cross contamination from raw broilers).

The model contains several assumptions that may be questioned. The basic data were from the two Swedish studies on broiler flocks, but in subsequent calculations the serving size was based on Danish data. Combining data from different countries without explicit 'meta-country' level modeling implies that the result cannot directly describe any specific country, taken literally. Also, the salad study concerns only one particular type of meal which cannot be generalized to other meals and other populations with different practices and behaviors. Therefore, generalizations of the results depend on the plausibility of the assumptions. Extrapolations to other populations without analysis of similar data and the underlying

conditions in those populations remain hypothetical. Moreover, there were assumptions about translating original measured concentrations from carcass rinse samples per ml to represent concentrations per gram, and about further translating this (reduction of 1 log) to concentration in consumed (raw) meat. Also these remain as assumptions and there are no concrete data to provide more informed estimates or distributions. These are examples of uncertainties that cannot be quantified based on the stated evidence that was available here. The assumed values of such quantities are part of the assumed model and could be treated as model uncertainty since the model itself is an assumption. Model assessment would be possible based on e.g. predictions of observable quantities and other statistical model assessment criteria which balance between model complexity and model fit.

However, there were many parameters within the (assumed) model for which uncertainty could be quantified based on the stated evidence. For all those parameters that *can be* probabilistically linked (via the model) to data, the posterior distribution is the probabilistically coherent representation of uncertainty (and this has not been analyzed before). The resulting distribution of parameters then represents holistically the uncertainties about that set of parameters, based on the given empirical evidence about them. Hence, we can reduce the need for hypothetical assumptions about them. This also provides the possibility to infer the results from each data set provided and to make comparisons e.g. between countries, based on country specific data. The different strength of evidence in each data is then reflected to the shape of the posterior distribution of the unknown parameters. This represents how well such unknown parameters can be known, given the stated concrete data. Further predictions can be based on this distribution.

5 Modeling MC

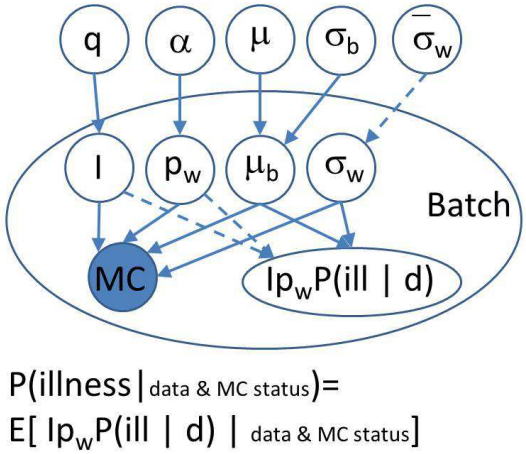

Figure 11: Simplified DAG with batch parameters and MC outcome. Upper level parameters are taken from the posterior distribution described earlier. Batch level observation of MC will update the distribution of the batch parameters. This is reflected on illness probability. Solid arrows denote conditional distribution, dashed arrows functional (algebraic) relationship.

Using the posterior predictive distribution for concentrations in carcasses of a batch, we can predict whether the Microbial Criteria (MC) is met or not. The MC is defined by a triplet of (1) test sample size, (2) allowed maximum number of samples exceeding (3) a given cfu-level ('n/c/m'). In this example we assume MC is set at 5/1/1000. For a batch, the 'acceptance' is modeled as

$$\text{MC} \sim \text{Bern}(I \times P(c \leq 1 | \mu_b, \bar{\sigma}_w^2, p_w) + (1 - I) \times 1)$$

where I is the true status of the batch ($I = 1$: batch contains at least one carcass with campylobacter, $I = 0$: batch is clean). Conditional to $I = 0$, the batch is accepted with probability one. Conditional to $I = 1$, the acceptance probability is $P(c \leq 1 | \mu_b, \bar{\sigma}_w^2, p_w)$ which is defined as

$$\begin{aligned}
&p_w^5(p_m^5 + 5p_m^4(1 - p_m))+ \\
&5p_w^4(1 - p_w)(p_m^4 + 4p_m^3(1 - p_m))+ \\
&10p_w^3(1 - p_w)^2(p_m^3 + 3p_b^2(1 - p_m))+ \\
&10p_w^2(1 - p_w)^3(p_m^2 + 2p_m(1 - p_m))+ \\
&5p_w(1 - p_w)^4 p_m + (1 - p_w)^5
\end{aligned}$$

where $p_m = \Phi((3 - \mu_b)/\bar{\sigma}_w)$. Batch parameters are drawn from the posterior distribution, giving a prediction for this particular batch, $\pi(I|q) = \text{Bern}(q)$, $\pi(p_w \mid \alpha) = \text{Beta}(\alpha, 2)$, $\pi(\mu_b \mid \mu, \sigma_b^2) = \text{N}(\mu, \sigma_b^2)$, and using $\bar{\sigma}_w^2$ for within batch variance.

The Bernoulli variable MC (zero or one) of the batch is predicted at every iteration step in the MCMC simulations, from the parameter values at the same iteration. From this we can analyze the probability that MC is met, given the previous model and conditional to all its data. In 50,000 iterations, this resulted to 0.957. With a high predictive probability, a batch is accepted. (Based on Hansson & Lindblad data).

With the Bayes model, we can also set the MC variable to a given value, pretending that this was an actual observation of the outcome of Microbial Criterion for this batch. A new posterior distribution can then be computed which becomes conditional to this set value. Two such posterior distributions were computed, one conditional to MC=1, and one conditional to MC=0. In the third distribution, no

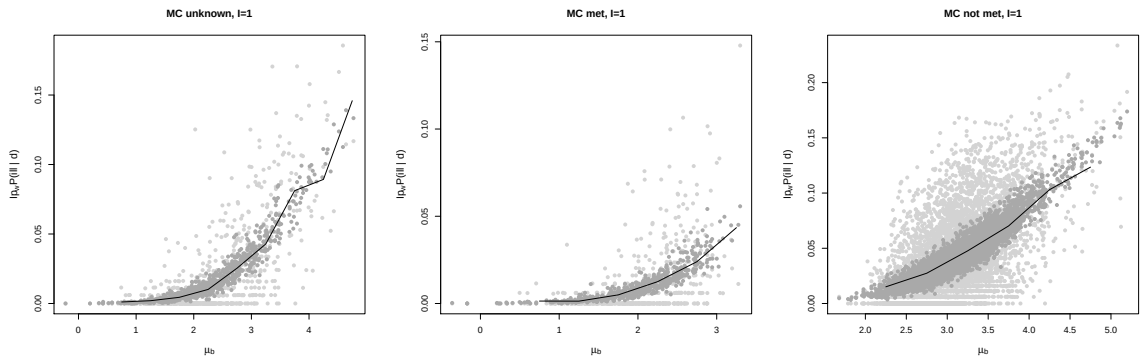

Figure 12: Posterior distribution of $(\mu_b, \text{P.ill}(I, p_w, d))$ and $(\mu_b, \overline{\text{P.ill}}_{20}(I, p_w))$, conditional to 'MC unknown' / 'MC met' / 'MC not met', *and conditional to the hidden variable* $I = 1$, concerning that batch. When MC for the batch is unknown or is met, hidden variable I can be either zero or one. When MC is not met, I must be one.

knowledge of MC outcome is assumed for the batch. Naturally, these distributions are also conditional to all the same data as previously explained. For the batch with observed or unobserved MC, we predict the resulting illness probability based on the new posterior distribution of parameters. *The prediction is then specific for this batch, for which we have added a new piece of evidence: the MC result.* This describes the risk resulting from all such batches for which the same evidence is given.

An average over 50,000 iterations of $\text{P.ill}(I, p_w, d) = Ip_wP(\text{ill} \mid d)$ gives the posterior probability of illness (a single number) in the given situation. By plotting the density of the conditional probability $\text{P.ill}(I, p_w, d)$ we can see how much uncertainty there is due to unknown I, p_w and d, concerning a consumer eating from that batch. This distribution can be plotted under different conditions: (1) MC is unknown, (2) MC is met, (3) MC is not met. The relationship between μ_b and $\text{P.ill}(I, p_w, d)$ is shown in Figure (12), but *conditional to the hidden variable* $I = 1$ for the batch.

Only if MC was not met, we can be sure that $I = 1$. When MC is met or unknown, the hidden status can be either zero or one. Therefore, the posterior of μ_b is given in Figure (13) to show how it depends on the hidden variable (which we cannot be sure of, except if MC was not met), and how the result looks like when we do not condition on its value.

Graphical comparisons of posterior distributions of batch specific $\text{P.ill}(I, p_w, d) = IP_wP(\text{ill} \mid d)$ are difficult because the distributions have a large peak at zero, and a long tail. Distributions that are conditional to the hidden variable $I = 1$ (in addition to the observable MC status) are shown in Figure (14). In the case of 'unknown MC' or 'MC met' the hidden variable is zero with a large probability so that unconditional to I these probabilities are even more peaked at zero. Numerical batch specific results are summarized in Table (1). For a batch with 'MC met' it is less probable that the batch was contaminated ($P(I = 1 \mid \text{data, MC met}) \approx 0.11$ vs $P(I = 1 \mid \text{data, MC unknown}) \approx 0.15$) and *in the case it was*, the contamination level is probably much smaller compared to a batch with 'MC unknown'. This stems from the difference between posterior distributions of batch specific μ_b given $I = 1$ in Figure (13). The difference of distributions for μ_b carries over to the posterior distribution of the dose d, given MC status, and given $I = 1$.

To conclude: in order to study the effect on prognosis due to the added information about MC status for a batch, we can compute predictions from that batch, based on the parameters of that batch. The batch specific parameters (μ_b, p_w, I) are uncertain, and this is described by the posterior distribution of these parameters. This distribution changes depending on the added information about the MC status. Hence, predictions were made of disease risk stemming from a batch for which the MC status is given. **If MC is met**, it becomes more probable that the batch was truly clean ($I = 0$) than if MC is unknown.

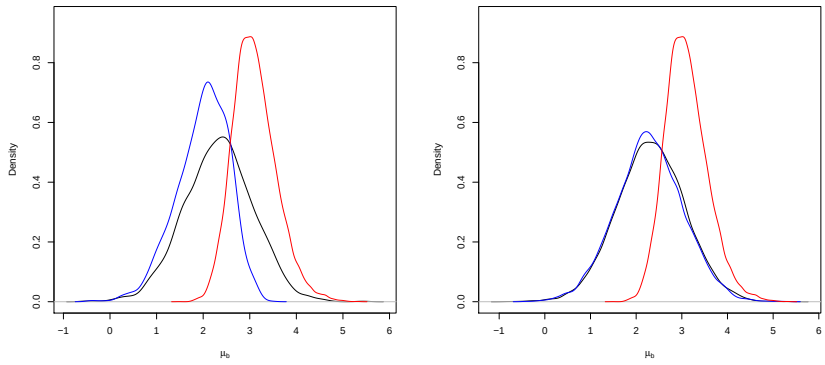

Figure 13: Left: posterior density of μ_b with no condition of MC (black), on condition that MC is met (blue), on condition that MC is not met (red). In all cases on the left, densities are conditional to the hidden variable $I = 1$. Right: without knowing the hidden variable I, the blue and black densities are nearly the same, with means and 95% CI: 2.28 [0.83, 3.75] vs 2.32 [0.85,3.80]. The hidden variable clearly makes a difference: given that the batch is truly contaminated ($I = 1$), the posterior distribution of μ_b in that batch becomes different under each MC status.

Moreover, in the case it yet happens to be contaminated ($I = 1$), the mean concentration μ_b of carcasses in the batch is more probable to be lower than in a contaminated batch for which MC is unknown. In total, this amounts to clearly lower risk for those batches for which MC is met compared to those for which MC is unknown. **If MC is not met**, we can infer that the batch obviously must have been truly contaminated ($I = 1$) and the predicted doses are highest but also prediction interval becomes large. The risk ratio between 'MC met' versus 'MC unknown' was 0.37. For global parameters, (outside the particular batch for which MC status is given), there is less influence from the MC outcome of a batch. This can be seen if we replace I (of the batch) by its conditional expectation q (global), in computing P.ill(q, p_w, d). This quantity would describe the probability concerning a different batch for which we do not have observed MC status, but we expect it to be contaminated with probability q, and apply the same values for $p_w P(\text{ill} \mid d)$ as for the observed batch. It matters at which level of the hierarchy we make comparisons. Also, it must be noted that the path from carcass to serving contained many assumptions, all diminishing the cfu-values along the path. In the extreme case, assuming perfect hygiene, any original amount of cfu would be wiped out from the plates. Then, initial cfu values in the slaughter stage would not matter whatever their levels would be and knowing MC would not be informative anymore.

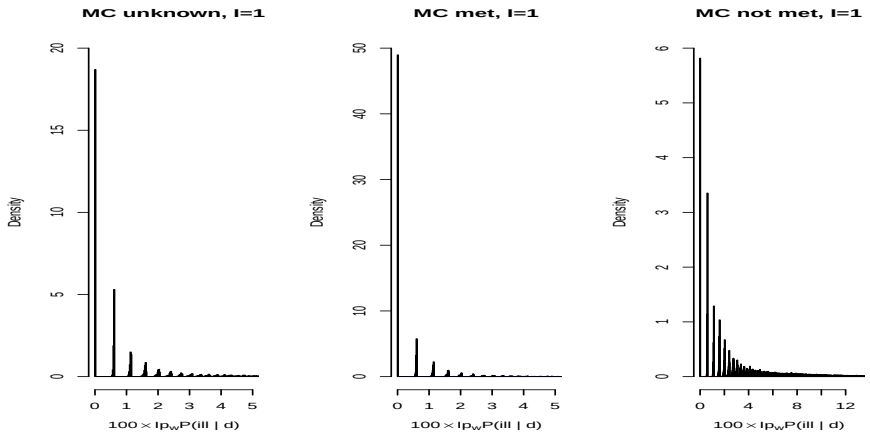

Figure 14: Comparing posteriors of $100 \times IP_wP(\text{ill} \mid d)$, given the MC status and hidden variable $I = 1$

$100 \times Ip_wP(\text{ill} \mid d)$	95% CI	mean	RR
$P(\cdot \mid$ data, MC unknown, $I = 1)$	[0,11.2]	**2.1**	1
$P(\cdot \mid$ data, MC met, $I = 1)$	[0,6.4]	**1.0**	0.5
$P(\cdot \mid$ data, MC not met, $I = 1)$	[0,13.7]	**4.3**	2.1
$P(\cdot \mid$ data, MC unknown)	[0,4.3]	**0.3**	1
$P(\cdot \mid$ data, MC met)	[0,1.6]	**0.1**	0.37
$P(\cdot \mid$ data, MC not met)	[0,13.7]	**4.3**	13.9
mean = posterior probability of illness (single number)			
$100 \times qp_wP(\text{ill} \mid d)$	95% CI	mean	RR
$P(\cdot \mid$ data, MC unknown)	[0,1.64]	0.30	1
$P(\cdot \mid$ data, MC met)	[0,1.58]	0.29	0.94
$P(\cdot \mid$ data, MC not met)	[0,2.07]	0.64	2.09
mean = posterior probability of illness (single number)			

Table 1: N.B. Absolute illness probabilities contain unknown bias and they are difficult to calibrate to epidemiological data (as was briefly discussed in section 4.1.) but RRs might be less affected by this. Therefore, the absolute values of probability should not be taken literally.

6 Model code, including MC and risk estimation

```
model{
# Lindblad data:
for(i in 1:NBpos){  # number of batches
for(j in 1:SB[i]){ # number of samples in batch
# concentration data in pos samples:
# conversion of cfu, subtract 2:
logcfu2[i,j] <- (logcfu[i,j]-2)
logcfu2[i,j] ~ dnorm(Lmub[i],mtau_w)
}
Lmub[i] ~ dnorm(mu,tau_b)
}
# pos/neg-data:
# probability to sample a positive carcass from a positive flock
NBpos ~  dbin(prpos,N);  prpos <- q*mpw
N <- NBpos+NBneg
##########################
# Hansson data:
```

```
for(i in 1:Nflocks){
# sample means reported as data in each batch:
# conversion of cfu, add log10(4):
mlogcfu2[i] <- (mlogcfu[i]+0.60206)
mlogcfu2[i] ~ dnorm(Hmub[i],tau_sample[i])
tau_sample[i] <- 1/(var_w[i]/n.carcass[i])
Hmub[i] ~ dnorm(mu,tau_b)
# sample SDs reported as data:
tau_w[i] ~ dgamma(alpha[i],beta[i])
sigma_w[i] <- 1/sqrt(tau_w[i]); var_w[i] <- 1/tau_w[i]
alpha[i] <- (n.carcass[i]-1)/2
beta[i] <- (n.carcass[i]-1)*pow(sdlogcfu[i],2)/2
pos.carcass[i] ~ dbin(pwithin[i],n.carcass[i])
}
msigma_w <- mean(sigma_w[])
mtau_w <- 1/(msigma_w*msigma_w)
############################
# Predict the outcome of  MC-criteria for a single batch,
# conditionally on  q, pw, mub, mtau_w
# criteria "n=5,c=1,m=1000", met if "c<=1"
pw ~ dbeta(apw,2); mub ~ dnorm(mu,tau_b); I ~ dbern(q)
MC ~ dbern(prmet)
   prmet<-I*(case1+case2+case3+case4+case5+case6)+(1-I)*1
case1 <-  pow(pw,5)*
                            (pow(pmet,5)+
                             5*pow(pmet,4)*(1-pmet))
case2  <-    5*pow(pw,4)*(1-pw)*
                            (pow(pmet,4)+
                             4*pow(pmet,3)*(1-pmet) )
case3  <-   10*pow(pw,3)*pow(1-pw,2)*
                           (pow(pmet,3)+
                             3*pow(pmet,2)*(1-pmet) )
case4 <-    10*pow(pw,2)*pow(1-pw,3)*
                           (pow(pmet,2)+2*pmet*(1-pmet))
case5 <-    5*pw*pow(1-pw,4)*
                            (pmet)
case6 <-    pow(1-pw,5)
pmet  <- phi((3-mub)*sqrt(mtau_w))  # P(m<1000)

# Generate prediction for resulting servings,
# conditional to parameters of the single batch,
# knowing the corresponding MC result generated above.
  logcfuMC ~ dnorm(mub,mtau_w)
# assume meat log(cfu) is 1 log smaller than skin log(cfu):
  theta  <- pow(10,logcfuMC-1)

# weight of a serving:
  wc ~ dlnorm(wmean,wtau); m <- 189; s2 <- 127
     wmean <- log(m)-0.5*log(1+s2/(m*m))
     wtau <- 1/(  log(1+s2/(m*m))  )

# expected dose in a positive serving,
 lambda <- wc*theta

 # predict a random dose in a positive serving
 NC ~ dpois(lambda)
```

```
# probability of bacteria to transfer in a salad making:
for(i in 1:55){
       ptr[i] <- pow(10,-minuslogptr[i])  # from salad experiment
       ptr[i] ~ dbeta(2,ptrb)
   }
   ptrb ~ dunif(1,10000)
   mptr <- mean(ptr[])   # 0.0004320062
# predict number of transferred bacteria in a salad making:
 ptrnew ~ dbeta(2,ptrb)  # predict a new salad making
 d ~ dbin(ptrnew,NC)  # final random dose in a salad

# probability of illness from a salad
# (dose response & prob of positive meat):
P.ill <- I*pw*0.33*
 (1- exp(loggam(a+b)+loggam(b+d)-loggam(b)-loggam(a+b+d)))
a <- 0.145; b <- 7.59   # assumed dose resp parameters

# priors:
q ~ dunif(0,1)
mu ~ dnorm(0,0.001)
tau_tot <- 1/(var_b+msigma_w*msigma_w)
var_b <- sigma_b*sigma_b; tau_b <- 1/var_b
sigma_b ~ dunif(0,100) # between batch SD (pos batches)
phi <- msigma_w*msigma_w/(var_b+msigma_w*msigma_w)
for(i in 1:Nflocks){pwithin[i] ~ dbeta(apw,2)}
apw ~  dunif(1,10000)
mpw <- apw/(apw+2)
}
```

References

[1] M Lindblad, H Lindmark, S T Lambertz, Lindqvist R: Microbiological baseline study of broiler chickens at Swedish slaughterhouses. J Food Prot. 2006, 69, (12), 2875-82.

[2] I Hansson, N Pudas, B Harbom, E Olsson Engvall: Within-flock variations of *Campylobacter* loads in caeca and on carcasses from broilers. Int. J. Food Microbiology. 141, (2010), 51-55.

[3] M J Nauta, M Sanaa, A H Havelaar: Risk based microbiological criteria for *Campylobacter* in broiler meat in the European Union. Int. J. Food Microbiology. 158, (2012), 209-217.

[4] A Ternhag, A Törner, Åke Svensson, J Giesecke, K Ekdahl: Mortality following *Campylobacter* infection: a registry-based linkage study. BMC Infectious Diseases. 2005, 5: 70. doi:10.1186/1471-2334-5-70