Performing a Quantitative Microbial Risk Analysis Using Second Order Monte Carlo Simulation

James Dawber[1], Jennifer Brown[1] and Beverley Horn[2]

[1]*University of Canterbury, Christchurch, NEW ZEALAND,* [2]*Institute of Environmental Science and Research, Christchurch, NEW ZEALAND*

Abstract

Quantitative microbial risk analysis modelling is increasingly being used in food safety as a tool to evaluate health risks. Accurately forming such models can be very difficult due to the uncertainty in the available data. Second order Monte Carlo simulation allows for the inclusion of this uncertainty and separates it from the variability incorporated in the model. This modelling process is illustrated by performing a simple risk assessment on the infection of Campylobacter during chicken preparation at a typical New Zealand barbecue.

Key words: Quantitative microbial risk analysis, *Campylobacter*, Variability, Uncertainty

1. Introduction

Risk analysis is the qualitative or quantitative assessment of the potential impact of risk [1]. Risk analysis is an important and widely used tool, and is vital in high risk fields such as the health and chemical industries. The accurate use of risk analysis can help in the understanding and management of risk which can lead to precautionary action and the protection of those at risk.

Quantitative microbial risk analysis (QMRA) is specifically used to estimate health risks that arise due to exposure to harmful microbes. Typically in a food safety related QMRA, the aim is to statistically model the transmission of a specific pathogen through a chain of processes [2]. The resulting model will acquire an estimate of the probability of adverse effects upon consumption. The model can then give information about which links in the chain lead to a higher abundance of the pathogen. This is very helpful when trying to prevent or limit the effect of the pathogen upon consumption.

Applying an accurate QMRA model can potentially be quite complex as there are often many details that need to acknowledged. Finding the abundance of the pathogens on certain surfaces is essential to building a QMRA model, but doing so is both time-consuming and inexact. Therefore, there is likely to be a large amount of uncertainty in these data sets, and thus in the models. Expressing this uncertainty in the models is an integral part of the QMRA.

The inability to precisely predict future outcomes is due to two components, variability and uncertainty [1]. Variability represents the naturally occurring random heterogeneity within a population [3], [4]. Uncertainty is the lack of knowledge, or level of ignorance, about the parameters representing the system being modelled [1]. The variability of a system cannot be reduced, however the uncertainty of a system can be reduced upon further study or measurement. The combination of variability and uncertainty is referred to as total uncertainty [1].

Keeping uncertainty and variability separate can be very important and useful to a risk analysis model [4]. Mixing the two together will bring about model outputs that are difficult to interpret, as they will include the mixture of both variability and uncertainty. By separating them, the effect that both the variability and uncertainty has on the model output can be deduced. If uncertainty turns out to be the dominating component, then one would know that the model can be improved by further collection of information [1]. Separating variability and uncertainty can aid in the understanding of the system being modelled, and give valuable insight into where in the model the total uncertainty can be reduced.

Probability distributions are used to represent both the variability of the population and the uncertainty of the parameters used in the variability distributions. The core structure of the risk analysis model is based around the variability distributions in the system being modelled. Once in place, the uncertainty distributions as-

sociated with the parameters representing the variability can be added. Monte Carlo simulation can then be applied by first generating an estimate of the parameters from the uncertainty distributions. These parameters are then fixed while large samples of realisations can be generated from the corresponding variability distributions. This looping procedure can then be repeated many times with different parameter estimates. This two-stage process is referred to as a second order Monte Carlo simulation.

It is widely recognised that a major cause of foodborne illnesses is a result of unsafe food practices in the home [5]. Cross-contamination is the transfer of pathogens from one object to another and is believed to be a major factor in causing these illnesses [6]. In particular, the pathogen *Campylobacter* is recognised as a substantial cause of these foodborne illnesses [7]. In fact it is amongst the top five pathogens which cause the most infections worldwide [8]. One of the leading causes of *Campylobacter* infections is through the consumption or handling of chicken products [7]. With this in consideration, it is understandable why food safety authorities may want to conduct a QMRA on the risk of infection of *Campylobacter* through the unsafe preparation of chicken in the home.

By performing a QMRA on the cross-contamination of *Campylobacter* during chicken preparation in the home, estimates can be gained about the number of colony-forming units (CFUs) of *Campylobacter* that are consumed, and hence probabilities of infection can be deduced. If these probabilities are substantially high (and the model is believed to be accurate), food safety authorities can then take appropriate action in reducing the risk of infection. The model can also be used to assess which parameters (which may represent cooking malpractices) are more influential in causing a higher contamination of *Campylobacter*. Clearly, cross-contamination can occur in many different forms which can make model fitting an arduous task, thus separating variability and uncertainty is an additional task that is commonly neglected [3]. For this reason, it is often beneficial to keep the model simple by emphasising the more important risk factors.

In this investigation a simple QMRA was performed to estimate the proportion of people who will be infected by *Campylobacter* in a common home barbecue setting. The model was constructed using second order Monte Carlo simulation to allow for the separation of the variability and uncertainty in the system. The purpose of the model is to provide a basis for using second order models in QMRA, and more specifically in food safety risk analyses.

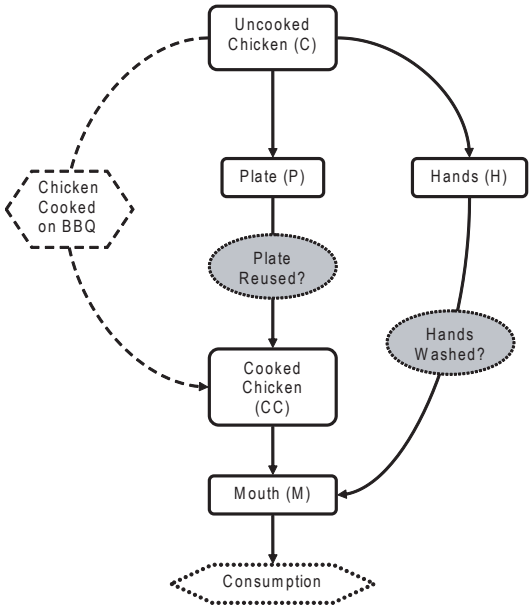

Figure 1: Model summary of the cross-contamination of *Campylobacter*. The solid arrows represent the transfer of *Campylobacter* from one surface to another.

2. Methods

The scenario used as a basis for the model is based on a very simple barbecue preparation process. Every step in the scenario is performed by the same person, who consumes the chicken on their own. The scenario is as follows:

1. Uncooked chicken is picked up by hand and placed on a plate.
2. The plate with the chicken is then taken to the barbecue for cooking.
3. The chicken is placed on the barbecue with tongs and cooked appropriately. The plate remains by the barbecue.
4. Once chicken has been cooked it is placed back on the same plate, using the tongs.
5. The cooked chicken is then eaten with unwashed hands.

There are two fundamental food preparation malpractices in this scenario. The first is that hands are not washed, the second is that the plate is reused. Transmission by tongs is omitted due to assumptions detailed below. The chain of possible cross-contaminations is summarised in Figure 1. There are two paths which can lead to a contaminated chicken being consumed. The first is that *Campylobacter* in the uncooked chicken (C) goes to

the hand (H), and then from the hand to the mouth (M). The other path is when *Campylobacter* transfers from the uncooked chicken to the plate (P), and then from the plate to the cooked chicken. This scenario now establishes a foundation in which a quantitative model can be formed.

In order to form a quantitative model, data first needs to be obtained. In terms of this model, there are three main parts that need to be quantified. There is the number of *Campylobacter* CFUs in the initial piece of uncooked chicken, the CFU transfer rates during cross-contamination and the probabilities of both a person washing their hands and reusing their plate. The data used to estimate these three parts of the model were provided by the Institute of Environmental Science & Research (ESR) in New Zealand. Because the model is primarily focused around cross-contamination it is the transfer rate data which is the most vital. Transfer rates (TR) are defined as the percentage of CFUs that transfer from one surface to another [9]. The probability distributions of TR are distinctly right-skewed, and is considered to be best represented as a normal distribution on $log_{10}(TR)$ with mean μ and variance σ^2 [9]. The distribution of the number of CFUs in the initial uncooked chicken, given it is contaminated, is also assumed to be a normal distribution of the log transformation. The parameters of this normal distribution were supplied by previously gathered estimates performed by ESR, therefore it is assumed there is no uncertainty around these parameters. Finally, the probabilities of a person washing their hand and reusing their plate were estimated from a survey on food preparation practices, conducted by ESR. With all the data available and a knowledge of all the necessary probability distributions associated with variability, the core structure of the model can be built.

Thus far, the uncertainty in the model has not been considered. Firstly, the parameters which have uncertainty must be identified. This is the μ and σ^2 parameters used in the normal distribution on $log_{10}(TR)$ and the probabilities related to unwashed hands (UWH) and reuse of plate (RUP). There are two primary objectives when representing uncertainty in a model, selecting which probability distribution to use, and then selecting the associated parameters.

Through bootstrapping the $log_{10}(TR)$ data, it was revealed that the distribution of the sample mean and sample standard deviation appear to be approximately bell-shaped, hence they were assumed to be normally distributed with parameters selected accordingly from the bootstrap samples. Furthermore the Central Limit Theorems provides a theoretical basis for representing the distribution of the parameter μ with the normal distribution. Using Bayesian inference with diffuse priors and the use of sample proportions, the distribution of the probabilities associated with unwashed hands and reuse of plate can be estimated using the beta distribution. The representation of these uncertainty distributions in the model, combined with the variability distributions are shown in Table 1.

Variable	Variability and Uncertainty Distributions
$log_{10}(TR_{X\rightarrow Y})$	$Normal[\mathbf{Normal}(\bar{\mathbf{x}}, \mathbf{s}^2_{\bar{\mathbf{x}}}), \mathbf{Normal}(\mathbf{s}^2, \mathbf{s}^2_{\mathbf{s}^2})]$
UWH	$Binomial(1, \mathbf{Beta}(\mathbf{43}, \mathbf{275}))$
RUP	$Binomial(1, \mathbf{Beta}(\mathbf{17}, \mathbf{288}))$

Table 1: Probability distributions representing the different variables. The uncertainty distributions are in bold.

In the scenario, certain assumptions must be made for simplicity in the model to be maintained. A major assumption is that the heat of the barbecue kills all *Campylobacter* in the chicken, and hand washing removes all *Campylobacter* on the hand. Another assumption is that the number of *Campylobacter* on the contaminated surfaces, (i.e. hand, plate, cooked chicken) remains constant. This means there is no growth or deaths throughout the procedures, except during hand washing. The cross-contamination from hands to cooked chicken will not be considered in the model, assuming that hands to mouth is the main cross-contamination in regards to hands.

The model quantifies the progression of CFUs from the initial number of CFUs in the uncooked chicken (CFU_C) through to the final number of CFUs that will be consumed (CFU_F). It is the progression throughout the cross-contamination events that will influence CFU_F. All the procedures that take place in the scenario are assumed independent and we assume all uncooked chicken is contaminated. The CFUs at each surface were calculated using:

- $CFU_H = CFU_C \times TR_{C-H}$
- $CFU_P = (CFU_C - CFU_H) \times TR_{C-P}$
- $CFU_{CC} = CFU_P \times TR_{P-CC}$
- $CFU_M = CFU_H \times TR_{H-M}$

And thus the final amount of CFUs that are consumed is given by:

$$CFU_F = CFU_{CC} \times RUP + CFU_M \times UWH$$

Once CFU_F is found, the corresponding probability that this number of CFUs will cause an infection can be deduced using a dose-response relation. From this the risk can then be formally assessed.

3. Results

To quantify the different effect variability and uncertainty have on the model consider the credible intervals in Table 2. The 95% credible intervals were calculated for the model with no uncertainty, and for the model with uncertainty. The credible interval for the proportion of people who are at risk of infection is (0.027, 0.045) with no uncertainty and (0.022, 0.053) with uncertainty. Naturally, the interval width for the latter is larger because it includes the uncertainty as well as the existing variability. Therefore the increase in the interval width is due to the uncertainty. The fourth column shows what proportion of the interval width is contributed by the variability. In this case it is 0.581 which suggests the impact of variability and uncertainty is reasonably similar. For the confidence intervals associated with the proportion of people who have higher probabilities of infection, the variability contribution to the credible interval seems to increase as the probability of infection increases.

x	Without Uncert.	With Uncert.	$\frac{\text{Interval Length of WU}}{\text{Interval Length of WOU}}$
0	(0.027, 0.045)	(0.022, 0.053)	0.581
0.05	(0.013, 0.027)	(0.011, 0.031)	0.700
0.1	(0.008, 0.020)	(0.007, 0.023)	0.857
0.25	(0.002, 0.010)	(0.002, 0.012)	0.800

Table 2: 95% credible intervals of the proportion of people, x, who are at risk of infection for models without (WOU) and with uncertainty (WU), for infection probabilities greater than 0, 0.05, 0.1 and 0.25.

After the level of risk has been assessed, it is important to know which variables in the model have the greatest impact on *Campylobacter* infections. The correlations between each variable in the model and the CFU_F were calculated. The variable with the highest correlation of 0.158 is CFU_C, followed by RUP with 0.140. All other correlations are less than 0.06. These correlations are very weak due to the many contributing variables that effect the CFU_F. The complexity of the cross-contamination process results in many interactions between the variables. This upsets the possibility of any strong correlations with the final output CFU_F.

4. Conclusion

The model presented here provides simple guidelines on how to use second order Monte Carlo simulation to assess risk in a food safety QMRA. A wide range of models can be adapted for a QMRA. The methods used to construct this model were primarily used because they are the most common or simplistic methods available.

Although the literature discusses ways of going about separating variability and uncertainty, it does not seem to discuss in much detail ways to go about analysing or quantifying the differing effects that this separation brings to the output of the second order Monte Carlo model. Often the importance of separating the two is stressed, but not how to summarise the implications of this separation. In our model, the effect of the separation was quantified by comparing the differing margins of error that resulted from a model without uncertainty and a model with uncertainty. This however, may not be the best way to go about this. Further work should investigate better ways of quantitatively expressing this separation.

Second order Monte Carlo simulation is an extremely helpful tool used in risk analysis, as it allows for the separation of variability and uncertainty which leads to more useful and interpretable model outputs. Hence the use of such models can lead to better risk management, particularly in the food safety field.

References

[1] D. Vose, Risk Analysis, John Wiley & Sons Ltd, Chichester, England, 2000.

[2] M. J. Naata, Separation of uncertainty and variability in quantitative microbial risk assessment models, International Journal of Food Microbiology 57 (2000) 9–18.

[3] F. Wu, Y. Tsang, Second-order monte carlo uncertainty/variability analysis using correlated model parameters: application to salmonid embryo survival risk assessment, Ecological Modelling 177 (2004) 393–414.

[4] R. Pouillot, I. Albert, M. Cornu, J.-B. Denis, Estimation of uncertainty and variability in bacterial growth using bayesian inference; application to listeria monocytogenes, International Journal of Food Microbiology 81 (2003) 87–104.

[5] T. Wong, S. Gilbert, R. Lake, B. Horn, Domestic Food Practices In New Zealand 2006-2007 Project Report, Institute of Environmental Science & Research Limited, Christchurch, New Zealand, 2008.

[6] Y. Chen, K. M. Jackson, F. P. Chea, D. W. Schaffner, Quantification and variability analysis of bacterial cross-contamination rates in common food service tasks, Journal of Food Protection 64 (1) (2001) 72–80.

[7] S. D. Mylius, M. J. Nauta, A. H. Havelaar, Cross-contamination during food preparation: A mechanistic model applied to chicken-borne *Campylobacter*, Risk Analysis 27 (4) (2007) 803–813.

[8] E. D. van Asselt, A. E. I. de Jong, R. de Jong, M. J. Nauta, Cross-contamination in the kitchen: Estimation of transfer rates for cutting boards, hands and knives, Journal of Applied Microbiology 105 (2008) 1392–1401.

[9] F. Pérez-Rodriguez, A. Valero, E. Carrasco, R. M. Garcia, G. Zurera, Understanding and modelling bacterial transfer to foods: A review, Trends in Food Science & Technology 19 (2008) 131–144.