Microbial Risk Analysis 4 (2016) 36–42

Contents lists available at ScienceDirect

Microbial Risk Analysis

journal homepage: www.elsevier.com/locate/mran

A Bayesian approach to estimating the uncertainty in the distribution of *Cronobacter* spp. in powdered infant formula arising from microbiological criteria test outcomes

Friedrich von Westerholt[a], Ursula Gonzales-Barron[b], Francis Butler[a,*]

[a] UCD School of Biosystems and Food Engineering, University College Dublin, Belfield, Dublin 4, Ireland
[b] CIMO Mountain Research Centre, School of Agriculture, Polytechnic Institute of Braganza, Portugal

ARTICLE INFO

Article history:
Available online 28 August 2016

Keywords:
Microbiological criteria
Cronobacter
Bayesian
Food safety
Sampling plan
Powdered infant formula

ABSTRACT

The application of microbiological criteria related to foods has become well established for the protection of public health. Sampling plans will more likely detect a microorganism when the level of contamination is high. However, as the concentration of the microorganism drops, detection becomes more and more infrequent. *Cronobacter* spp. is an opportunistic pathogen that can occur infrequently and in low concentrations in powdered infant formula (PIF) with a distribution that is typically heterogeneous. This paper developed a Bayesian approach to quantify the uncertainty in the concentration of *Cronobacter* spp. clusters that may be present in a batch of PIF depending on the outcome of a sampling plan. Two approaches were developed. The first was a Bayesian methodology using a spreadsheet approach to develop the appropriate likelihood and posterior distributions based on an uninformed prior distribution. The second approach was similar but used an algebraic approach rather than a spreadsheet numerical approximation to characterise the uncertainty. Different sampling plans were considered based on the EC Microbiological Criteria for *Cronobacter* spp. When a zero positive test was the outcome of the sampling plans considered, the Bayesian analysis indicated that while the most likely outcome for all the sampling plans considered was zero clusters present, the analysis indicated that the true number of clusters present could be as high as several thousand clusters per tonne of powder depending on the sampling plan. The algebraic approach demonstrated that for zero or one positive tests, the uncertainty distribution could be approximated by a gamma distribution. Choice of the prior distribution influenced the level of uncertainty. The Bayesian approach demonstrates that even when zero positives are detected for a given sampling plan, there remains a considerable uncertainty in the true number of microorganisms that may be present undetected in a consignment of powder.

1. Introduction

The application of microbiological criteria related to foods has become well established for the protection of public health. Codex Alimentarius (CAC 1997) defines a microbiological criterion as "a risk management metric which indicates the acceptability of a food, or the performance of either a process or a food safety control system following the outcome of sampling and testing for microorganisms, their toxins/metabolites or markers associated with pathogenicity or other traits at a specified point in the food chain" . Microbiological criteria are a potential tool for evaluating food safety and risk management systems, both for the industry and for food regulatory authorities. CODEX sets out that a microbiological criterion consists of the following components (CAC 1997):

- The purpose of the microbiological criterion;
- The food, process or food safety control system to which the microbiological criterion applies;
- The specified point in the food chain where the microbiological criterion applies;
- The microorganism(s);
- The microbiological limits;
- A sampling plan defining the number of sample units to be taken (n), the size of the analytical unit and where appropriate, the acceptance number (c);
- Depending on its purpose, an indication of the statistical performance of the sampling plan; and
- Analytical methods and their performance parameters.

* Corresponding author.
E-mail address: f.butler@ucd.ie (F. Butler).

http://dx.doi.org/10.1016/j.mran.2016.08.002
2352-3522/

The effectiveness of a sampling plan can be assessed by using operating characteristic curves, provided the distribution of microorganisms is known. An operating characteristic curve relates the probability of lot acceptance in terms of the number of samples tested, and the proportion of the sampled units that exceed the specified acceptable level (van Schothorst et al. 2007). Sampling plans are more likely to detect a microorganism when the level of contamination is high. However, as the concentration of the microorganism drops, detection becomes more and more infrequent. Consequently, many sampling plans will fail to detect a microorganism when it is present at very low levels, particularly if the sample number and size are relatively small. This can be further exacerbated if the distribution of the microorganism is heterogeneously distributed through the food product (Jongenburger et al., 2011; Gonzales-Barron et al. 2013). In a manufacturing environment, it is conceivable that for a given sampling plan, there may well be an extended period where the microorganism of interest is not detected even though the organism continues to be present, perhaps intermittently, in low concentrations in the product.

Cronobacter spp. is an opportunistic pathogen widely associated with powdered infant formula (PIF) (Healy et al. 2009; Jongenburger 2012; Yan et al. 2012). Because *Cronobacter* spp. can cause neonatal death, it is a pathogen of concern for the PIF industry (Jongenburger et al. 2011). Jongenburger et al. (2011) has discussed the heterogeneous distribution of *Cronobacter* spp. in PIF and how the spatial distribution of microorganisms affects the outcome of sampling plans. Mussida et al. (2013) further explored the heterogeneous distribution of *Cronobacter* spp. in PIF and used mixture distributions such as the Poisson-logarithmic distribution to characterise the distribution of microorganisms in a food. That study assumed that microorganisms are present in dairy powder as clusters or groups of individual cells and these clusters are Poisson distributed. The number of cells in a cluster varies and can be described by a distribution such as a logarithmic distribution. Subsequent intensive experimental analysis of powder samples from eleven production lots detected by the industry as positive for *Cronobacter* spp. demonstrated that in the powder tested, the clusters were randomly distributed (Mussida 2013). The current investigation will also assume that clusters of *Cronobacter* spp. are randomly distributed in a batch of PIF in a Poisson process. Moreover, it is assumed that for a presence/absence type test for the detection of *Cronobacter* spp. in PIF such as ISO 22,964 (ISO, 2006), detection is triggered by the presence of one or more clusters in the test sample. The actual (unknown) number of cells in the cluster is of no real significance providing the test is sufficiently sensitive to detect low number of organisms (as low as one cell is possible) that some clusters may contain.

The microbiological criteria for foodstuffs in the EU are set out in Commission Regulation (EC) 2073/2005 (EC, 2005). For *Cronobacter* spp. in PIF, each batch must be tested with thirty 10 g samples and all samples must show absence. This is a particularly stringent criteria given the high risk associated with this pathogen in PIF. The majority of the other sampling plans in EC 2073/2005 only have a sample size of five. PIF manufacturers carry out a considerable amount of monitoring for *Cronobacter* spp. With the introduction of enhanced biosecurity and cleaning processes, detection rates have become more infrequent with zero detects routinely reported across large number of batches. However, sampling by its nature only tests a very small proportion of the total production and there is always the possibility of defective material slipping through. A Bayesian approach allows the possibility of quantifying the underlying uncertainty associated with a test outcome (including a non-detect) to give a more realistic understanding of the true level of the microorganism that may present in the foodstuff depending on the test outcome. The objective of this paper is to use a Bayesian approach to quantify the uncertainty in the concentration of clusters of a microorganism that may be present in a batch of product such as PIF when the results of a sampling plan indicates presence/absence of the microorganism.

2. Method 1: model development

Two approaches were developed to quantify the uncertainty associated with microbiological criterion sampling for *Cronobacter* spp. in PIF. The first used a Bayesian methodology based on a number of spreadsheet approaches developed by Vose (2008) for assessing animal import risks such as estimating the number of animals in a group or population that are actually infected given a positive test outcome for a limited number of animals. The second approach was similar but used an algebraic approach rather than a spreadsheet numerical approximation to characterise the uncertainty. Bayesian inference is the use of Bayes' theorem for using data to improve an estimate of a parameter (Vose, 2008). It essentially involves three steps: Firstly, determine a prior, which is a confidence distribution based on prior belief of the parameters; secondly, find an appropriate likelihood function for observed data; and thirdly, calculate the posterior confidence distribution based on multiplying the prior and likelihood functions and normalising the output. In the absence of better information, Vose conservatively assumed a prior distribution that was a discrete uniform distribution over all non-negative integers. The likelihood for the range of discrete values considered possible was then calculated. This was multiplied by the prior to form the posterior distribution, and then normalised to form the normalised posterior and hence the confidence of each individual value occurring.

3. Bayesian spreadsheet model

The model assumes that, for presence/absence microbiological testing for *Cronobacter* spp. in PIF, a positive test is triggered by the presence of one or more contaminated clusters in a test sample. Vose's Bayesian spreadsheet models (Vose, 2008) were discrete distributions (0, 1, 2, 3 etc. animals). The equivalent discrete distribution in the current study is the number of *Cronobacter* spp. clusters that may be present in a consignment of powder. The consignment or unit chosen was a tonne of powder. Commercial producers commonly fill multiple one tonne bags of PIF during a production batch as an intermediate step prior to final filling into retail packs. The choice of a tonne of powder also results in a substantial range (see the results section, for example, Fig. 1 or 2) in the number of *Cronobacter* spp. clusters that are estimated to be potentially present.

An Excel spreadsheet (Table 1) was created to run the Bayesian model. Column B represents all possible values for the number of *Cronobacter* spp. clusters that are potentially present in a tonne of powder. The values are discrete non-negative integers (0, 1, 2, etc.) up to some very high number. The choice of this number is such that the confidence value for this number in the Bayesian posterior distribution is close to zero (C/F Fig. 1). Column C represents the equivalent number of clusters per test sample. Following Vose's approach, in the absence of better information, the model conservatively assumes an uninformed prior distribution that is a discrete uniform distribution over all non-negative integers (Column D). For the uninformed prior, all values for the number of clusters/tonne are considered to be equally likely to occur and are assigned a value of one.

The model assumes that on the scale being considered, the *Cronobacter* spp. clusters are randomly distributed and can be represented by a Poisson distribution with parameter λ, the mean number of clusters in the sample tested. Column C serves as the λ input for a Poisson function in column E, which calculates the probability of a zero positive sample for each of the *Cronobacter*

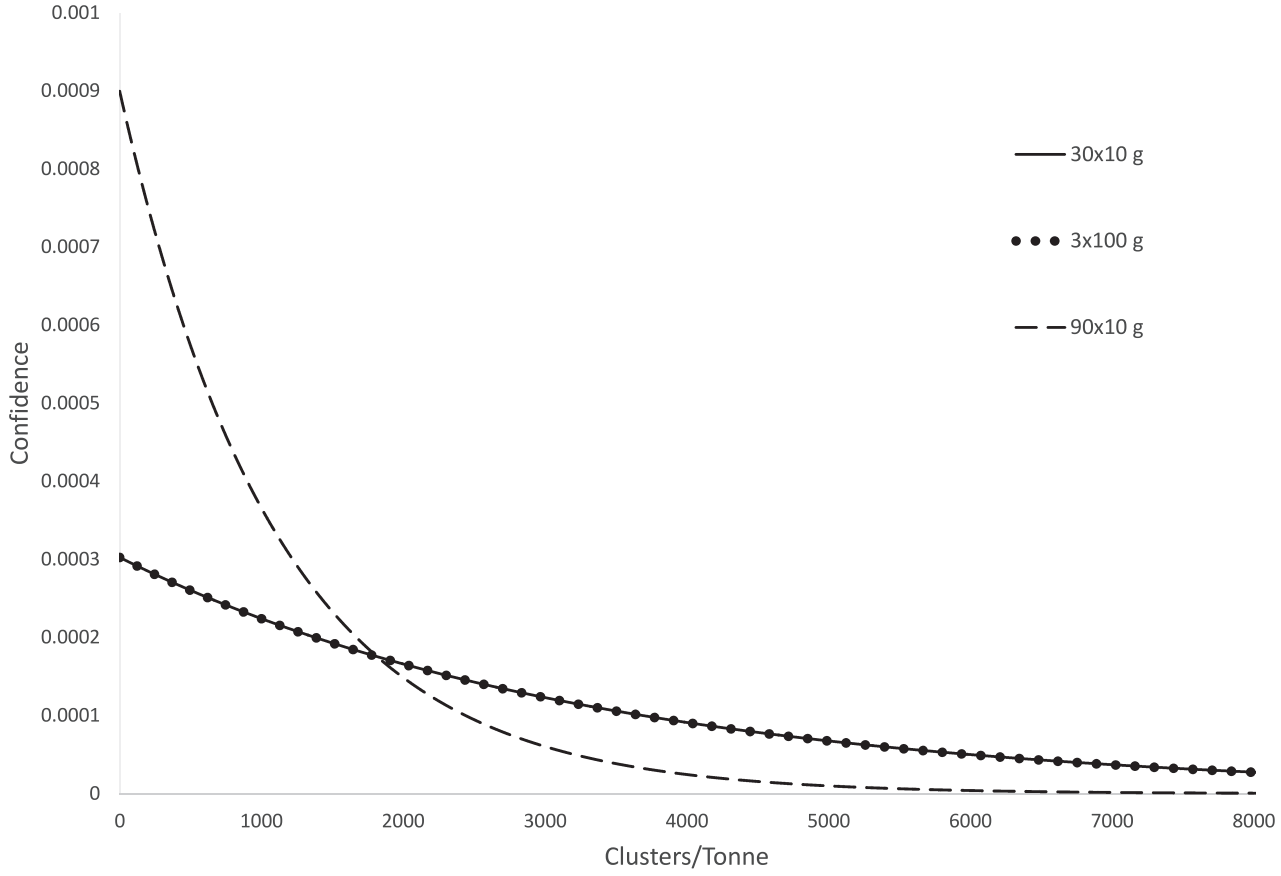

Fig. 1. Uncertainty using the Bayesian spreadsheet approach in the number of *Cronobacter* spp. clusters present in a tonne of powder associated with a zero positive result (The gamma function approximations coincide with the spreadsheet derived plots and are consequently not shown).

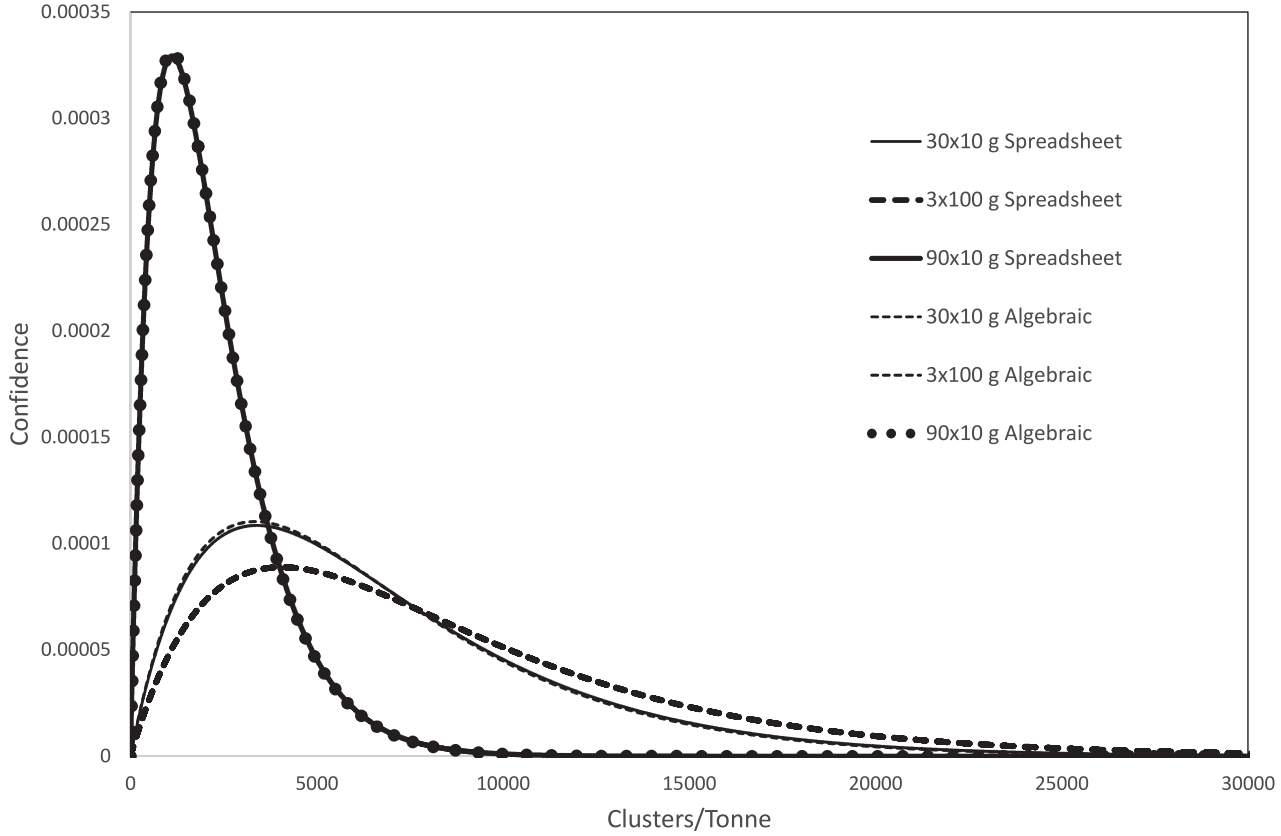

Fig. 2. Uncertainty using the Bayesian spreadsheet and algebraic approaches in the number of *Cronobacter* spp. clusters present in a tonne of powder associated with a single positive result (The plots for the gamma function approximations for thirty 10 g samples and three 100 g samples are identical).

Table 1
Spreadsheet Bayesian model to quantify the uncertainty associated with sampling for *Cronobacter* spp. in powder infant formula.

	A	B	C	D	E	F	G	H	I	J
1										
2								**Formulae Table**		
3								D5:D7	Constants	
4			**Parameters**				Clusters/Tonne	B17:B16024	{0,...,16000}	
5		Number of Samples	n	30			Clusters/m	C17:C16024	={0,...,16000}*m/1e6	
6		Number of Positive Samples	s	0			Prior	D17:C16024	1	
7		Sample Weight	m	10			Probablility of 0	E17:E16024	=POISSON.DIST(0,{C17:C16024},0)	
8							Probablity >0	F17:F16024	=1-{E17:E16024}	
9							Likelihood	G17:G16024	=BINOM.DIST(D6,D5,{F17:F16024},0)	
10							Posterior	H17:H16024	={D17:D16024}*{G17:G16024}	
11							Sum of Posterior	H14	=Sum(H17:H16024)	
12							Normalised Posterior	I17:I16024	={H17:H16024}/H14	
13										
14							Sum of Posterior	3305.569664		
15										
16		**Clusters/Tonne**	**Clusters/m**	**Prior**	**Probablility of 0**	**Probablity >0**	**Likelihood**	**Posterior**	**Normalised Posterior**	
17		0	0	1	1.000000	0.000000	1.000000	1.000000	0.000303	
18		1	0.00001	1	0.999990	0.000010	0.999700	0.999700	0.000302	
19		2	0.00002	1	0.999980	0.000020	0.999400	0.999400	0.000302	
20		3	0.00003	1	0.999970	0.000030	0.999100	0.999100	0.000302	
16022		15998	0.15998	1	0.852161	0.147839	0.008235	0.008235	0.000002	
16023		15999	0.15999	1	0.852152	0.147848	0.008232	0.008232	0.000002	
16024		16000	0.16	1	0.852144	0.147856	0.008230	0.008230	0.000002	
16025										

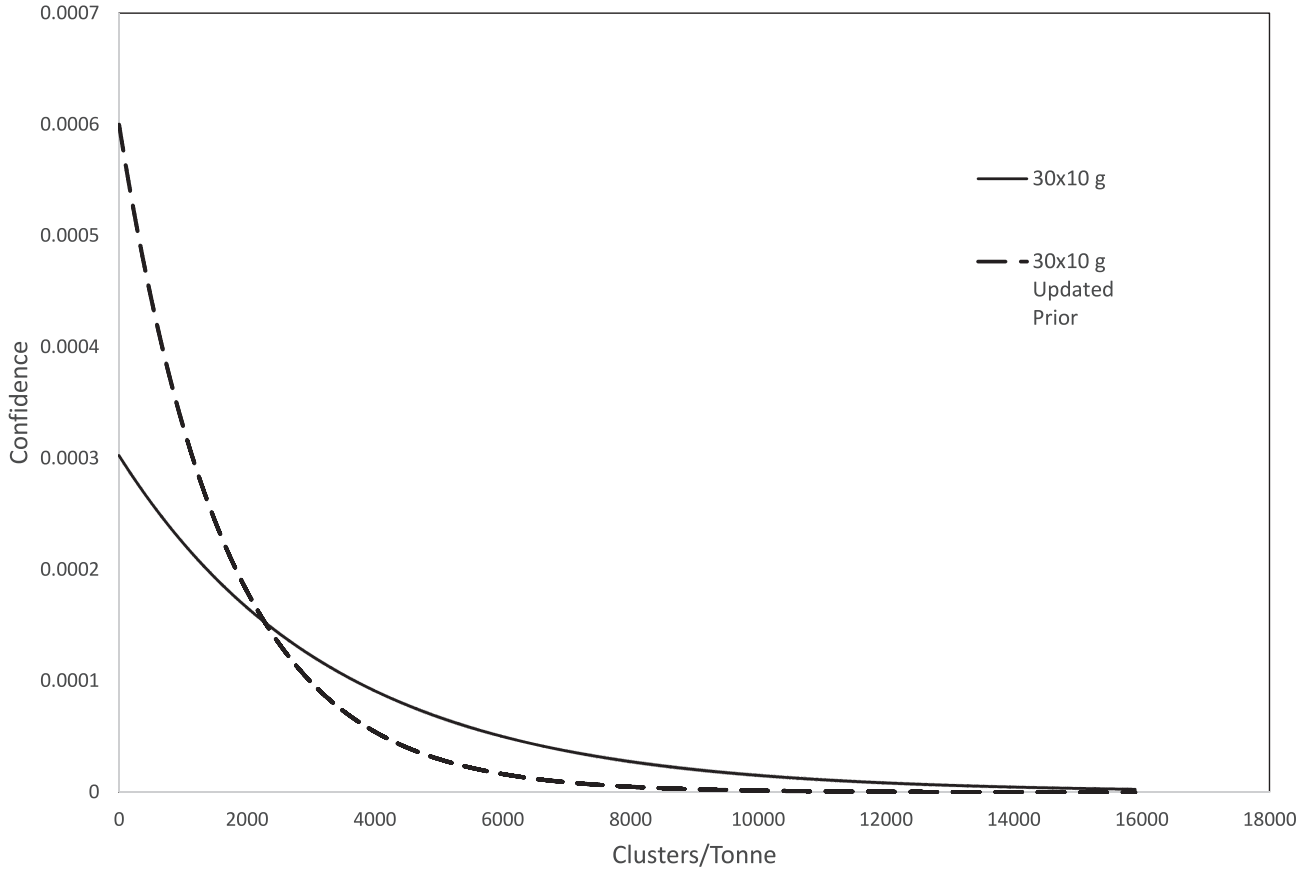

Fig. 3. Uncertainty using the Bayesian spreadsheet approach in the number of *Cronobacter* spp. clusters present in a tonne of powder associated with a non-informed uniform prior and a second iteration where the normalised posterior from the first iteration formed the prior for the second iteration (Thirty 10 g samples).

spp. clusters values set out in column B. Subtracting column E from one, gives the probability of the number of clusters in the test sample being greater than one (Column F), i.e. a positive test is recorded.

The model uses a binomial distribution to characterise the probability of s successes (0, 1, 2, etc.) occurring from n samples depending on the probability of success parameter, p. Column F supplies the probability of success parameter, p, for the binomial distribution (Column G) which is the Bayesian likelihood function for the probability of s successes. A Bayesian posterior distribution (Column H) is created by multiplying the likelihood function (Column G) by the uninformed prior (Column D). The posterior is then normalised (Column I) by dividing each value in Column H by the sum of all posterior values (cell H14). The highest number of clusters per tonne considered has to be adjusted to ensure that cell H14 approaches a stable finite value. The normalised posteriors represent a confidence value for each discrete value of clusters per tonne considered.

4. Method 2: algebraic approach

The algebraic approach followed the same general approach as the spreadsheet model. The details of the algebraic solution are given in the Appendix. In summary, the model considers all possible values for the number of *Cronobacter* spp. clusters that are potentially present in a tonne of powder. The values are discrete non-negative integers (0, 1, 2, etc.). As before, the model assumes a non-informed discrete uniform prior of 1 for all values of x. Assuming that the clusters are randomly distributed throughout the powder, the number of clusters in a test sample would follow a Poisson distribution, with λ representing the mean number of clusters per sample (Eq. 5). For any x, assuming a Poisson process, the probability of zero clusters in the sample is calculated (Eq. 6). Assuming one or more clusters in a sample will trigger a detection, then the probability of detection is (1- probability of zero clusters) (Eq. 7). As before, the model uses a binomial distribution to characterise the probability of s successes occurring from n samples depending on the probability of success. Accordingly, the Bayesian likelihood function is given by the probability mass function of the Binomial distribution (Eq. 8). The model is developed for zero defects detected (s= 0) and for one defect detected (s = 1). Assuming a non-informed discrete uniform prior of 1 for all values of x, the posterior distribution has the same value as the likelihood function which is then normalised. For zero positives, the normalised posterior is an exponential decreasing function (Eq. 17). If the total amount of powder tested (w*n) is small in comparison to the unit of powder being considered (one tonne in this example), then through a Taylor series approximation, the normalised posterior can be approximated by a Gamma distribution with parameters (for zero defects):

$$\alpha = 1 \tag{1}$$

$$\beta = \frac{\mathrm{m}}{wn} \tag{2}$$

For one positive detected, the algebraic solution follows a similar approach and again it is shown that the normalised posterior can be approximated by a Gamma distribution with parameters

$$\alpha = 2 \tag{3}$$

$$\beta = \frac{\mathrm{m}}{wn} \tag{4}$$

Both the spreadsheet and algebraic models were run to estimate the uncertainty associated with sampling for *Cronobacter* spp. in PIF for the following scenarios: Zero and one positive results from thirty 10 g samples (the EC microbiological criteria for *Cronobacter* spp. in PIF); three 100 g samples (pooling the thirty 10 g samples into three samples) and ninety 10 g samples (three times the EC microbiological criteria).

5. Results

Fig. 1 shows the uncertainty calculated using the Bayesian spreadsheet approach in the number of clusters present in a tonne of powder associated with a zero positive result arising from a sampling plan of either thirty 10 g samples, three 100 g samples or ninety 10 g samples. While the most likely outcome (the mode) for all sampling plans is zero clusters present, the figure indicates that the true number of clusters present could be as high as several thousand clusters per tonne for all three sampling strategies. The uncertainty associated with the thirty 10 g samples and the three 100 g samples is identical. This is to be expected as the algebraic solution for the normalised posterior is an exponential decreasing function (Eq. 13) which only contain constant parameters in the form of (w n / m). So there is direct equivalence between sampling schemes when the total amount of powder sampled (w n) is the same where the outcome is a zero positive result. Increasing the sampling scheme to ninety 10 g samples (i.e. three times the EC Micro Criteria requirements for a batch) considerably reduces the uncertainty, however even with this very stringent testing scheme, up to 3000 clusters (90th percentile: 2558 clusters/tonne) could be present in each tonne and yet the test result for all ninety samples is zero. Fig. 1 does not show the uncertainty calculated using the algebraic approach as the plots are effectively coincident with the plots generated by the spreadsheet approach.

Fig. 2 shows the uncertainty in the number of clusters present in a tonne of powder associated with a single positive result arising from the three sampling plans considered. In contrast to the distributions shown in Fig. 1, as a single positive result has been detected, then by definition, the probability of detecting zero clusters in the tonne of powder is zero. The uncertainty in the number of clusters present is much higher (90th percentile: 12,966 clusters/tonne for 30×10 g) when there is one positive result compared to when zero positives are detected (90th percentile: 7675 clusters/tonne for 30×10 g). As before, increasing the sampling scheme to ninety 10 g samples considerably reduces the uncertainty compared to the other two sampling schemes considered. Figure two demonstrates that there is no longer a direct equivalence in the uncertainty associated with the thirty 10 g sample plan and the three 100 g sample plan. This can be explained as there is now a $w^2n(n-1)$ term in the normalised posterior (Eq. 28). The gamma distribution approximation of the uncertainty agrees well with the spreadsheet estimation when the sample number is thirty or ninety. However, when the sample number is three, Fig. 2 shows that there is quite a variance between the gamma distribution approximation and the spreadsheet estimation. This arises from the assumption made, ie. $n(n-1) \approx n^2$ which introduces an error when n is small.

6. Discussion

The Bayesian approach developed in this paper demonstrates that even when zero positives are detected for a given sampling plan, there remains a considerable uncertainty in the true number of clusters that may be present in a consignment of powder. The present focus was based on the EC microbiological criterion for *Cronobacter* spp. in powdered infant formula (EC, 2005). This criterion is particularly stringent in having a value for n (the number of units comprising the sample) of 30. For the majority of the other pathogen / product combinations, n is typically 5. In this case the resulting uncertainty would be considerably greater and demonstrates that great care is needed in developing sampling

plans in situations where the concentration of the contaminant is generally low. In general to reduce the uncertainty associated with a zero positive outcome, as Eq. 13 demonstrates, the option is to increase the total sample weight taken, either by increasing the sample weight, w, or by increasing the number of samples taken, n.

There are a number of published estimates for the concentration of *Cronobacter* spp. in powdered infant formula which puts the magnitude of uncertainty estimated in the present study into context. The FAO/WHO (2006) collated data from 29 studies where there was a positive occurrence of *Cronobacter* spp. in PIF. As all the studies reported occurrence rather than concentration, the FAO/WHO study converted occurrence data to likely concentration values based on the assumption that sampling is a Poisson process. The estimated concentrations ranged from -5.24 to -2.79 log(cfu/g). This equates to concentrations between 6 and 1620 cfu per tonne. While the FAO/WHO study reported cfu, the authors of this paper are of the view that the correct unit is clusters rather than cfu given the Poisson assumption used to convert occurrence data to likely concentration values. These reported concentrations are well within the uncertainty in the number of clusters present in a tonne of powder as presented in Figs. 1 and 2. For this study we make the assumption that it is a cluster of *Cronobacter* spp. that triggers the detection of an absence/presence test and that the distribution of cells in the clusters is of no importance provided the test is sufficiently sensitive to detect low number of organisms. There is currently no experimental knowledge as to what the distribution of cells in these clusters might be. From a food safety perspective, the distribution of cells in a cluster is of major significance as it influences the number of organisms potentially ingested. However, further consideration of this is outside the scope of this paper.

The prior used in this work for the Bayesian estimation of the uncertainty associated with a test outcome was an uninformed prior. An uninformed prior assumes no previous knowledge and assigns equal likelihood for all values. In practice, food companies are routinely testing batches of product following some defined sampling plan. Rather than having just a single test result, a company will, over time, have the outcomes of numerous tests carried out on multiple batches. Bayesian inference allows iterations of its methodology. For example, to carry out a second iteration of the Bayesian estimate shown in Fig. 1, the discrete uniform prior can be replaced by the posterior distribution from the first outcome. Fig. 3 demonstrates the outcome from this Bayesian iteration for a sampling plan of thirty 10 g samples per batch in the case where two consecutive zero positives were recorded. In the first iteration, using a uniform prior, the outcome is the same as that portrayed in Fig. 1. This distribution then becomes the prior for the second posterior distribution. As Fig. 3 shows, the uncertainty in the true number of clusters present in a tonne of powder dramatically reduces. In theory, this iterative procedure could be continued if routine testing recorded ongoing negative results with further and further reduction in the uncertainty in the number of clusters present. The underlying assumption to allow this iteration is that samples are collected from an invariant population. However, in sampling theory, it is assumed that an individual batch is independent of any other batch (FAO, 2014) thus nullifying the idea of the invariant population. In practice, the concentration of *Cronobacter* spp. clusters present in the powder will vary from batch to batch again nullifying this assumption. Extreme caution is needed in iterating the Bayesian estimation of the uncertainty a large number of times, even when routine testing records ongoing negative results. Increased confidence can be achieved by using more informed priors. However, Bayesian inference ultimately requires a subjective judgement on the choice of the most appropriate prior.

7. Conclusions

A Bayesian estimation of the uncertainty associated with a test outcome is a robust approach to assessing the stringency of a sampling plan and augments the more classical approach of using operating characteristic curves that are commonly used to assess the efficiency of a sampling plan. The use of an uninformed prior is probably overly conservative in the estimation of the uncertainty. However care is needed in interpreting the outcome of multiple Bayesian iterations when routine sampling returns a prolonged sequence of zero positives. The Bayesian approach demonstrates that for a pathogen like *Cronobacter* spp. in PIF, even when a sampling scheme returns a zero positive outcome, there is a real likelihood that low concentrations of the pathogen could be present in the powder. When pathogen concentrations are low, most conventional sampling plans will fail to detect non conforming product. While the focus of this paper was on sampling schemes for *Cronobacter* spp. in PIF, the Bayesian approach developed in this paper can be adapted for other product / pathogen combinations. This Bayesian approach can be used with other measures of sampling plan efficiency such as an operating characteristic curve to assess the effectiveness of sampling plans being used to control pathogens in foods.

Declaration of interest statement

None

Acknowledgements

This work was funded under the Irish Department of Agriculture, Food and the Marine Food Institutional Research Measure (Research contract number13/F/423).

Appendix

Algebraic Solution

x	Number of clusters in the consignment unit, m (Discrete, 0,1,2,3,…)
w	Sample weight (g)
n	Number of samples taken from a batch, (Discrete, 0,1,2,3,…)
s	Number of positive samples detected, (Discrete, 0,1,2,3,…)
m	The mass of the consignment being considered in grams (In this work, this was taken to be one tonne, i.e. 10^6 g)

Assume a non-informed discrete uniform prior of 1 for all values of x.

Assume that the clusters are randomly distributed throughout the powder. The number of clusters in a sample would be a Poisson process, where λ represents the mean number of clusters per sample.

$$\lambda(x) = \frac{wx}{m} \tag{5}$$

For any x, assuming a Poisson process, the probability of zero clusters in the sample is:

$$P(0) = e^{\left(\frac{-wx}{m}\right)} \tag{6}$$

Assuming one or more clusters in a sample will trigger a detection, then the probability of detection:

$$P(>0) = 1 - e^{\frac{-wx}{m}} \tag{7}$$

The likelihood function is given by the probability mass function of the Binomial distribution:

$$l(s|x) = \binom{n}{s}\left(1 - e^{\frac{-wx}{m}}\right)^{s}\left(e^{\frac{-wx}{m}}\right)^{n-s} \tag{8}$$

Zero defects detected (s = 0)

$$l((s=0)|x) = \binom{n}{0}\left(1-e^{\frac{-wx}{m}}\right)^{0}\left(e^{\frac{-wx}{m}}\right)^{n} \quad (9)$$

$$l((s=0)|x) = e^{\frac{-wnx}{m}} \quad (10)$$

Assuming a non-informed discrete uniform prior of 1 for all values of x, the posterior distribution, f(x | (s=0)) has the same form, i.e.

$$f(x|(s=0)) = e^{\frac{-wnx}{m}} \quad (11)$$

To normalise the posterior, the above equation has to be summed for all values of x:

$$\sum_{x=0}^{\infty} e^{\frac{-wnx}{m}} = \frac{1}{e^{\frac{wn}{m}}-1} \quad (12)$$

(C/F Wolfram (2016))

The normalised posterior, f' (x | (s=0)) is given by:

$$f'(x|(s=0)) = \left(e^{\frac{wn}{m}}-1\right)e^{\frac{-wnx}{m}} \quad (13)$$

Which is an exponential decreasing function.

If (w n) is small in comparison to m, then the Taylor series expansion for the exponential function e^{ax} can be approximated by the first two Taylor series terms if a is very small. i.e.

$$e^{ax} \approx 1 + \frac{ax}{1!} \quad (14)$$

Or alternatively:

$$e^{ax} - 1 \approx ax \quad (15)$$

Accordingly:

$$\left(e^{\frac{wn}{m}}-1\right) \approx \frac{wn}{m} \quad (16)$$

Then the normalised posterior can be approximated by:

$$f'(x|(s=0)) \approx \left(\frac{wn}{m}\right)e^{\frac{-wnx}{m}} \quad (17)$$

Which is a Gamma distribution with parameters $\alpha = 1$, $\beta = \frac{m}{wn}$.

One defect detected (s = 1)

$$l((s=1)|x) = \binom{n}{1}\left(1-e^{\frac{-wx}{m}}\right)^{1}\left(e^{\frac{-wx}{m}}\right)^{n-1} \quad (18)$$

$$l((s=1)|x) = n\left(1-e^{\frac{-wx}{m}}\right)e^{\frac{-wx(n-1)}{m}} \quad (19)$$

$$l((s=1)|x) = n\left(e^{\frac{-wx(n-1)}{m}} - e^{\frac{-wxn}{m}}\right) \quad (20)$$

Rearranging

$$l((s=1)|x) = ne^{-wxn)/m}(e^{(wx/m)}-1) \quad (21)$$

Assuming a non-informed discrete uniform prior of 1 for all values of x, the posterior distribution, f(x | (s=1)) has the same form, i.e.

$$f(x|(s=1)) = ne^{\frac{-wxn}{m}}\left(e^{\frac{wx}{m}}-1\right) \quad (22)$$

To normalise the posterior, the above equation has to be summed for all values of x:

$$\sum_{x=0}^{\infty} n\left(e^{\frac{-wx(n-1)}{m}} - e^{\frac{-wxn}{m}}\right) = n\sum_{x=0}^{\infty}\left(e^{\frac{-wx(n-1)}{m}}\right) - n\sum_{x=0}^{\infty}\left(e^{\frac{-wxn}{m}}\right) \quad (23)$$

$$\sum_{x=0}^{\infty} n\left(e^{\frac{-wx(n-1)}{m}} - e^{\frac{-wxn}{m}}\right) = \frac{n}{e^{\frac{w(n-1)}{m}}-1} - \frac{n}{e^{\frac{wn}{m}}-1} \quad (24)$$

The normalised posterior, f' (x | (s=1)) is given by:

$$f'(x|(s=1)) = \left(\frac{1}{\frac{n}{e^{\frac{w(n-1)}{m}}-1} - \frac{n}{e^{\frac{wn}{m}}-1}}\right)ne^{\frac{-wxn}{m}}\left(e^{\frac{wx}{m}}-1\right) \quad (25)$$

As before,

$$e^{ax} - 1 \approx ax \quad (26)$$

The normalised posterior can be approximated by:

$$f'(x|(s=1)) \approx \left(\frac{1}{\frac{n}{\frac{w(n-1)}{m}} - \frac{n}{\frac{wn}{m}}}\right)ne^{\frac{-wxn}{m}}\left(\frac{wx}{m}\right) \quad (27)$$

Rearranging

$$f'(x|(s=1)) \approx \left(\frac{w^2n(n-1)x}{m^2}\right)e^{\frac{-wxn}{m}} \quad (28)$$

When $n(n-1) \approx n^2$ (i.e. n > 10 approximately), then the normalised posterior can be approximated by a Gamma distribution with parameters $\alpha = 2$ and $\beta = \frac{m}{wn}$

References

Codex Alimentarius Commission, 1997. CAC/GL 21-1997, Principles for the establishment and application of microbiological criteria for foods. www.fao.org/input/download/standards/394/CXG_021e.pdf.

EC, 2005. Commission regulation (EC) No 2073/2005 on microbiological criteria for foodstuffs. http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32005R2073.

FAO/WHO, 2006. *Enterobacter sakazakii* and Salmonella in Powdered Infant Formula. Microbiological Risk Assessment Series no. 10, Rome, p.95. Meeting Report (Table 20).

FAO/WHO, 2014. Risk Manager's Guide to the Statistical Aspects of Microbiological Criteria Related to Foods. Available at http://www.fao.org/food/food-safety-quality/home-page/en/.

Gonzales-Barron, U., Zwietering, M.H., Butler, F., 2013. A novel derivation of a within-batch testing regime based on a Poisson-gamma model characterising low microbial counts in foods. Int. J. Food Microbiol. 161 (2), 84–96.

Healy, B., Cooney, S., O'Brien, S., Iversen, C., Whyte, P., Nally, J., Callanan, J.J., Fanning, S., 2009. *Cronobacter* (*Enterobacter sakazakii*): an opportunistic foodborne pathogen. Foodborne Pathogens and Disease 7 (4), 339–350.

ISO, 2006. Milk and Milk Products -Detection of *Enterobacter sakazakii* ISO/TS 22964:2006 (IDF/RM 210: 2006). International Organization for Standardization, Geneva.

Jongenburger, I., 2012. Distributions of Microorganisms in Foods and their Impact on Food Safety PhD thesis (Proefschrift Wageningen), available: http://edepot.wur.nl/196895 (accessed 10 May 2016).

Jongenburger, I., Reij, M.W., Boer, E.P.J., Gorris, L.G.M., Zwietering, M.H., 2011. Actual distribution of *Cronobacter* spp. in industrial batches of powdered infant formula and consequences for performance of sampling strategies. Int. J. Food Microbiol. 151 (1), 62–69.

Mussida, A., 2013. Performance of microbiological criteria for the detection of Cronobacter spp. in powder infant formula as affected by theoretical mixture distributions. Thesis (PhD) - University College Dublin.

Mussida, A., Gonzales-Barron, U., Butler, F., 2013. Effectiveness of sampling plans by attributes based on mixture distributions characterising microbial clustering in food. Food Control 34 (1), 50–60.

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.

Vose, D., 2008. 'Risk Analysis: A Quantitative Guide', 3rd ed Hoboken, NJ: Wiley, Chichester, England.

Wolfram 2016. Wolfram Research, Available http://functions.wolfram.com/ElementaryFunctions/Exp/23/01/ (Accessed 25 April 2016).

Yan, Q.Q., Condell, O., Power, K., Butler, F., Tall, B.D., Fanning, S., 2012. *Cronobacter* species (formerly known as *Enterobacter sakazakii*) in powdered infant formula: a review of our current understanding of the biology of this bacterium. J. Appl. Microbiol. 113 (1), 1–15.