Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Food Science 1 (2011) 979 – 986

Procedia Food Science

11th International Congress on Engineering and Food (ICEF11)

Operating characteristic curves for single, double and multiple fraction defective sampling plans developed for *Cronobacter* in powder infant formula

Arianna Mussida *, Ursula Gonzales-Barron, Francis Butler

Biosystems Engineering, UCD School of Agriculture, Food Science and Veterinary Medicine, University College of Dublin, Belfield, Dublin 4, Ireland

Abstract

The microbiological criteria established in the EC 2073/2005 for *Cronobacter* in powder infant formula (PIF) are based on two-class attribute sampling plans, where the sample results are qualitative (sample indicates presence or absence) and the lot is rejected if any sample is positive. The performance of a sampling plan is revealed by its Operating Characteristic (OC) curve which plots the probability of acceptance against possible values of proportion defective. The objective of this study is to generate several OC curves for single, double and multiple sampling plans assuming different statistical distributions of *Cronobacter* in PIF in order to determine and compare the probabilities of rejecting/accepting the lot and the respective level of contamination. In this study, the microbial distribution of *Cronobacter* in the PIF is described by assuming to be Poisson-lognormal (PLN), Poisson-gamma (PG), Zero-inflated Poisson-lognormal (ZIPLN) and Zero-inflated Poisson-gamma (ZIPG). For each distribution the proportion defective of the population is estimated in order to determine the probability of acceptance for single, double and multiple sampling plans. Furthermore, the effect of clustering of the bacteria on the probability of acceptance of the lot was assessed through a Poisson-logarithmic (PLOG) and a PLogn distribution. Probabilities of accepting the lot at a given level of contamination change drastically according to the statistical distribution assumed and by changing the values of its own parameters. In addition, a procedure of multiple sampling plan based on resampling reveals that the producer's risk (α risk) can be reduced without affecting the consumer's risk (β risk).

Selection and/or peer-review under responsibility of 11th International Congress on Engineering and Food (ICEF 11) Executive Committee.

Keywords: OC curves; Cronobacter; Poisson; Zero-inflated; Clusters

* Corresponding author. Tel.: +353-1-7161815;
E-mail address: arianna.mussida@ucd.ie.

Selection and/or peer-review under responsibility of 11th International Congress on Engineering and Food (ICEF 11) Executive Committee.

doi:10.1016/j.profoo.2011.09.147

1. Introduction

Cronobacter represents a significant risk to the health of neonates. This bacterium is an emerging opportunistic pathogen that is associated with rare but life-threatening cases of meningitis, necrotizing enterocolitis, and sepsis in premature and full-term infants [1-5]. Although, the organism natural habitat is currently unknown, molecular typing methods have identified PIF as a source and vehicle of neonatal infection [2, 4, 6]. The recovery of *Cronobacter* from numerous locations in powdered milk production facilities [7, 8] suggests that contamination of the final product occurs via the environment of the processing facility. Monitoring the contamination in the manufacturing environment, along the production chain and in the final product using an appropriate sampling plan represents an important first step in reducing the risk of contaminating PIF product. In safety and compliance testing, an acceptance number of zero is particularly desirable, since, to the uninformed, it would appear that the use of any greater acceptance number implies passing lots which have been shown to have defectives in them [9]. The main objective of this study is to evaluate the performance of the sampling plan established in the EC 2073/2005 [10] (absence in 10 g, 30 samples per unit) for *Cronobacter* in PIF. As the results of the monitoring system are qualitative (presence/absence), it is hard to define a statistical distribution and the level of contamination of this bacterium in the product. In this study, several statistical distributions (PLN, ZIPLN, PG, ZIPG, PLOG and PLogn) are assumed to describe *Cronobacter* in the PIF and, based on each of them, OC curves are generated in order to determine the probabilities of accepting/rejecting a lot at a given level of contamination. Furthermore, as zero-acceptance plans are unfavourable to the producer [9, 11, 12], an example of a multistage resampling scheme is presented and the advantages for the producer are shown.

Nomenclature

Pa	Proportion acceptable of the sample
Pd	Proportion defective of the sample
α	Shape parameter of the Gamma distribution
π_o	Probability that a zero count arises from the fixed-zero group in a zero-inflated distribution
w	Sample weight
n	Number of samples

2. Materials and Methods

The proportion acceptable (Pa) in each sample, so the probability of finding a zero count at a given level of contamination, for each distribution was calculated as follows:

2.1 Poisson-lognormal (PLN)

We assume that the mean level of contamination follows a lognormal distribution in the powder, while the distribution of the sample follows a Poisson process, with the intensity given by the random lognormal concentration where the sample is taken.

$$Pa = \exp(-\lambda * w) \quad (1)$$

Where λ = *Lognormal* (μ, σ)

A further assumption is given, such that the distribution of the bacteria follows a lognormal distribution between and within lot with a given between-lot standard deviation (σ_b) across lots and within-lot standard deviation (σ_w) for individual PIF lots. Simulations are run with @Risk software.

2.2 Zero-inflated Poisson-lognormal (ZIPLN)

We assume that the level of contamination follows two different groups, one that will always be zero (we sample from zero-counts) and one that follows a PLN.

$$Pa = \pi_o + (1-\pi_o) * \exp\left(\left(\frac{-\lambda}{(1-\pi_o)}\right) * w\right) \quad (2)$$

Where λ = *Lognormal* (μ, σ)

2.3 Poisson-Gamma (PG)

We assume that the mean level of contamination follows a gamma distribution in the powder, while the distribution of the sample follows a Poisson process.

$$Poisson\left(\lambda * \Gamma\left(\alpha, \frac{1}{\alpha}\right)\right) \quad (3)$$

Where Γ (α, 1/α) has mean= 1 and variance α, so the Gamma distribution is adding variation about the expected number of bacteria (λ).

$$Pa = \left(\frac{\alpha}{\alpha + \lambda * w}\right)^{\alpha} \quad (4)$$

2.4 Zero-inflated Poisson-Gamma (ZIPG)

We assume that the level of contamination follows two different groups, one that will always be zero and one that follows a PG.

$$Pa = \pi_o + (1-\pi_o) * \left(\frac{\alpha}{\alpha + \mu * w}\right)^{\alpha} \quad (5)$$

μ = is the mean of the non- zero mixed PG distribution

2.5 Modelling clusters

We assume that *Cronobacter* is present in PIF in colonies or group of individuals randomly distributed in a Poisson process. Notice that clustering is modelled within the contagious process framework that produces "aggregate distributions" which is conceptually different from the heterogeneous Poisson framework used above [13]. If the numbers of individuals within the colonies are distributed independently in a logarithmic distribution we obtain a PLOG or negative binomial distribution [14-16].

PLOG= *Poisson* (λ) ^ *Logarithmic* (θ)

θ = the shape parameter of the logarithmic distribution
λ = number of clusters or colonies in the powder

The aggregation of these two distributions is done mathematically and the parameters calculated. If the numbers of individuals within the colonies are distributed following a lognormal distribution we obtain a PLogn. Due to mathematical complexity, in order to aggregate these two distributions simulations are needed and are run with ModelRisk software (version 3.0, Belgium).

PLogn = *Poisson* (λ) ^ *Lognormal* (μ, σ)

μ = mean of individuals within each cluster
σ = variation of bacteria within different clusters

For each distribution considered the proportion defective (Pd) of each sample was calculated:
Pd = 1- Pa

Finally the probability of acceptance (P accept) of the lot was obtained with the excel function BINOMDIST:
P accept = BINOMDIST (0, n, Pd, 1)

2.6 Resampling scheme

The sampling plan in Table 1 is proposed and all the probability calculations are performed and analyzed for the PLN, ZIPLN, PG, ZIPG distributions. Several values for the number of samples (n) and sample weight (w) at each stage are considered for the calculation of the P accept.

Table 1. Resampling scheme

Stage	n	w	Accept	Reject
1	n_1	w_1	0	*
2	n_2	w_2	0	*
3	n_3	w_3	0	*
i	n_i	w_i	0	For any +

*resampling allowed for any positive sample found at each stage

3. Results and Methods

A first set of results (Table 2) is obtained by comparing the level of contamination of the lot for the producer's (0.99 and 0.90 P accept) and consumer's risk (0.5 and 0.1 P accept) assuming the same standard deviation for PLN, ZIPLN, PG and ZIPG distributions.

Table 2. Mean level of contamination (CFU/g) of the lot and conditional probability of acceptance assuming a PLN, ZIPLN, PG, ZIPG and fixed standard deviation (σ =10 and 100 CFU/g).

P accept	PLN	ZIPLN (π = 0.1)	PG	ZIPG (π = 0.1)
99 %	0.00015-0.00028	0.0003-0.00075	0.00678-0.06	0.0084-0.0665
90%	0.002-0.004	0.0055-0.013	0.0241-0.203	0.0318-0.2427
5%	0.083-0.202	1.5-6.5	0.153-1.218	0.2967-2.1367
1%	0.135-0.34	*	0.195-1.53	**

*the β risk of 0.1 is never reached by the OC Curve.
** Shape parameter (α) from the variance could not be derived.

The PLN is the distribution that gives the best confidence of rejecting the lot at a given level of contamination. It is important to notice that assuming the same standard deviation at any given level of contamination the shape parameter (α) of the PG distribution increases by increasing the mean level of contamination. It appears more natural that the standard deviation should not be fixed for any mean level of contamination, but instead should be correlated with the mean. The results shown in Table 3 are obtained for a PG distribution with a fixed shape parameter. As we can see the standard deviation increases accordingly to the level of contamination of the lot.

Table 3. Mean level of contamination and standard deviation (CFU/g) of the lot and conditional probability of acceptance assuming a PG distribution with fixed values of α (0.01, 0.05, and 0.025).

P accept	α= 0.01	α= 0.05	α= 0.025
99%	0.00003/0.018	0.00003/0.018	0.00003/0.018
90%	0.00035/0.06	0.00036/0.06	0.00037/0.06
5%	0.017/0.68	0.032/1.52	0.13/8.35
1%	0.036/1.3	0.097/4.45	1.08/68.86

It is evident from the results in Table 3 that, decreasing the value of the shape parameter, the consumer's risk of a lot being accepted with a higher level of contamination increases.

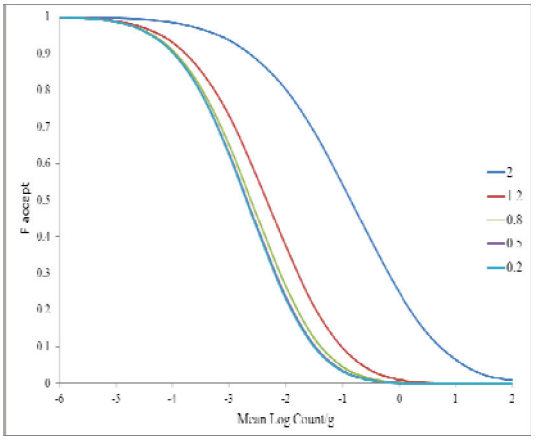

Fig. 1. OC curves based on a PLN distribution with σ_b=0.8 and σ_w= 2, 1.2, 0.8, 0.5, 0.2

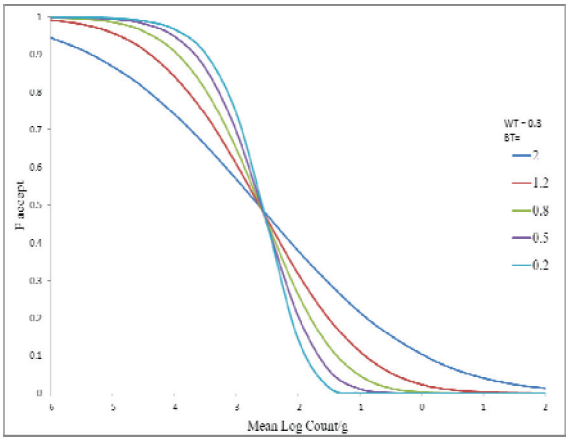

Fig. 2. OC curves based on a PLN distribution with σ_w= 0.8 and σ_b=2, 1.2, 0.8, 0.5, 0.2

The OC curves presented graphically (Figures 1 & 2) are obtained based on a PLN distribution and simulating for several values of within and between-lot variability.

As we can see from Figure 1 and 2, obtained by simulation, the within and between-lot variability employed for the PLN distribution play a crucial role in determining the confidence levels of accepting/rejecting a lot at a given level of contamination, and hence determines the real effectiveness of an acceptance sampling plan.

A further study is conducted assuming that *Cronobacter* contaminates PIF in random clusters, each of them with different sizes. If we assume that the distribution within each cluster follows a logarithmic fashion we obtain a PLOG, also referred to as negative binomial. Several values of the shape parameter θ (from 0.001 to 0.9999) are employed in order to describe different sizes of clusters while the number of clusters (λ) is assumed constant (Table 4). The results obtained, show that the size of the clusters does not have any effect on the proportion acceptable of the sample and on the probability of acceptance of the lot, which is affected only by the distribution of the clusters in the powder.

Table 4. Mean level of contamination of the powder with 5 and 95% probability of acceptance assuming an aggregated Poisson-logarithmic distribution.

Logarithmic distribution			PLOG distribution			
			λ= 0.00017 (P accept 95%)		λ= 0.01 (P accept 5%)	
θ	μ (number of bacteria/cluster)	σ	μ (number of bacteria/10 g)	σ	μ (number of bacteria/10 g)	σ
0.001	1.0005	0.02	0.00168	0.0411	0.1	0.32
0.1	1.05	0.24	0.00178	0.0444	0.1	0.34
0.9	3.9	4.88	0.0066	0.26	0.4	1.98
0.99	21.5	41.08	0.36	1.9	2.15	14.66
0.9999	1085	3110	1.83	135	109	1042

From Table 4 we can see that if the mean level of contamination of clusters is 0.00017/g (1 cluster in 6000 g), the number of bacteria within each cluster will not influence the P accept (95%). It means that in

the worst case scenario (θ equal to 0.9999) we are 95% confident of accepting a lot with a mean level of contamination of 1.83 cells/g and standard deviation 135 cells. The number of bacteria/g increases if we increase the number of cluster/g (λ= 0.01, 1 cluster in 100 g).

A similar study is proposed by assuming this time a lognormal distribution of microorganisms within each cluster (PLogn). As in the case of the PLOG distribution the size of the cluster does not influence the Pa of the sample and the probability of acceptance of the lot (Figure 3). The calculations are performed for any values of μ and σ in order to describe the size of the clusters and the variation between them.

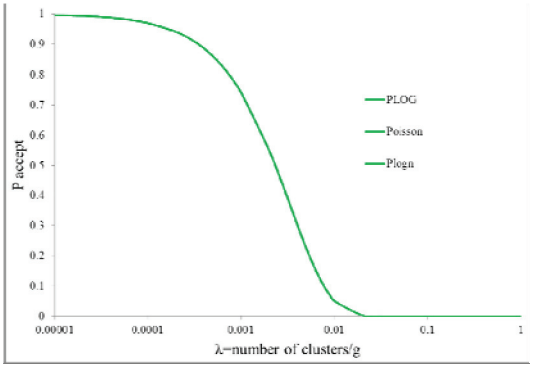

Fig. 3. Identical OC curves developed assuming a PLogn, PLOG and Poisson distribution of *Cronobacter* in PIF

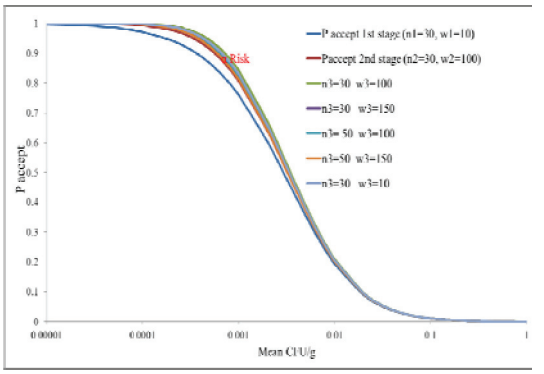

Fig. 4. OC curves of double and multiple resampling plans with 3 stages based on a PG distribution with shape parameter (α) 0.05

It is common practice in the manufacturing environment the pooling of the samples. Although this practice makes testing of the product much quicker and cheaper, we lose information concerning the number of samples that are contaminated. In order to overcome this lack of knowledge in the multistage sampling plan proposed in Table 1, a rejection number cannot be determined but a resampling procedure is applied for any positive sample that might contaminate the pooled sample. The Pa is calculated based on each distribution considered and on the weight of the sample employed at each stage. The probability of acceptance of the lot at each stage is given by the number of samples employed at each stage, the probabilities of acceptance of the previous stage and the probabilities of resampling from the previous stage. Finally the probability of acceptance of the lot at the final stage is calculated. As an example, in this paper is graphically presented a resampling scheme with 3 stages assuming a PG distribution and applying five different combinations of weight and number of samples used at the last stage (Figure 4). The reader can visually realize how the producer's risk, so the probability of rejecting a lot at a low mean level of contamination, decreases without altering the consumer's risk. The decrease of the producer's risk by using such a scheme can be observed for each distribution considered in this study.

4. Conclusion

The variety of scenarios developed in this study in order to describe the statistical distribution of *Cronobacter* in PIF, reveal that the sampling plan implemented by EC 2073/2005 ensures a consumer's quality level (CQL) at a high mean level of contamination. In fact, the best case scenario is described by the PLN (Table 2) where we are 95% confident of rejecting a lot with mean level of contamination 0.083 or 0.202 CFU/g (assuming standard deviation 10 and 100 CFU respectively), while in other scenarios described by ZIPLN, PG, ZIPG distributions, the probabilities of rejecting the lot at the same mean level of contamination decrease. The results also show the importance of defining the values of the parameters involved in each distribution considered due to their impact on the probabilities of acceptance/rejection.

Further studies are needed in order to determine if the bacterium contaminates the powder in clusters [17] and, if this is the case, it is important- in terms of food safety to determine the distribution and number of bacteria within each cluster.

Furthermore, it is well known [9, 11, 12] that the resubmission of lots in the event of non-acceptance is a provision allowed to protect the interest of the producer, especially in the case of zero acceptance numbers since a single positive sample could lead to the rejection of the entire lot. The OC curves developed on the basis of the multiple resampling plan proposed in this paper, show that the producer's risk of rejecting a lot at a good quality level can be reduced without affecting the CQL.

References

[1] Arseni A., Malamou-Ladas E., Koutsia C., Xanthou M. & Trikka E. 1987. Outbreak of colonization of neonates with Enterobacter sakazakii. J. Hosp. Infect. 9, 143–150.

[2] Biering G., Karlsson S., Clark N.C., Jonsdottir K.E., Ludvigsson P. & Steingrimsson O. 1989. Three cases of neonatal meningitis caused by Enterobacter sakazakii in powdered milk. J. Clin. Microbiol. 27, 2054–2056.

[3] Nazarowec-White M. & Farber J.M. 1999. Phenotypic and genotypic typing of food and clinical isolates of Enterobacter sakazakii. J. Med. Microbiol. 48, 559–567.

[4] van Acker J., de Smet F., Muyldermans G., Bougatef A., Naessens A. & Lauwers S. 2001. Outbreak of necrotizing enterocolitis associated with Enterobacter sakazakii in powdered milk formula. J. Clin. Microbiol. 39, 293–297.

[5] Drudy D., Mullane N.R., Quinn T.,Wall P.G. & Fanning S. 2006a. Enterobacter sakazakii: an emerging pathogen in powdered infant formula. Clin. Infect. Dis. 42, 996–1002.

[6] Clark N.C., Hill B.C., O'Hara C.M., Steingrimsson O. & Cooksey, R.C. 1990. Epidemiologic typing of Enterobacter sakazakii in two neonatal nosocomial outbreaks. Diagn. Microbiol. Infect. Dis. 13, 467–472.

[7] Kandhai M.C., Reij M.W., van Puyvelde K., Guillaume-Gentil O., Beumer R.R. & van Schothorst M. 2004a. A new protocol for the detection of Enterobacter sakazakii applied to environmental samples. J. Food Prot. 67, 1267–1270.

[8] Kandhai M.C., Reij M.W., Gorris L.G., Guillaume-Gentil O. & van Schothorst M. 2004b. Occurrence of Enterobacter sakazakii in food production environments and households. Lancet 363, 39–40.

[9] Schilling E.G. & Neubauer D.V. 2009. Acceptance sampling in quality control. Statistics: A Series of Textbooks and Monographs. Chapman & Hall/CRC publishing.

[10] Commission Regulation (EC) No 2073/2005 (OJ L338, p1, 22/12/2005) of 15 November 2005 on microbiological criteria for foodstuffs.

[11] Govindaraju K. & Ganesalingam S. 1998. Zero acceptance number quick switching system for compliance sampling. Journal of Applies Statistics, 25, 399-407.

[12] Govindaraju K. & Ganesalingam S. 1997. Sampling inspection for resubmitted lots. Communication in Statistics-Simulation and Computation, 26 (3), 1163-1176.

[13] Gonzales-Barron U. & Butler F., 2011. A comparison between the discrete Poisson-Gamma and Poisson-Lognormal distributions to characterise microbial counts in foods. Food Control, doi:10.1016/j.foodcont.2011.01.029.

[14] Anscombe F. J. 1950. Sampling theory of the negative binomial and logarithmic series distributions. Biometrika, 37 (3/4), 358-382.

[15] Bliss C. I., 1953. Fitting the negative binomial distribution to biological data. Biometrics, 9 (2), 176-200.

[16] Quenouille M. H., 1949. A relation between the logarithmic, Poisson, and negative binomial series. Biometrics, 5 (2), 162-164.

[17] Mussida A., Butler F. & Fanning S., 2010. Statistical aspects of C. Sakazakii in powder infant formula. IAFP Sixth European Symposium on Food Safety, UCD, Dublin, Ireland, 9-11 June, 2010

Presented at ICEF11 (May 22-26, 2011 – Athens, Greece) as paper MFS717.