International Journal of Food Microbiology 383 (2022) 109935

Contents lists available at ScienceDirect

International Journal of Food Microbiology

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

Critical comparison of statistical methods for quantifying variability and uncertainty of microbial responses from experimental data

Alberto Garre [a], Annemarie Pielaat [b], Marcel H. Zwietering [a], Heidy M.W. den Besten [a,*], Joost H. Smid [b]

[a] Food Microbiology, Wageningen University & Research, P.O. Box 17, 6700 AA Wageningen, the Netherlands
[b] Unilever R&D, 6708 WJ Wageningen, the Netherlands

ARTICLE INFO

Keywords:
Variance analysis
Robust statistics
QMRA
Monte Carlo simulations
In silico experiments

ABSTRACT

Variability and uncertainty are important factors for quantitative microbiological risk assessment (QMRA). In this context, variability refers to inherent sources of variation, whereas uncertainty refers to imprecise knowledge or lack of it. In this work we compare three statistical methods to estimate variability in the kinetic parameters of microbial populations: mixed-effect models, multilevel Bayesian models, and a simplified algebraic method previously suggested. We use two case studies that analyse the influence of three levels of variability: (1) between-strain variability (different strains of the same species), (2) within-strain variability (biologically independent reproductions of the same strain) and, at the most nested level, (3) experimental variability (species independent technical lab variability resulting in uncertainty about the population characteristic of interest) on the growth and inactivation of *Listeria monocytogenes*. We demonstrate that the algebraic method, although relatively easy to use, overestimates the contribution of between-strain and within-strain variability due to the propagation of experimental variability in the nested experimental design. The magnitude of the bias is proportional to the variance of the lower levels and inversely proportional to the number of repetitions. This bias was very relevant in the case study related to growth, whereas for the case study on inactivation the resulting insights in variability were practically independent of the method used. The mixed-effects model and the multilevel Bayesian models calculate unbiased estimates for all levels of variability in all the cases tested. Consequently, we recommend using the algebraic method for initial screenings due to its simplicity. However, to obtain parameter estimates for QMRA, the more complex methods should generally be used to obtain unbiased estimates.

1. Introduction

Most food products are rich media where pathogenic microorganisms can survive (and in some cases multiply), potentially resulting in a risk for the consumer. Currently, Quantitative Microbiological Risk Assessment (QMRA) is a common methodology used by governmental agencies and food industries to analyse the microbial risk of food consumption (FAO and WHO, 2021). Mathematical models that predict the variations in the microbial count (e.g. growth during storage or inactivation during the farm-to-fork chain) are an essential part of QMRA.

One of the main challenges of mathematical modelling for QMRA is that the models describe a system that is highly variable (Pouillot and Guillier, 2020). One of the most relevant sources of variation is induced by biological differences, e.g. between species, between microbial strains of the same species, between independent reproductions of the same strain (den Besten et al., 2018). Although variability is often focused on genetic differences, changes in the physiological state of cells can also affect their growth and/or inactivation kinetics (Arvaniti and Skandamis, 2022; Clemente-Carazo et al., 2021; Muñoz-Cuevas et al., 2013).

The different aspects of biological variability can be investigated in experiments and, subsequently, used in QMRA to estimate the impact of these different sources of variability (Pouillot and Guillier, 2020). This is an additional complication for model development because experimental errors, referring to any experimental limitation as defined by Box et al. (2005), are unavoidable under laboratory conditions where only a sample of the total population is investigated, and their effects confound with those of biological variability. For this reason, in the context of

* Corresponding author.
E-mail address: heidy.denbesten@wur.nl (H.M.W. den Besten).

https://doi.org/10.1016/j.ijfoodmicro.2022.109935
Received 6 April 2022; Received in revised form 9 September 2022; Accepted 14 September 2022
Available online 17 September 2022

QMRA, the variation in experimental data (e.g. differences in observed microbial counts between experiments) is separated in two sources: variability and uncertainty (Haas et al., 2014; Nauta, 2000; Schendel et al., 2018; Vose, 2008). Variability includes the variation in the results due to inherent factors of the system. For instance, it includes the impact of the biological differences between the microbial cells. Because variability is inherent to the system, its contribution to the variation of the response cannot be reduced: gathering additional data describes it better but does not reduce variability. That is a common rule of thumb to separate variability from uncertainty that, in this context, refers to the variation due to the use of imperfect or incomplete information. For example, a common source of uncertainty is the difference in microbial counts in the serial dilutions and plating, whose impact on variability in estimating a population characteristic can be reduced by increasing the quality or volume of the experimental data (Garre et al., 2019a; Jongenburger et al., 2010).

International guidelines for QMRA often recommend a separation between variability and uncertainty in the model output, in order to better understand the results and make appropriate recommendations for intervention strategies (EFSA Scientific Committee et al., 2018; Nauta, 2000; Thompson, 2002). However, this separation is hard to realise in practice due to several issues. The variation in any experiment will be a combination of different sources of variation with nested effects (Garre et al., 2021). Therefore, advanced data analysis methods must be applied to quantify and separate the sources of variation. Multilevel models (also called hierarchical or mixed-effects models) are probably the most common to describe and analyse this type of variation (Garre et al., 2020; Jaloustre et al., 2012; Pennone et al., 2021; van Boekel, 2021; Wijnands et al., 2017). These methods, albeit precise, are complex to develop, requiring the use of advanced statistical software and, potentially, long computation times. Furthermore, post-processing and interpretation can be challenging, especially if a Bayesian model was used for parameter estimation (van Boekel, 2020).

An approximate method for the quantification and separation of variability in QMRA has been suggested (Aryani et al., 2015a) (called the *Aryani method* from here onwards and described in depth in the next section). It is an algebraic method, so it is simpler to carry out and can be easily implemented in spreadsheets (e.g. MS Excel or Google Sheets). However, whereas the robustness of mixed-effects and Bayesian models have been thoroughly studied (Gelman and Hill, 2007; McElreath, 2016), that is not the case for the Aryani method. In this article, we compare these three different methods for quantifying and separating uncertainty and variability using both numerical simulations and statistical analysis. The comparison is supported by two case studies, one about the growth and one about the inactivation of *Listeria monocytogenes*.

2. Materials and methods

2.1. Statistical methods to quantify variability and uncertainty

2.1.1. The Aryani method

The Aryani method has been applied in several, recent research papers to evaluate the relevance of variability in microbial kinetics (e.g. Aryani et al., 2015a; Aryani et al., 2015b; Wells-Bennik et al., 2019). This method defines three sources of variability: "between-strain variability" (differences in the kinetic parameters between strains of the same species; called "strain variability" by Aryani et al. (2015a)), "within-strain variability" (differences in the kinetic parameters between biologically independent reproductions of the same strain; called "biological variability" by Aryani et al. (2015a)) and "experimental variability" (differences in the kinetic parameters observed between experiments using the same bacterial population). For this purpose, the Aryani method employs a nested experimental design. To estimate differences between strains, the design includes a relatively large number of strains of the same bacterial species (~20). For each one of them, several (3) bacterial cultures were prepared independently, so the variability between cultures of the same strain could be described. Then, a growth or inactivation experiment was performed several times in parallel (2) for each individual culture, enabling an estimation of the experimental variability. This resulted in a family of growth or inactivation curves.

Once the data has been generated, the contribution of each source of variability was estimated by a post-hoc statistical analysis. First, a primary model (growth or inactivation, depending on the case studied) was fitted to the experimental data. Then, the sample variance (s^2) of the kinetic parameters of interest (e.g. the specific growth rate for growth or the log D-value for inactivation) was used as an estimate of variability at each level (Eqs. (1)–(3)[1]).

$$s_B^2 = \frac{\sum_{B=1}^{n_B}(\overline{X}_B - \overline{X})^2}{df_B}, \tag{1}$$

$$s_W^2 = \frac{\sum_{B=1}^{n_B}\sum_{W=1}^{n_W}(\overline{X}_{WB} - \overline{X}_B)^2}{df_W}, \tag{2}$$

$$s_E^2 = \frac{\sum_{B=1}^{n_B}\sum_{W=1}^{n_W}\sum_{E=1}^{n_E}(X_{EWB} - \overline{X}_{WB})^2}{df_E}, \tag{3}$$

In Eqs. (1)–(3), X_{EWB} are the observed data (for example: measured maximal growth rates or decimal reduction times) for experiment E, within strain reproduction W and between strain B. Variables $\overline{X}_{WB}$ and $\overline{X}_B$ are computed from these data by taking averages across levels of experimental and within-strain variability, respectively. For instance, $\overline{X}_{WB}$ equals the mean of the n_e observations for within-strain replicate W of strain B ($\overline{X}_{WB} = \frac{1}{n_e}\sum_E X_{EWB}$), and $\overline{X}_B$ equals the mean of the n_W observations for strain B ($\overline{X}_B = \frac{1}{n_W \cdot n_e}\sum_E\sum_W X_{EWB}$). $\overline{X}$ is the grand mean across all data. The terms df_B, df_W and df_E stand for the number of degrees of freedom at the three different levels, calculated as the number of data points minus the number of model parameters.

The parameters s^2_B, s^2_W, s^2_E quantify variability at between-strain, within-strain and experimental levels, respectively. These are estimated from experimental data and, as such, affected by uncertainty. This point was not considered in the original work by Aryani, who only reported expected (mean) values for each variability factor. Therefore, in this work we expand the work by Aryani by proposing a method for estimating the uncertainty associated with them. According to Cochran's theorem, the ratio between the sample variance (s^2) times the number of degrees of freedom (df) and the population variance (σ^2), where n is the sample size, follows a Chi-square distribution with n-1 degrees of freedom (χ^2_{df}).

$$\frac{df \cdot s^2}{\sigma^2} \sim \chi^2_{df} \tag{4}$$

Based on this result, a confidence interval (CI) can be constructed for σ^2 as shown in Eq. (5), where $X^2_{df;\,\alpha/2}$ is the $\alpha/2$ percentile of the Chi-square distribution with df degrees of freedom. This CI (Eq. (5)) can be used to describe the uncertainty associated with the three estimates of variability considered in this method, s^2_B, s^2_W, s^2_E.

$$CI\ for\ \sigma^2 : \left(\frac{df \cdot s^2}{X^2_{df;1-\alpha/2}}\ \frac{df \cdot s^2}{X^2_{df;\alpha/2}}\right) \tag{5}$$

2.1.2. Mixed-effects models

Linear models assume that the residuals (the deviation between the empirical observations and the model fittings) are independent and

[1] Note that Aryani et al. (2015a, 2015b) use the term *MSE*, but the summation of the squared error divided by the number of degrees of freedom is more correctly known as the sample variance. Only the summation of the squared error divided by the number of data points is actually the *MSE*.

identically normally distributed random variables (Bates and Watts, 2007). Multilevel models extend linear models by enabling errors to remain constant within groups (Gelman and Hill, 2007; McElreath, 2016). Although the nomenclature is somewhat inconsistent in the scientific literature, multilevel models fitted using a frequentist approach are usually called mixed-effects models.

Mixed-effects models are very flexible because they can describe (linear) effects that may differ per group and can be used in a variety of scenarios. In the case of the experimental design of the Aryani method, a mixed-effects model separating the contribution of the three sources of variability can be written as shown in Eq. (6):

$$X_{EWB} = \overline{X} + \varepsilon_B + \varepsilon_W + \varepsilon_E \quad (6)$$

where the notation of subindexes is analogous to that in Eqs. (1)–(3) and variation with respect to the grand mean X is generated through the error terms ε. The errors are assumed to be random draws from normal distributions: $\varepsilon_B \sim N(0,\sigma_B)$; $\varepsilon_W \sim N(0,\sigma_W)$; $\varepsilon_E \sim N(0,\sigma_E)$ where $\sigma_{B;\ W;\ E}$ represent between-strain, within-strain and experimental variability, respectively.

The mixed-effects models were fitted to the data (the family of specific growth rates or log D-values) by restricted maximum likelihood using the *lme4* package (Bates et al., 2015) for the R programming language version 3.5.3 (R Core Team, 2016). Confidence intervals for the variance parameters were calculated under the same assumptions as for the Aryani method, using Eq. (5).

2.1.3. Bayesian modelling

Bayesian inference is a type of statistical inference in which Bayes' theorem is used to update the probability distribution of the model variables when more evidence or information becomes available (Gelman, 2014). In a Bayesian inference model, the joint probability distribution is formed across all parameters, processes and data. This method facilitates the inclusion of relatively complex likelihoods, such as those with hierarchical structures. The Bayesian method provides great flexibility, because it enables the definition of prior information about the parameters.

The proposed Bayesian model for our purpose has the same structure as the mixed-effects model but includes prior distributions for the σ parameters. We use uninformative priors, reflecting our prior lack of knowledge about levels of variability. There is some discussion regarding the definition of uninformative prior distributions for hierarchical models (Gelman, 2006). In this work, we compared two families of prior distributions. The first one is InverseGamma priors, as suggested by Vidakovic (2011):

$$\sigma_B^2 \sim InvGamma(\epsilon_B, \epsilon_B)$$

$$\sigma_W^2 \sim InvGamma(\epsilon_W, \epsilon_W)$$

$$\sigma_E^2 \sim InvGamma(\epsilon_E, \epsilon_E)$$

Here ϵ_B, ϵ_W and ϵ_E should be close to zero for non-informative priors. Calculations were repeated for three values of these parameters (0.01, 0.001 and 0.0001). The credible intervals (for this purpose of this article, the Bayesian equivalent to frequentist confidence intervals) of the posterior distributions of the model parameters were compared to ensure that the prior was actually non-informative, concluding that $\epsilon_B = \epsilon_W = \epsilon_E = 0.0001$ represents non-informativeness. Moreover, an exponential family of distributions was proposed:

$$\sigma_B^2 \sim Exponential(\lambda_B)$$

$$\sigma_W^2 \sim Exponential(\lambda_W)$$

$$\sigma_E^2 \sim Exponential(\lambda_E)$$

In this case, the parameters λ_B, λ_W, λ_E should be large to ensure that priors are non-informative. Calculations were repeated assigning to these coefficients several values (0.1, 0.01 and 0.001) and the impact on the credible intervals of the model parameters was assessed concluding that a value of $\lambda_B = \lambda_W = \lambda_E = 0.1$ represents non-informativeness. For the grand population mean ($\overline{X}$), a non-informative Normal (0, 1000) prior was used.

The prior distributions are updated using the data by an inference approach. Markov chain Monte Carlo (MCMC) methods, comprising a class of algorithms for sampling the joint probability distribution, were used for this task. We used the MCMC sampler JAGS (Just Another Gibbs Sampler) (Plummer, 2003), implemented in the *rjags* package for R (Plummer, 2019) for this task. The convergence and mixing of the Markov chain was evaluated using common guidelines (Brooks et al., 2011) requiring 1000 iterations for convergence. The spread of posterior distributions of the variance parameters represents remaining uncertainty about their true values, after fitting the model to the data.

2.2. Assessment of model performance using numerical simulations

The robustness of the Aryani method and the mixed-effects model for estimating the values of the variance parameters has been evaluated using numerical simulations (Garre et al., 2019b). We devised a computational scheme that simulates the microbial response that would be observed according to the hypotheses proposed by Aryani et al. (2015a). The simulation scheme is described in Box 1. We consider that between-strain variability is a perturbation (ε_B) with respect to a hypothetical population grand mean ($\overline{X}$) that follows a normal distribution with mean zero and variance $\sigma^2{}_B$. Then, within-strain variability is an additional perturbation (ε_W) also following a normal distribution mean zero, but with variance $\sigma^2{}_w$. Finally, experimental variability is an additional perturbation (ε_E) with a normal distribution with mean zero and variance $\sigma^2{}_e$. Therefore, the observation X_{EWB} corresponding to between-strain B and within-strain replicate W and experimental replicate E is $X_{EWB} = \overline{X} + \varepsilon_B + \varepsilon_W + \varepsilon_E$. Note that, simulating the nested experimental designs by Aryani et al. (2015a), we simulated several experiments for the same strain. This is represented in the algorithm by using the same value of ε_B for every simulation corresponding to a given strain. The same type of nesting was also applied for within-strain variability with parameter ε_W.

The robustness of the methods was evaluated based on the classical theory of biased estimators from frequentist probability theory (Voinov and Nikulin, 2012), which defines an unbiased estimator as the one whose difference between its expected value and the true value of the estimated quantity tends to zero. Accordingly, the algorithm in Box 1 was used to simulate 1000 studies (each with $n_b \cdot n_w \cdot n_e$ experiments). Then, the Aryani method and the mixed-effects model were applied to each simulated study, resulting in 1000 estimates for each source of variation (between-strain, within-strain and experimental variability). If the methods were robust (and the number of simulated studies was high enough), according to the definition of biased estimators, the distribution of the parameter estimates would be centred around the value used to generate the simulations (Garre et al., 2019b). This was checked by visually inspecting scatter and density plots, and based on the mean of the parameter estimates obtained in the simulations. Calculations were repeated, increasing the number of simulated studies, without observing any influence on the results. The algorithm was implemented in R version 3.5.3 and is available on the GitHub page of one of the co-authors (https://github.com/albgarre/comparison-variability-methods).

2.3. Description of the case studies

We used data from two published datasets (Aryani et al., 2015a; Aryani et al., 2015b) as case studies to compare the three proposed methods. These were selected due to the large differences in the

Box 1
Algorithm for simulating the impact of nested within-strain, between strain and experimental variability in kinetic experiments.

1. For s in 1 to $n_{studies}$:

1.1 Assign a value to the grand mean ($\overline{X}$).

1.2 For B in 1 to n_s:

1.2.1 Calculate the mean value of between-strain replicates ($\overline{X}_B$) by adding a random number from a normal distribution with variance ${\sigma^2}_b$ to $\overline{X}$.

1.2.2 For W in 1 to n_b:

1.2.2.1 Calculate the mean value for the within-strain replicates ($\overline{X}_{WB}$) by adding a random number from a normal distribution with variance ${\sigma^2}_w$ to $\overline{X}_B$.

1.2.2.2 For E in 1 to n_e:

1.2.2.2.1 Simulate the observed value (X_{EWB}) by adding a random number from a normal distribution with variance σ_e to $\overline{X}_{WB}$.

1.3 Save the vector, X_{EWB}, of simulated observations for study s.

2. Return the matrix of simulated observations for the $n_{studies}$ studies, Y.

relevance of each variability source reported in each one of them.

2.3.1. Case study I: microbial growth of Listeria monocytogenes

The goal of the first study (Aryani et al., 2015b) was the quantification of between-strain, within-strain and experimental variability in growth of *Listeria monocytogenes*. Briefly, this study analysed the microbial growth under a plethora of conditions (different media pH, water activity, lactic acid concentrations and temperatures). For each condition, experiments were performed for 20 different *L. monocytogenes* strains, accounting for differences between strains of the same species. For each strain, experiments were reproduced on three different days with freshly prepared cultures to observe differences in growth of biologically independent reproductions of the same strain. The experiments were done in duplicate at the same time using the same culture to analyse the experimental variability. Therefore, 120 growth experiments (20 strains × 3 biologically independent reproductions × 2 experimental repetitions) were done in total for each growth condition.

For each experiment, the maximum specific growth rate (μ) was estimated using the 2-fold dilution method. Hence, for each condition, a table of 120 values of μ was reported. Although the complete dataset has been analysed and the results were reported as supplementary material, for brevity, only the growth data obtained in medium supplemented with 5 % sodium chloride (NaCl) is discussed in this article.

2.3.2. Case study II: microbial inactivation of Listeria monocytogenes

As a second case study, the data on isothermal inactivation of *L. monocytogenes* obtained by Aryani et al. (2015a) was used. This study had a similar goal and experimental design as the one used as Case Study I, although it focused on the variation of microbial inactivation with respect to between-strain, within-strain and experimental variability. The experimental design included the same 20 strains of *L. monocytogenes* as for the growth experiments. For each strain, experiments were done on 3 different days with freshly prepared cultures and each inactivation experiment was done in duplicate at the same time using the same culture, resulting in 120 experiments (20 strains × 3 biologically independent reproductions × 2 experimental repetitions) in total per temperature. In that study, the authors fitted a re-parameterized version of the Weibull inactivation model (Metselaar et al., 2013) for each repetition, estimating the treatment time required to cause 6 log-reductions in the microbial count (t_{6D}). Therefore, for each treatment temperature, a table of 120 values of t_{6D} was reported. Before applying the statistical methods, they were divided by six to calculate equivalent D-values. Although the study by Aryani et al. (2015a) reports experiments at three different temperatures (55, 60 and 65 °C), only the data at 55 °C are analysed here. The data obtained at the other two temperatures were also analysed, reaching the same conclusions as for 55 °C (not shown).

3. Results

3.1. Assessment of the robustness of each method

The robustness of the different methods for the quantification of variability included in our study was evaluated using numerical simulations mimicking the experimental design of Aryani et al. (2015a, 2015b). In this study, the inactivation of 20 bacterial strains was

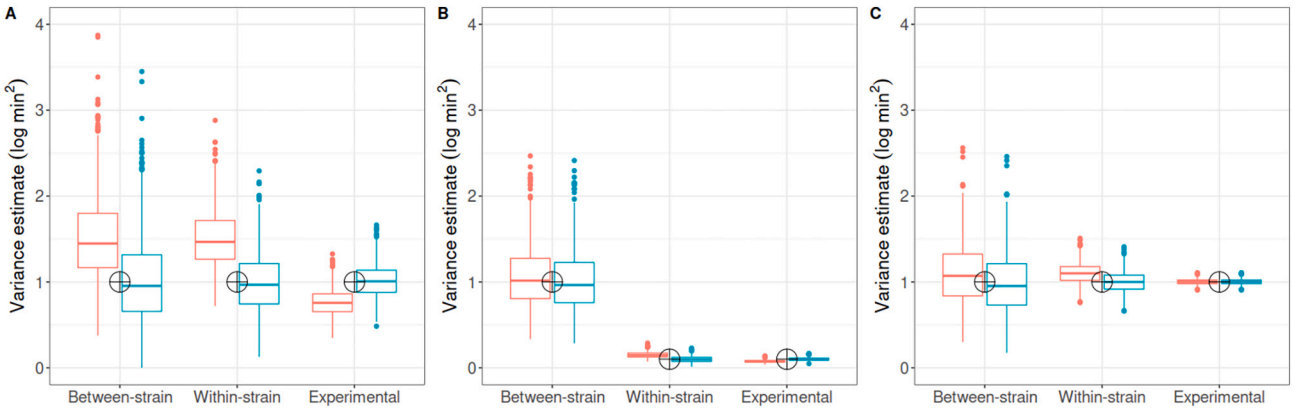

Fig. 1. Comparison between the variance estimates for the 6D reduction time (in log minutes) for simulated data obtained using the Aryani method (red boxes) and the mixed-effects model (blue boxes) in 1000 simulated inactivation experiments with different variances: (A) $\sigma^2_b = \sigma^2_w = \sigma^2_e = 1$log min^2; $n_b = 20$, $n_w = 3$, $n_e = 2$; (B) $\sigma^2_b = 1$log min^2; $\sigma^2_w = \sigma^2_e = 0.01$log min^2; $n_b = 20$, $n_w = 3$, $n_e = 2$; (C) $\sigma^2_b = \sigma^2_w = \sigma^2_e = 1$log min^2; $n_b = 20$, $n_w = 10$, $n_e = 10$. The crossed circles illustrate the values of variance used to generate the simulations for each condition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

observed for 3 biologically independent reproductions and 2 experimental repetitions. The outcome variable was the time for 6 log-reductions divided by 6 to calculate an average *D*-value, which were log-transformed to study the variability because preliminary exploratory analysis showed that this parameter was approximately log-normally distributed over temperature.

First, we made simulations under the (unrealistic) assumption that the three variance components take the same value ($\sigma^2_b = \sigma^2_w = \sigma^2_e = 1$ log min^2). Fig. 1A illustrates the estimates of variance obtained by the Aryani method in 1000 simulated studies. As shown in this plot, in these conditions, the Aryani method overpredicts the between-strain (mean variance (s^2_b) of 1.51 log min^2) and within-strain (mean s^2_w of 1.50 log min^2) variances with respect to the values used for the simulations, whereas the experimental variance is underpredicted (mean s^2_e of 0.766 log min^2). On the other hand, the mixed-effects model calculates unbiased estimates for every variance component (mean values of 1.01, 0.99 and 1.02 log min^2 for the between-strain, within-strain and experimental variances, respectively; practically equal to those used for generating the data).

The bias in the Aryani method is caused by the nested experimental design that is not accounted for by the statistical method. The data points are not totally independent because the same bacterial culture is used several times in the same experiment. When the mean squared errors are calculated, the deviations of the particular replicate with respect to the grand mean are propagated between levels. This represents a violation of the statistical hypotheses of the Aryani method. The Aryani method used the sample variance (s^2) as an estimator of the population variance. However, this is only valid when every observation is independent. As demonstrated in Appendix I for the case of two levels, the nested experimental design is a violation of this hypothesis, introducing a bias in the parameter estimates. Consequently, the estimates of the Aryani method are not unbiased estimators according to the classical definition from frequentist statistics (Voinov and Nikulin, 2012).

According to the calculations in Appendix I, in case of two variability levels, the bias would be proportional to the variance of the lower level (experimental variability). An interpretation of this result is based on the

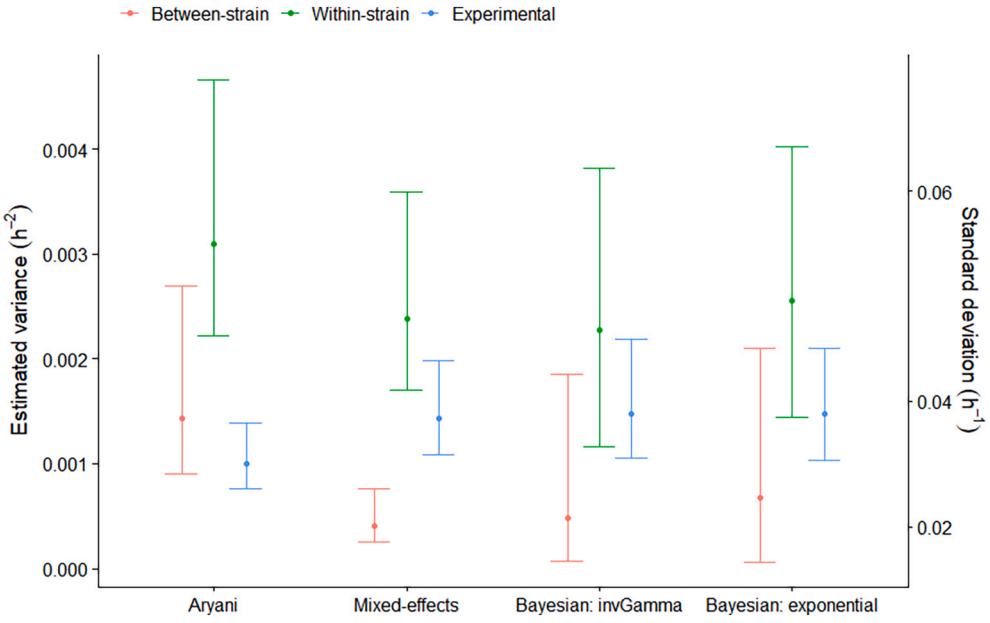

Fig. 2. Comparison of the variance quantification obtained for each method for the maximum specific growth rate (in log CFU/h) of *Listeria monocytogenes*. The dots represent the expected value for the variance and the bars the 95 % confidence intervals (for the Aryani method and the mixed-effects model) and the 95 % credible intervals (for the Bayesian models). The bars are coloured according to the source of variation they refer to.

nested experimental design: the method should account for the levels of variation that have already been explained (e.g. by variation at the experimental level) when studying variation at a higher level (e.g. by intrinsic differences at reproduction level). Failing to do so results in an overestimation of the between-strain and within-strain variability. To inspect this hypothesis, simulations have been repeated setting both the experimental and the within-strain variability to 0.01 log min^2 (i.e. a variance 100 times smaller than the between-strain variability). The distributions of the variance estimates are shown in Fig. 1B, confirming that, as expected based on the theoretical results, the bias of the Aryani method is reduced when the variance of the lower levels is smaller than the one of the highest level. Namely, a mean variance of 1.05, 0.15 and 0.08 log min^2 were obtained for the between-strain, within-strain and experimental variances, respectively. By the mixed-effects model, unbiased estimates are obtained for the variance components (mean values of 1.00, 0.10 and 0.10 log min^2 for the between-strain, within-strain and experimental variances, respectively).

In the case studied here, the two lowest levels of variation are the within-strain and experimental variability. The contribution of these components to the total variability can theoretically be reduced by the experimental design. As shown in the supplementary material, using a simulation scheme similar to the one by Garre et al. (2019b), the use of 20 time points per treatment instead of 5 can cause almost a two-fold reduction of the experimental variability (0.056 log minutes down to 0.035). On the other hand, the within-strain variability is associated with a biological variation (it is the upshot of genetic/physiological differences between cells of the same strain) and, as such, cannot be reduced through experimental design. According to the analytical result, the bias in the Aryani method is inversely proportional to the number of data points (in this case, *D*-values) taken at the lower levels of the model. Consequently, simulations have been repeated considering that 10 biologically independent and 10 technical replicates were made for each one of the 20 observed strains (Fig. 1C). This experimental design results in unbiased estimates for the experimental variability estimated using the Aryani model (mean s^2_e of 1.00 log min^2) and strongly reduces the bias for the other two variance components related to variability (mean of 1.10 log min^2 for both the within-strain and between-strain variances). Again, the mixed-effects model produced unbiased estimates for every variance component (mean values of 0.99, 1.00 and 1.00 log min^2 for the between-strain, within-strain and experimental variances, respectively).

3.2. Case study I: variability in growth of L. monocytogenes

Fig. 2 shows the estimated experimental, biological and strain variability using the Aryani method, the mixed-effect model and the Bayesian model with two types of prior distributions for growth of *L. monocytogenes* in high salt conditions. According to the Aryani method, the experimental variability ($\sigma_e = 0.032$ h^{-1}) has the smallest contribution to the total variation, which is dominated by within- and between-strain variability ($\sigma_w = 0.057$ h^{-1}; $\sigma_b = 0.038$ h^{-1}). The mixed-effects model resulted in very different variance estimates: by this method, between-strain variability ($\sigma_b = 0.020$ h^{-1}) has the lowest contribution to the total variation and experimental variability ($\sigma_e = 0.038$ h^{-1}) is close to within-strain variability ($\sigma_w = 0.049$ h^{-1}).

The numerical studies (Section 3.1) and the statistical analysis (Appendix I) anticipated that the Aryani method overestimates the contribution of between-strain and within-strain variability, and underestimates the experimental variability, whereas the mixed-effects model would be unbiased. The results are consistent with those predictions, showing a bias in the Aryani method due to the nested design, and the estimates of the mixed-effect model are more reliable than the ones of the Aryani method.

Regarding the Bayesian models, both calculated rather similar expected values for every parameter, indicating that the priors defined for these models are poorly informative (as intended). These expected values are similar to those estimated using the mixed-effects model. This is also expected, as the underlying Bayesian model is in structure identical to the mixed-effects model. However, the Bayesian credible intervals for the within-strain and between-strain variability are broader than the confidence intervals calculated by the mixed-effects model. The confidence intervals in the mixed-effects model have been calculated based on the hypothesis that the ratio between the population and sample variance follows a Chi-square distribution. Therefore, it only considers the number of degrees of freedom of the model and the estimated value for the variance for that level of variability (experimental, within-strain or between-strain). However, due to the hierarchical structure of the system, the uncertainty about the true variability of the lower levels of the model should be propagated to higher levels, inflating the uncertainty associated with the variance estimates. This is illustrated

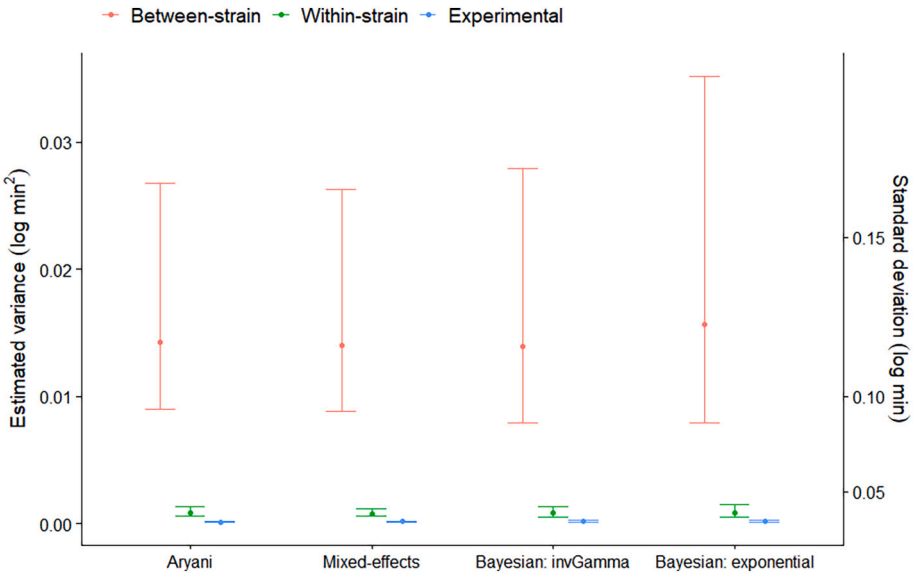

Fig. 3. Comparison of the variance quantification obtained for each method for the inactivation of *Listeria monocytogenes*. The dots represent the expected value for the variance and the bars the 95 % confidence intervals (for the Aryani method and the mixed-effects model) and the 95 % credible intervals (for the Bayesian models). The bars are coloured according to the source of variation they refer to.

in the numerical simulations, as the parameter estimates of the higher level are more spread than those of the lower levels (Fig. 1A). As a result, the hypotheses for building the confidence intervals are violated, resulting in too narrow confidence intervals for the mixed-effects model. Hence, the confidence intervals calculated for the mixed-effects model should be considered as an approximation. As a general recommendation, for a precise description of their uncertainty, Bayesian models should be used because they provide a more adequate framework to include variability in models (van Boekel, 2020).

These calculations were repeated for every growth condition analysed by Aryani et al. (2015b). The results were similar to those illustrated in this section, with a consistent bias in the Aryani model when the experimental variability and/or within-strain variability were large. The parameter estimates obtained for each case are reported in Supplementary Table 1.

3.3. Case study II: variability in inactivation of L. monocytogenes

Fig. 3 shows the estimates for between-strain, within-strain and experimental variability estimated from the data on thermal inactivation gathered by Aryani et al. (2015a). In this case, there are practically no differences between values estimated by the three models. As concluded in Section 3.1 based on numerical simulations, the bias of the Aryani method is proportional to the variance of the lower levels. Because both experimental and within-strain variabilities are much smaller than between-strain variability in this study, the bias in the Aryani method is also small and results are very similar to the other models.

The Bayesian credible intervals are again broader than those estimated using the mixed-effects model (especially for the exponential prior). However, here the differences are smaller than for Case Study I. As argued in Section 3.2, uncertainty is propagated from the lowest level (experimental variability, in this case) to the higher levels. Because the uncertainty around experimental variability is relatively small in this case study, the inflation of the credible intervals of the higher levels due to uncertainty propagation is also small. As a result, the difference between the CI of the Bayesian model and those calculated using the mixed-effects models are closer in this case than for Case Study I. When using an exponential prior in the Bayesian model, we obtained a credible interval for the between-strain variability that are broader than the other three methods tested, showing the impact of prior choice on parameter estimates. The calculations were repeated from the other two temperatures tested by Aryani (60 and 65 °C). Also for these temperatures there were no major differences between the results of the four models (results not shown).

Table 1
Advantages and disadvantages of the three methods for quantification of variability and uncertainty considered in this study.

	Advantages	Disadvantages
Aryani method	• Does not require advanced statistical knowledge.	• Biased estimates, although its magnitude can be small depending on the biological characteristic of interest. • Calculation of confidence intervals requires post-hoc calculations. • Cannot easily deal with unbalanced designs or missing data. • Can only analyse nested sources of variation. • Not-robust estimation of the uncertainty associated with the sources of variation on the highest levels.
Mixed-effects model	• Unbiased estimates. • Short computation times. • Convergence of the fitting algorithm can be easily evaluated. • Estimates of variability include a measure of uncertainty. • Can deal with unbalanced designs or missing data. • Can analyse some types of non-nested sources of variation.	• Requires specific statistical software. • Calculation of confidence intervals requires post-hoc calculations. • May fail to converge for complex (non-linear) models. • Limited definition of probability distributions for model variables. • Underestimation of the uncertainty of the higher levels
Bayesian model	• Can be applied for complex (non-linear) models. • Flexibility for defining distributions of model variables. • Prior distributions grant flexibility. • Can deal with unbalanced designs or missing data. • Provides flexibility to model many types of non-nested sources of variation. • Accurate estimation of the uncertainty associated with the sources of variation on the highest levels.	• Requires expertise on Bayesian methods. • Requires specific statistical software. • Potentially long running times. • The convergence of the Markov Chain must be assessed. • Prior distributions may impact the parameter estimates, especially when data is poorly informative.

4. Discussion

The consideration of variability is very relevant for quantitative microbiological risk assessment (EFSA Scientific Committee et al., 2018; Pouillot and Guillier, 2020). This is certainly complex, as even simple QMRA are affected by a large number of sources of variability with very different nature (variability in the climate, variability in logistic parameters, batch variability, variability in the microbial response…). One of the sources of variability that has received more attention during the last years is that in the microbial response (reviewed by den Besten et al., 2018; Koutsoumanis and Aspridou, 2017). This variability can be simplified as the variance in the kinetic parameters between different microbial populations, a parameter that can be estimated from experimental data (Aryani et al., 2015a; Garre et al., 2020; Jaloustre et al., 2012). However, differences in experimental results are rarely due to a single source of variability (Garre et al., 2021). Therefore, the use of methods able to account for complex variance forms is essential for the incorporation of variability in QMRA models.

In this article, we have compared three different methods for this task: the Aryani method, mixed-effect models, and Bayesian models. Table 1 summarises the advantages and disadvantages of each method considered based on the results of this investigation. The main advantage of the Aryani method is its simplicity. It can be implemented in spreadsheets (e.g. MS Excel) and it does not require the use of statistical software (e.g. the R programming language). However, as demonstrated in this article, it does not account for the nesting of the experimental design. Therefore, it introduces a bias in the estimates of variance on the higher levels that is proportional to the variance of the lower levels and inversely proportional to the number of data points per group (Appendix I). In contraposition, the mixed-effect and Bayesian models calculate unbiased estimates of the variance estimates. However, both approaches require the use of advanced statistical software tools and the definition of a statistical model describing the different error terms. Consequently, the application of mixed-effects or Bayesian models requires an increase in the complexity of model definition, implementation and interpretation.

Besides the biased parameter estimates, one main limitation of the Aryani method is that it is restricted to balanced datasets (as those applied in the cases studied here). It is based on hypotheses regarding the distribution of the sample variance that only hold if the degrees of freedom of the different groups are balanced. For instance, in Case Study I, this means that the number of biologically independent reproductions per strain remains constant. Otherwise, the method would add an

additional bias, whose contribution is hard to quantify. This restriction to balanced designs imposes a restriction in the design that can be very relevant for growth experiments near the growth/no-growth boundary. Although methods applied for correcting the degrees of freedom in unbalanced designs in ANOVA analysis (Gelman and Hill, 2007) could potentially be applied to the method of Aryani, they can be quite complex. Hence, their application to a method whose main advantage is its simplicity can be hard to justify.

Another limitation of the Aryani method is that it is only applicable to relatively simple experimental designs. It can only be used to describe cases where the sources of variability are related to the same experimental outcome. For instance, in Case Study I, the three levels of variability considered (between-strain, within-strain and experimental) are defined as the variance of the square root of the specific growth rate. If the purpose of the study was analysing the variability of several parameters (e.g. specific growth rate and lag phase), the Aryani method could only analyse the variability of each parameter independently. To account for any co-dependence between parameter estimates, more complex models would be needed.

Mixed-effect and Bayesian models provide more flexible approaches for the quantification of variability. Both can be used to analyse the variation of several model parameters in the same model (e.g. the slope and the curvature of the survivor curve), accounting for co-dependencies (McElreath, 2016). However, although this was not the case in the simple case studies analysed here, mixed-effects models may fail to converge for complex models (especially non-linear models) due to identifiability issues (Gelman and Hill, 2007; Nie, 2007).

Moreover, mixed-effects models require a post-hoc calculation to compute uncertainty associated with the parameter estimates, such as the one proposed in this article (Eq. (5)). Since this equation does not account for uncertainty propagation, it can underestimate the width of the confidence intervals for the higher levels of the model (between-strain and within-strain variability in the case studies in this study). Also, in mixed-effect models only a limited range of probability distributions for the model parameters (normal distributions) can be used. Therefore, in complex cases, a more flexible Bayesian approach for parameter estimation (Gelman and Hill, 2007) may be needed.

Although Bayesian models provide further flexibility for model definition, they are also more complex to implement and parameterize. First, Bayesian models need prior distributions, which are updated using data through Bayesian inference. In complex models (e.g. multilevel models), the impact of the prior distributions can be hard to determine without a sensitivity analysis, where the effect of choice of prior on the posterior distributions is analysed. In case that the priors strongly affect the posteriors, there should be good reasons for their choice. Second, Bayesian methods for model calibration require numerical techniques to estimate the posterior distributions of the model parameters, usually based on algorithms using MCMC simulations (Csilléry et al., 2010; McElreath, 2016). This requires a careful diagnostic of the fitted model, to ensure that the Markov Chain converged and mixed correctly (Brooks et al., 2011). Otherwise, the model may converge to spurious model parameters resulting in uninformative inferences (McElreath, 2016). Therefore, the use of a Bayesian method increases the work required for model definition and analysis, and the statistical skills of the researcher.

The three methods analysed in this study have all advantages and disadvantages. The simplicity of the Aryani method makes it ideal for an initial screening, providing a general, first insight in the effects of variability that can support a decision for further analysis. However, if the scope of the study is to obtain a quantitative estimation of variability that can be used as input for a mathematical model to support risk assessment, then a more advanced approach may be needed. The bias in the Aryani method can be somewhat mitigated through experimental design, e.g. by increasing the number of data points on the lower levels of the model. Nevertheless, this may be impractical as it would require a large number of experiments (e.g. Fig. 1C). Hence, in cases where the contribution of the source of variation in the lowest (experimental) level is high, a mixed-effects or Bayesian model may be more practical than a change of the experimental design. To conclude, the choice of model should be justified according to the research question under investigation, the required accuracy of the model and the statistical skills of the involved researchers. In this sense, the development of software applications that provide a user-friendly interface to complex statistical methods can facilitate their use by applied scientists, especially when these tools are distributed Open Access (Possas et al., 2022). This is relevant because, as illustrated in this article, the interpretation of the experimental results and the conclusions drawn from the study can be affected by the underlying statistical method.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijfoodmicro.2022.109935.

Conflict of interest

All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.

This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

Data availability

Data will be made available on request.

Acknowledgments

Alberto Garre was supported by a postdoctoral granted from the Fundación Seneca (20900/PD/18) and by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Individual Fellowship grant No 844423 (FANTASTICAL).

Appendix I. Demonstration of the bias in the Aryani method in the case of only two levels

In the case when there are only two sources of variation (e.g. in an experimental design with n_b biologically independent reproductions and n_e experimental replicates), Eq. (6) from the main manuscript can be re-written as

$$X_{EB} = \overline{X} + \varepsilon_B + \varepsilon_E \tag{A.1}$$

$$\varepsilon_B \sim N(0, \sigma_B)$$

$$\varepsilon_E \sim N(0, \sigma_E)$$

According to the method suggested by Aryani, s^2_b is used as an estimator of σ^2_b. This estimator is defined as

$$s_b^2 = \frac{1}{df}\sum_{E=1}^{n_e}\sum_{B=1}^{n_b}(X_{EB} - \overline{X}_B)^2 \tag{A.2}$$

where df is the number of degrees of freedom, n_e is the number of experiments, and n_b is the number of biologically independent reproductions. The term X_{EB} is the observed quantity and $\overline{X}_B$ is the mean of all of the observations corresponding to biological replicate B. The latter is defined as

$$\overline{X_B} = \frac{1}{n_e}\sum_{E=1}^{n_e} X_{EB}; \forall i \in B_j \tag{A.3}$$

where n_e is the number of observations taken for biologically independent reproduction B. Substituting Eq. (A.1) into Eq. (A.3) yields:

$$\overline{X_B} = \frac{1}{n_e}\sum_{E=1}^{n_e} X_{EB} = \frac{1}{n_e}\sum_{E=1}^{n_e}(\overline{X} + \varepsilon_B + \varepsilon_e) = \frac{1}{n_e}\left(n_e\overline{X} + n_e\varepsilon_B + \sum_{E=1}^{n_e}\varepsilon_e\right) = \overline{X} + \varepsilon_B + \frac{1}{n_e}\sum_{E=1}^{n_e}\varepsilon_e \tag{A.4}$$

That is, $\overline{X_B}$ equals the expected value of our outcome variable for biologically independent reproduction B ($\overline{X_B} = \overline{X} + \varepsilon_B$) plus the term $\frac{1}{n_e}\sum_{E=1}^{n_e}\varepsilon_e$. Substituting this results in Eq. (A.2) yields:

$$s_b^2 = \frac{1}{df}\sum_{E=1}^{n_e}\sum_{B=1}^{n_b}\left(X_{EB} - \overline{X_B} - \frac{1}{n_e}\sum_{E=1}^{n_e}\varepsilon_e\right)^2 \tag{A.5}$$

However, the unbiased estimator of the population variance is the sample variance (Box et al., 2005), defined for σ^2_b as:

$$\sigma_b^2 = \frac{1}{df}\sum_{E=1}^{n_e}\sum_{B=1}^{n_b}(X_{EB} - \overline{X_B})^2 \tag{A.6}$$

Hence, the s^2_b deviates from the unbiased estimator of σ^2_b by the factor $\frac{1}{n_e}\sum_{E=1}^{n_e}\varepsilon_e$. Considering that the model hypotheses state that all the errors are normally distributed, the Central Limit Theorem can be applied to estimate that this term follows a normal distribution with mean zero and standard error $\frac{\sigma_\varepsilon}{\sqrt{n_e}}$. Accordingly, the Aryani method overpredicts the contribution of the higher level of variance (see also Fig. 1a). The magnitude of the bias will depend on the variance of the lower source of variation (the experimental variation) and the number of samples taken per strain. In this case, the bias in the estimate for the reproduction variability is small when the experimental variability is small.

References

Arvaniti, M., Skandamis, P.N., 2022. Defining bacterial heterogeneity and dormancy with the parallel use of single-cell and population level approaches. Curr. Opin. Food Sci. 44, 100808 https://doi.org/10.1016/j.cofs.2022.100808.

Aryani, D.C., den Besten, H.M.W., Hazeleger, W.C., Zwietering, M.H., 2015a. Quantifying variability on thermal resistance of listeria monocytogenes. Int. J. Food Microbiol. 193, 130–138. https://doi.org/10.1016/j.ijfoodmicro.2014.10.021.

Aryani, D.C., den Besten, H.M.W., Hazeleger, W.C., Zwietering, M.H., 2015b. Quantifying strain variability in modeling growth of listeria monocytogenes. Int. J. Food Microbiol. 208, 19–29. https://doi.org/10.1016/j.ijfoodmicro.2015.05.006.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 https://doi.org/10.18637/jss.v067.i01.

Bates, D.M., Watts, D.G., 2007. Nonlinear Regression Analysis and Its Applications, 1edition. Wiley-Interscience, New York, NY.

Box, G.E.P., Hunter, J.S., Hunter, W.G., 2005. Statistics for Experimenters: Design, Innovation, and Discovery. Wiley-Blackwell, Hoboken, N.J.

Brooks, S., Gelman, A., Jones, G.L., Meng, X.-L., 2011. Handbook of Markov Chain Monte Carlo, Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press, Taylor & Francis, Boca Raton London.

Clemente-Carazo, M., Leal, J.-J., Huertas, J.-P., Garre, A., Palop, A., Periago, P.M., 2021. The different response to an acid shock of two salmonella strains marks their resistance to thermal treatments. Front. Microbiol. 12, 2616. https://doi.org/10.3389/fmicb.2021.691248.

Csilléry, K., Blum, M.G.B., Gaggiotti, O.E., François, O., 2010. Approximate bayesian computation (ABC) in practice. Trends Ecol. Evol. 25, 410–418. https://doi.org/10.1016/j.tree.2010.04.001.

den Besten, H.M.W., Wells-Bennik, M.H.J., Zwietering, M.H., 2018. Natural diversity in heat resistance of bacteria and bacterial spores: impact on food safety and quality. Annu. Rev. Food Sci. Technol. 9, 383–410. https://doi.org/10.1146/annurev-food-030117-012808.

Scientific Committee, E.F.S.A., Benford, D., Halldorsson, T., Jeger, M.J., Knutsen, H.K., More, S., Naegeli, H., Noteborn, H., Ockleford, C., Ricci, A., Rychen, G., Schlatter, J. R., Silano, V., Solecki, R., Turck, D., Younes, M., Craig, P., Hart, A., Von Goetz, N., Koutsoumanis, K., Mortensen, A., Ossendorp, B., Martino, L., Merten, C., Mosbach-Schulz, O., Hardy, A., 2018. Guidance on uncertainty analysis in scientific assessments. EFSA J. 16 https://doi.org/10.2903/j.efsa.2018.5123.

FAO, WHO, 2021. Microbiological Risk Assessment - Guidance for Food. Microbiological Risk Assessment Series No. 36. Rome. https://doi.org/10.4060/cb5006en.

Garre, A., den Besten, H.M.W., Fernandez, P.S., Zwietering, M.H., 2021. Not just variability and uncertainty; the relevance of chance for the survival of microbial cells to stress. Trends Food Sci. Technol. 118, 799–807. https://doi.org/10.1016/j.tifs.2021.10.033.

Garre, A., Egea, J.A., Esnoz, A., Palop, A., Fernandez, P.S., 2019a. Tail or artefact? Illustration of the impact that uncertainty of the serial dilution and cell enumeration methods has on microbial inactivation. Food Res. Int. 119, 76–83. https://doi.org/10.1016/j.foodres.2019.01.059.

Garre, A., Peñalver-Soto, J.L., Esnoz, A., Iguaz, A., Fernandez, P.S., Egea, J.A., 2019b. On the use of in-silico simulations to support experimental design: a case study in microbial inactivation of foods. PLoS ONE 14. https://doi.org/10.1371/journal.pone.0220683.

Garre, A., Zwietering, M.H., den Besten, H.M.W., 2020. Multilevel modelling as a tool to include variability and uncertainty in quantitative microbiology and risk assessment. Thermal inactivation of Listeria monocytogenes as proof of concept. Food Res. Int. 137, 109374 https://doi.org/10.1016/j.foodres.2020.109374.

Gelman, A., 2014. Bayesian data analysis. In: Chapman & Hall/CRC Texts in Statistical Science, Third edition. CRC Press, Boca Raton.

Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by browne and Draper). Bayesian Anal. 1, 515–534. https://doi.org/10.1214/06-BA117A.

Gelman, A., Hill, J., 2007. Cambridge University Press, Cambridge; New York.

Haas, C.N., Rose, J.B., Gerba, C.P., 2014. Quantitative Microbial Risk Assessment, 2nd ed. John Wily & Sons Inc, Hoboken, New Jersey.

Jaloustre, S., Guillier, L., Morelli, E., Noël, V., Delignette-Muller, M.L., 2012. Modeling of Clostridium perfringens vegetative cell inactivation in beef-in-sauce products: a meta-analysis using mixed linear models. Int. J. Food Microbiol. 154, 44–51. https://doi.org/10.1016/j.ijfoodmicro.2011.12.013.

Jongenburger, I., Reij, M.W., Boer, E.P.J., Gorris, L.G.M., Zwietering, M.H., 2010. Factors influencing the accuracy of the plating method used to enumerate low numbers of viable micro-organisms in food. Int. J. Food Microbiol. 143, 32–40. https://doi.org/10.1016/j.ijfoodmicro.2010.07.025.

Koutsoumanis, K.P., Aspridou, Z., 2017. Individual cell heterogeneity in Predictive Food Microbiology: Challenges in predicting a “noisy” world. Int. J. Food Microbiol. 240, 3–10. https://doi.org/10.1016/j.ijfoodmicro.2016.06.021.

McElreath, R., 2016. Statistical Rethinking: A Bayesian Course With Examples in R and Stan, Chapman & Hall/CRC Texts in Statistical Science Series. CRC Press/Taylor & Francis Group, Boca Raton.

Metselaar, K.I., den Besten, H.M.W., Abee, T., Moezelaar, R., Zwietering, M.H., 2013. Isolation and quantification of highly acid resistant variants of Listeria monocytogenes. Int. J. Food Microbiol. 166, 508–514. https://doi.org/10.1016/j.ijfoodmicro.2013.08.011.

Muñoz-Cuevas, M., Guevara, L., Aznar, A., Martínez, A., Periago, P.M., Fernández, P.S., 2013. Characterisation of the resistance and the growth variability of Listeria monocytogenes after high hydrostatic pressure treatments. Food Control 29, 409–415. https://doi.org/10.1016/j.foodcont.2012.05.047.

Nauta, M.J., 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. Int. J. Food Microbiol. 57, 9–18.
Nie, L., 2007. Convergence rate of MLE in generalized linear and nonlinear mixed-effects models: theory and applications. J. Stat. Plann. Infer. 137 (6) https://doi.org/10.1016/j.jspi.2005.06.010.
Pennone, V., Gonzales-Barron, U., Hunt, K., Cadavez, V., McAuliffe, O., Butler, F., 2021. Omnibus modeling of Listeria monocytogenes growth rates at low temperatures. Foods 10, 1099. https://doi.org/10.3390/foods10051099.
Plummer, M., 2019. rjags: Bayesian Graphical Models Using MCMC.
Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling.
Possas, A., Valero, A., Pérez-Rodríguez, F., 2022. New software solutions for microbiological food safety assessment and management. Curr. Opin. Food Sci. 100814 https://doi.org/10.1016/j.cofs.2022.100814.
Pouillot, R., Guillier, L., 2020. Understanding uncertainty and variability in risk assessment. In: Risk Assessment Methods for Biological and Chemical Hazards in Food. CRC Press, pp. 165–190.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Schendel, T., Jung, C., Lindtner, O., Greiner, M., 2018. Guidelines for uncertainty analysis: Application of the respective documents of EFSA and BfR for exposure assessments: Framework Partnership Agreement GP/EFSA/AMU/2016/01 (agreement 2). EFSA Support. Publ. 15 https://doi.org/10.2903/sp.efsa.2018.EN-1472.
Thompson, K.M., 2002. Variability and uncertainty meet risk management and risk communication. Risk Anal. 22, 647–654. https://doi.org/10.1111/0272-4332.00044.
van Boekel, M.A.J.S., 2021. To pool or not to pool: That is the question in microbial kinetics. Int. J. Food Microbiol. 354, 109283 https://doi.org/10.1016/j.ijfoodmicro.2021.109283.
van Boekel, M.A.J.S., 2020. On the pros and cons of Bayesian kinetic modeling in food science. Trends Food Sci. Technol. 99, 181–193. https://doi.org/10.1016/j.tifs.2020.02.027.
Vidakovic, B., 2011. Bayesian inference using Gibbs sampling – BUGS project. In: Vidakovic, B. (Ed.), Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support, Springer Texts in Statistics. Springer, New York, NY, pp. 733–745. https://doi.org/10.1007/978-1-4614-0394-4_19.
Voinov, V.G., Nikulin, M.S., 2012. Unbiased estimators and their applications: volume 1: univariate case. In: Mathematics and Its Applications Series No. 263, 1st ed. Springer.
Wells-Bennik, M.H.J., Janssen, P.W.M., Klaus, V., Yang, C., Zwietering, M.H., Den Besten, H.M.W., 2019. Heat resistance of spores of 18 strains of Geobacillus stearothermophilus and impact of culturing conditions. Int. J. Food Microbiol. 291, 161–172. https://doi.org/10.1016/j.ijfoodmicro.2018.11.005.
Vose, D., 2008. Risk Analysis - A Quantitative Guide, 3rd ed. Wiley.
Wijnands, L.M., Teunis, P.F.M., Kuijpers, A.F.A., Delfgou-Van Asch, E.H.M., Pielaat, A., 2017. Quantification of Salmonella survival and infection in an in vitro model of the human intestinal tract as proxy for foodborne pathogens. Front. Microbiol. 8, 1139. https://doi.org/10.3389/fmicb.2017.01139.