Italian Journal of Food Safety

Italian Journal of Food Safety 2026; 15:14544
doi:10.4081/ijfs.2026.14544

The Bayesian β-binomial model to improve the inference based on the results of sampling plans: the case study of the biotoxin monitoring plan of *Chamelea gallina* natural beds

Cesare Ciccarelli,[1] Angela Marisa Semeraro,[1] Vittoria Di Trani,[1] Melina Leinoudi,[2] Guglielmo D'Aurizio,[3] Francesca Barchiesi,[4] Sara Gentili,[5] Elena Ciccarelli[6]

[1]Local Public Health Service - Ascoli Piceno, San Benedetto del Tronto, Italy; [2]Chemist, Thessaloniki, Greece; [3]Regional Public Health Agency of Marche Region, Ancona, Italy; [4]Experimental Zooprophylactic Institute of Umbria and Marche "Togo Rosati", Ancona, Italy; [5]Veterinarian, Montemonaco, Italy; [6]Biologist, Reading, United Kingdom

Abstract

Sampling is a cornerstone of food safety monitoring, yet traditional frequentist approaches can yield unreliable inferences in the presence of rare or absent events, particularly when maximum likelihood estimates lie on the boundary of the parameter space and asymptotic approximations are not justified due to insufficient event counts. This study introduces the Bayesian β-binomial model as an effective framework to improve statistical inference in sampling plans, integrating prior knowledge with observed data to yield robust uncertainty estimates of contamination prevalence. The model was applied to ten years (2015-2024) of monitoring data on algal biotoxins in *Chamelea gallina* striped clams from classified harvesting areas in the Marche Region, Italy. Given the scarcity of historical data, three prior scenarios were tested – optimistic [β(1,10)], non-informative [β(1,1)], and pessimistic [β(5,10)] – to evaluate sensitivity to prior assumptions. Analytical determinations for five toxin groups (domoic acid, saxitoxin, azaspiracid, okadaic acid, and yessotoxin) showed that all results, except one sample below the maximum legal limit, were under quantification thresholds. Bayesian posterior estimates confirmed a very low probability of biotoxin accumulation above limits, even under pessimistic assumptions. The findings demonstrate the model's capacity to generate stable, interpretable estimates and credible intervals, which is particularly valuable when the prevalence is not exactly zero but may be extremely small. Overall, the Bayesian β-binomial approach strengthens evidence-based decision-making in food safety surveillance, providing a transparent and adaptable tool for risk assessment and regulatory management.

Key words: Bayesian inference, β-binomial model, sampling plans, risk assessment, zero-positive problem.

Correspondence: Cesare Ciccarelli, Local Public Health Service - Ascoli Piceno, via L. Manara n. 3/7 - 63074 San Benedetto del Tronto, Italy.
E-mail: cesare.ciccarelli@sanita.marche.it Tel.: +39.0735.7937474.

Introduction

Food safety is a critical component of public health protection and trade policy. Surveillance systems are implemented to monitor biological and chemical hazards in food products. However, due to cost, logistical constraints, and the perishable nature of many products, only a subset of items can be tested. As it is often impractical or impossible to observe entire populations, sampling is essential, and statistical inference from sample data plays a pivotal role in food safety decision-making (EFSA, 2012).

Traditional analyses generally rely on binomial models and frequentist estimation methods. However, these can yield unreliable estimates when few or no contaminated units are observed, particularly in low-prevalence settings (Van der Voet and Slob, 2007). The Bayesian β-binomial model provides an effective alternative by integrating prior knowledge with observed data, producing more stable and informative estimates (Spiegelhalter *et al.*, 2004; Johnson *et al.*, 2022).

The Bayesian framework's core strength lies in its natural knowledge-building process: prior beliefs are continuously updated and refined as new data become available, providing great flexibility in uncertainty quantification through probability theory (Gelman *et al.*, 2013; Johnson *et al.*, 2022). Bayesian methods have been increasingly proposed as an alternative framework for food safety risk assessment. Within this context, the Bayesian binomial model has become a natural tool to support quantitative microbial risk assessment (Williams *et al.*, 2011; Ranta *et al.*, 2015), to predict prevalence and evaluate sampling schemes (Benschop *et al.*, 2010), and to compare prevalence under outbreak vs. baseline conditions (Khan *et al.*, 2023). This study illustrates the theoretical foundations and practical application of the Bayesian β-binomial model to enhance statistical inference and decision reliability in sampling-based food safety assessments. As a case study, the model was applied to biotoxin monitoring data (2015-2024) in *Chamelea gallina* (striped clam) collected from natural harvesting areas in the Marche Region, Italy, where only limited data are available.

Materials and Methods

In statistical inference, the goal is to estimate an unknown quantity under uncertainty – typically expressed as a parameter of a probability distribution. Within the Bayesian framework, a probability distribution is inferred for the parameter given observed data. The Bayesian β-binomial model is mathematically tractable and offers a clear understanding of the core components of Bayesian inference: the prior distribution, the likelihood function, and the posterior distribution (Johnson *et al*., 2022).

Let θ denote the parameter space of interest, expressing the probability of compliance in a sampling plan. Our prior beliefs about θ are expressed by a probability distribution p (θ), the prior distribution. After observing data y, we update these beliefs *via* Bayes' theorem [Eq. 1]:

$$p(\theta|y) = \frac{p(y|\theta)\,p(\theta)}{p(y)} \qquad \text{[Eq. 1]}$$

Here, p (y│θ) is the likelihood function, expressing the probability of the observed data given parameter θ, and p(y) is the normalizing constant, or marginal likelihood [Eq. 2]:

$$p(y) = \int_{\Theta} p(y,\theta)d\pi = \int_{\Theta} p(y|\theta)p(\theta)d\theta \qquad \text{[Eq. 2]}$$

Assuming a binary outcome – for instance, meeting or not meeting a specific criterion – each sampling result can be modelled as a Bernoulli trial with success probability θ. The number of compliant results Y among n samples follows a binomial distribution [Eq. 3]:

$$p(Y = y|\Theta = \theta) = p(y|\theta) = \binom{n}{y}\theta^{y}(1-\theta)^{(n-y)} \qquad \text{[Eq. 3]}$$

The prior distribution is specified as a B (α, β) distribution [Eq. 4]:

$$p(\theta) = \left(\frac{1}{B(\alpha,\beta)}\right)\theta^{(\alpha-1)}(1-\theta)^{(\beta-1)} \qquad \text{[Eq. 4]}$$

Combining Eqs. 1, 3 and 4, the posterior distribution is proportional to Eq. 5:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto \frac{\theta^{(\alpha-1)}(1-\theta)^{(\beta-1)}}{B(\alpha,\beta)}\binom{n}{y}\theta^{y}(1-\theta)^{(n-y)} \propto \theta^{y}(1-\theta)^{(n-y+\beta-1)} \qquad \text{[Eq. 5]}$$

Thus, the posterior is also a β distribution [Eq. 6]:

$$p(\theta|y) = C\theta^{(\alpha'-1)}(1-\theta)^{(\beta'-1)} = \left(\frac{1}{B(\alpha',\beta')}\right)\theta^{(\alpha'-1)}(1-\theta)^{(\beta'-1)} \qquad \text{[Eq. 6]}$$

where α′ = α + y and β′ = β + n – y

Prior specification

The choice of B (α, β) is a crucial modeling decision, particularly in food safety contexts where regulatory or economic implications are significant. The prior encapsulates our beliefs about contamination probability before new data are observed (Johnson *et al*., 2022). Three common scenarios are: i) non-informative priors (Gelman *et al.* 2013), *e.g*., β (1, 1) or Jeffreys' β (0.5, 0.5), are used when little prior information is available; ii) weakly informative priors, reflecting general expectations while remaining modestly influential; iii) informative priors, derived from historical or empirical data, suitable for ongoing monitoring programs. When used appropriately, the prior enhances transparency by making assumptions explicit rather than implicit – a major strength of Bayesian analysis.

Case study: *Chamelea gallina* biotoxin monitoring

The model was applied to algal biotoxin monitoring data collected between 2015 and 2024 along the Marche Region's coast (Adriatic Sea, Italy). This area spans approximately 180 km and includes 68 classified harvesting zones and 70 sampling stations (Giunta della Regione Marche, 2024). Samples were analyzed by the *Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche* for domoic acid (DA), saxitoxin (STX), azaspiracid (AZA), okadaic acid (OA), and yessotoxin (YTX). Data were obtained from the Italian Ministry of Health's SINVSA information system. Table 1 resumes analytical methods, related limits of quantification (LoQ) and maximum limits (ML) following the Regulation CE/853/2004 (European Parliament, 2004). Because no local data were available before 2015 and only limited literature exists on biotoxin accumulation in *C. gallina* (Poletti, 2014; Fernández *et al*., 2019; Accoroni *et al*., 2024; Barnouin *et al.,* 2025), three scenarios were compared: i) optimistic – low probability of accumulation, β (1, 10); ii) clueless (non-informative) – uniform prior, β (1, 1); iii) pessimistic: high probability of accumulation, β (5, 10). The first and third scenarios are based on two opposite extreme assumptions and could be classified as informative priors although not based on real historical data. Analyses were conducted in R software (version 4.4.0) using the bayesrules, tidyverse, and janitor packages.

Results and Discussion

All analytical determinations, as reported in Table 2, were classified by year and categorized into three classes of values: <LoQ, ≥LoQ and ≤ML, and >ML. All results were below the LoQ, except for a single 2015 sample, which exceeded the LoQ but remained below the ML. Samples were collected from 49 different stations, corresponding to 70% of the 70 identified sites. Of the total, 69% of samples were collected regularly between 2015 and 2023, while the remaining 31% were obtained in 2024 during an algal bloom event characterized by OA bioaccumulation in mussels across all classified areas of the Marche Region. *Supplementary Table 1* summarizes the distribution of samples by year and sampling station. Assuming that samples were independent and that each had an identical probability θ of being below the LoQ – the parameter space of interest θ, where p(θ=θ) – the corresponding probability density function (PDF) was calculated. Figure 1 presents, for OA, the PDFs of the prior, likelihood, and posterior distributions under the three evaluated scenarios (optimistic, clueless, and pessimistic) at the end of 2015, 2023, and 2024, respectively. In Figure 1, the first column shows the PDFs obtained by processing the results of the first year of monitoring (only 10 results, one of which with OA>LoQ): for all three hypothesized scenarios, the posterior (the green area) appears shifted to the right, so the probability of obtaining results with OA>LoQ appears non-negligible. In the second column, obtained by processing the results obtained up to and including 2023, the posterior PDFs appear significantly shifted to the left and with higher vertices, mainly in the first scenario; therefore, it seems unlikely that OA>LoQ values will occur. Finally, in the third column, obtained by also processing the results obtained in 2024, under favorable conditions for the detection of algal biotoxins as a consequence of the intense algal blooms that occurred in that year, all the posterior PDFs, including that of the pessimistic scenario, appear flattened on the left, then with a limited probability of finding OA>LoQ values.

Figure 2 shows analogous PDFs for the remaining biotoxins: DA, STX, AZA, and YTX, with a more evident shifting to the left of posterior PDFs. Table 3 reports the 95% Bayesian credible intervals for θ at the end of 2024, calculated for OA≥LoQ and for each of the other biotoxins under the three prior scenarios.

The results indicate that the probability of observing OA concentrations above the LoQ in *C. gallina* striped clams from the classified harvesting areas of the Marche Region is extremely low, regardless of whether a low prior probability (optimistic scenario) or a non-informative prior (clueless scenario) is assumed. As illustrated in Figure 1 and the first column of Table 3, the probability of detecting values above the LoQ is very small, even under the pessimistic scenario, an unrealistic assumption given the available data. For the other biotoxins, the likelihood of detecting values

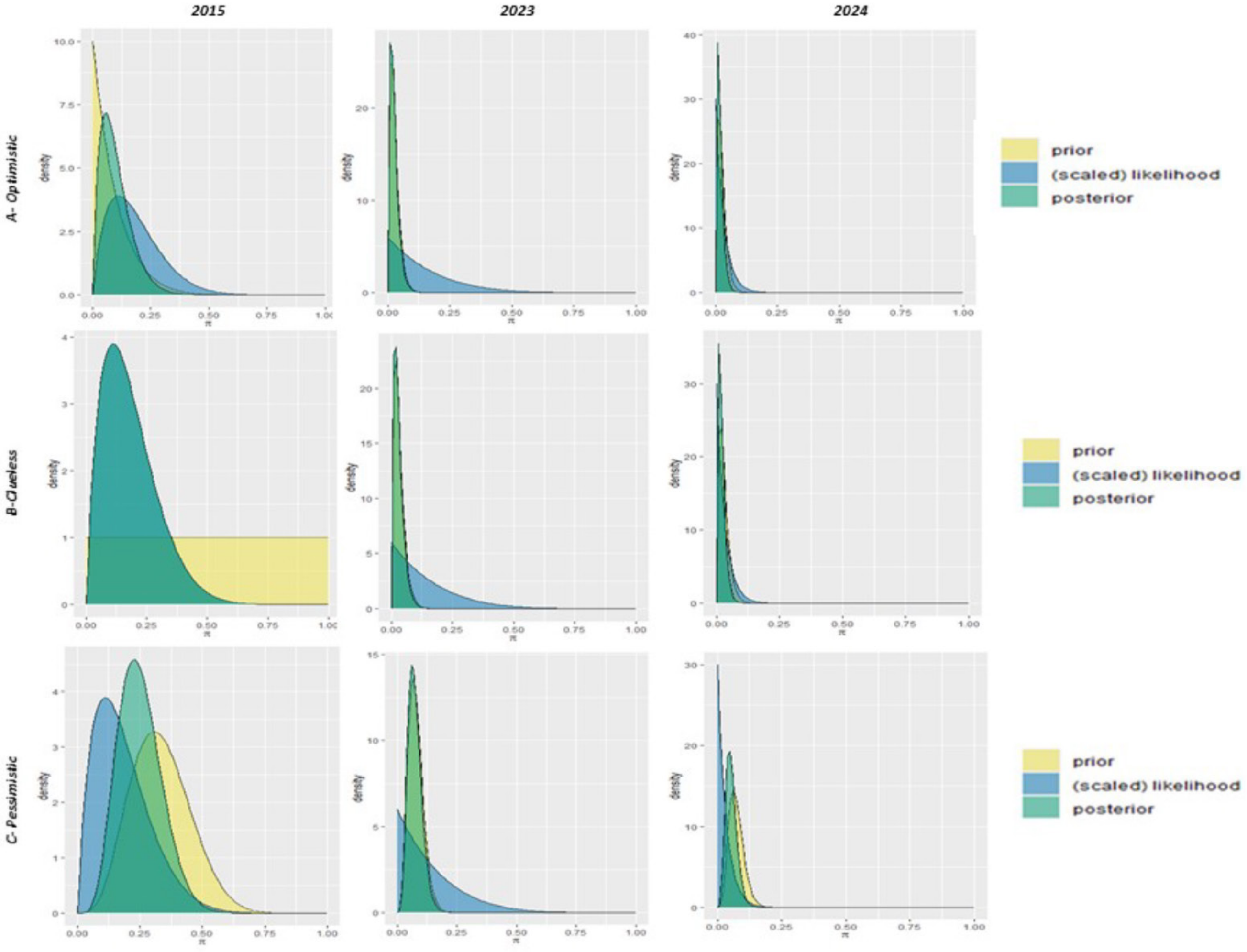

Figure 1. Probability density function related to okadaic acid, for the scenarios A (optimistic), B (clueless) and C (pessimistic), calculated for prior, likelihood (scaled) and posterior, respectively, at the end of 2015, 2023, and 2024.

Table 1. Analytical methods, related limits of quantification and maximum limits following the Regulation CE/853/2004.

	Analytical methods	Expressed as	LoQ	ML
Domoic acid group toxins	HPLC – AESAN EU-RL-MB Domoic acid, ver 1 2008	DA mg /kg	0.2	20
Saxitoxins	HPLC-FLD method EURLMB SOP – June 2020	µg STX 2HCL eq./kg	600	800
Azaspiracid-group toxins	LC-MS/MS - AESAN EU-RL-MB Lipophilic toxins, ver 5 2015	µg AZA eq./kg	30	160
Okadaic acid group toxins		µg AO eq./kg	40	160
Yessotoxin group toxins		mg YTX eq./kg	0.2	3.75

LOQ, limit of quantification; ML, maximum limit; HPLC, high performance liquid chromatography; AESAN, *Agencia Española de Seguridad Alimentaria y Nutrición*; EU-RL-MB, European Union Reference Laboratory for Monitoring of Marine Biotoxins; HPLC-FLD, high performance liquid chromatography-spectrofluorometric detection; EURLMB SOP, European Union Reference Laboratory for Monitoring of Marine Biotoxins Standard Operating Procedures; LC-MS/MS, liquid chromatography-tandem mass spectrometry; DA, domoic acid; STX, saxitoxin; AZA, azaspiracid; OA, okadaic acid; YTX, yessotoxin.

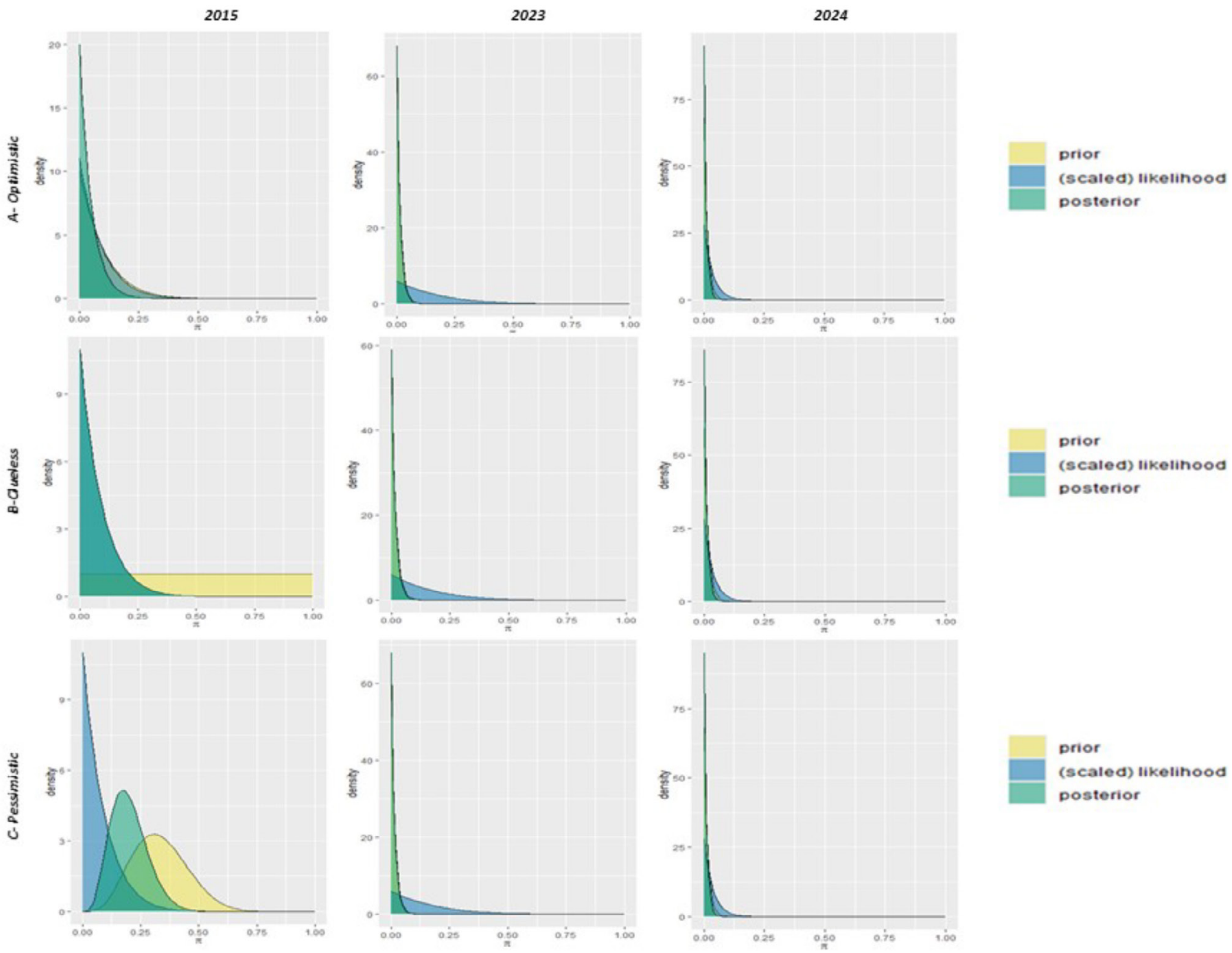

Figure 2. Probability density function related to domoic acid, saxitoxin, azaspiracid and yessotoxin, for the scenarios A (optimistic), B (clueless) and C (pessimistic), calculated for prior, likelihood (scaled) and posterior, respectively, at the end of 2015, 2023 and 2024.

Table 2. Analytical determinations grouped by year and divided into three classes of values: <LoQ, ≥LoQ and ≤ML, >ML.

Analytical determinations	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	Total	<LoQ	≥LoQ, ≤ML	>ML
Domoic acid group toxins	10	7	8	7	6	8	6	6	5	27	90	90	0	0
Muscle-paralysing toxin	10	7	7	5	6	7	6	5	5	27	85	85	0	0
Azaspiracid-group toxins	10	7	8	7	6	9	8	6	5	29	95	95	0	0
Okadaic acid group toxins	10	7	8	7	6	9	8	6	5	29	95	94	1	0
Yessotoxin group toxins	10	7	8	8	6	9	8	6	5	29	96	96	0	0
Total	50	35	39	34	30	42	36	29	25	141	461	460	1	0

LoQ, limit of quantification; ML, maximum limit.

Table 3. 95% posterior credible intervals at the end of 2024, calculated for okadaic acid ≥ limit of quantification and for each of the remaining biotoxins and related to the evaluated scenarios.

Scenarios	95% credible interval for OA≥LoQ	95% credible interval for DA, STX, AZA, YTX≥LoQ
Optimistic	0-0.045	0-0.031
Clueless	0-0.049	0-0.03
Pessimistic	0.011-0.099	0.006-0.094

OA, okadaic acid; LoQ, limit of quantification; DA, domoic acid; STX, saxitoxin; AZA, azaspiracid; YTX, yessotoxin.

above the LoQ was even lower, as shown in Figure 2 and the second column of Table 3.

Finally, data from the 10 years of monitoring on *C. gallina* striped clams, from classified harvesting areas along the central-western Adriatic coast of the Marche Region, confirmed that the probability of algal biotoxins accumulation above the analytical limits of quantification is exceedingly low, even considering the limited dataset and temporal variability.

Conclusions

The fundamental strength of the Bayesian framework lies in its intuitive and rigorous capacity to update probabilistic beliefs in light of new evidence. The β-binomial model, as shown above with the case study of *C. gallina* biotoxins monitoring, is able to enhance statistical inference from food safety sampling plans by providing: i) robust uncertainty estimates even in the presence of sparse or zero-positive data, such as in the case of bivalve mollusks with limited biotoxins' bioaccumulation ability; ii) quantitative uncertainty assessments aligned with regulatory thresholds as in the case of biotoxins limits stated by European Union regulation; iii) flexibility to incorporate prior knowledge and accommodate complex sampling designs as shown above.

Integrating this model into routine food safety surveillance and risk assessment enables authorities to make more transparent, evidence-based, and scientifically sound decisions.

Key advantages for food safety inference include:

- improved risk communication: Bayesian credible intervals are easily interpretable. There is a 95% probability that the true prevalence lies within the interval, given the data and prior assumptions;
- flexibility: the Bayesian β-binomial model could be extended to a hierarchical framework that accommodates variable sample sizes and clustering structures, such as by batch or producer, and uncertainties in prior information, which can be explored through sensitivity analyses;
- enhanced decision support: posterior probabilities and credible intervals directly inform regulatory and risk management decisions;
- computational accessibility: β-binomial models can be efficiently implemented in R (*e.g.*, with the bayesrules package).

Overall, while Bayesian binomial models are well established in food safety risk assessment, their implications for sampling plan inference under rare or zero contamination events remain underexplored. This paper aims to fill that gap.

References

Accoroni S, Cangini M, Angeletti R, Losasso C, Bacchiocchi S, Costa A, Di Taranto A, Escalera L, Fedrizzi G, Garzia A, Longo F, Macaluso A, Melchiorre N, Milandri A, Milandri S, Montresor M, Neri F, Piersanti A, Rubini S, Suraci C, Susini F, Vadrucci MR, Mudadu AG, Vivaldi B, Soro B, Totti C, Zingone A, 2024. Marine phycotoxin levels in shellfish—14 years of data gathered along the Italian coast. Harmful Algae 131:102560.

Barnouin G, El Hourany R, Record NR, Bajc Z, Kennedy J, Lafitte A, Dechraoui Bottein MY, 2025. Marine biotoxins in Mediterranean seafood: environmental occurrence and food safety concerns. Ecotox Environ Safety 303:119032.

Benschop J, Spencer S, Alban L, Stevenson M, French N, 2010. Bayesian zero-inflated predictive modelling of herd-level Salmonella prevalence for risk-based surveillance. Zoonoses Pub Health 57:60-70.

EFSA, 2012. Scientific Opinion on the development of a risk ranking framework on biological hazards. EFSA J 10:2724.

European Parliament, 2004. Regulation (EC) No 853/2004 of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for food of animal origin. In: Official Journal, L 139/55, 30/04/2004.

Fernández R, Mamán L, Jaén D, Fernández Fuentes L, Ocaña MA, Gordillo MM, 2019. Dinophysis species and diarrhetic shellfish toxins: 20 years of monitoring program in Andalusia, south of Spain. Toxins 11:189.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB, 2013. Bayesian data analysis (3rd ed.). CRC Press, Boca Raton, FL, USA.

Giunta della Regione Marche, 2024. Delibera n. 857 del 04/06/2024, Riclassificazione sanitaria delle zone di produzione/stabulazione dei molluschi bivalvi vivi ai sensi del Regolamento UE del 15 marzo 2019, n. 627; revoca della Deliberazione di Giunta del 20 giugno 2022, n. 773. BUR Marche LV(53), 10644.

Johnson AA, Ott MQ, Dogucu M, 2022. Bayes rules! An introduction to applied Bayesian modeling. (1st ed.) Chapman & Hall, Boca Raton, FL, USA.

Khan MA, Bruce BB, Bottichio L, Wise M, 2023. A Bayesian method for exposure prevalence comparison during foodborne disease outbreak investigations. Foodborne Pathog Dis 20:414-8.

Poletti R, 2014. Biotossine marine: origine, diffusione e controllo. Centro Ricerche Marine di Cesenatico. Available from: https://cipesalute.org/cedo/allegati/4500-BiotossineMarine.pdf.

Ranta J, Lindqvist R, Hansson I, Tuominen P, Nauta M, 2015. A Bayesian approach to the evaluation of risk-based microbiological criteria for Campylobacter in broiler meat. Ann Appl Statistics 9:1415-32.

Spiegelhalter DJ, Abrams KR, Myles JP, 2004. Bayesian approaches to clinical trials and health-care evaluation. John Wiley & Sons, Ltd, Hoboken, NJ, USA.

Van der Voet H, Slob W, 2007. Integration of probabilistic exposure assessment and probabilistic hazard characterization. Risk Anal 27:351-71.

Williams MS, Ebel ED, Hoeting JA, 2011. Bayesian analysis for food-safety risk assessment: evaluation of dose-response functions within WinBUGS. J Stat Software 43:1-14.

Online supplementary material
Supplementary Table 1. Samples classified by year and sampling station.

Received: 15 October 2025; Accepted: 13 January 2026; Early view: 3 February 2026.

Contributions: Cesare Ciccarelli: study concept, statistical analysis and writing - original draft. Angela Marisa Semeraro, Vittoria Di Trani, Guglielmo D’Aurizio, Francesca Barchiesi, Sara Gentili: writing – review and editing. Melina Leinoudi: methodology, writing – review and editing. Elena Ciccarelli: study concept, writing – review and editing.

Conflict of interest: the authors declare no potential conflict of interest.

Availability of data and materials: the dataset used during the current study is available upon reasonable request from the corresponding author.

Conference presentation: XXXIV National Conference of the Italian Association of Veterinary Food Hygienists (AIVI) | Parma, 10-12 September 2025.