doi: 10.4081/ijfs.2026.15116

SUPPLEMENTARY MATERIAL

Integrating uncertainty into official control decisions: Bayesian risk-based decision support for the safe reopening of live bivalve mollusk harvesting areas

Cesare Ciccarelli,[1] Angela Marisa Semeraro,[1] Vittoria Di Trani,[1]
Michela Conquista,[1] Elena Ciccarelli[2]

[1]Local Public Health Service - Ascoli Piceno, San Benedetto del Tronto, Italy; [2]Reading, UK

Correspondence: Cesare Ciccarelli, Local Public Health Service - Ascoli Piceno, via L. Manara n. 3/7 - 63074 San Benedetto del Tronto, Italy.
Tel.: (+39) 0735 7937474. E-mail: cesare.ciccarelli@sanita.marche.it

Key words: Bayesian decision model, shellfish safety, harvest area reopening, probabilistic risk assessment, pathogen monitoring.

Supplementary Table 1. Mathematical foundations of Bayesian framework.

Let $Y_i \in \{0,1\}, \quad i = 1, \dots, n$ denote the outcome of the i-th analytical sample collected during the post-closure health survey, where:

- $Y_i = 1$: non-compliant or hazard detected
- $Y_i = 0$: compliant result

Conditional on the underlying contamination probability θ, the sampling process is modelled as:

$$Y_i|\theta \sim \text{Bernoulli}(\theta)$$

Assuming conditional independence between samples, the total number of non-compliant samples

$$X = \sum_{i=1}^{n} Y_i \qquad \text{Eq. 1}$$

follows a binomial distribution: $X|\theta \sim \text{Binomial}(n, \theta)$. This likelihood represents routine microbiological monitoring data (e.g. *E. coli*, Salmonella, or other relevant indicators).

Prior uncertainty about θ is represented using a Beta distribution: $\theta \sim \text{Beta}(\alpha_0, \beta_0)$

where:

- α_0 : prior pseudo-count of non-compliant observations
- β_0 : prior pseudo-count of compliant observations

The hyperparameters (α_0, β_0) are derived from multiple evidence sources:

- historical monitoring data from the same harvesting area,
- epidemiological records of previous contamination events,
- Sanitary Survey findings,
- peer-reviewed knowledge on pathogen persistence and accumulation,
- precautionary assumptions reflecting regulatory practice.

If historical data consist of x_h non-compliant results out of n_h samples, a conjugate prior may be defined as:

$$\alpha_0 = x_h + a, \qquad \beta_0 = n_h - x_h + b \qquad \text{Eq. 2}$$

where *a, b* represent baseline precautionary parameters ensuring conservative inference when data are limited.

For example, consider a harvesting area with:

- 4 historical exceedances;
- 70 compliant samples;
- no recent outbreaks;
- resolved sewage overflow event;
- moderate environmental vulnerability.

Historical monitoring therefore yielded: $x_h=4 \quad n_h=74$

To incorporate precautionary uncertainty, conservative pseudo-counts: $a=0, b=2$ were adopted.

Based on Eq. 2, the resulting prior became: $\theta \sim Beta(4,72)$ with prior mean: $E(\theta) = 0.053$

reflecting a low but non-negligible expected contamination probability consistent with the environmental history of the area.

Given observed monitoring results $X = x$, Bayesian updating yields the posterior distribution:

$$\theta | x \sim Beta(\alpha_0 + x,\ \beta_0 + n - x) \qquad \text{Eq. 3}$$

This distribution represents updated knowledge about the true non-compliance probability in the harvesting area after considering all available evidence.

The posterior mean and variance are:

$$E[\theta | x] = \frac{\alpha_0 + x}{\alpha_0 + \beta_0 + n} \qquad \text{Eq. 4}$$

$$\mathrm{Var}[\theta | x] = \frac{(\alpha_0 + x)(\beta_0 + n - x)}{(\alpha_0 + \beta_0 + n)^2(\alpha_0 + \beta_0 + n + 1)} \qquad \text{Eq. 5}$$

Credible intervals for θ are obtained directly from the Beta posterior distribution.

When multiple harvesting areas or time periods are analysed, a hierarchical structure is introduced.

Let θ_j denote the non-compliance probability for area j. Then:

$$X_j | \theta_j \sim \mathrm{Binomial}(n_j\ , \theta_j) \qquad \text{Eq. 6}$$

$$\theta_j | \alpha, \beta \sim Beta(\alpha\ , \beta) \qquad \text{Eq. 7}$$

$$\alpha, \beta \sim p(\alpha, \beta) \qquad \text{Eq. 8}$$

This structure allows pooling information across areas while preserving area-specific variability.

To operationalise the regulatory concept of residual consumer risk, a tolerable contamination threshold is defined: θ^*, representing the maximum acceptable non-compliance probability compatible with consumer protection. The key risk metric is:

$$P_{risk} = P(\theta > \theta^* | x) \qquad \text{Eq. 9}$$

For the Beta posterior:

$$P_{risk} = 1 - F_{\mathrm{Beta}}(\theta^*;\ \alpha_0 + x,\ \beta_0 + n - x\ ) \qquad \text{Eq. 10}$$

where F_{Beta} is the cumulative distribution function.

This probability represents the likelihood that LBM still pose an additional risk to human health at the time of re-opening.

Uncertainty arising from limited sample size, measurement error, prior specification and environmental variability is explicitly represented through posterior distributions.

The posterior probability that the harvesting area meets microbiological safety expectations is:

$$P_{safe} = P(\theta \leq \theta^* | x) = 1 - P_{risk} \qquad \text{Eq. 11}$$

This metric supports regulatory interpretation of whether the precept "*does not present any other risk to human health*" is satisfied.

To better reflect operational regulatory practice under conditions of uncertainty, the Bayesian framework was implemented using a precautionary tolerance level for residual risk τ, representing the maximum acceptable posterior probability of unacceptable contamination, based on a three-zone precautionary decision rule rather than a strictly binary classification.

Three regulatory interpretation zones were defined as:

- Low residual risk zone: $P_{risk} \leq \tau_1$ supporting re-opening;
- Intermediate uncertainty zone: $\tau_1 < P_{risk} < \tau_2$ supporting additional investigative sampling and environmental assessment;
- High residual risk zone: $P_{risk} \geq \tau_2$ supporting continued closure or re-closure.

In the numerical simulations presented in this study $\tau_1 = 0.05$ and $\tau_2 = 0.20$ were selected to reflect precautionary regulatory interpretation.

The intermediate uncertainty zone acknowledges that reopening decisions frequently occur under incomplete information and that additional investigative samples may reduce posterior uncertainty before definitive regulatory action is taken.

Accordingly, the Bayesian decision rule was operationalised as follows:

- Re-open if: $P(\theta > \theta^*|x) \leq 0.05$
- Request additional investigative sampling if: $0.05 < P(\theta > \theta^*|x) < 0.20$
- Maintain closure if: $P(\theta > \theta^*|x) \geq 0.20$

This three-zone structure was considered more consistent with practical official-control decision-making than a purely binary rule.

The precautionary tolerance threshold τ can be derived from a loss function with asymmetric consequences – that is, false re-opening (L_{FP} = public health risk) and unnecessary continued closure (L_{FN}= economic loss), with $L_{FP} \gg L_{FN}$, .

The precautionary tolerance threshold τ can be interpreted as a function of asymmetric regulatory losses.

Let:

- L_{FP}: loss associated with false re-opening;
- L_{FN}: loss associated with unnecessary continued closure.

Following Bayesian decision theory: $\tau = \frac{L_{FN}}{L_{FP}+L_{FN}}$

For example:

- if public health consequences are considered ten times more severe than economic losses:

$$L_{FP}=10,\ L_{FN}=1 \quad \text{then: } \tau = \frac{1}{11} = 0.91$$

- Conversely, under a more precautionary interpretation:

$$L_{FP}=20,\ L_{FN}=1 \quad \text{yields : } \tau = \frac{1}{21} = 0.048$$

This value is approximately consistent with the precautionary tolerance threshold $\tau_1 = 0.05$ adopted in the present study.

Sensitivity analyses are conducted by varying: $(\alpha_0, \beta_0), \theta^*, \tau$ and evaluating their influence on $P(\theta > \theta^*|x)$, whereas *Robustness* of decisions is assessed by examining whether conclusions remain stable across plausible parameter ranges.

Supplementary Table 2. Monte Carlo simulation framework and reproducible R code.

Bayesian LBM Re-opening Decision Model

A Monte Carlo simulation approach was used to propagate uncertainty in contamination probability and to estimate regulatory decision metrics based on posterior distributions. This approach is particularly advantageous when closed-form posterior solutions are unavailable, as in models incorporating non-conjugate priors, hierarchical structures or latent variables.

Let:

- θ = probability that a harvested batch exceeds acceptable contamination threshold
- Prior: $\theta \sim \text{Beta}(\alpha_0, \beta_0)$
- Monitoring data: $x|\theta \sim \text{Binomial}(n, \theta)$
- Posterior: $\theta|data \sim \text{Beta}(\alpha_0 + x, \beta_0 + n - x)$

Decision parameters:

- Unacceptable contamination threshold: (θ^*= 0.10)
- Precautionary tolerance: ($\tau = 0.05$)

Key decision probability: $P(\theta > \theta^*|xdata)$

Decision rule:

Condition	Outcome
$P(\theta > \theta^*\|xdata)$ < 0.05)	Re-open
0.05–0.20	Request more data
>0.20	Maintain closure

Monte Carlo Simulation Design:

- Monte Carlo draws: (N = 100,000) (draws from the posterior distribution)
- For each scenario:
  - Draw samples from the posterior Beta distribution.
  - Estimate:
    - Posterior mean
    - 95% credible interval (CI)
    - Probability of exceeding threshold θ^*
  - Map these metrics to a regulatory decision.

Monte Carlo simulation is used in this model because:

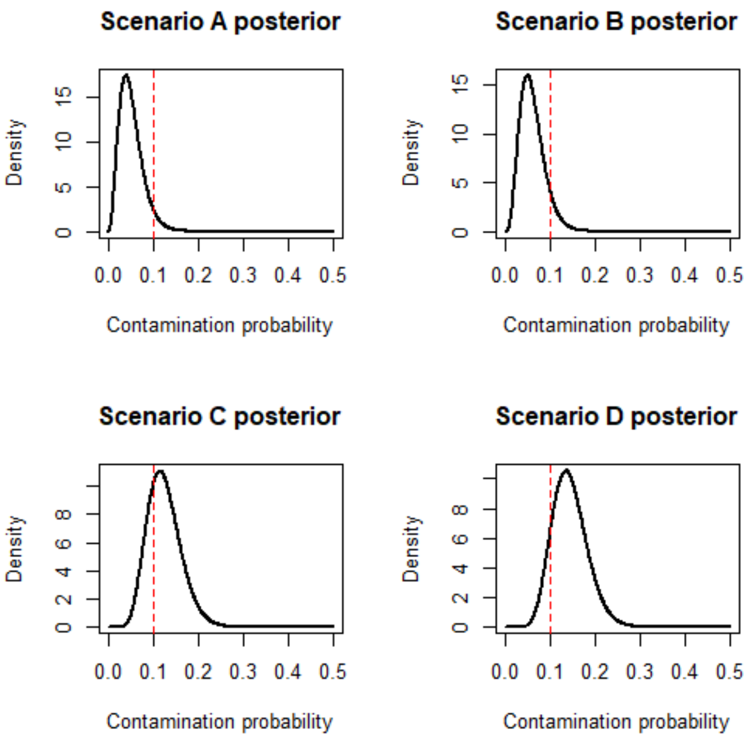
1. The exact analytical solutions exist for simple conjugate models like β-Binomial, but in practice, more complex models with non-conjugate priors, hierarchies, or unobserved latent variables might require approximation methods like Monte Carlo.
2. The Monte Carlo simulation allows flexibility in sampling from posterior distributions even in the absence of closed-form solutions.
3. The uncertainty propagation is handled by simulating multiple possible posterior outcomes to quantify the variability of decisions.

Scenario Definitions:

Scenario	α_0	β_0	n	x	Context
A	4	72	1	0	Favourable monitoring
B	4	72	5	1	Limited positives
C	10	60	10	0	Pollution event
D	12	56	15	0	Environmental pressure

Monte Carlo Simulation Results:

Scenario	Posterior α	Posterior β	Posterior Mean	95% CI Lower	95% CI Upper	P(θ > 0.10)	Decision
A	4	73	0.0520	0.0146	0.1109	0.0462	Re-open
B	5	76	0.0618	0.0206	0.1233	0.0878	Request more data
C	10	70	0.1249	0.0625	0.2055	0.7346	Remain closed
D	12	71	0.1447	0.0781	0.2274	0.8861	Remain closed



Conclusions:

The Monte Carlo simulation confirmed that posterior risk estimates were highly sensitive to both prior assumptions and observed exceedances. Under favourable monitoring conditions (Scenario A), even after a single favourable result, the posterior probability of unacceptable contamination remained below the precautionary tolerance (4.6%), supporting re-opening.

In contrast, scenarios incorporating even limited exceedances or adverse environmental priors resulted in posterior probabilities exceeding regulatory tolerance thresholds, leading to continued closure or requests for additional data.

High-risk scenarios following pollution events yielded posterior probabilities of unacceptable contamination exceeding 88%, providing strong quantitative justification for maintaining closure. These findings demonstrate the ability of the Bayesian framework to translate heterogeneous evidence into explicit, probabilistically justified regulatory outcomes.

Reproducible R code. The following script reproduces all simulations and results presented above

```
################################################################
# Bayesian LBM Re-opening Decision Model
# Monte Carlo simulation for posterior analysis
################################################################
```

```r
set.seed(84735)

# Monte Carlo size
N <- 100000

# Decision parameters
theta_star <- 0.10
tau <- 0.05

# Function to run scenario
run_scenario <- function(alpha0, beta0, n, x, name){

  # Posterior parameters
  alpha_post <- alpha0 + x
  beta_post  <- beta0 + n - x

  # Monte Carlo draws
  theta <- rbeta(N, alpha_post, beta_post)

  # Metrics
  post_mean <- mean(theta)  # Posterior mean of sampled values
  ci_lower <- quantile(theta, 0.025)  # 2.5% quantile (lower bound of credible
interval)
  ci_upper <- quantile(theta, 0.975)  # 97.5% quantile (upper bound of
credible interval)
  prob_exceed <- mean(theta > theta_star)  # Probability of exceeding the
threshold

  # Decision rule
  decision <- if(prob_exceed < 0.05){
    "Re-open"
  } else if(prob_exceed < 0.20){
    "Request more data"
  } else{
    "Remain closed"
  }

  data.frame(
    Scenario = name,
    Prior_alpha = alpha0,
    Prior_beta = beta0,
    n = n,
    x = x,
    Post_alpha = alpha_post,
    Post_beta = beta_post,
    Posterior_mean = round(post_mean, 4),
    CrI_2.5 = round(ci_lower, 4),
    CrI_97.5 = round(ci_upper, 4),
    P_exceed = round(prob_exceed, 4),
    Decision = decision
  )
}

############################################################
# Run scenarios
############################################################

A <- run_scenario(4,72,1,0,"A favourable")
B <- run_scenario(4,72,5,1,"B low positives")
C <- run_scenario(10,60,10,0,"C pollution")
D <- run_scenario(12,56,15,0,"D environmental risk")

results <- rbind(A,B,C,D)
print(results)
```

```
############################################################
# Plot posterior distributions
############################################################

par(mfrow=c(2,2))

plot_density <- function(alpha, beta, title){
  x <- seq(0, 0.5, length=500)
  y <- dbeta(x, alpha, beta)
  plot(x, y, type="l", lwd=2,
       main=title,
       xlab="Contamination probability",
       ylab="Density")
  abline(v=theta_star, lty=2, col="red")
}

# Plot posterior distributions based on the updated alpha_post and beta_post
plot_density(4+0, 72+1, "Scenario A posterior")  # Posterior parameters for
Scenario A
plot_density(4+1, 72+4, "Scenario B posterior")  # Posterior parameters for
Scenario B
plot_density(10+0, 60+10, "Scenario C posterior")  # Posterior parameters for
Scenario C
plot_density(12+0, 56+15, "Scenario D posterior")  # Posterior parameters for
Scenario D
```