

Inferring leakage in imports of frozen seafood allowing for censoring, testing accuracy and a minimum positive prevalence

Sumonkanti Das¹, Robert G. Clark^{1,*}, Mahdi Parsa², Belinda Barnes²

¹Research School of Finance, Actuarial Studies and Statistics, The Australian National University, Canberra 2601, Australia

²Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra 2601, Australia

*Corresponding author: Robert G. Clark, Research School of Finance, Actuarial Studies and Statistics, The Australian National University, Canberra 2601, Australia (robert.clark@anu.edu.au).

ABSTRACT

Many countries screen import consignments to guard against the entry of exotic pests, contaminants, and pathogens. A widely used strategy is to pool individual units into groups and test each group for presence or absence of contamination. Consignments are typically rejected if there are any detections. Screening samples are commonly designed to give a high chance of detection assuming a design prevalence. What is less common, however, is to analyze the history of testing outcomes to infer how many accepted consignments contain contaminated units, and the number of contaminated units—jointly referred to as “leakage.” We build on existing censored beta-binomial models to answer these questions for the importation of frozen seafood into Australia, allowing for unknown test sensitivity and specificity. We present new theory and empirical results demonstrating that specificity is identifiable from test data in this context, but sensitivity is not. We also develop a new class of models in which consignment propensity is either zero or above a minimum positive prevalence threshold, motivated by our case study but applicable more widely. Past testing data are modeled under multiple scenarios using both hierarchical Bayes and maximum likelihood methods, revealing new insights into the risk of leakage.

Keywords: acceptance sampling, beta-binomial distribution, biosecurity, group testing, multi-level models, rare events

1 Introduction

Group Testing of Import Consignments

Flows of goods and people across national borders come with the inevitable risk that contamination, disease, or infection are introduced into a country (Trouvé et al., 2024). Border screening is commonly applied to slow or prevent incursions and to verify the fidelity and suitability of pre-border risk mitigation. A widespread class of designs for the importation of agricultural and other goods involves inspecting or testing a sample of units from each consignment or shipment and then disallowing the whole consignment if any contaminant is detected, or if the number of detections is above some threshold. To save costs and enable a larger sample to be tested, it is common to group units and then to test each group, with a presence or absence result. Examples include seed lots imported into Australia (Constable et al., 2018), a variety of imports into the United States (Ma et al., 2025), and

tubers into the European Union (EU, 1993, §3.1). We refer to contamination for simplicity, noting that disease, infection or other binary outcome may be substituted as applicable.

Groups are usually assumed to be non-overlapping simple random samples of units selected from the consignment. This relies on sound operational practices avoiding possible clustering of contamination within groups, although this may not always be fully achievable, particularly when the grouping relates to operational units such as bags or boxes, as in our case study. Sensitivity analysis can be conducted to understand the impact of clustering, but its extent cannot be directly estimated from typically available data (Clark et al., 2023).

The number of units in each group is usually fixed by operational or test performance concerns. The number of groups, or total units tested, is often chosen so that the probability of detecting contamination in the sample reaches a high value (eg, 95%) at an assumed design prevalence for a notional typ-

Received: 25 December 2025. Revised: 15 May 2026. Accepted: 6 August 2026

© The Author(s) 2026. Published by Oxford University Press on behalf of The International Biometric Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

ical contaminated consignment (eg, §5 of IPPC, 2008; Trouvé et al., 2024). Testing costs, and importer costs when tests are destructive, may also influence the choice of the sample and group sizes.

If a consignment contains contamination, but it is not detected due to test inaccuracy or because only a sample of units is grouped and tested, then contaminated units will enter the country. This is referred to as leakage, and we will refer to the leakage rate of consignments (proportion of contaminated consignments that evade detection) and of units (proportion of units that are both contaminated and cross the border). The centrality of the concept of leakage is one of the features that distinguish group testing of imports from the modeling of rare events in general. Leakage is also relevant in public health screening, where a subpopulation or locality may be deemed infection-free based on testing a sample of individuals, and in audit sampling, where a balance sheet or business may be deemed compliant based on auditing a sample of records, transactions, or items.

Multi-level Models for Import Screening

The assumption of a design prevalence to underpin the sampling strategy naturally prompts the questions: what are typical prevalence levels in consignments, and how might these affect an assessment of risk? Can the distribution of consignment-level prevalence be inferred, even though estimates of prevalence for individual consignments may be highly uncertain? What are the implications for leakage rates, both at the consignment level and the unit level? Can historical testing results shed light on these questions? These considerations motivate modeling a distribution of consignment-level propensities, leading to a multilevel framework that captures both between-consignment variability and within-consignment sampling error. The relevance of such models is further supported by the observed distribution of test counts across consignments, which frequently exhibit zero inflation and overdispersion (eg, Clark et al., 2023; Trouvé and Robinson, 2024), including in the case study analyzed in this paper. Analysis is further complicated by censoring—only a positive or negative result is available for each group—and by test sensitivity and specificity.

If consignment-level propensity is assumed to follow a beta distribution, and individual units are sampled and tested, positive test counts follow a beta-binomial distribution. Applications of beta-binomial distributions in biosecurity screening include IPPC (2008) and Trouvé et al. (2022). When units are tested in groups, we instead observe censored beta-binomial observations, for which the likelihood can be readily derived. See for example Cowling et al. (1999) who also allowed for testing errors, although the beta distribution was used as a prior rather than a latent distribution with shape parameters requiring estimation. Clark et al. (2023) derived distributions for the leakage of units, and the probability of leakage, under group testing with latent beta propensities, fitting these models from historical testing results for cucurbit seed imports. They also investigated non-random group formation using sensitivity analysis due to the non-identifiability of within-group clustering. An advantage of beta models is that exact or approximate analytical results are often tractable for simple settings (see Trouvé and Robin-

son, 2021 for a non-beta multilevel model in a related context).

Group testing from a single population without the first level model for propensity has also been investigated (eg, Hepworth, 2019 and references therein).

Outline and Contributions of This Paper

We consider the importation of frozen prawns into Australia. Around 800 consignments are imported annually with around 300 000 prawns per consignment, so this represents a significant import pathway with concomitant biosecurity risks. Prawns are tested in 13 groups of 5 from each consignment, which is intended to achieve a 95% chance of obtaining at least one positive test result when infection comprises 5% of a consignment (DAFF, 2023, §17.2.3). We analyze historical test outcomes for whitespot syndrome virus (WSSV) by fitting multilevel models with latent beta variables.

Motivated by the case study, we develop several major extensions to existing models for multilevel group testing data. In particular, we address the statistical problem of estimating rare-event prevalence and leakage using censored, grouped observations, where test sensitivity or specificity may be unknown. These contributions include:

- (i) New findings regarding the identifiability of model parameters in multilevel group testing with potentially unknown sensitivity and specificity, based on approximations assuming rare events. Our focus is on group testing, but the results apply even to groups of size one, ie, to individual item testing.
- (ii) A minimum positive prevalence framework, accounting for the biological implausibility of very small positive prevalences. This approach has not been considered in previous literature.
- (iii) Analytical and approximate formulas for leakage at both the consignment and unit levels. These are new as existing formulas either do not allow for imperfect testing (eg, Clark et al., 2023) or consignment-level heterogeneity.

Section 2 introduces notation and a two-level model with beta-distributed consignment propensities, extending hierarchical beta-binomial frameworks similar to those used by IPPC (2008), Trouvé et al. (2022), and Clark et al. (2023). In particular, we extend Clark et al. (2023) by incorporating imperfect testing (sensitivity and specificity) and a minimum positive prevalence threshold. The latter reflects biological considerations, where the latent propensity is zero with some probability or, if contamination is present, exceeds a positive lower bound due to (for example) the virulence of a virus. Analytical approximations to the likelihood suggest that specificity and a shape parameter controlling heterogeneity are identifiable, but sensitivity and mean prevalence are not, assuming that contamination is rare and heterogeneous.

In Section 3, we apply the new models to the frozen prawn data using maximum likelihood, Markov Chain Monte Carlo sampling, and an approximate Bayes method. We find that profile likelihoods are relatively flat with respect to sensitivity but strongly identify specificity near 1, consistent with our theoretical results in Section 2. The minimum positive prevalence models indicate that the threshold parameter is weakly

Table 1 Summary of notation.

Symbol	Definition
m	Number of items in each group
B	Number of groups in a consignment
b	Number of groups selected for testing from each consignment
N	Total number of units (eg, prawns) in a consignment, $N = mB$
n	Total number of units (eg, prawns) selected in each consignment, $n = mb$
$X_{ij\ell}$	Indicator for contamination of unit ℓ in group j of consignment i
X_{ij}	Number of contaminated units in group j of consignment i , values $0, 1, \dots, m$
Y_{ij}	Presence of contamination in group j of consignment i , equal to $I(X_{ij} > 0)$
$\tilde{Y}_{ij}$	Test result in group j of consignment i equals Y_{ij} except for measurement error
T_{Xi}	Total number of contaminated units in consignment i , $T_{Xi} = \sum_{j=1}^B X_{ij}$
t_{xi}	Number of contaminated units among sampled groups, $t_{xi} = \sum_{j=1}^b X_{ij}$
t_{yi}	Number of positive groups among sampled groups, $t_{yi} = \sum_{j=1}^b Y_{ij}$
$\tilde{t}_{yi}$	Number of positive tests among sampled groups, $\tilde{t}_{yi} = \sum_{j=1}^b \tilde{Y}_{ij}$
L_i	Leakage in consignment i , defined as $L_i = T_{Xi}I(\tilde{t}_{yi} = 0)$

identified. Together, the models enable insights into leakage risks for policy makers, including the effect on leakage when varying the number of groups tested.

Section 4 presents conclusions and future directions. The proposed models and methods are implemented in an R package and are being used to inform decisions by the Australian Department of Agriculture, Fisheries, and Forestry. We anticipate that similar analyses of historical test data will become increasingly common in import biosecurity practice globally.

2 Censored Beta-binomial Models with Test Sensitivity and Specificity

2.1 Notation and measurement error setup

Each consignment is assumed to contain N units (in our case, prawns). We assume that the N units are randomly divided into B groups each containing $m = N/B$ units. Of these B groups, b are selected for testing. We will index groups from $1, \dots, B$, with groups $1, \dots, b$ being the sampled groups and groups $b+1, \dots, B$ being the remainder. The non-sampled groups may not actually be constructed in practice—it is not unusual for b groups of equal size to be extracted and tested without any subdivision made in the remainder. However, we still define notional non-sampled groups for convenience. Consignments may be of unequal size in practice, although we ignore this for simplicity, noting that generalizing to varying sized consignments would be straightforward in principle.

We define $X_{ij\ell}$ as indicating the presence (1) or absence (0) of contamination in a unit ℓ of group j in consignment i , where $j = 1, \dots, b$ (group) and $\ell = 1, \dots, m$ (unit within group). We write X_{ij} for the number of contaminated units in group j of consignment i , $j = 1, \dots, B$, and so X_{ij} may take on values $0, 1, \dots, m$.

We do not directly observe the number of contaminated units X_{ij} even for sampled groups. Instead, all that is available is the binary test result for each selected group. We define $Y_{ij} = I(X_{ij} > 0)$ which equals 1 if any units are contaminated in the group and 0 otherwise. We write $T_{Xi} = \sum_{j=1}^B X_{ij}$

and $t_{xi} = \sum_{j=1}^b X_{ij}$ for the population and sample number of contaminated units, respectively, in consignment i . We let $t_{yi} = \sum_{j=1}^b Y_{ij}$, which is the number of sampled groups containing contaminated units from consignment i , with domain $0, 1, \dots, b$.

Tests for pathogens or contaminants are usually imperfect. We assume constant sensitivity Δ , the probability of a positive test given infection in a group, and specificity Λ , the probability of a negative test given no infection. We let $\tilde{Y}_{ij}$ be the test result for group j in consignment i , with 1 denoting a positive result (apparent contamination) and 0 denoting a negative result (apparently contamination-free). It follows that $\mathbb{P}(\tilde{Y}_{ij} = 1 | Y_{ij}) = \Delta Y_{ij} + (1 - \Lambda)(1 - Y_{ij})$. The available data are the values of $\tilde{t}_{yi} = \sum_{j=1}^b \tilde{Y}_{ij}$.

We define the *leakage* to equal $L_i = T_{Xi}I(\tilde{t}_{yi} = 0)$, which is the number of contaminated units (prawns) imported in consignment i , reflecting the fact that consignments are blocked if there are any positive test results. The expected leakage $\mathbb{E}(L_i)$ and probability of leakage $\mathbb{P}(L_i > 0)$ are of particular interest. Table 1 summarizes our notation.

2.2 Assumed two-level beta model

We consider the following model:

$$X_{ij}|p_i \stackrel{\text{i.i.d.}}{\sim} \text{Bin}(m, p_i); \text{ and} \quad (1)$$

$$p_i \stackrel{\text{i.i.d.}}{\sim} \text{Beta}(\alpha, \beta) \quad (2)$$

for consignments i and groups j , where α and β are shape parameters, and p_i is a latent consignment propensity. Different consignments are assumed to be independent. See, for example, IPPC (2008), Trouvé et al. (2022) and Clark et al. (2023) for related models. The model implies that $T_{Xi} \sim \text{BB}(N, \alpha, \beta)$, where BB denotes the beta-binomial distribution. We write $\mu = \alpha / (\alpha + \beta)$ so that $\mathbb{E}(T_{Xi}/N) = \mathbb{E}(p_i) = \mu$. We extend (1)–(2) to incorporate test sensitivity and specificity. Firstly note that $P(Y_{ij} = 1 | p_i) = \phi_i$ where $\phi_i = 1 - (1 - p_i)^m$. The distributions of the test results with error directly follow:

$$\tilde{Y}_{ij}|p_i \stackrel{\text{i.i.d.}}{\sim} \text{Ber}(\tilde{\phi}_i) \quad (3)$$

$$\tilde{t}_{yi}|p_i \stackrel{\text{indep}}{\sim} \text{Bin}(b, \tilde{\phi}_i) \quad (4)$$

where $\tilde{\phi}_i = \Delta\phi_i + (1 - \Lambda)(1 - \phi_i) = (\Delta + \Lambda - 1)\phi_i + 1 - \Lambda$ can be expressed as

$$\tilde{\phi}_i = (\Delta + \Lambda - 1)[1 - (1 - p_i)^m] + 1 - \Lambda \quad (5)$$

and Ber refers to the Bernoulli distribution. Note that $\tilde{\phi}_i = \tilde{\phi}(p_i)$ is a function of p_i .

The marginal probability of t_{yi} immediately follows by integrating out p_i :

$$\mathbb{P}(\tilde{t}_{yi}) = \int \left(\frac{b}{\tilde{t}_{yi}} \right) \tilde{\phi}(p_i)^{\tilde{t}_{yi}} [1 - \tilde{\phi}(p_i)]^{b - \tilde{t}_{yi}} p_i^{\alpha - 1} (1 - p_i)^{\beta - 1} dp_i / B(\alpha, \beta) \quad (6)$$

where B is the beta function. If observations of $\tilde{t}_{yi}$ for a sample of consignments are available, then the likelihood is given by $\prod_i \mathbb{P}(\tilde{t}_{yi})$. Numerical integration is required—for details see [Web Appendix A](#).

2.3 Approximation to the model and identifiability of sensitivity and specificity

We will show empirically in Section 3 that when model (6) is fitted to our case study with α , β , Λ , and Δ all treated as unknown parameters, the first shape parameter α and the specificity Λ are reasonably well identified, but the second shape parameter and the sensitivity Δ are not, with the log-likelihood almost flat with respect to β and Δ . To understand this result and whether it applies more widely, we now develop an approximation to the distribution of $\tilde{t}_{yi}$ in (6). The approximate model will be useful to explain identifiability of model parameters, and will also provide a less computationally intensive means of fitting the model.

We assume that β is large so contamination is rare. Under the model defined in Section 2.2, $\tilde{t}_{yi}|p_i \sim \text{Bin}(b, \tilde{\phi}_i)$ and we then assume that a Poisson approximation applies with

$$\tilde{t}_{yi}|p_i \stackrel{a}{\sim} \text{Poisson}(b\tilde{\phi}_i). \quad (7)$$

This holds in distribution limit when $\tilde{\phi}_i \rightarrow 0$ with $b\tilde{\phi}_i$ tending to a constant. Although $b = 13$ in our case study, Hodges and Le Cam (1960) demonstrate that the maximum absolute error in the Poisson approximation to the binomial cumulative distribution function can be acceptable even when the binomial population size is small, provided that the probability parameter is close enough to zero. Further calculations to check this approximation are detailed in [Web Appendix B](#).

Using a well known property of beta random variables, (2) means that as $\beta \rightarrow \infty$, βp_i tends in distribution to $\text{Gamma}(\alpha, \alpha)$, using the shape-mean parameterisation of the Gamma family. Thus we can approximate the distribution of p_i as $\text{Gamma}(\alpha, \alpha/\beta)$. We use the asymptotically equivalent approximation

$$p_i \stackrel{a}{\sim} \text{Gamma}(\alpha, \alpha(\alpha + \beta)^{-1}) \quad (8)$$

so that the first moment of the approximate distribution is exact.

Now, large β implies that p_i are close to zero with high probability. A first order Taylor approximation of ϕ_i in (5) with respect to p_i gives $\tilde{\phi}_i \approx (\Delta - 1 + \Lambda)mp_i + 1 - \Lambda$. Using

the properties of the Gamma distribution, we can therefore conclude from (8) that

$$\tilde{\phi}_i \stackrel{a}{\sim} 1 - \Lambda + \text{Gamma}[\alpha, m\alpha(\alpha + \beta)^{-1}(\Delta - 1 + \Lambda)] \quad (9)$$

We now strike a fork in the road where $\Lambda = 1$ and $\Lambda < 1$ (perfect vs imperfect specificity) lead to very different approximate distributions of $\tilde{t}_{yi}$. We obtain from (7) and (9) that

$$\tilde{t}_{yi} \stackrel{a}{\sim} \text{Poisson}\{b[1 - \Lambda + \text{Gamma}(\alpha, m\alpha(\alpha + \beta)^{-1}(\Delta - 1 + \Lambda))]\} \quad (10)$$

For $\Lambda = 1$ (perfect specificity), this simplifies to

$$\tilde{t}_{yi} \stackrel{a}{\sim} \text{NB}(\alpha, bm\alpha(\alpha + \beta)^{-1}\Delta) \quad (11)$$

using a shape-mean parameterisation for the negative binomial (NB) family, because Gamma-Poisson mixtures are NB. Approximation (11) is related to the result of Teerapabolarn (2015) that beta-binomial distributions tend to NB when the second shape parameter tends to infinity. Using the same result from Teerapabolarn (2015), we further approximate (11) by

$$\tilde{t}_{yi} \stackrel{a}{\sim} \text{BB}(b, \alpha, \beta m^{-1}\Delta^{-1}) \quad (12)$$

which is a more convenient model as it respects the constraint that $\tilde{t}_{yi} \leq b$.

By inspection of (10), it is clear that β and Δ are not separately identifiable, even if Λ is known. Approximation (12) confirms that this is true even if $\Lambda = 1$. In other words, it is not possible to distinguish between rarity of true contamination and imperfect test sensitivity.

In contrast, (10) suggests that the specificity, Λ , is identifiable when contamination is rare. In effect, $1 - \Lambda$ acts as a floor on the values of $\tilde{t}_{yi}$ (to be precise, on the conditional mean of $\tilde{t}_{yi}$). There is one caveat, however. In the limit when both α and β tend to infinity at the same rate, with $\mu = \alpha/(\alpha + \beta)$ held fixed, the Gamma distribution in (10) tends to a point mass at $m\mu(\Delta - 1 + \Lambda)$ so that (10) becomes

$$\tilde{t}_{yi} \stackrel{a}{\sim} \text{Poisson}\{b[1 - \Lambda + m\mu(\Delta - 1 + \Lambda)]\} \quad (13)$$

This limit corresponds to zero variance in p_i , so there is no inter-consignment heterogeneity. We can hardly identify anything. The quantity $[1 - \Lambda + m\mu(\Delta - 1 + \Lambda)]$ can be estimated but we cannot estimate Λ , Δ or μ , even if one of these parameters is assumed to be known.

We conclude that under the approximate model, provided that $\alpha < \infty$ (ie, some heterogeneity between consignments), the specificity and the first shape parameter are identifiable, but the sensitivity and the second shape parameter are not. A more detailed numerical evaluation in [Web Appendix D](#) suggests that for $b = 13$ and m equal to either 5 or 10, the approximation is tolerably accurate provided μ is 0.01 or less, and in some cases even for larger μ . Although this supplementary investigation focused on specific values of b and m , the theoretical results are not based on specific design settings, and can be expected to apply more generally where contamination is rare and heterogeneous.

2.4 Models incorporating a minimum positive prevalence

The virus of interest in the case study is highly infectious, so very small non-zero values are unlikely to be biologically

Table 2 Distribution of the number of positive detections per consignment out of 13.

Positive test results:	0	1	2	3	4	5	6	8	9	10
No. of consignments:	2815	9	10	6	1	3	2	1	2	1
% of consignments:	98.77	0.32	0.35	0.21	0.04	0.11	0.07	0.04	0.07	0.04

possible—if the virus is present at all, then it is likely to spread beyond a minimum positive value. This conflicts with the latent beta model developed in Section 2.2, which can have high density of p_i for values close to zero. Motivated by this discrepancy, we develop models which have positive probability attached to a zero value for p_i , and then a continuous distribution for $k \leq p_i \leq 1$ where $k \geq 0$ is a threshold called the minimum positive prevalence. We define the new model by:

$$p_i^* = p_i \cdot I(p_i \geq k), \quad \text{with } p_i \sim \text{Beta}(\alpha, \beta), \quad (14)$$

where $I(\cdot)$ is the indicator function. Correspondingly, the group-level contamination probability, under imperfect testing with sensitivity Δ and specificity Λ , becomes:

$$\tilde{\phi}_i^* = \Delta \phi_i^* + (1 - \Lambda)(1 - \phi_i^*), \quad (15)$$

where $\phi_i^* = 1 - (1 - p_i^*)^m$. The observed number of positive group tests, $\tilde{t}_{yi}$, then follows a Binomial distribution conditional on p_i with $\tilde{t}_{yi}|p_i \sim \text{Bin}(b, \tilde{\phi}_i^*)$. Numerical integration is required to calculate the resulting likelihood.

2.5 Leakage

The leakage rate is a key metric of risk for an import pathway. We define the leakage for consignment i as $L_i = T_{Xi}I(\tilde{t}_{yi} = 0)$, where $\tilde{t}_{yi}$ is subject to testing error (Section 2.2). A minimum prevalence k may also apply (Section 2.4). Of interest will be the distribution of L_i , $\mathbb{P}(L_i > 0)$ (the proportion of consignments that are contaminated and evade detection), and $\mathbb{E}(L_i)$ (the expected number of infected units entering the country per consignment).

In general, $\mathbb{P}(L_i > 0)$ and $\mathbb{E}(L_i)$ can be obtained by simulating values of T_{Xi} and $\tilde{t}_{yi}$ for given parameter values of α , β , Δ , Λ and k . However, when testing is perfect and a minimum positive prevalence k applies, exact expressions for $\mathbb{P}(L_i > 0)$ and $\mathbb{E}(L_i)$ can be derived. We let $B_{(a,b)}(u, v) = \int_a^b t^{u-1}(1-t)^{v-1}dt$ be the incomplete beta function. Then, when $\Delta = \Lambda = 1$,

$$\mathbb{P}[L_i > 0] = \frac{B_{(k,1)}(\alpha, \beta + mb)}{B(\alpha, \beta)} - \frac{B_{(k,1)}(\alpha, \beta + N)}{B(\alpha, \beta)} \quad (16)$$

$$\mathbb{E}(L_i) = (b - b)m \cdot \frac{B_{(k,1)}(\alpha + 1, \beta + bm)}{B(\alpha, \beta)}. \quad (17)$$

When testing is imperfect, exact closed forms are unavailable; however, the approximation (12) from Section 2.3 under perfect specificity ($\Lambda = 1$) yields approximate expressions for $\mathbb{P}(L_i > 0)$ and $\mathbb{E}(L_i)$, allowing for $\Delta \in [0, 1]$ and $k \in [0, 1]$. Then:

$$\mathbb{P}[L_i > 0] \approx \frac{B_{(k,1)}(\alpha, \beta m^{-1} \Delta^{-1} + b)}{B(\alpha, \beta m^{-1} \Delta^{-1})} - \frac{B_{(k,1)}(\alpha, \beta + N)}{B(\alpha, \beta)} \quad (18)$$

$$\mathbb{E}(L_i) \approx (b - b)m \cdot \frac{B_{(k,1)}(\alpha + 1, \beta + bm\Delta)}{B(\alpha, \beta)} \quad (19)$$

where $k_1 = \Delta(1 - (1 - k)^m)$ and the approximations are based on rare events (ie, β large). See [Web Appendix C](#) for details of the derivations of (16)–(19).

3 Case Study

3.1 Background and data

For consistency, we use the term consignments throughout this case study, although in practice some are subdivided. Each consignment consists (on average) of about 800 boxes, with each box containing 10 bags of approximately 40 prawns (units). For simplicity, we assume constant consignment sizes. A sample of 13 boxes is selected from each consignment and one bag is selected from each box. From each of the 13 selected bags, a group of 5 prawns is randomly taken for testing. Tissue from each group of 5 prawns is pooled and tested using polymerase chain reaction (PCR) methods for WSSV or yellowhead virus genotype 1 (YHV1) (no YHV1 has been detected, so all positives refer to WSSV). Each test yields a presence/absence result for the group. The PCR test is thought to have approximately 80% sensitivity and more than 99% specificity, although both are uncertain. The consignment outcome is the number of positive tests out of 13.

Group size was set at 5 to avoid loss of PCR sensitivity. The total sample of 65 per consignment was designed to give a 95% probability of at least one positive detection, assuming a design prevalence of 5% and 90% test sensitivity (here we use a more conservative 80%). It is also assumed that infected units are randomly assorted among bags. The whole consignment is blocked from importation if any of the 13 tests are positive. If infected prawns are present but no positives are detected, leakage occurs, ie, infected prawns enter Australia.

We will focus on testing results from 2850 consignments over an approximately 4-year period between 2015 and 2022. Table 2 shows the data as a frequency table. There are 0/13 positive test results for 98.8% of consignments and the remaining positive results range from 1 to 10, with 1 and 2 the most frequent. Thus, positive test results are extremely rare, and counts are very overdispersed, with quartiles equal to 1.5, 2.0, and 4.5.

The virus WSSV is highly infectious in farmed prawns, with prevalence when detected ranging from 48–90% in an Australian Import Risk Analysis (Biosecurity Australia, 2009). More recent prevalences ranged from 12–66% (DAFF, 2023 and references therein). Discussions with Australian biosecurity policy managers suggest that consignments would not normally be sourced from more than ten distinct sources, suggesting that, if infection is present at all in a consignment, it affects at least 1–2% of prawns.

Table 3 Posterior estimates with 95% confidence intervals (CI) of beta-binomial parameters (α , β , and μ) for t_{yi} under varying sensitivity (Δ) and specificity (Λ). The exact model is developed using frequentist maximum likelihood (ML) and Bayesian Metropolis-Hastings (MH) algorithms, while the approximate model is derived using the Bayesian Hamiltonian Monte Carlo (HMC) algorithm implemented in the `brms` R package. The posterior estimates of leakage parameters $E(L_i)$ (equation 17, 19) and $P(L_i > 0)$ (equations 16, 18) are obtained accordingly. CIs for ML are based on the profile log-likelihood, while MH and HMC use Bayesian credible intervals.

Parameter	Estimators (95% CI)			
	ML $\Delta = 0.8$	MH $\Delta = 0.8$	HMC $\Delta = 0.8$	MH $\Delta \sim U(.75, .85)$
$\alpha \times 10^3$	4.68 (2.99, 7.12)	4.63 (2.89, 6.96)	5.13 (3.29, 7.59)	4.58 (2.88, 6.80)
β	4.40 (2.30, 6.87)	4.31 (1.94, 7.78)	6.88 (3.83, 11.15)	4.14 (1.85, 7.34)
$\mu \times 10^3$	1.06 (0.64, 1.88)	1.16 (0.67, 1.92)	0.77 (0.50, 1.15)	1.19 (0.69, 2.00)
$E(L_i)$	26.2 (17.3, 34.1)	25.9 (16.7, 37.4)	27.4 (18.4, 39.1)	25.7 (16.6, 37.1)
$P(L_i > 0) \times 10^2$	3.75 (2.39, 4.78)	3.68 (2.23, 5.56)	4.15 (2.66, 6.09)	3.63 (2.22, 5.43)

It is noted that even if leakage occurs, multiple conditions must be met before it leads to a significant biosecurity event such as spread to wild aquatic populations. PCR can detect dead or non-viable virus. For infection to reach wild populations, contaminated prawns would typically need to be used as bait by recreational fishers, whereas the vast majority are instead prepared and consumed by humans. While leakage is the focus here, the probability of actual biosecurity breaches is much lower than the probability of leakage (DAFF, 2023).

3.2 Identifiability of specificity, sensitivity, and minimum positive prevalence

We assessed identifiability of the parameters α , β , Δ and Λ in the model from Section 2.2 using profile likelihoods. Details are provided in Web Appendix E. Specificity is well identified, with MLE on the boundary $\Lambda = 1$ and the log-likelihood decreasing precipitously as Λ decreases even slightly below 1. Sensitivity Δ was very poorly identified, motivating the theoretical approximations in Section 2.3, which confirm that for rare events with inter-consignment heterogeneity, specificity is identifiable but sensitivity is not. We therefore consider models with sensitivity fixed at $\Delta = 0.8$ or assigned a prior distribution $U(0.75, 0.85)$. Although specificity could be treated as unknown or assigned a diffuse prior, it is well identified from the data and is therefore fixed at $\Lambda = 1$, as it is extremely close to 1.

We suspected that the minimum positive prevalence (k) in the models from Section 2.4 would also be difficult to estimate, given the paucity of positive test results. The profile likelihood peaks at its MLE (≈ 0.01). Of particular interest is whether there is evidence that k is positive. We calculated a p-value of 0.084 for the null hypothesis $k = 0$ against the alternative $k \geq 0$. This provides moderate evidence that a (non-zero) minimum positive prevalence applies. Web Appendix E1 contains details, including an approximate 90% confidence interval for k which is indeed wide. Due to this uncertainty, we did not estimate k and instead conducted sensitivity analyses with k fixed at 0, 0.01, and 0.02.

3.3 Model fitting approaches

We employed three model fitting approaches. In the first, the approximate beta-binomial model (12) was fitted using the Hamiltonian Monte Carlo (HMC) method, as implemented in

the `brms` package (Bürkner, 2021). Note that fitting the approximate model results in posterior distributions for α and β^* rather than α and β where $\beta^* = \beta m^{-1} \Delta^{-1}$, so β can be estimated by multiplying β^* by $m\Delta$ (assuming for this purpose that $\Delta = 0.8$). The `brms` default priors, $\text{logit}(\mu) \sim t(3, 0, 2.5)$ and $\theta = \alpha + \beta^* \sim \text{Gamma}(0.01, 0.01)$ were used.

For the second approach, Bayesian posterior distributions were obtained by an implementation of the Metropolis-Hastings (MH) algorithm, using the exact likelihood calculated by numerical integration. Two different sets of priors were used, one with $\Delta = 0.8$ and the other with Δ uniformly distributed between 0.75 and 0.85.

The MH algorithm iterated on parameters θ_1 , θ_2 and (when Δ was not fixed at 0.8) θ_3 . We defined $\theta_1 = \log(\alpha)$ with prior $\theta_1 \sim U(-A, A)$ for large A . We defined $\mu = \text{expit}(\theta_2)$ and, when Δ was not fixed, we let $\Delta = 0.75 + 0.1 \text{expit}(\theta_3)$. The prior distributions of θ_2 and θ_3 were defined such that the priors of μ and Δ were $\mu \sim U(0, 1)$ and $\Delta \sim U(0.75, 0.85)$. Further specifications are described in Web Appendix F1.

Finally, maximum likelihood (ML) estimation was implemented using the exact likelihood obtained by numerical integration for a given sensitivity and specificity. Wald intervals were somewhat unreliable, so we opted to construct confidence intervals using profile likelihoods which, while computationally slow, enabled inference about functions of the parameters such as leakage measures. In the ML and MH methods, likelihoods were computed by numerically integrating over p_i using evenly spaced beta quantiles. Leakage measures, $P(L_i > 0)$ and $E(L_i)$, were calculated from the estimated parameters using (18) and (19).

3.4 Model fitting results

Table 3 shows model estimates and confidence/credible intervals from the ML, HMC, and MH methods when there is no minimum positive prevalence ($k = 0$). The MH method was applied both assuming $\Delta = 0.8$ and with a prior $\Delta \sim U(0.75, 0.85)$. The ML and MH parameter estimates, which are both calculated using the exact likelihood, are very similar. This suggests that the MH results are not overly sensitive to the priors used. The HMC method, based on an approximate likelihood, gave similar estimate of α , but β appears to have been over-estimated by about 50% (relative to the exact methods), while μ was under-estimated by about 25%.

Table 4 Posterior mean estimates with 95% credible interval (CI) of thresholded beta-binomial parameters (α , β , and μ) for t_{yi} under imperfect test sensitivity ($\Delta = 0.8$, $\Lambda = 1.0$), and minimum positive prevalence ($k = 0.0, 0.01, 0.02$). The corresponding posterior estimates of leakage parameters are calculated using (18) and (19), respectively.

Parameter	Minimum positive prevalence, k		
	0%	1%	2%
	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
$\alpha \times 10^3$	4.63 (2.90, 6.97)	6.37 (3.79, 10.00)	9.11 (5.09, 15.34)
β	4.31 (1.93, 7.77)	5.64 (2.61, 10.03)	7.29 (3.43, 12.86)
$\mu \times 10^3$	1.16 (0.67, 1.92)	1.19 (0.74, 1.87)	1.30 (0.85, 1.91)
$\mathbb{E}(L_i)$	25.9 (16.7, 37.5)	19.4 (12.4, 28.4)	14.5 (9.1, 21.3)
$\mathbb{P}(L_i > 0) \times 10^2$	3.68 (2.23, 5.57)	0.31 (0.19, 0.46)	0.15 (0.10, 0.23)

Table 3 also shows that the corresponding estimates of $\mathbb{E}(L_i)$ and $\mathbb{P}(L_i > 0)$ remain reasonably comparable across all four methods, underscoring the robustness of the modeling approaches in capturing leakage dynamics. Some differences arise in the β estimates, with the approximate models consistently yielding higher values than those obtained using the exact models. These discrepancies suggest a trade-off between computational efficiency and estimation accuracy when opting for approximate over exact modeling approaches.

The confidence and credible intervals for all parameter estimates were wide but still allow useful conclusions, particularly about leakage. $\mathbb{P}(L_i > 0)$ is estimated to be between 2.2% and 5.6% of consignments, and $\mathbb{E}(L_i)$ between 16 and 38 infected prawns per consignment. The latter interval refers to expected leakage noting this is an average over both zero and positive values where the latter may considerably exceed $\mathbb{E}(L_i)$ for particular consignments.

The fitted models were consistent with the observed data, based on randomized quantile residual plots (Dunn and Smyth, 1996), formal tests, and (for MH) posterior predictive checks, with no evidence of systematic lack of fit. Further details are provided in Web Appendix F2.

The results discussed so far assume no minimum positive prevalence, (ie, $k = 0$). However, expert opinion suggests that a non-zero minimum prevalence should apply, and in Section 3.2 we found that k is significantly different from 0 at the 0.1 (but not 0.05) level. The ML estimate of k was approximately 0.01. Table 4 reports results under alternative values of k (0%, 1%, and 2%). The first case reproduces the results in Table 3. For simplicity, these models were fitted using the MH method with sensitivity fixed at 0.8. As expected, increasing k shifted the posterior distribution of the beta parameters to the right, indicating greater concentration of consignment-level prevalence at higher values. However, the overall mean prevalence ($\mu = \alpha/(\alpha + \beta)$) remained similar across scenarios, indicating that k mainly affects the lower tail rather than the central tendency.

Table 4 also shows the associated leakage quantities, using (18) and (19) for $\mathbb{P}(L_i > 0)$ and $\mathbb{E}(L_i)$, respectively. When a positive k is assumed, estimated $\mathbb{P}(L_i > 0)$ and $\mathbb{E}(L_i)$ are both considerably reduced, particularly $\mathbb{P}(L_i > 0)$ which reduces from 3.7% when $k = 0$ is assumed to 0.3% when $k = 0.01$ is assumed. Expected leakage drops more modestly from 26 to 19 units per consignment. We suggest that policy makers

tentatively assume $k = 0.01$ as this aligns with both expert advice and ML estimation of this parameter.

3.5 Communicating the risk

The practical implications of the assumed model and parameter estimates for biosecurity policy and risk are not straightforward to extract. While $\mathbb{E}(L_i)$ and $\mathbb{P}(L_i > 0)$ have intuitive interpretations, they provide limited insight into the full distribution of leakage, particularly given that $L_i = 0$ with high probability. To address this, we adopt a more interpretable, distributional summary: histograms of the number of infected items per consignment (T_{Xi}), partitioned into true positives (TP), false negatives (FN), and true negatives (TN), where leakage corresponds to FN consignments.

To make the results more concrete, we consider expected outcomes for a notional set of 1000 consignments (about one year of imports). Results were obtained by simulating both T_{Xi} and $\tilde{t}_{yi}$ from the assumed model, using parameters obtained from their MH posterior distribution with $\Delta = 0.8$. We consider minimum prevalence values $k = 0, 1\%$, and 2% , with additional results in Web Appendix G.

Figure 1 shows the results. Although information-dense, feedback from biosecurity policy managers suggests it is useful in visualizing and understanding model implications. Each bar represents the expected number of consignments (out of 1000) across ranges of T_{Xi} . A different scale is used for $T_{Xi} = 0$, which dominates. Bars are color-coded by outcome (TP, TN, FN): when $T_{Xi} = 0$, all consignments are TN; when $T_{Xi} > 0$, outcomes are TP ($\tilde{t}_{yi} > 0$) or FN ($\tilde{t}_{yi} = 0$), shown as stacked bars. False positives are not possible under perfect specificity.

For example, when $k = 0$ (top panel), over 900 of the 1000 consignments are true negatives. Around 22 consignments contain 1-100 infected prawns, almost all false negatives (ie, leakage). This reflects very low prevalence ($\leq 0.03\%$ of 320 000 prawns per consignment), making detection unlikely. As T_{Xi} increases, detection probability rises; once infections exceed about 5000 ($\approx 1.5\%$), true positives dominate, indicating the sampling is effective. The picture changes when k is greater than zero, eliminating very small consignment infection rates and reducing false negatives. Consequently, there is less red in the second and third panels, as ruling out tiny infection levels makes the sampling regime highly effective. We advocate presenting all three panels to planners, reflecting pessimistic, balanced, and optimistic scenarios.

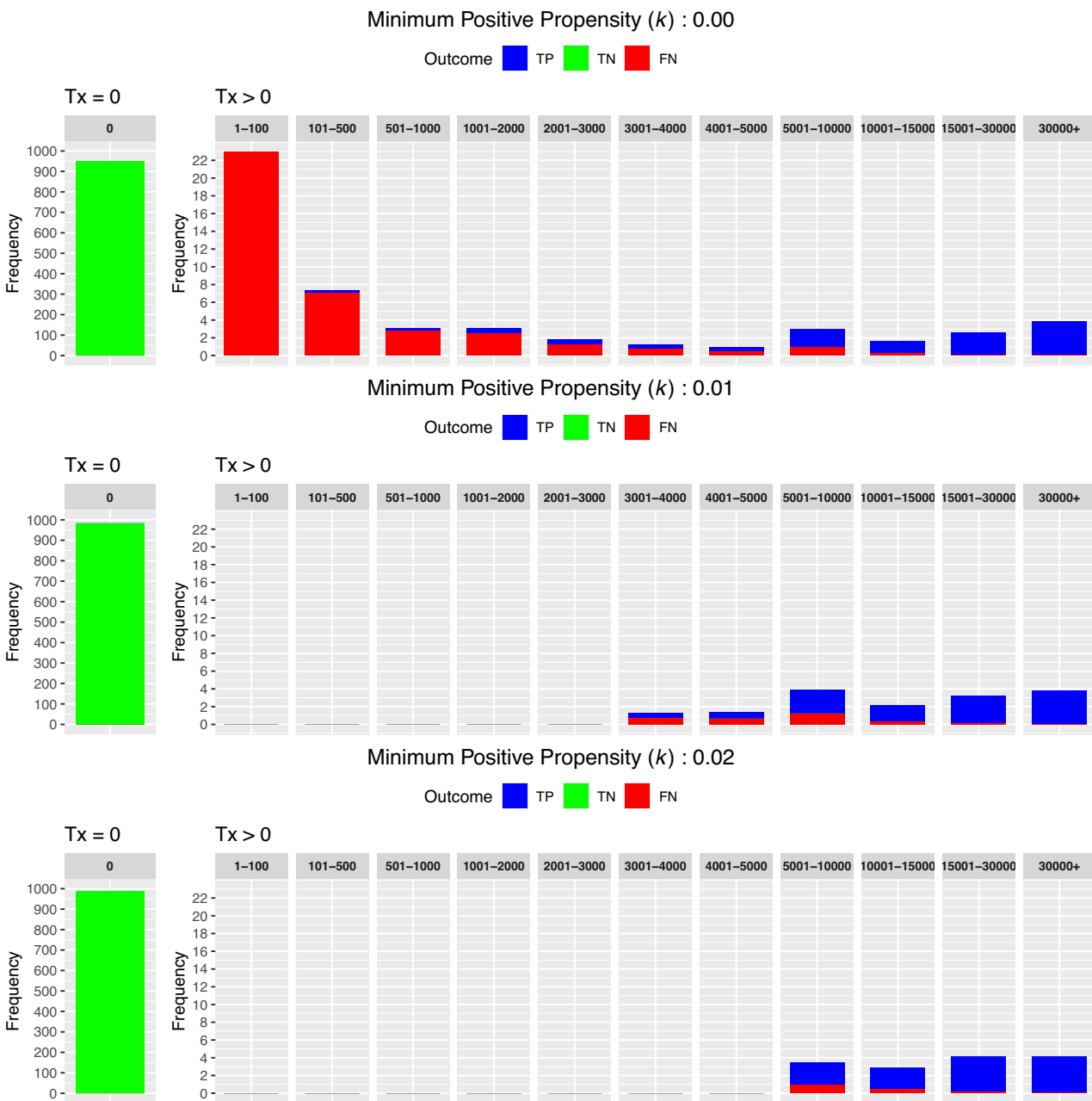


Figure 1 Expected frequencies of consignments out of a notional set of 1000 consignments, broken down by the number of contaminated prawns per consignment and the testing outcomes in three different scenarios based on minimum propensity, $k = (0, 1, 2)\%$. PCR test sensitivity is assumed to be $\Delta = 0.80$.

Our framework can also assess the impact of changing the number of tests, a key lever for managing leakage risks. With group size fixed at $m = 5$ (to maintain sensitivity), we vary the number of groups (b) from 13 to 20 or 26 and simulate outcomes with $k = 0.01$. Figure 2 shows that increasing b sharply reduces false negatives: $P(L_i > 0)$ falls by about two-thirds and expected leakage by half when b goes from 13 to 26. Biosecurity managers need to maintain a balance between operational costs, trade access, the risk of incursions and the impact of incursions. Results like those in Figure 2 allow decisions on the testing protocol to be made based on clear quantified measures of the effect of sample size on leakage risk. We

recommend periodic modeling exercises of historical testing data using the models like those developed here, feeding into ongoing review of surveillance scheme design.

4 Conclusions

Biosecurity sampling regimes are often designed to achieve a high chance of detecting contamination or infection under an assumed fixed design prevalence. While useful for setting long-term inspection regimes, this approach simplifies the reality that prevalences are highly heterogeneous across consignments and does not account for the amount of leakage at

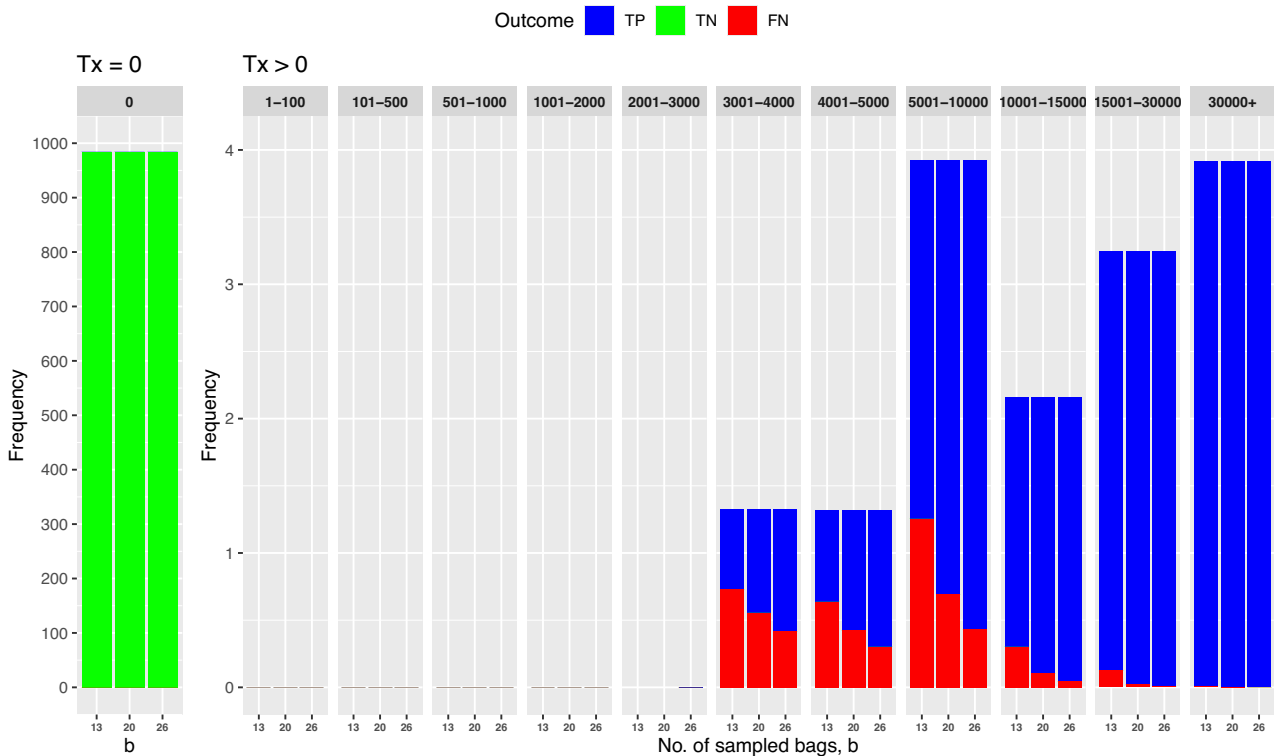


Figure 2 Expected frequencies of consignments out of a notional set of 1000 consignments, broken down by the number of contaminated prawns per consignment and their testing outcomes for three different numbers of sampled groups, $b = (13, 20, 26)$. Minimum positive prevalence is assumed to be $k = 1\%$ and the test sensitivity is $\Delta = 0.80$.

low prevalences. Historical inspection results can instead be modeled using latent consignment propensities, enabling estimation of leakage rates, key measures of importation risk. Modeling such data is challenging due to sparsity, predominance of zero outcomes, imperfect testing, and grouped observations. The importation of prawns into Australia exemplifies these issues.

To address these challenges, this article develops new modeling approaches and results. A key issue is parameter identifiability under sparse and grouped data. Specificity is shown to be identifiable and is near perfect in our case study. In contrast, sensitivity is unidentifiable and is therefore handled by either fixing it at 0.8 or assigning a prior over 0.75–0.85, reflecting expert advice. We also introduce the minimum positive prevalence, reflecting the view that extremely small but non-zero consignment prevalence is biologically implausible. This parameter appears empirically identifiable (albeit not very precisely), with evidence of a non-zero value in the case study.

These developments yield insights that are not otherwise available. In the prawn import case study, we find that although the sampling regime is highly effective, non-negligible leakage may still occur in large consignments even when infection prevalence and downstream risk are very low. This highlights the importance of assumptions about minimum prevalence, as the estimated probability of leakage decreases substantially when this parameter is positive. The models also allow leakage distributions to be evaluated under alternative

sample sizes, informing decisions about whether increased testing effort is justified. For example, doubling the number of groups tested would roughly halve infected prawns entering Australia and reduce consignments with non-zero leakage by about two-thirds.

These findings also suggest several directions for further work. Other multilevel models such as generalized linear mixed models are worth investigating for this class of problems, as they are more widely used and better supported by software than the latent beta models developed here. The censored grouped data structure and a minimum positive prevalence would still need to be incorporated. Fixed or random effects covariates could also be included for country of origin, importing company, or other consignment characteristics—these were not appropriate for a simple biosecurity process but would be relevant in other settings. Future work could embed the modeling framework into a continuous sampling plan with dynamically adapted inspection regimes (Trouvé et al., 2024), or allow sensitivity to vary with prevalence, viral load, or group size rather than assuming it fixed; this would be worthwhile where supported by data or theory, though not supported in the present case study. Another avenue would be to incorporate testing validation data if available (eg, from gold-standard testing), which would provide independent information about specificity or sensitivity.

The model allows leakage rates to be estimated for different values of b , which could inform design changes. Future research should consider optimal design more generally, in-

cluding the choice of m . This was beyond the scope of this paper, because a realistic treatment would require a model for test sensitivity as a function of pool size as well as quantification of the relative costs associated with the number of units extracted and the number of tests conducted. The models developed here would be an essential ingredient in such an investigation.

Beyond the case study, we establish a general result for rare-event multilevel binary data: specificity and heterogeneity (ie, α) can be identified even when sensitivity and mean prevalence cannot, provided there is heterogeneity between consignments (or other clusters). This applies to both group and individual testing, since the approximations developed in Section 2.3 are based on rarity of the events, and do not make assumptions about the group size. Other possible applications include public health (eg, Gelman and Carpenter, 2020 reanalyze SARS-CoV-2 antibody data clustered by zip code where sensitivity and specificity are uncertain), environmental sampling (eg, Furlan et al., 2016 consider a form of group testing where sensitivity is known approximately but not exactly and sites may act as clusters in place of consignments), and surveillance schemes for animal health (eg, USDA-APHIS, 2024, p28 refer to clustering in premises and imperfect sensitivity albeit assumed known). Further applications are discussed in Web Appendix H.

Overall, analyzing historical test result data in a biosecurity context presents unique challenges but also important opportunities to quantify risks that are otherwise implicit. This paper shows how routinely collected inspection data can be used to estimate leakage rates, assess the consequences of imperfect testing, and evaluate alternative sampling regimes.

Acknowledgments

We thank the Australian Department of Agriculture, Fisheries and Forestry for releasing the anonymized frozen seafood testing results data and for support as industry partner of the ARC grant. Anonymous reviewers made extensive comments which significantly improved the paper.

Supplementary materials

Supplementary material is available at *Biometrics* online.

Web Appendices, Tables, and Figures referenced in Sections 2, 3, and 4 and R code which reproduce the results are available with this paper at the Biometrics website on Oxford Academic. An R package which implements some of the proposed methods is at <https://github.com/sumon148/groupedBBMH>.

Funding

The research was supported by an Australian Research Council (ARC) Linkage Project (LP220100003).

Conflict of interest

Two of the authors were employed by the Department of Agriculture, Fisheries and Forestry of Australia (DAFF), who also

supported the grant funding this research as the industry partner.

Data availability

The data analyzed in this paper are available at Clark (2023).

References

- Biosecurity Australia, (2009). Final import risk analysis report for prawns and prawn products. Canberra, Australia: Biosecurity Australia. https://www.agriculture.gov.au/sites/default/files/sitecollectiondocuments/ba/memos/2009/Final_prawn_IRA_report_7_Oct_09.pdf. [Accessed 20 December 2025].
- Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, 100, 1–54.
- Clark, R. G. (2023). *Deidentified Frozen Seafood Dataset*. Data Set Version 1.0. <https://doi.org/10.5281/zenodo.7762164>.
- Clark, R. G., Barnes, B. and Parsa, M. (2023). Clustered and unclustered group testing for biosecurity. *Journal of Agricultural, Biological and Environmental Statistics*, 29, 193–211.
- Constable, F., Daly, A., Terras, M. A., Penrose, L. and Dall, D. (2018). Detection in Australia of Cucumber green mottle mosaic virus in seed lots of cucurbit crops. *Australasian Plant Disease Notes*, 13, 18.
- Cowling, D., Gardner, I. and Johnson, W. (1999). Comparison of methods for estimation of individual-level prevalence based on pooled samples. *Preventive Veterinary Medicine*, 39, 211–225.
- Department of Agriculture, Fisheries and Forestry (DAFF), (2023). Review of the Biosecurity Risks of Prawns Imported from All Countries for Human Consumption. Canberra, Australia: Department of Agriculture, Fisheries and Forestry. <https://www.agriculture.gov.au/sites/default/files/documents/prawn-final-report.pdf>. [Accessed 20 December 2025].
- Dunn, P. K. and Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5, 236–244.
- EU (1993). Council Directive 93/85/EEC of 4 October 1993 on the control of potato ring rot. *Official Journal of the European Communities*, L 259, 1–25. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31993L0085> [Accessed 21 August 2025].
- Furlan, E. M., Gleeson, D., Hardy, C. M. and Duncan, R. P. (2016). A framework for estimating the sensitivity of eDNA surveys. *Molecular Ecology Resources*, 16, 641–654.
- Gelman, A. and Carpenter, B. (2020). Bayesian analysis of tests with unknown specificity and sensitivity. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 69, 1269–1283.
- Hepworth, G. (2019). Bias correction of estimated proportions using inverse binomial group testing. *Australia and New Zealand Journal of Statistics*, 61, 51–60.

- Hodges, J. L. and Le Cam, L. (1960). The Poisson approximation to the Poisson binomial distribution. *The Annals of Mathematical Statistics*, 31, 737–740.
- IPPC (2008). International standards for phytosanitary measures: methodologies for sampling of consignments. International Plant Protection Convention. CPM-3 Report, ISPM No. 31. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO). <http://www.fao.org/3/cb2570en/cb2570en.pdf>. [Accessed 20 December 2025].
- Ma, H., Leibowicz, B. D. and Hasenbein, J. J. (2025). Optimal sampling strategy for probability estimation: An application to the Agricultural Quarantine Inspection Monitoring program. *Risk Analysis*, 45, 1410–1427.
- Teerapabolarn, K. (2015). Negative Binomial approximation to the Beta Binomial distribution. *Int J Pure Appl Math*, 98, 39–43.
- Trouvé, R., Arthur, A. D. and Robinson, A. P. (2022). Assessing the quality of offshore binomial sampling biosecurity inspections using onshore inspections. *Ecological Applications*, 32, e2595.
- Trouvé, R., Bland, L. M., Robinson, A. P., Ducey, M. J. and Hester, S. M. (2024). Screen: designing sampling schemes for border inspection. In Hester, S. M., Bland, L. M. and Arndt, E., and others, editors. *Biosecurity: A Systems Perspective*. Boca Raton: CRC Press.
- Trouvé, R. and Robinson, A. P. (2021). Estimating consignment-level infestation rates from the proportion of consignment that failed border inspection: possibilities and limitations in the presence of overdispersed data. *Risk Analysis*, 41, 992–1003.
- Trouvé, R. and Robinson, A. P. (2024). Efficient border biosecurity inspection leverages superspreading to reduce biological invasion risk. *Risk Analysis*, 44, 1839–1849.
- USDA-APHIS (2024). Guidance on how to design a surveillance scheme. Fort Collins, CO: U.S. Department of Agriculture, Animal and Plant Health Inspection Service. <https://www.aphis.usda.gov/sites/default/files/guidance-design-surveillance.pdf>. [Accessed 20 December 2025].