



Estimating herd prevalence on the basis of aggregate testing of animals

Christel Faes, Marc Aerts and Saskia Litière,
Universiteit Hasselt, Diepenbeek, Belgium

Estelle Méroc and Yves Van der Stede
Veterinary and Agrochemical Research Center, Ukkel, Belgium

and Koen Mintiens
Vose Consulting, Gent, Belgium

[Received September 2009. Final revision March 2010]

Summary. It is common practice that some or all animals in a group of animals, e.g. a herd, are tested for their health status by using a diagnostic test to investigate whether the herd is infected by a disease. Several obstacles complicate the estimation of herd prevalence on the basis of test results of the animals. First, diagnostic tests are often imperfect, resulting in a misclassification of the animal's disease status. It is well known how to correct the animal's apparent prevalence by using the diagnostic sensitivity and specificity of the animal test, but the effects on herd prevalence are less clear. Second, in practice, a herd is often defined as positive when at least one sampled animal tested positively. This definition is ambiguous and is also different from the herd prevalence that is based on having at least one diseased animal in the herd. The paper provides a discussion of these aspects and proposes a method to estimate the true herd prevalence on the basis of the health status of (all or a sample of) animals within a herd corrected for the sensitivity and specificity of the individual test, the number of animals that are tested in the herd and the uncertainty of the diagnostic test characteristics.

Keywords: Animal prevalence; Beta–binomial model; Diagnostic tests; Herd prevalence; Sensitivity; Specificity

1. Introduction

The surveillance of emerging diseases in animals is a topic of great significance in veterinary epidemiology. Examples of recent emerging diseases that are related to the animal population are foot-and-mouth disease, classical swine fever and blue tongue (BT). Both the quantification of occurrence of the disease and the description of risk factors affecting the health status in the population are important characteristics in disease surveillance. A key feature in veterinary epidemiology is that animals are grouped within herds. A possible way to quantify a disease pattern in such a population is via animal prevalence, describing the probability that an individual from the population is diseased. However, control programmes for infectious animal diseases are typically based on the disease status of a herd rather than of individual animals. Therefore, interest is also in herd prevalence, describing the probability that a randomly selected

Address for correspondence: Christel Faes, Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Universiteit Hasselt, Agoralaan 1, Diepenbeek 3590, Belgium.
E-mail: christel.faes@uhasselt.be

herd contains at least one animal which is affected by the disease. Although this is common practice, there are several issues which complicate the estimation, interpretation and analysis of herd prevalence.

The first complexity is the *grouping of animals* in the population. We can think of animals grouped in a herd, holding, farm, flock, pen, litter or any other form of clustering. In a different context we can think of individuals grouped in households or municipalities. Throughout the text, the term 'herd' will be used for any aggregation of individuals. Because of the same environment and same herd characteristics, animals in the same herd are possibly correlated. We cannot ignore the possibility that animals within herds are more alike than between herds and, consequently, the intraclass correlation should be addressed to obtain valid statistical interest (Aerts *et al.*, 2002; Molenberghs and Verbeke, 2005). This will be discussed in Section 2.

The second complexity when studying the prevalence of disease is that *diagnostic tests are often imperfect*. Possibly, some of the diseased animals give false negative diagnostic test results, whereas some of the healthy animals might falsely react positively to the test. The capacity of a diagnostic test to identify the disease status of an animal correctly is indicated by the diagnostic sensitivity and specificity of the test, being the probability that a diseased animal would give a positive test result and that a non-diseased animal would give a negative test result respectively. Diagnostic sensitivity and specificity can be used to calculate the true animal prevalence (the proportion of animals that are actually infected) from the apparent animal prevalence (the proportion of animals that tested positively) taking the test misclassification into account (Rogan and Gladen, 1978). Of course, a misclassification of the individual's health status implies also a misclassification of the health status of the herd. However, the effect of the individual's test misclassification on herd prevalence is less clear. Herd prevalence based on aggregate testing was studied by for example Jordan (1996), Baggesen *et al.* (1996), Christensen and Gardner (2000) and Humphry *et al.* (2004), but these methods do not account for within-herd correlation. A herd can be classified as positive because one of the animals sampled was correctly tested positively, or one of the healthy animals falsely tested positively. A herd can only be negative when all healthy animals were correctly tested negatively, and all diseased animals falsely tested negatively. This is the topic of discussion in Sections 3 and 4. We also extend the concepts of sensitivity and specificity of individual test results to the herd level sensitivity and specificity in Section 5.

The third difficulty complicating the estimation of herd prevalence is the sampling design of the survey. Often, a *two-stage sampling design* is used. At the first stage, herds are selected at random from a sampling frame. At the second stage, animals from the selected herds are randomly sampled. In such a design, we can test either all the animals in a herd (one-stage cluster sampling) or a (small) fraction of animals in the herd. Often, a (small) fixed number of animals per herd are tested. In the latter case, the definition of herd positivity is ambiguous. It can be interpreted as having at least one diseased animal in the sample of animals or, as before, as having at least one diseased animal in the full herd, resulting in two different estimates of prevalence. This is discussed in Section 4.

In this paper, a method to estimate the true herd prevalence based on the test results of animals is proposed. The methods proposed are used to study the disease distribution of BT in Belgium as a result of the outbreak of the disease in 2006. BT is an insect-borne viral disease, affecting all species of ruminants. In August 2006, very unexpectedly, BT was for the first time notified in the Netherlands, Belgium and Germany. Later on during the epizootic episode, related cases were also declared in France and Luxembourg. The virus was identified as BT virus (BTV) serotype 8 (European Commission Reference Laboratory, 2006), which before this epizootic episode had occurred only in Africa, Central America, Malaysia and India–Pakistan (Herniman

et al., 1980; Hassan, 1992; Mo *et al.*, 1994; Daniels *et al.*, 2004; Gerdes, 2004). During the winter of 2006–2007, it was assumed that climate conditions were unfavourable for further propagation of BTV. A serological and virological cross-sectional survey (BT winter screening) targeting all Belgian ruminants was undertaken in January–February 2007 to establish the true final dispersion of the virus across the country (Méroc *et al.*, 2008). The objective of the study was to provide unbiased estimates of BT prevalence in Belgium based on the winter screening. Both animal prevalence and herd prevalence are of interest. This is illustrated in Section 6. We end this paper with a discussion on this topic (Section 7).

2. Beta–binomial model to estimate animal and herd prevalence

In this section, the basic approach to estimate simultaneously the animal and herd prevalence on the basis of an aggregate of tests applied to all animals in a herd is explained. This basic model assumes that the diagnostic test is perfect and that all animals in a herd are sampled (one-stage cluster sampling). In Sections 3 and 4, the effects of violating these assumptions are described and extensions of the basic method are proposed.

We shall first introduce some notation, of which an overview is given in Table 1. Assume that there are n_i animals in herd (group) i ($i = 1, \dots, N$), of which z_i tested positively for some disease. Let π^A denote animal prevalence, describing the probability that an animal is diseased. π^H denotes herd prevalence, defined as the probability of having at least one diseased animal in a herd.

In this paper, we propose the use of the beta–binomial model to estimate both the animal and the herd prevalence (Skellam, 1948; Kleinman, 1973; Williams, 1982). This model allows that animals within the same herd are more alike than animals from different herds. Other models for correlated binomial data are candidate models as well. A possible alternative is the Bahadur (1961) model, which is also known as the correlated binomial distribution, which leads to very elegant closed form expressions for the likelihood but suffers from computational drawbacks due to the highly restrictive form of the parameter space (Declerck *et al.*, 1998; Aerts *et al.*, 2002). The beta–binomial model has some advantages in terms of analytical expression, marginal interpretation and use of conjugate priors making it robust against misspecification of the prior (Bolstad, 2005). The beta–binomial model assumes that the number of positively tested animals Z_i follows a binomial distribution with probability θ_i . To account for clustering of animals in a

Table 1. Explanation of the notation used in the text

<i>Notation</i>	<i>Description</i>
n_i	Number of animals in herd i
z_i	Number of positively tested animals in herd i
π^A	True animal prevalence
π^H	True herd prevalence
Se	Sensitivity of a single test
Sp	Specificity of a single test
p^A	Apparent animal prevalence (proportion of animals tested positively)
p^H	Apparent herd prevalence (proportion of herds tested positively)
Se ^H	Herd sensitivity
Sp ^H	Herd specificity

herd, the underlying disease probabilities θ_i are assumed to vary within a herd according to a beta distribution $\text{beta}(\alpha, \beta)$ with parameters $\alpha = \pi^A(1/\rho - 1)$ and $\beta = (1 - \pi^A)(1/\rho - 1)$. As a result, the beta distribution has mean π^A and determines the strength of within-herd correlation ρ . This leads to the beta-binomial distribution of the number of positive test results Z_i in herd i with probability density function given by

$$f(z_i; \pi^A, \rho, n_i) = \binom{n_i}{z_i} \frac{B\{\pi^A(\rho^{-1} - 1) + z_i, (1 - \pi^A)(\rho^{-1} - 1) + n_i - z_i\}}{B\{\pi^A(\rho^{-1} - 1), (1 - \pi^A)(\rho^{-1} - 1)\}}, \quad (1)$$

where $B(\cdot, \cdot)$ denotes the beta function (Skellam, 1948; Kleinman, 1973; Williams, 1982), and can be rewritten as

$$f(z_i; \pi^A, \rho, n_i) = \binom{n_i}{z_i} \prod_{k=0}^{z_i-1} \left(\pi^A + \frac{k\rho}{1-\rho} \right)^{n_i-z_i-1} \prod_{k=0}^{n_i-z_i-1} \left(1 - \pi^A + \frac{k\rho}{1-\rho} \right)^{n_i-1} \left(1 + \frac{k\rho}{1-\rho} \right)^{-1}.$$

The moments of the beta-binomial distribution are

$$E(Z_i) = n_i \pi^A, \\ \text{var}(Z_i) = n_i \pi^A (1 - \pi^A) \{1 + \rho(n_i - 1)\}.$$

Note that the beta-binomial model reduces to the familiar binomial model when there is no within-herd correlation ($\rho = 0$). When for example a stratified sampling design with unequal sampling probabilities is used to select the herds, it might be necessary to account for the sampling design. A weighted log-likelihood can then be used where each contribution from a herd is weighted with the inverse of the sampling probabilities. Estimation of this model can be performed by using maximum likelihood, where the maximum likelihood can be derived directly from equation (1), or can be handled in the Bayesian framework, where the binomial likelihood is combined with a beta distribution as a prior. Because of its flexibility, we have used the Bayesian framework for illustration of the application in this paper and programmed all models in the software package WinBUGS (Lunn *et al.*, 2000). WinBUGS performs Bayesian inference by using Gibbs sampling. Some of the code is given in Appendix A.

On the basis of the beta-binomial model, it is easy to derive the herd prevalence π^H from the animal prevalence π^A and the within-herd correlation ρ . The probability that at least one animal in a herd of size n is affected by the disease is equal to

$$P(Z > 0 | n) = 1 - P(Z = 0 | n) \\ = 1 - \frac{B\{\pi^A(\rho^{-1} - 1), (1 - \pi^A)(\rho^{-1} - 1) + n\}}{B\{\pi^A(\rho^{-1} - 1), (1 - \pi^A)(\rho^{-1} - 1)\}}.$$

This expression can be written as (Aerts *et al.*, 2002)

$$P(Z > 0 | n) = 1 - \prod_{k=0}^{n-1} \left\{ 1 - \pi^A + \frac{k\pi^A\rho}{1 + (k-1)\rho} \right\}. \quad (2)$$

Note that this probability depends on the herd size n , the animal prevalence π^A and the within-herd correlation ρ . For a fixed animal prevalence and within-herd correlation, the probability that at least one animal is infected increases with the size of herd. The herd prevalence π^H is obtained from this formula by integrating out the herd size n . Considering all possible herd sizes n_i in the population of herds, with corresponding probabilities $P(n_i)$, this leads to

$$\begin{aligned} \pi^H &= \sum_{n_i=0}^{\infty} P(n_i) P(Z > 0 | n_i) \\ &= \sum_{n_i=0}^{\infty} P(n_i) \left[1 - \prod_{k=0}^{n_i-1} \left\{ 1 - \pi^A + \frac{k\pi^A\rho}{1 + (k-1)\rho} \right\} \right], \end{aligned} \tag{3}$$

where π^A is the animal prevalence corrected for the within-herd correlation ρ in the data, on the basis of the beta–binomial model. A plug-in estimate of π^H can be obtained by plugging in the maximum likelihood estimates of π^A and ρ , or, in the Bayesian framework, by sampling from the posterior distribution. An advantage of the Bayesian framework is that we account for the full posterior distributions of π^A and ρ to obtain the posterior distribution of π^H .

In a regression context where for example we wish to correct the prevalence for known herd characteristics, the model needs to incorporate covariates. To model the marginal parameters π^A and ρ we use a composite link function. An appropriate choice is given by the following generalized linear regression relationships:

$$\begin{aligned} \ln\left(\frac{\pi_i^A}{1 - \pi_i^A}\right) &= X_{1i}\beta_1, \\ \ln\left(\frac{1 + \rho_i}{1 - \rho_i}\right) &= X_{2i}\beta_2, \end{aligned}$$

with β_1 and β_2 the vectors of unknown regression coefficients, and X_1 and X_2 design matrices corresponding to the covariate values of the herd. Subscript i reflects the dependence on herd-specific covariates of herd i . Derivations of the herd prevalence can be conducted in a similar way as before.

3. Correcting for misclassification of diagnostic test

Typically, the tests that are used to classify whether an animal is diseased or not are imperfect. In this section, it is explained how the true animal and herd prevalence can be derived from the sensitivity and specificity of the diagnostic test at the individual level.

We model the number of positively tested animals z_i out of n_i animals in herd i by using a beta–binomial model (Section 2):

$$z_i | n_i \sim \text{beta-binomial}(p^A, \rho, n_i),$$

with p^A the apparent animal prevalence and ρ the correlation of the test results from animals in the same herd. The true animal prevalence π^A can be derived from the apparent animal prevalence p^A by using the equation (Rogan and Gladen, 1978)

$$p^A = \text{Se}^A \pi^A + (1 - \text{Sp}^A)(1 - \pi^A),$$

or

$$\pi^A = \frac{p^A + \text{Sp}^A - 1}{\text{Se}^A + \text{Sp}^A - 1}, \tag{4}$$

where Se^A and Sp^A are the sensitivity and specificity of the test at the animal level. To be a valid estimate, the Rogan–Gladen estimate requires that the probability of a positive test result is larger for a diseased individual than for an undiseased individual ($\text{Se}^A > 1 - \text{Sp}^A$), that the probability of a positive test result for a diseased individual is larger than the probability of a positive test result ($\text{Se}^A \geq p^A$) and that the probability of a negative test result for a

non-diseased individual is larger than the probability of a negative test result ($Sp^A \geq 1 - p^A$) (Rogan and Gladen, 1978).

On the basis of the beta-binomial model, the true herd prevalence π^H can be derived from the apparent animal prevalence p^A by inserting equation (4) into equation (3). This yields

$$\begin{aligned} \pi^H &= \sum_{n_i=0}^{\infty} P(n_i) \left(1 - \prod_{k=0}^{n_i-1} \left[1 - \frac{p^A + Sp^A - 1}{Se^A + Sp^A - 1} + \frac{k(p^A + Sp^A - 1)\rho}{\{1 + (k-1)\rho\}(Se^A + Sp^A - 1)} \right] \right) \\ &= \sum_{n_i=0}^{\infty} P(n_i) \left(1 - \prod_{k=0}^{n_i-1} \left[1 + \frac{(p^A + Sp^A - 1)(\rho - 1)}{(Se^A + Sp^A - 1)\{1 + (k-1)\rho\}} \right] \right). \end{aligned} \tag{5}$$

This equation depends on the apparent animal prevalence p^A , the animal’s test sensitivity Se^A and specificity Sp^A and the within-herd correlation of the disease status ρ . An estimate of the herd prevalence π^H can be made by plugging in estimates of all these parameters. Using a Bayesian analysis, the uncertainty of the parameters is taken into account by plugging in the posterior distributions of p^A and ρ . This is performed by calculating π^H in WinBUGS and is illustrated in Appendix A. Note that we make the assumption that the correlation ρ between the disease status of different animals in a herd is equal to the correlation between the test result of different animals in a herd. Although it is not possible to estimate the correlation between the true disease status of different animals, because the true disease status is a latent variable, the correlation between the test results can be estimated from the data, by using the beta-binomial model. Fig. 1 shows the apparent (broken curve) and true (full curve) probability for a herd to be infected as a function of the size of the herd, corresponding to an apparent animal prevalence of 0.20 (Fig. 1(a)) and 0.60 (Fig. 1(b)), test sensitivity and specificity 0.80 and 0.90 respectively and within-herd correlation 0.30. These are constructed by using expressions (2) directly (apparent prevalence) and inserting equation (4) in equation (2) (true prevalence).

3.1. Accounting for the uncertainty of sensitivity and specificity

Typically, however, the sensitivity and specificity are not known fixed values (Bollaerts *et al.*, 2009). Often, confidence bounds for the sensitivity and specificity are known from the literature. It is not always clear how to use such confidence bounds, and how to account for the uncertainty of the sensitivity and specificity in the analysis. The strength of the Bayesian modelling framework is that we can impose a prior distribution on the sensitivity and specificity, instead of assuming a fixed value for the diagnostic test characteristics.

A possible way to quantify the prior knowledge on the sensitivity (or specificity) of the test is the use of a beta distribution $\text{beta}(a, b)$. Parameters of the beta distribution can then be selected to represent best the bounds of the sensitivity (or specificity) and the most probable value of the sensitivity (or specificity). The parameters a and b are estimated by using the minimal (m_1), maximal (m_3) and most probable (m_2) value of the sensitivity (or specificity), using the equations (three-point estimation; Grubbs (1962))

$$\begin{aligned} a &= \frac{\mu^2(1-\mu) - \sigma^2\mu}{\sigma^2}, \\ b &= a \frac{1-\mu}{\mu} \end{aligned}$$

where the mean μ and standard error σ of the beta-binomial distribution are approximated with

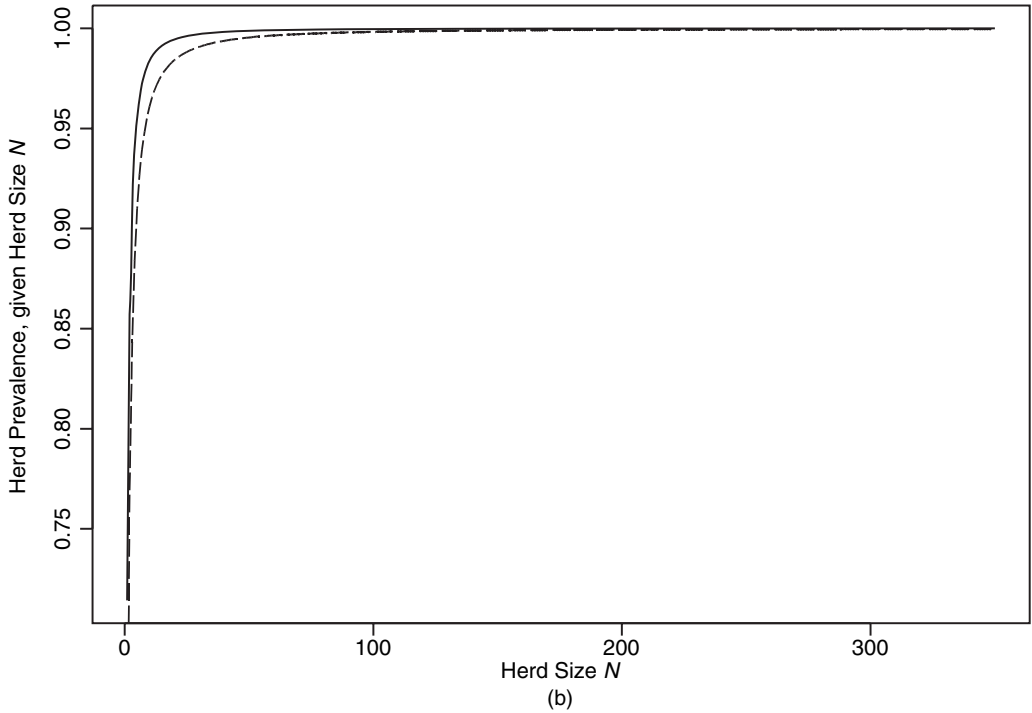
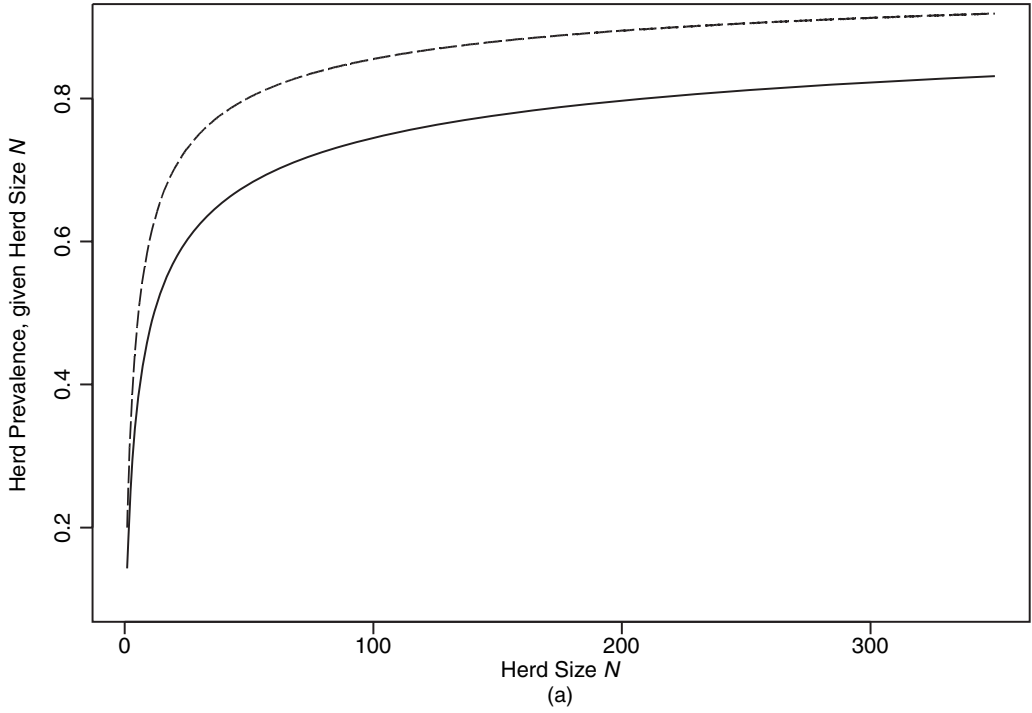


Fig. 1. True (—) and apparent (---) herd prevalence corresponding to a specific size of herd for an apparent animal prevalence of (a) 0.20 and (b) 0.60 with test sensitivity and specificity 0.80 and 0.90 respectively and a within-herd correlation of 0.30

$$\mu = \frac{m_1 + 4m_2 + m_3}{6}$$

and

$$\sigma = \frac{m_3 - m_1}{6}.$$

Although this method is used in this paper, other methods are possible as well to obtain a prior distribution describing the knowledge on the sensitivity and specificity.

4. Herd prevalence based on two-stage sampling design

So far, it has been assumed that all animals in a sampled herd are tested for the disease of interest, i.e. a one-stage sampling design is adopted. In surveillance programmes, this is often too costly and, typically, a small number of animals in a herd is sampled. In these surveys, a two-stage sampling design is often used. In the first stage, N herds are sampled. At the second stage, m_i animals are sampled from the selected herds of size n_i . Often, $m_i \equiv m$. It is important to understand the effects of such a sampling design on animal prevalence and herd prevalence.

To study the performance of the estimates based on a two-stage *versus* one-stage sampling design, a simulation study was performed, simulating 100 herds of size $n_i \equiv 100$. This corresponds to the average size of herd of a large cattle herd (more than 60 animals) in Belgium. The disease status of each animal in this herd was simulated by using a beta-binomial model with a prespecified animal prevalence and within-herd correlation. The values (0.03, 0.10, 0.30) for π^A correspond to low, moderate and more severe disease occurrence scenarios. Having ρ in the set (0.2, 0.4, 0.6) corresponds to low, medium and high within-herd correlations. These simulated data correspond to a one-stage sampling design, in which all animals from the herd are being tested for the presence of the disease. In the second step, a random sample of m animals is taken from each herd. The number of sampled animals (m) varies between 3 and 10. This corresponds to a two-stage sampling design. A perfect test is assumed in this simulation ($p^A = \pi^A$; $\text{Se}^A = \text{Sp}^A = 1$). The animal and herd prevalence corresponding to each simulated data set are estimated by using a beta-binomial model. Maximum likelihood was used to obtain estimates of the parameters ρ and π^A from the beta-binomial model. The herd prevalence is estimated from equation (3) by plugging in the maximum likelihood estimates of ρ and π^A , and with n_i replaced by the size of the sample m . Note that the herd prevalence that is obtained from a sample of animals within a herd defines the probability of having at least one positive sample, whereas the prevalence that is obtained from the full herd defines the probability of having at least one positive animal in the herd. Tables 2 and 3 summarize the results. They show the average estimate and the empirical standard errors over the 1000 simulations.

It can be concluded that the animal prevalence is very close to the true (assumed) animal prevalence, even when a small number of animals within herds is sampled. There is, however, a clear drop in the precision of the estimate when using a random sample of animals. This corresponds to the larger standard errors in the two-stage sample design. On the basis of Table 3, we see that there is a clear discrepancy between the estimated herd prevalence that is based on either the sample or the full herd. As expected, the prevalence of at least one positive sample of a herd is not a good representative for the prevalence of at least one positive animal in the herd. Also, the herd prevalence that is based on the two-stage sample design increases as the number of animals sampled increases. Thus, the definition of a positive herd defined as a positive sample should be interpreted with great care and depends on the number of sampled animals in the herd.

Table 2. Summary of the simulation study: estimated animal prevalence, based on a sample of the herd (two-stage sampling) or based on all animals in the herd (one-stage sampling)†

ρ (true)	π^A (true)	Sample size	Estimate for two-stage sampling	Estimate for one-stage sampling
0.2	0.03	3	0.0289 (0.0110)	0.0291 (0.0074)
0.2	0.03	5	0.0306 (0.0100)	0.0299 (0.0078)
0.2	0.03	10	0.0294 (0.0088)	0.0293 (0.0075)
0.2	0.10	3	0.1004 (0.0189)	0.0995 (0.0134)
0.2	0.10	5	0.0998 (0.0168)	0.0995 (0.0131)
0.2	0.10	10	0.0986 (0.0147)	0.0989 (0.0133)
0.2	0.30	3	0.2984 (0.0311)	0.2989 (0.0208)
0.2	0.30	5	0.2965 (0.0277)	0.2982 (0.0214)
0.2	0.30	10	0.2987 (0.0237)	0.2998 (0.0205)
0.4	0.03	3	0.0302 (0.0128)	0.0294 (0.0101)
0.4	0.03	5	0.0293 (0.0121)	0.0290 (0.0103)
0.4	0.03	10	0.0288 (0.0116)	0.0285 (0.0103)
0.4	0.10	3	0.0990 (0.0232)	0.0985 (0.0183)
0.4	0.10	5	0.1005 (0.0213)	0.0991 (0.0184)
0.4	0.10	10	0.0997 (0.0207)	0.0990 (0.0186)
0.4	0.30	3	0.2979 (0.0349)	0.2987 (0.0273)
0.4	0.30	5	0.2988 (0.0329)	0.2987 (0.0283)
0.4	0.30	10	0.2990 (0.0305)	0.2995 (0.0273)
0.6	0.03	3	0.0303 (0.0150)	0.0292 (0.0125)
0.6	0.03	5	0.0287 (0.0120)	0.0262 (0.0105)
0.6	0.03	10	0.0297 (0.0133)	0.0289 (0.0123)
0.6	0.10	3	0.0997 (0.0259)	0.0994 (0.0220)
0.6	0.10	5	0.1003 (0.0246)	0.0998 (0.0219)
0.6	0.10	10	0.0970 (0.0234)	0.0969 (0.0211)
0.6	0.30	3	0.2976 (0.0383)	0.2982 (0.0322)
0.6	0.30	5	0.2997 (0.0391)	0.2994 (0.0342)
0.6	0.30	10	0.3025 (0.0368)	0.3017 (0.0339)

†The herds have a fixed size of 100 animals. Average estimates and empirical standard errors (in parentheses) over 1000 simulations are given.

We can, however, obtain an estimate of the prevalence of at least one positive animal in the herd from the two-stage sample (the corrected estimate column in Table 3). This is obtained from

$$P(\widehat{Z}_i > 0) = 1 - \prod_{k=0}^{n_i-1} \left\{ 1 - \widehat{\pi}^A + \frac{k\widehat{\pi}^A\widehat{\rho}}{1 + (k-1)\widehat{\rho}} \right\}, \tag{6}$$

where n_i is the number of animals in a herd, $\widehat{\pi}^A$ is animal prevalence based on the sample of size m_i and $\widehat{\rho}$ is the estimated correlation between the samples. From Table 3 we see that this method seems to work well when the sample is not too small. When the animal prevalence is small, larger samples in the herd are needed to obtain a good estimate of the herd prevalence. The price to pay is precision, which is much smaller compared with the estimates that are based on the full herd.

5. Herd level sensitivity and specificity

Typically, the sensitivity and specificity of a single test are known, but less information is available on the sensitivity and specificity corresponding to the aggregate testing of animals within a herd. Therefore, the sensitivity and specificity of a test for the disease status of a group of individuals based on tests on the individuals separately is the topic of discussion in this section.

Table 3. Summary of the simulation study: estimated herd prevalence†

ρ (true)	π^A (true)	m_i	π^H (true)	Results for two-stage sampling		Estimate‡ for one-stage sampling
				Estimate‡	Corrected estimate§	
0.2	0.03	3	0.335	0.067 (0.025)	0.443 (0.239)	0.333 (0.048)
0.2	0.03	5	0.335	0.102 (0.029)	0.364 (0.138)	0.334 (0.048)
0.2	0.03	10	0.335	0.145 (0.034)	0.334 (0.088)	0.333 (0.047)
0.2	0.10	3	0.747	0.227 (0.037)	0.754 (0.127)	0.747 (0.041)
0.2	0.10	5	0.747	0.305 (0.040)	0.745 (0.089)	0.745 (0.040)
0.2	0.10	10	0.747	0.417 (0.044)	0.733 (0.065)	0.745 (0.041)
0.2	0.30	3	0.986	0.570 (0.046)	0.990 (0.020)	0.985 (0.006)
0.2	0.30	5	0.986	0.691 (0.041)	0.981 (0.014)	0.985 (0.006)
0.2	0.30	10	0.986	0.826 (0.031)	0.983 (0.010)	0.985 (0.006)
0.4	0.03	3	0.187	0.059 (0.023)	0.243 (0.169)	0.187 (0.039)
0.4	0.03	5	0.187	0.074 (0.026)	0.200 (0.088)	0.186 (0.040)
0.4	0.03	10	0.187	0.098 (0.030)	0.190 (0.063)	0.186 (0.038)
0.4	0.10	3	0.502	0.186 (0.038)	0.498 (0.111)	0.499 (0.048)
0.4	0.10	5	0.502	0.238 (0.041)	0.504 (0.086)	0.502 (0.050)
0.4	0.10	10	0.502	0.301 (0.045)	0.496 (0.067)	0.501 (0.048)
0.4	0.30	3	0.886	0.494 (0.047)	0.875 (0.052)	0.885 (0.027)
0.4	0.30	5	0.886	0.582 (0.045)	0.879 (0.039)	0.884 (0.027)
0.4	0.30	10	0.886	0.681 (0.042)	0.878 (0.034)	0.885 (0.027)
0.6	0.03	3	0.112	0.048 (0.022)	0.134 (0.108)	0.113 (0.031)
0.6	0.03	5	0.112	0.056 (0.020)	0.116 (0.047)	0.107 (0.029)
0.6	0.03	10	0.112	0.069 (0.024)	0.114 (0.040)	0.112 (0.031)
0.6	0.10	3	0.331	0.156 (0.037)	0.328 (0.084)	0.328 (0.048)
0.6	0.10	5	0.331	0.183 (0.039)	0.329 (0.070)	0.332 (0.048)
0.6	0.10	10	0.331	0.213 (0.042)	0.319 (0.062)	0.326 (0.048)
0.6	0.30	3	0.716	0.424 (0.049)	0.702 (0.072)	0.715 (0.042)
0.6	0.30	5	0.716	0.482 (0.050)	0.710 (0.060)	0.713 (0.044)
0.6	0.30	10	0.716	0.547 (0.049)	0.709 (0.053)	0.716 (0.045)

†The herds have a fixed size of 100 animals. Average estimates and empirical standard errors (in parentheses) over 1000 simulations are given.

‡The prevalence of a positive herd defined as having at least one infected animal in the set of tested animals.

§The prevalence of a positive herd defined as having at least one animal in the herd being infected.

Assume that either all animals in a herd or a group (random sample) of animals from a herd are investigated. At the group or herd level, the sensitivity and specificity refer to respectively the probability that at least one of the animals tested positively if the group or herd is infected and the probability that none of the animals in the group or herd tested positively if the group or herd is not diseased. The group or herd sensitivity and specificity are functions of the number of animals in the group or herd (Martin *et al.*, 1992). In this section, we shall always refer to a herd, but this can also be a group of animals from a herd.

The herd specificity corresponding to m tested animals is, under the assumption of conditional independence, equal to

$$\begin{aligned}
 Sp^H &= P(Z=0|\text{none of the animals is infected}) \\
 &= \prod_{i=1}^m P(\text{test of animal } i \text{ is negative}|\text{animal } i \text{ is not infected}) \\
 &= \prod_{i=1}^m Sp^A = Sp^{Am}.
 \end{aligned}
 \tag{7}$$

As a result, the larger the number of animals in a herd, the smaller the herd specificity that is based on all individual test results. As $m \rightarrow \infty$, the herd specificity $Sp^H \rightarrow 0$. Fig. 2 shows herd specificity as a function of the size of herd (i.e. the number of tested animals), for an animal's test specificity of 0.99 (full curve), 0.95 (broken curve) and 0.90 (dotted curve). It can be seen that a decrease in the animal's test specificity drastically decreases the herd specificity of the test.

Herd sensitivity is less easy to calculate because of the definition of a herd test, which is positive if at least one of the test results of the animals is positive. However, the following equation can be used to derive the herd sensitivity:

$$p^H = Se^H \pi^H + (1 - Sp^H)(1 - \pi^H),$$

with Se^H the herd sensitivity, Sp^H the herd specificity, p^H the apparent herd prevalence and π^H the true herd prevalence, or

$$Se^H = \frac{p^H - (1 - Sp^{Am})(1 - \pi^H)}{\pi^H}. \tag{8}$$

This is similar to formula (4) where the animal characteristics are replaced by herd characteristics corresponding to the test results of m animals. This function can be rewritten in terms of the true animal prevalence, animal test sensitivity and test specificity by first rewriting the herd prevalences p^H and π^H in terms of the animal prevalences p^A and π^A by using expression (2), and then writing the animal prevalence p^A in terms of π^A by using the Rogan and Gladen estimate. This yields

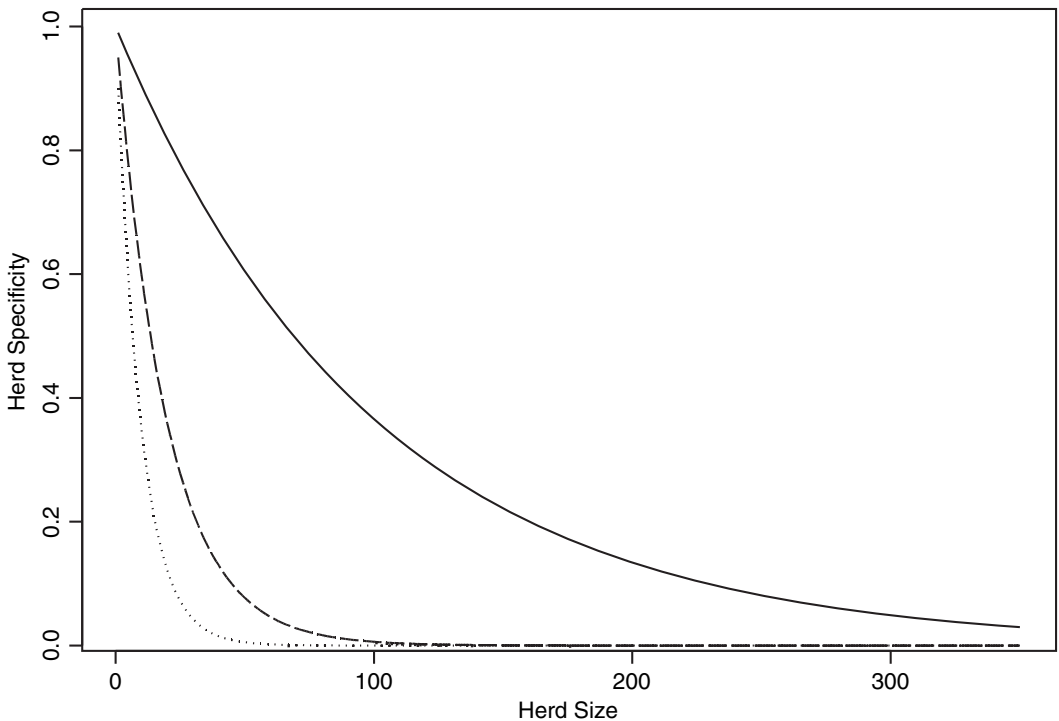


Fig. 2. Herd specificity for a herd test based on individual tests from animals in the herd: —, individual's test specificity of 0.99; - -, individual's test specificity of 0.95; ·····, individual's test specificity of 0.90

$$\text{Se}^{\text{H}} = \frac{1 + \prod_{k=0}^{m-1} \left\{ 1 + \frac{(\rho-1)\pi^{\text{A}}}{1 + \rho(k-1)} \right\} (\text{Sp}^m - 1) - \prod_{k=0}^{m-1} \left[1 + \frac{(\rho-1)\{\text{Se}\pi^{\text{A}} + (1-\text{Sp})(1-\pi^{\text{A}})\}}{1 + \rho(k-1)} \right]}{1 - \prod_{k=0}^{m-1} \left\{ 1 + \frac{(\rho-1)\pi^{\text{A}}}{1 + \rho(k-1)} \right\}}$$

For example, if the true animal prevalence is 0.25, the within-herd correlation is 0.20, the sensitivity of a single test is 0.97 and the specificity of a single test is 0.98 and 10 animals are being tested, then $\text{Se}^{\text{H}} = 0.94$, i.e. 94% of the affected herds will have at least one positively tested animal in the sample of 10 animals (from either a truly positive animal or a false positive animal). Thus, in general, the herd sensitivity is a complex function of the number of animals tested (m), the animals's test sensitivity Se^{A} and specificity Sp^{A} , and also of the animal's prevalence π^{A} and the correlation ρ . Some special situations can be considered.

- When the number of animals tested is 1, the herd sensitivity Se^{H} reduces to the animal sensitivity Se^{A} .
- If there is no correlation in the data, the formula reduces to equation (8) where we replace ρ^{H} by $1 - (1 - p^{\text{A}})^n$ and π^{H} by $1 - (1 - \pi^{\text{A}})^m$.
- When the specificity of a single test is 100% the herd sensitivity is equal to

$$\text{Se}^{\text{H}} = \frac{1 - \prod_{k=0}^{m-1} \left\{ 1 + \text{Se} \frac{(\rho-1)\pi^{\text{A}}}{1 + \rho(k-1)} \right\}}{1 - \prod_{k=0}^{m-1} \left\{ 1 + \frac{(\rho-1)\pi^{\text{A}}}{1 + \rho(k-1)} \right\}} = \frac{\rho^{\text{H}}}{\pi^{\text{H}}}.$$

In this situation, the true animal prevalence is always greater than the apparent animal prevalence, and this inequality also applies to the herd prevalence, and as a result $\text{Se}^{\text{H}} \leq 1$, as expected.

- When both the specificity and the sensitivity of a single test are 100%, the herd sensitivity is also 100%.

In Fig. 3 the sensitivity of a herd test is shown as a function of the sensitivity and specificity of a single test, the animal prevalence and within-herd correlation, and the number of tests performed. Figs 3(a) and 3(d), 3(b) and 3(e), and 3(c) and 3(f) correspond to herd sensitivity based on four, 10 and 50 tests respectively. In general, the sensitivity tends to increase with the number of tests performed. Figs 3(a)–3(c) show the dependence on the sensitivity and specificity of a single test. The sensitivity of a single test has a large effect on the herd sensitivity when only a small-to-moderate number of tests is used, with an increase in the herd sensitivity when the sensitivity of a single test increases. The effect of the specificity of a single test is not monotone and depends on the interplay between the number of animals tested and the sensitivity of a single test. Figs 3(d)–3(f) show the effect of within-herd correlation and the animal prevalence on herd sensitivity. Larger correlations and lower animal prevalences imply a smaller herd sensitivity.

6. Estimating prevalence of blue tongue virus in Belgium

To obtain unbiased estimates of prevalence of BT in Belgium, a cross-sectional serological survey targeting all Belgian ruminants was undertaken during the vector-free season. The study population of the winter screening consisted of dairy cattle that were more than 2 years old which were housed in dairy farms with on-farm delivery of dairy products. A one-stage sampling design

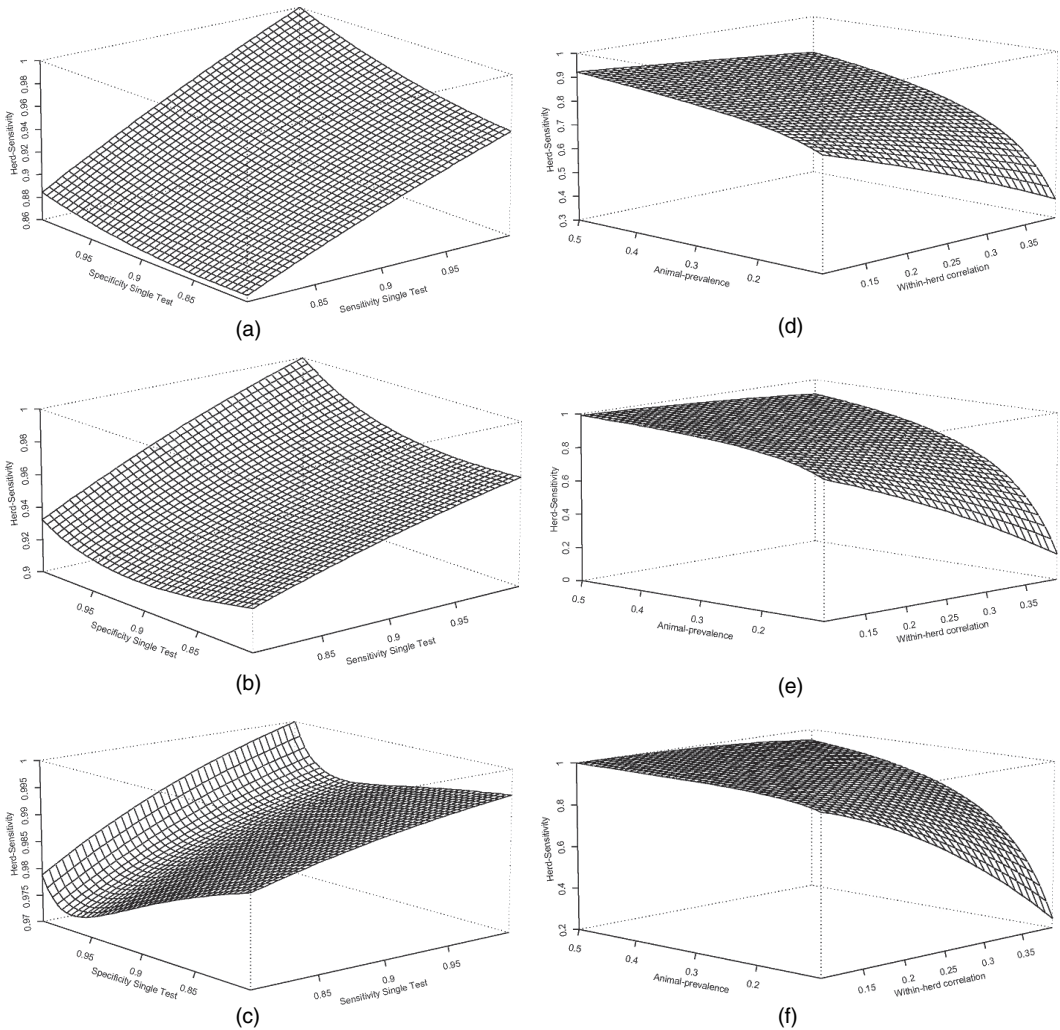


Fig. 3. Herd sensitivity for a herd test based on multiple individual tests, for (a)–(c) herd sensitivity as a function of sensitivity and specificity of a single test (the animal prevalence and within-herd correlation are kept fixed at 0.30 and 0.20 respectively) and (d)–(f) herd sensitivity as a function of animal prevalence and within-herd correlation (the sensitivity and specificity of a single test are kept fixed at 0.90): (a), (d) four tests; (b), (e) 10 tests; (c), (f) 50 tests

was performed with stratification of the herds by province and proportional allocation according to province area. An overview of the number of herds per province is given in Table 4. A total of 344 farms were sampled and all animals that were older than 2 years old were blood tested for the presence of antibodies against BTV serotype 8. Fig. 4 shows the distribution of herd sizes of the farms sampled. The serum samples were assayed by using a commercially available competitive enzyme-linked immuno-sorbent assay test kit.

Using realtime quantitative polymerase chain reaction as reference test, the diagnostic sensitivity and specificity of the competitive enzyme-linked immuno-sorbent assay test was estimated at 0.874 (95% credibility interval 0.835–0.904) and 0.990 (95% credibility interval 0.972–0.996) respectively (Vandenbussche *et al.*, 2007). On the basis of this information, beta distributions on

Table 4. Overview of the sampling design of BT winter screening

<i>Province</i>	<i>Number of herds</i>	<i>Number of sampled herds</i>	<i>Weight</i>
Antwerpen	4410	35	1.08141
Brabant-Wallon	707	10	0.60679
Hinaut	4156	50	0.71339
Limburg	2972	31	0.82282
Liège	3797	44	0.74064
Luxembourg	2968	34	0.74921
Namur	2518	45	0.48025
Oost-Vlaanderen	8095	34	2.04342
Vlaams-Brabant	2789	24	0.99737
West-Vlaanderen	7669	37	1.77892
Belgium	40081	344	

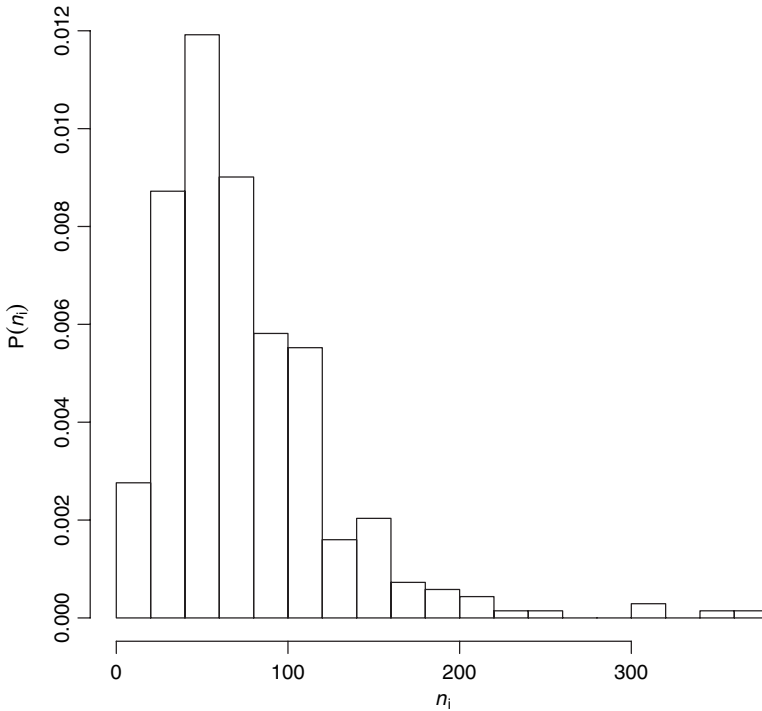


Fig. 4. Observed herd size probabilities based on BT winter screening

the sensitivity and specificity parameters were constructed (as explained in Section 3.1), yielding $Se \sim \text{beta}(733, 107)$ and $Sp \sim \text{beta}(731, 9)$.

The numbers of positive test results per herd (Z_i) are used to estimate the apparent and true animal and herd prevalences. It is assumed that Z_i follows a beta-binomial distribution with parameters p^A and ρ , as in equation (1). Because of the unequal sampling probabilities, some of the provinces might be overrepresented in the sample compared with the other provinces.

Therefore, a weighted log-likelihood with weights equal to the inverse of the sampling probabilities is used. The sampling probabilities are given by

$$w_i^{-1} = \frac{k_i \sum_i K_i}{K_i \sum_i k_i},$$

where k_i denote the number of herds sampled and K_i the total number of herds in province i . In this way, provinces which are underrepresented by the sample will receive a high weight, whereas provinces which were overrepresented in the sample will receive a small weight. The weights that were assigned to each province are given in the last column of Table 4. The apparent herd prevalence is obtained from the estimated parameters, using equation (3), and the true animal and herd prevalences are derived from the apparent herd prevalence and sensitivity and specificity of an animal’s test result, using equations (4) and (5). This analysis was conducted in a Bayesian framework (using WinBUGS), yielding not only point estimates of the animal and herd prevalences derived, but also full posterior distributions from which a mean and credibility interval for the prevalence can be estimated. This framework also allows us to specify a prior distribution for the parameters Se and Sp easily, instead of using a fixed value for them.

Results for the overall prevalence estimates in Belgium are displayed in the top panel of Table 5. The posterior mean and 95% credibility intervals are displayed. The correlation between the test results of two animals in a herd was estimated as 0.30 with a credibility interval (0.27,0.34). It is clear that the apparent animal and herd prevalence underestimate the true prevalence of BTV, and that such a correction cannot be ignored. The weighted and unweighted analyses give very similar results. Thus, in this setting, it seems that the sampling design does not have a large effect.

By inclusion of the province as a covariate in the model, it is possible to estimate the prevalence of BT per province, and as such to obtain an idea of the distribution of the disease in different parts of the country. The results are given in Table 6. The within-herd correlation is estimated as 0.16 with credibility interval (0.14,0.19). There is a large variation in the animal and

Table 5. Overall prevalence estimates of BTV in Belgium, based on the winter screening of 2006–2007†

Prevalence		Mean (%)	Credibility interval (lower,upper) (%)
<i>Without sampling design</i>			
Animal	Apparent p^A	21.57	(19.26,24.07)
	True π^A	23.67	(20.83,26.69)
Herd	Apparent p^H	82.60	(79.03,85.73)
	True π^H	85.27	(81.65,88.39)
<i>With sampling design</i>			
Animal	Apparent p^A	21.62	(19.30,24.12)
	True π^A	23.73	(20.90,26.81)
Herd	Apparent p^H	82.57	(79.03,85.76)
	True π^H	85.27	(81.66,88.49)

†The posterior mean and 95% credibility intervals of the apparent (uncorrected) and true (corrected) animal and herd prevalences are shown. The panels show the estimates without or with accounting for the sampling design.

Table 6. True prevalence estimates of BTV per province, based on the winter screening of 2006–2007†

Province	Animal prevalence (%)		Herd prevalence (%)		Population mean herd size
	Mean	Credibility interval (lower, upper)	Mean	Credibility interval (lower, upper)	
Antwerpen	31.66	(24.54,38.90)	97.94	(95.80,99.11)	87.94
Brabant-Wallon	29.72	(17.82,42.48)	97.39	(90.82,99.68)	82.50
Hinaut	4.64	(2.44,7.22)	46.42	(28.79,62.16)	71.34
Limburg	56.37	(47.84,65.03)	99.95	(99.83,99.99)	70.42
Liège	51.52	(44.19,58.93)	99.47	(99.17,99.68)	71.55
Oost-Vlaanderen	26.53	(20.24,33.53)	95.78	(91.58,98.31)	59.97
Vlaams-Brabant	26.55	(19.03,34.67)	95.03	(90.35,97.59)	67.54
West-Vlaanderen	7.70	(4.45,11.37)	60.57	(43.01,75.00)	56.57
Luxembourg	3.88	(1.38,6.95)	40.58	(17.87,60.65)	94.18
Namur	15.20	(11.02,19.97)	87.80	(79.69,93.39)	89.78

†The posterior mean and 95% credibility intervals of the true animal and herd prevalences are shown.

herd prevalences between the provinces. The herd prevalence is a function of both the animal prevalence and the population of herds (which is reflected in $P(n_i)$) per province. For most of the provinces, the herd prevalence (the probability of an infected herd) is very large. The animal prevalence is much lower, although in some provinces also the probability that an animal is infected is quite high.

In this survey, a one-stage sampling design was used. To reduce costs, it is of interest to know the effect of reducing the number of animals sampled per herd. We performed a simulation study, randomly selecting a fixed number of animals from each herd. As before, the animal prevalence and herd prevalence can then be calculated from this sample. The correction as proposed in Section 3 is used to obtain an estimate of the herd prevalence, which is defined as the probability of having a herd with at least one affected animal, and which can be compared with the prevalence that is obtained under the one-stage sampling design. Results are summarized in Table 7. It can be seen that the animal prevalence is very well estimated on the basis of the two-stage sampling design. For the two-stage design sampling five animals per herd,

Table 7. True prevalence estimates of BTV per province, based on the winter screening of 2006–2007†

Prevalence		Results for 5 animals		Results for 15 animals		Results for 30 animals	
		Mean	Credibility interval (lower, upper)	Mean	Credibility interval (lower, upper)	Mean	Credibility interval (lower, upper)
Animal	p^A	21.95	(18.94,24.87)	21.54	(18.96,24.27)	21.34	(18.98,24.16)
	π^A	24.11	(20.51,27.61)	23.64	(20.49,26.89)	23.41	(20.49,26.71)
Herd	p^H	49.98	(44.66,54.96)	67.52	(62.68,72.13)	75.11	(70.72,79.27)
	π^H	53.57	(47.53,59.10)	71.06	(65.84,75.93)	78.34	(73.76,82.59)
	$p^{H,corr}$	81.05	(74.53,86.45)	81.89	(77.56,85.89)	81.63	(77.59,85.31)
	$\pi^{H,corr}$	83.86	(77.61,88.92)	84.62	(80.22,88.55)	84.37	(80.34,87.97)

†The posterior mean and 95% credibility intervals of the true animal and herd prevalences are shown.

the credibility intervals are somewhat larger compared with those from the one-stage sampling design, but the difference is negligible when 15 or 30 animals are sampled per herd. This is in line with our expectations, since the amount of information in a clustered data setting is bounded (Faes *et al.*, 2009). We can conclude that, in this setting, a sample of size 15 has almost the same amount of information on the animal prevalence as all animals. Herd prevalence based on the definition of at least one positive sample increases with the sample size and underestimates the herd prevalence based on the one-stage sampling design. The corrected herd prevalence is very close to the herd prevalence that is based on the one-stage sample. Although almost no differences are seen in the point estimate, the precision increases as the number of animals tested increases. Similarly to the result for animal prevalence, the difference in credibility interval with the herd prevalence based on a one-stage design seems negligible when 15 or 30 animals are sampled, suggesting that sampling 15 animals would be sufficient to estimate the animal and herd prevalence appropriately.

On the basis of the results obtained for animal prevalence and correlation, we can estimate the performance of the multiple-testing procedure, by using formulae (8) and (7). On average, 75 animals were tested per herd. With an apparent herd prevalence of 82.57% and a true herd prevalence of 85.27%, the herd specificity is $Sp^H = 0.99^{75} = 0.47$ and the herd sensitivity is $Se^H = 0.88$. We note the large drop in herd specificity, with more than half of the non-diseased herds testing positively. In comparison, the herd sensitivity stays almost constant, with 12.31% of the herds affected falsely testing negatively.

7. Discussion

In this paper, use of the beta–binomial model is proposed to estimate the individual level and group level prevalence simultaneously. This model allows us to correct for

- (a) the *within-group correlation*,
- (b) the *diagnostic characteristics* of the used test and
- (c) the *sampling design*.

Often, the diagnostic characteristics are known only for the performance of a single test. Sometimes, the sensitivity and specificity of the test are used to correct the herd prevalence directly by using a Rogan–Gladen estimator. This paper shows that the test characteristics of a single test are not the same as for a group of tests, and this *ad hoc* procedure would lead to biased results. Using the beta–binomial model makes it possible to correct the group level prevalence, which is obtained from aggregate testing of all (or a subgroup of) individuals in the group, using the individual level characteristics of the test.

The sampling design also has a great effect on the herd prevalence. Indeed, direct calculation of at least one positive sample in a one-stage sampling design is different from direct calculation of at least one positive sample in a two-stage sampling design. The second calculation depends on the number of individuals sampled per group as well as on the number of individuals per group. This dependence makes it very difficult to compare the herd prevalences on the basis of a two-stage sampling design, especially if unequal group sizes are present in the population. Because of this difficulty, we recommend calculation of the herd prevalence defined as the probability of having at least one affected individual in the group from a two-stage sample. This method corrects for the number of individuals in the group and allows us to compare herd prevalence estimates from different sampling designs.

This paper raises questions concerning the design of a surveillance programme. In this paper it was shown that by the use of appropriate methods one could use a two-stage sampling design

instead of a one-stage sampling design to estimate herd prevalence. It was shown that one can reduce costs, without loss in precision in the animal and herd prevalence, by applying a two-stage sampling design instead of a one-stage design, as long as the number of animals sampled is sufficiently large. The appropriate sample size to estimate both animal prevalence and herd prevalence is an issue of interest. This concerns both the number of herds to be sampled and the number of animals within a herd. One could opt for a sampling design with many animals sampled in a small amount of herds, a small number of animals sampled in a large number of herds or some design in between. Using the effective sample size, representing the amount of information in a clustered data set (Faes *et al.*, 2009), it can be investigated how one can optimize the design to estimate both animal and herd prevalence. This is a topic of further research.

In this paper, we have classified a herd as positive when at least one animal in the herd (or in the sample) tested positively. In some settings, other cut-off values are used, e.g. assuming that a positive herd corresponds to at least two animals in the herd being positive. Although it was not discussed in the text, the method proposed can be easily generalized to the use of other cut-off values.

Acknowledgements

We thank the referees for their valuable suggestions that have led to an improved version of the manuscript. This work was supported by the Interuniversity Attraction Pole research network P6/03 of the Belgian Government (Belgian Science Policy) and Research Foundation Flanders.

Appendix A

The beta-binomial model can easily be programmed in a Bayesian framework, using the software package WinBUGS. The core code used in this paper is as follows.

```
# Use beta-binomial model to estimate apparent animal and herd prevalence
model{
  for (i in 1:344)
  {
    # Specification of beta-binomial model
    Z[i] ~ dbin(theta[i],N[i])
    theta[i] ~ dbeta(a,b)
    # Derivation of apparent herd prevalence
    for (k in 0:N[i])
    {
      hlp[i, (k+1)] <- log(1-paa + (k*paa*rho) / (1+(k-1)*rho))
    }
    q[i] <- 1 - exp(sum(hlp[i, 1:N[i]]))
  }
  a <- paa * (1/rho - 1)
  b <- (1-paa) * (1/rho - 1)
  # Estimation of intraclass correlation
  rho <- (exp(kappa) - 1) / (exp(kappa) + 1)
  # Estimation of apparent animal prevalence
  paa <- exp(beta) / (1 + exp(beta))
  # Estimation of apparent herd prevalence
  pha <- mean(q[1:344])
  # Specification of prior distributions for regression parameters
  beta ~ dnorm(0, 0.0001)
  kappa ~ dnorm(0, 0.0001)
}
```

When for example a stratified sampling design with unequal sampling probabilities is used to select the herds, it might be necessary to account for the sampling design. A weighted log-likelihood can then be used where each contribution of a herd is weighted with the inverse of the sampling probabilities. This can be done in WinBUGS by replacing the code

```
Z[i] ~ dbin(theta[i],N[i])
in
zeros[i] <- 0
logL[i]<-weight[i]*(Z[i]*log(theta[i])+(N[i]-Z[i])*log(1-theta[i]))
phi[i] <- -logL[i] + C
zeros[i] ~ dpois(phi[i])
```

where the weights are specified in the vector `weight[i]`.

On the basis of the above programs, the apparent animal and herd prevalence is obtained, corrected for the correlation of the test outcomes in animals from the same herd and for the sampling design of the surveillance programme. From the apparent animal and herd prevalence, we would like to derive the true animal and herd prevalence, correcting for possible misclassification of the test results. When Se and Sp are assumed to be known, the true animal and herd prevalence can be derived by adding the following equation to the code.

```
# Fixing the sensitivity and specificity
Se<-0.9
Sp<-0.9
# Derivation of true animal prevalence
pat<-(paa+Sp-1)/(Se+Sp-1)
# Derivation of true herd prevalence
for (i in 1:344)
{
for (k in 0:N[i])
{
hlp2[i,(k+1)]<-log(1-pat+(k*pat*rho)/(1+(k-1)*rho))
}
q2[i]<-1-exp(sum(hlp2[i,1:N[i]]))
pht<-mean(q2[1:344])
}
```

The true animal prevalence is derived from the Rodan–Gladen estimates (equation (4)), and the true herd prevalence is obtained from inserting the true animal prevalence into equation (3).

When the sensitivity and specificity are not known fixed values, a prior distribution on the sensitivity and specificity can be used instead. This can be done by replacing the code for the derivation of the true animal prevalence in the previous code by

```
# Derivation of true animal prevalence
pat<-(paa+Sp.cut-1)/(Se.cut+Sp.cut-1)
# Specification of prior distribution on Se and Sp
Se.cut<-cut(Se)
Sp.cut<-cut(Sp)
Se ~ dbeta(733.0418,107.1207)I(0.5,1)
Sp ~ dbeta(731.12,8.88)I(0.5,1)
```

The cut function is used because we wish to use the prior information on the sensitivity and specificity on the estimation of the animal prevalence but do not wish to estimate the sensitivity nor the specificity on the basis of the test results.

References

Aerts, M., Geys, H., Molenberghs, G. and Ryan, L. M. (2002) *Topics in Modeling of Clustered Data*. Boca Raton: Chapman and Hall.

- Baggesen, D. L., Wegener, H. C., Bager, F., Stege, H. and Christensen, J. (1996) Herd prevalence of *Salmonella enterica* infections in Danish slaughter pigs determined by microbiological testing. *Prev. Veter. Med.*, **26**, 201–213.
- Bahadur, R. R. (1961) A representation of the joint distribution of responses to n dichotomous items. In *Studies in Item Analysis and Prediction* (ed. H. Solomon). Stanford: Stanford University Press.
- Bollaerts, K., Aerts, M., Hens, N., Shkedy, Z., Faes, C., Van der Stede, Y. and Beutels, P. (2009) Estimating epidemiological parameters by using mixture-models for serological data. To be published.
- Bolstad, W. M. (2005) *Introduction to Bayesian Statistics*. Hoboken: Wiley.
- Christensen, J. and Gardner, I. A. (2000) Herd-level interpretation of test results for epidemiologic studies of animal diseases. *Prev. Veter. Med.*, **45**, 83–106.
- Daniels, P. W., Sendow, I., Pritchard, L. I., Sukarsih, and Eaton, B. T. (2004) Regional overview of bluetongue viruses in South-East Asia: viruses, vectors and surveillance. *Veter. Ital.*, **40**, 94–100.
- Declerck, L., Aerts, M. and Molenberghs, G. (1998) Behaviour of the likelihood ratio test statistic under a Bahadur model for exchangeable binary data. *J. Statist. Comput. Simul.*, **61**, 15–38.
- European Commission Reference Laboratory (2006) Bluetongue virus in the Netherlands identified as serotype 8 by Institute for Animal Health, Promed, August 28. European Commission Reference Laboratory. (Available from <http://www.promedmail.org>, archive 20060828.2448.)
- Faes, C., Molenberghs, G., Aerts, M., Verbeke, G. and Kenward, M. G. (2009) The effective sample size and a novel small sample degrees of freedom method. *Am. Statist.*, **63**, 389–399.
- Gerdes, G. H. (2004) A South African overview of the virus, vectors, surveillance and 472 unique features of bluetongue. *Veter. Ital.*, **40**, 39–42.
- Grubbs, F. E. (1962) Attempts to validate certain PERT statistics or picking on PERT. *Ops Res.*, **10**, 912–915.
- Hassan, A. (1992) Epidemiology of bluetongue virus infection in Malaysia. In *Bluetongue, African Horse Sickness and Related Orbiviruses* (eds T. E. Walton and B. I. Osburn), pp. 155–161. Boca Raton: CRC Press.
- Herniman, K. A., Gumm, I. D., Owen, L., Taylor, W. P. and Sellers, R. F. (1980) Distribution of Bluetongue viruses and antibodies in some countries of the eastern hemisphere. *Bull. Off. Int. Epizoot.*, **92**, 581–586.
- Humphry, R. W., Cameron, A. and Gunn, G. J. (2004) A practical approach to calculate sample size for herd prevalence surveys. *Prev. Veter. Med.*, **65**, 173–188.
- Jordan, D. (1996) Aggregate testing for the evaluation of Johne's disease herd status. *Aust. Veter. J.*, **36**, 187–209.
- Kleinman, J. C. (1973) Properties with extraneous variance: single and independent samples. *J. Am. Statist. Ass.*, **68**, 46–54.
- Lunn, D. J., Thomas, A., Best, N. and Spiegelhalter, D. (2000) WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Statist. Comput.*, **10**, 325–337.
- Martin, S. W., Shoukri, M. and Thorburn, M. A. (1992) Evaluating the health status of herds based on tests applied to individuals. *Prev. Veter. Med.*, **14**, 33–43.
- Méroc, E., Faes, C., Herr, C., Verheyden, B., Van Binst, T., Vandenbussche, F., Hooyberghs, J., Aerts, M., De Clercq, K. and Mintiens, K. (2008) Establishing the spread of bluetongue virus during the 2006 epizootic in Belgium. *Veter. Microbiol.*, **131**, 133–144.
- Mo, C. L., Thompson, L. H., Homan, E. J., Oviedo, M. T., Greiner, E. C., Gonzales, J. and Saenz, M. R. (1994) Bluetongue virus isolations from vectors and ruminants in Central America and the Caribbean. *Am. J. Veter. Res.*, **55**, 211–215.
- Molenberghs, G. and Verbeke, G. (2005) *Models for Discrete Longitudinal Data*. New York: Springer.
- Rogan, W. J. and Gladen, B. (1978) Estimating prevalence from the results of a screening test. *Am. J. Epidemiol.*, **107**, 71–76.
- Skellam, J. G. (1948) A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. *J. R. Statist. Soc. B*, **10**, 257–261.
- Vandenbussche, F., Vanbinst, T., Verheyden, B., Van Dessel, W., Demeestere, L., De Poorter, G., Houdart, P., Bertels, G., Praet, N., Berkvens, D., Mintiens, K., Goris, N. and De Clercq, K. (2007) Evaluation of antibody-563 ELISA and real time RT-PCR for the diagnosis and profiling of bluetongue virus serotype 8 during the epidemic in Belgium in 2006. *Veter. Microbiol.*, to be published.
- Williams, D. A. (1982) Extra-binomial variation in logistic linear models. *Appl. Statist.*, **31**, 144–148.