

1 A modelling framework for the prediction of  
2 the herd-level probability of infection from  
3 longitudinal data

4 Aurélien Madouasse<sup>1</sup>, Mathilde Mercat<sup>1</sup>, Annika van Roon<sup>2</sup>,  
5 David Graham<sup>3</sup>, Maria Guelbenzu<sup>3</sup>, Inge Santman Berends<sup>2, 4</sup>,  
6 Gerdien van Schaik<sup>2, 4</sup>, Mirjam Nielen<sup>2</sup>, Jenny Frössling<sup>5</sup>, Estelle  
7 Ågren<sup>5</sup>, Roger Humphry<sup>6</sup>, Jude Eze<sup>6</sup>, George Gunn<sup>6</sup>, Madeleine  
8 Henry<sup>6</sup>, Jörn Gethmann<sup>7</sup>, Simon J. More<sup>8</sup>, and Christine  
9 Fourichon<sup>1</sup>

10 <sup>1</sup>BIOEPAR, INRA, Oniris, La Chantrerie, Nantes 44307, France

11 <sup>2</sup>Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht  
12 University, PO Box 80151, 3508, TD Utrecht, the Netherlands

13 <sup>3</sup>Animal Health Ireland, Unit 4/5, The Archways, Bridge St.,  
14 Carrick-on-Shannon, Co. Leitrim N41 WN27, Ireland

15 <sup>4</sup>GD Animal Health, PO Box 9, 7400 AA, Deventer, the Netherlands

16 <sup>5</sup>Department of Disease Control and Epidemiology, National Veterinary Institute  
17 (SVA), 751 89 Uppsala, Sweden

18 <sup>6</sup>Scotland's Rural College, Kings Buildings, West Mains Road, Edinburgh, EH9  
19 3JG, United Kingdom

20 <sup>7</sup>Institute of Epidemiology, Friedrich-Loeffler-Institute, Südufer 10, 17493  
21 Greifswald, Germany

22 <sup>8</sup>Centre for Veterinary Epidemiology and Risk Analysis, UCD School of  
23 Veterinary Medicine, University College Dublin, Belfield, Dublin D04 W6F6,  
24 Ireland

25 July 20, 2020

26

## Abstract

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

For many infectious diseases of farm animals, there exist collective control programmes (**CPs**) that rely on the application of diagnostic testing at regular time intervals for the identification of infected animals or herds. The diversity of these CPs complicates the trade of animals between regions or countries because the definition of freedom from infection differs from one CP to another. In this paper, we describe a statistical model for the prediction of herd level probabilities of infection from longitudinal data collected as part of CPs against infectious diseases of cattle. The model was applied to data collected as part of a CP against infections by the bovine viral diarrhoea virus (**BVDV**) in Loire-Atlantique, France. The model represents infection as a herd latent status with a monthly dynamics. This latent status determines test results through test sensitivity and test specificity. The probability of becoming status positive between consecutive months is modelled as a function of risk factors (when available) using logistic regression. Modelling is performed in a Bayesian framework. Prior distributions need to be provided for the sensitivities and specificities of the different tests used, for the probability of remaining status positive between months as well as for the probability of becoming positive between months. When risk factors are available, prior distributions need to be provided for the coefficients of the logistic regression in place of the prior for the probability of becoming positive. From these prior distributions and from the longitudinal data, the model returns posterior probability distributions for being status positive in all herds on the current months. Data from the previous months are used for parameter estimation. The impact of using different prior distributions and model settings on parameter estimation was evaluated using the data. The main advantage of this model is its ability to predict a probability of being status positive on a month from inputs that can vary in terms of nature of test, frequency of testing and risk factor availability. The main challenge in applying the model to the BVDV CP data was in identifying prior distributions, especially for test characteristics, that corresponded to the latent status of interest, i.e. herds with at least one persistently infected (**PI**) animal. The model is available on Github as an R package (<https://github.com/AurMad/STOCfree>).

## 62 1 Introduction

63 For many infectious diseases of farm animals, there exist collective control  
64 programmes that rely the application of diagnostic testing at regular time  
65 intervals for the identification of infected animals or herds. In cattle, such dis-  
66 eases notably include infection by the bovine viral diarrhoea virus (**BVDV**)  
67 or by *Mycobacterium avium* subspecies *paratuberculosis* (MAP). These con-  
68 trol programmes (**CP**)s are extremely diverse. Their objective can range  
69 from decreasing the prevalence of infection to eradication. Participation in  
70 the CP can be voluntary or compulsory. The qualification of herds regarding  
71 infection can be based on a wide variety of testing strategies in terms of the  
72 nature of the tests used (identification of antibodies vs. identification of the  
73 agent), the groups of animals tested (e.g. breeding herd vs. young animals),  
74 number of animals tested, frequency of testing (once to several times a year,  
75 every calf born...). Even within a single CP, surveillance modalities may  
76 evolve over time. Such differences in CPs were described by [van Roon \*et al.\*](#)  
77 ([2020b](#)) for programmes targeting BVDV infections and by [Whittington \*et al.\*](#)  
78 ([2019](#)) for programmes against MAP.

79 Differences in surveillance modalities can be problematic when purchas-  
80 ing animals from areas with different CPs because the free status assigned  
81 to animals or herds might not be equivalent between CPs. A standardised  
82 method for both describing surveillance programmes and estimating confi-  
83 dence of freedom from surveillance data would be useful when trading animals  
84 across countries or regions. While inputs can vary between programmes, the  
85 output needs to be comparable across programmes. This is called output-  
86 based surveillance ([Cameron, 2012](#)). Probabilities measure both the chance  
87 of an event and the uncertainty around its presence/occurrence. If well de-  
88 signed, a methodology to estimate the probability of freedom from infection  
89 would meet the requirements of both providing a confidence of freedom from  
90 infection as well as of being comparable whatever the context.

91 Currently, the only quantitative method used to substantiate freedom  
92 from infection to trading partners is the scenario tree method ([Martin \*et al.\*,](#)  
93 [2007](#)). The method is applied to situations where there is a surveillance  
94 programme in place, with no animals or herds confirmed positive on testing.  
95 Scenario trees are based on the premise that it is impossible to prove that  
96 a disease is totally absent from a territory unless the entire population is  
97 tested with a perfect test. What is estimated with the scenario tree method  
98 is the probability that the infection would be detected in the population if it

99 were present at a chosen *design prevalence*. The output from this approach  
100 is the probability that the infection prevalence is not higher than the design  
101 prevalence given the negative test results (Cameron, 2012). Therefore, this  
102 method is well suited for those countries that are free from infection and that  
103 want to quantify this probability of freedom from infection for the benefit of  
104 trading partners (Norström *et al.*, 2014).

105 The scenario tree method is not adapted to countries or regions where  
106 there is a CP against an infectious disease which is still present. In such  
107 a context, only herds that have an estimated probability of freedom from  
108 infection that is deemed sufficiently high or, equivalently, a probability of  
109 infection that is deemed sufficiently low, would be safe to trade with. Identifying  
110 these herds involves estimating a probability of infection for each herd  
111 in the CP and then defining a decision rule to categorise herds as uninfected  
112 or infected based on these estimated probabilities.

113 In this paper, we propose a method to estimate herd level probabilities  
114 of infection from heterogeneous longitudinal data generated by CPs. The  
115 method predicts herd-month level probabilities of being latent status positive  
116 from longitudinal data collected in CPs. The input data are test results, and  
117 associated risk factors when available. Our main objective is to describe this  
118 modelling framework by showing how surveillance data are related to the  
119 *probabilities of infection* (strictly speaking, *probabilities of being latent status*  
120 *positive*) and by providing details regarding the statistical assumptions that  
121 are made. A secondary objective is to estimate these probabilities of being  
122 latent status positive, using different definitions for the latent status, from  
123 surveillance data collected as part of a CP against the infection by the BVDV  
124 in Loire-Atlantique, France. The challenges of defining prior distributions  
125 and the implications of using different prior distributions are discussed. R  
126 functions to perform the analyses described in this paper are gathered in an  
127 R package which is available from GitHub ([https://github.com/AurMad/](https://github.com/AurMad/STOCfree)  
128 [STOCfree](https://github.com/AurMad/STOCfree)).

## 129 2 Materials and methods

### 130 2.1 Description of the model

#### 131 2.1.1 Conceptual representation of surveillance programmes

132 Surveillance programmes against infectious diseases can be seen as imper-  
133 fect repeated measures of a true status regarding infection. In veterinary  
134 epidemiology, the issue of imperfect testing has traditionally been addressed  
135 using latent class models. With this family of methods, the true status re-  
136 garding infection is modelled as an unobserved quantity which is linked to  
137 test results through test sensitivity and specificity. Most of the literature on  
138 the subject is on estimating both test characteristics and infection prevalence  
139 (Collins & Huynh, 2014). For the estimations to work, the same tests should  
140 be used in different populations (Hui & Walter, 1980), the test characteristics  
141 should be the same among populations and test results should be condition-  
142 ally independent given the infection status (Toft *et al.*, 2005; Johnson *et al.*,  
143 2009). Latent class models can also be used to estimate associations between  
144 infection, defined as the latent class, and risk factors when the test used is  
145 imperfect (Fernandes *et al.*, 2019). In the study by Fernandes *et al.* (2019),  
146 the latent class was defined using a single test, through the prior distribu-  
147 tions put on sensitivity and specificity. When using latent class models with  
148 longitudinal data, the dependence between successive test results in the same  
149 herds must be accounted for. In the context of estimating test characteristics  
150 and infection prevalence from 2 tests in a single population from longitudi-  
151 nal data, Nusinovici *et al.* (2015) proposed a Bayesian latent class model  
152 which incorporated 2 parameters for new infection and infection elimination.  
153 The model we describe below combines these different aspects of latent class  
154 modelling into a single model.

155 We propose to use a class of models called Hidden Markov Models (HMM,  
156 see Zucchini *et al.* (2017)). Using surveillance programmes for infectious dis-  
157 eases as an example, the principles of HMMs can be described as follows:  
158 the latent status (*class*) of interest is a herd status regarding infection. This  
159 status is evaluated at regular time intervals: HMMs are discrete time mod-  
160 els. The status at a given time only depends on the status at the previous  
161 time (Markovian property). The status of interest is not directly observed,  
162 however, there exists some quantity (such as test results) whose distribu-  
163 tion depends on the unobserved status. HMMs have been used for decades

164 in speech recognition (Rabiner, 1989) and other areas. They have also been  
165 used for epidemiological surveillance (Le Strat & Carrat, 1999), although not  
166 with longitudinal data from multiple epidemiological units such as herds. The  
167 model we developed is therefore a latent class model that takes into account  
168 the time dynamics in the latent status. The probability of new infection  
169 between consecutive time steps is modelled as a function of risk factors.

170 Figure 1 shows how surveillance programmes are represented in the model  
171 as a succession of discrete time steps. The focus of this model is a latent  
172 status evaluated at the herd-month level. This latent status is not directly  
173 observed but inferred from its causes and consequences incorporated as data.  
174 The consequences are the test results. Test results do not have to be available  
175 at every time step for the model to work. The causes of infection are risk  
176 factors of infection. In the application presented below, the latent status  
177 will be either herd seropositivity or presence of a PI animal in the herd,  
178 depending on the testing scheme as well as on the prior distributions put  
179 on the characteristics of the tests used. The model estimates this latent  
180 status monthly, and predicts it for the last month of data. These herd-  
181 month latent statuses will be estimated/predicted from test results (BTM  
182 ELISA testing or confirmatory testing) and risk factors (cattle introductions  
183 or local seroprevalence) recorded in each herd.

### 184 2.1.2 Modelling framework, inputs and outputs

185 The model is designed to use longitudinal data collected as part of surveil-  
186 lance programmes against infectious diseases. In such programmes, each herd  
187 level status is re-evaluated when new data (most commonly test results, but  
188 may also be data related to risk factors) are available. The model mimics  
189 this situation by predicting the probability of a positive status for all herds  
190 in the CP on the last month of available data. Data from all participating  
191 herds up to the month of prediction are used as historical data for parameter  
192 estimation (Figure 1).

193 The estimation and prediction are performed within a Bayesian frame-  
194 work using Markov Chain Monte Carlo (MCMC) in the JAGS computer pro-  
195 gramme (Plummer, 2017). The model encodes the relationships between all  
196 the variables of interest in a single model. Each variable is modelled as drawn  
197 from a statistical distribution. The estimation requires prior distributions for  
198 all the parameters in the model. These priors are a way to incorporate either  
199 existing knowledge or hypotheses in the estimation. For example, we may

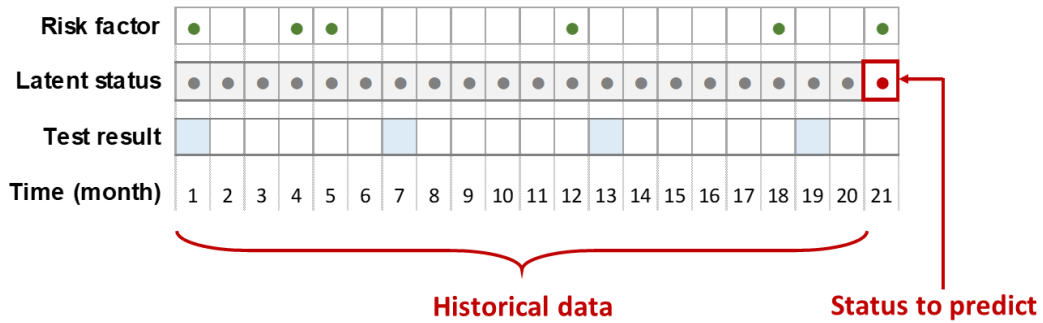


Figure 1: Conceptual representation of the implementation of a surveillance programme within a herd. The focus of the model is the latent status regarding infection, which is modelled at the herd-month level. This status partly depends on risk factors and determines test results. In this diagram, risk factors are represented as green dots when present and available test results as blue shaded squares. The model predicts a probability of infection for the most recent month in the surveillance programme using all the data collected for the estimation of model parameters.

200 know that the prevalence of herds infected with BVDV in our CP is probably  
201 lower than 20%, certainly lower than 30% and greater than 5%. Such con-  
202 straints can be specified with a Beta distribution. The Beta distribution is  
203 bounded between 0 and 1, with 2 parameters  $\alpha$  and  $\beta$  determining its shape.  
204 With the constraints specified above, we could use as a prior distribution a  
205  $Beta(\alpha = 15, \beta = 100)$ <sup>1</sup>. If we do not know anything about this infection  
206 prevalence (which is rare), we could use a  $Beta(\alpha = 1, \beta = 1)$  prior, which is  
207 uniform between 0 and 1. From the model specification, the prior distribu-  
208 tions and the observed data, the MCMC algorithm draws samples from the  
209 posterior distributions of all the variables in the model. These posterior dis-  
210 tributions are the probability distributions for the model parameters given  
211 the data and the prior distributions. MCMC methods are stochastic and  
212 iterative. Each iteration is a set of samples from the joint posterior distri-  
213 butions of all variables in the model. The algorithm is designed to reach the  
214 target joint posterior distribution, but at any moment, there is no guarantee  
215 that it has done. To overcome this difficulty, several independent instances  
216 of the algorithm (i.e. several chains) are run in parallel. For a variable, if  
217 all the MCMC draws from the different chains have the same distribution, it  
218 can be concluded that the algorithm has reached the posterior distribution.  
219 In this case, it is said that the model has converged.

220 The focus of our model is the monthly latent status of each herd. This  
221 latent status depends on the data on occurrence of risk factors and it affects  
222 test results. The data used by the model are the test results and risk factors.  
223 At each iteration of the MCMC algorithm, given the data and priors, a herd  
224 status (0 or 1) and the coefficients for the associations between risk factors,  
225 latent status and test results are drawn from their posterior distribution.

226 In the next 3 sections, the parameters for which prior distributions are  
227 required, i.e. test characteristics, status dynamics and risk factor parameters,  
228 are described. The outputs of Bayesian models are posterior distributions for  
229 all model parameters. Specifically, in our model, the quantities of interest  
230 are the herd level probabilities of being latent status positive on the last  
231 test month in the dataset as well as test sensitivity, test specificity, infection  
232 dynamic parameters and parameters for the strengths of association between  
233 risk factors and the probability of new infection. This is described in the

---

<sup>1</sup>The  $Beta(\alpha = 15, \beta = 100)$  distribution has a mean of 0.13 and a standard deviation of 0.03. In R, it can be plotted using the following instructions `curve(dbeta(x, 15, 100))`



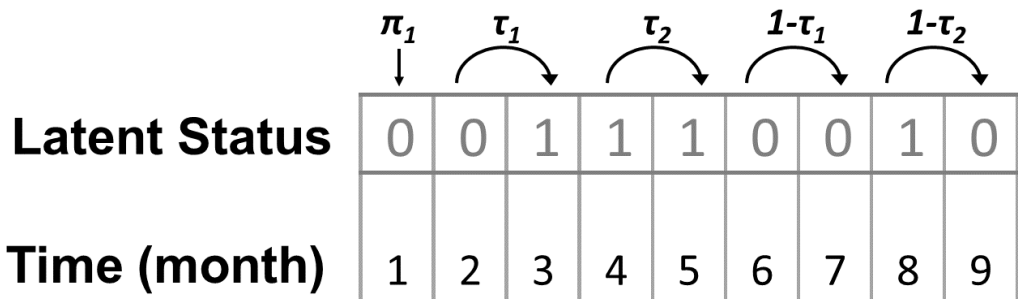


Figure 2: Modelling of infection dynamics. The diagram shows hypothetical latent statuses (0 for negative; 1 for positive) as a function of time in month, with examples of all possible transitions.  $\pi_1 = p(S_1^+)$  is the probability of being status positive at the first point in time,  $\tau_1 = p(S_t^+|S_{t-1}^-)$  is the probability of becoming status positive and  $\tau_2 = p(S_t^+|S_{t-1}^+)$  is the probability of remaining status positive.

234 corresponding sections.

### 235 2.1.3 Latent status dynamics

236 Between test events, uninfected herds can become infected and infected herds  
 237 can clear the infection. The model represents the probability of having a  
 238 positive status at each time step as a function of the status at the previous  
 239 time step (Figure 2). For the first time step when herd status is assigned,  
 240 there is no previous status against which to evaluate change. From the second  
 241 time step when herd status is assigned, and onwards, herds that were status  
 242 negative on the previous time step have a certain probability of becoming  
 243 status positive and herds that were status positive have a certain probability  
 244 of remaining status positive.

245 These assumptions can be summarised with the following set of equa-  
 246 tions<sup>2</sup>. The status on the first time step ( $S_1^+$ ) is a Bernoulli event with a  
 247 Beta prior on its probability of occurrence:

$$S_1^+ \sim \text{Bernoulli}(p(S_1^+)) \quad (1)$$

<sup>2</sup>Statuses are estimated/predicted at the herd-month level. Herd is omitted from the notation to facilitate reading.  $S_t^+$  should be read as  $S_{ht}^+$  where  $h$  represents the herd.

248

$$p(S_1^+) \sim \text{Beta}(\pi_{1a}, \pi_{1b}) \quad (2)$$

249 From the second time step when herd status is assigned, and onwards,  
250 a positive status is also a Bernoulli event ( $S_t^+$ ) with a probability of occur-  
251 rence that depends on the status at the previous time step as well as on  
252 the probability of becoming status positive and the probability of remaining  
253 status positive. In this case, the probability of becoming status positive is  
254  $\tau_1 = p(S_t^+ | S_{t-1}^-)$  and the probability of remaining positive is  $\tau_2 = p(S_t^+ | S_{t-1}^+)$ .

$$S_t^+ \sim \text{Bernoulli}(p(S_t^+)) \quad (3)$$

255

$$p(S_t^+) = (1 - S_{t-1}^+) \tau_1 + S_{t-1}^+ \tau_2 \quad (4)$$

256

$$\tau_1 \sim \text{Beta}(\tau_{1a}, \tau_{1b}) \quad (5)$$

257

$$\tau_2 \sim \text{Beta}(\tau_{2a}, \tau_{2b}) \quad (6)$$

258 Therefore, the status dynamics can be completely described by  $p(S_1^+)$ ,  $\tau_1$   
259 and  $\tau_2$ .

#### 260 **2.1.4 Incorporation of information on risk factors for new infec-** 261 **tion**

262 The probability of new infection is not the same across herds. For example,  
263 herds that introduce a lot of animals or are in areas where infection preva-  
264 lence is high could be at increased risk of new infection (Qi *et al.*, 2019).  
265 Furthermore, the association between a given risk factor and the probability  
266 of new infection could be CP dependent. For example, the probability of  
267 introducing infection through animal introductions will depend on the infec-  
268 tion prevalence in the population from which animals are introduced. As a  
269 consequence, estimates for these associations (as presented in the literature)  
270 could provide an indication about their order of magnitude, but their preci-  
271 sion may be limited. On the other hand, the CPs which are of interest in this  
272 work usually generate large amounts of testing data which could be used to  
273 estimate the strengths of association between risk factors and new infections  
274 within a given CP. The variables that are associated with the probability of  
275 new infection could increase the sensitivity and timeliness of detection.

276 When risk factors for new infection are available, the model incorporates  
277 this information by modelling  $\tau_1$  as a function of these risk factors through  
278 logistic regression, instead of the prior distribution for  $\tau_1$ .

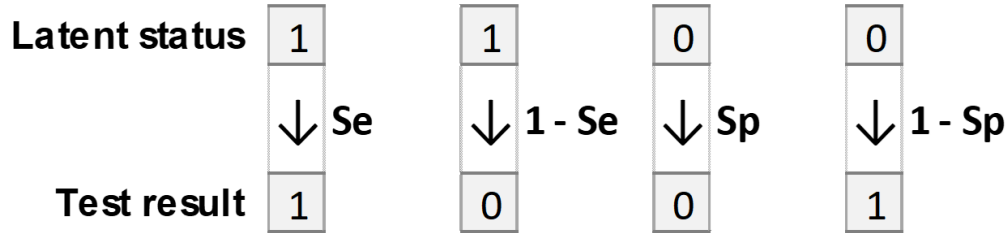


Figure 3: Relation of the model latent status to test result. Sensitivity is the probability of a positive test result in a status positive herd. Specificity is the probability of a negative test result in a status negative herd.

$$\text{logit}(\tau_{1ht}) = X_{ht}\theta \quad (7)$$

279 where  $X_{ht}$  is a matrix of predictors for herd  $h$  at time  $t$  and  $\theta$  is a vector  
 280 of coefficients. Normal priors are used for the coefficients of the logistic  
 281 regression.

$$\theta_i \sim \text{Normal}(\mu_i, \sigma_i) \quad (8)$$

### 282 2.1.5 Test characteristics

283 The model allows the inclusion of several test types but for the sake of clarity,  
 284 we show the model principles for only one test type. These principles can be  
 285 extended to several tests by specifying prior distributions for all tests.

286 Tests are modelled as imperfect measures of the latent status (Figure 3).  
 287 Test sensitivity is the probability of a positive test result given a positive  
 288 latent status ( $Se = p(T^+|S^+)$ , refers to true positives) and test specificity  
 289 is the probability of a negative test result given a negative latent status  
 290 ( $Sp = p(T^-|S^-)$ , refers to true negatives).

291 Test result at time  $t$  is modelled as a Bernoulli event with probability  
 292  $p(T_t^+)$  of being positive.

$$T^+ \sim \text{Bernoulli}(p(T_t^+)) \quad (9)$$

293 The relation between the probability of testing positive, the probability  
294 of a positive status, test sensitivity and test specificity is the following:

$$p(T_t^+) = S_t^+ Se + (1 - S_t^+)(1 - Sp) \quad (10)$$

295 Information or hypotheses regarding test characteristics are incorporated  
296 in the model as priors modelled by Beta distributions:

$$Se \sim Beta(Se_a, Se_b) \quad (11)$$

297

$$Sp \sim Beta(Sp_a, Sp_b) \quad (12)$$

298 It is important to note that the prior distributions used for sensitivity  
299 and specificity will determine what the latent status is. As an example, we  
300 consider the detection of BVDV infection with a test that detects BVDV  
301 specific antibodies in bulk tank milk. BVDV infection is associated with a  
302 long lasting antibody production. There can be cows that are seropositive  
303 long after the last PI animal has left the herd. In this situation, using a value  
304 of 1 for specificity will define the latent status as any herd with antibody  
305 positive cows. However, the herd-level specificity of the test, defined as the  
306 probability of a negative test result in a herd with no PI animals, is lower  
307 than the animal-level specificity defined as the probability of a negative test  
308 result in a sample from an non-PI animal. The specificity of interest, i.e. the  
309 detection of farms with PI animals, will depend on the proportion of antibody  
310 positive lactating dairy herds that are in farms with PI animals. In turn, this  
311 will depend on many factors that are CP dependent such as the prevalence of  
312 infection or the proportion of farms that use vaccination against the BVDV.  
313 With antibody testing alone, it is therefore difficult to define accurate prior  
314 distributions for sensitivity and specificity for the detection of farms with PI  
315 animals.

316 However, it is possible to align the meaning of the latent status with the  
317 status of interest. In most CPs, positive routine tests will be followed by con-  
318 firmatory testing. The objective of routine testing is to detect any potentially  
319 infected herd. The tests used for routine testing should be sensitive. The  
320 objective of confirmatory testing is to identify truly infected herds among  
321 herds positive in routine testing. The testing procedure used for confirma-  
322 tory testing should be both specific and sensitive. With our model, if these  
323 conditions are met and if prior distributions that reflect these hypotheses are  
324 used, the posterior distributions for the characteristics of both testing phases

325 should be more accurate. A useful property of HMMs is that accounting for  
 326 the status dynamics makes the results of tests performed on different months  
 327 in the same herd conditionally independent, because the conditional time de-  
 328 pendence between statuses is modelled with the dynamics part of the model.  
 329 For example, if a herd tests positive during routine testing, it will have a  
 330 higher than average prior probability of infection in subsequent confirmatory  
 331 testing. As a further consequence of this, the posterior distribution for the  
 332 specificity of routine testing will depend on the proportion of herds that are  
 333 confirmed positive in confirmatory testing.

### 334 2.1.6 Prediction of a probability of infection

335 In explaining how predictions are performed we use the following notation:  
 336  $\tilde{y}$  is the predicted value for  $y$ ,  $\hat{\beta}$  is the estimated value for  $\beta$ . The equation  
 337  $\tilde{y} = \hat{\beta}.x$  means that the predicted value for  $y$  is equal to  $x$  (data) times the  
 338 estimated value for  $\beta$ .

339 The model predicts herd-level probabilities of infection on the last month  
 340 in the data mimicking regular re-evaluation as new data come in. If there  
 341 is no test result available on this month, the predicted probability of being  
 342 status positive (called  $p(\tilde{S}_t^{+*})$ ) is the predicted status on the previous month  
 343 times  $\tilde{\tau}_{1t}$  if the herd was predicted status negative or times  $\hat{\tau}_2$  if the herd was  
 344 predicted status positive (Table 1)<sup>3</sup>. This can be written as:

$$p(\tilde{S}_t^{+*}) = p(\tilde{S}_t^+ | \hat{S}_{t-1}^+, \tilde{\tau}_{1t}, \hat{\tau}_2) = (1 - \hat{S}_{t-1}^+) \cdot \tilde{\tau}_{1t} + \hat{S}_{t-1}^+ \cdot \hat{\tau}_2 \quad (13)$$

345 where:

$$\tilde{\tau}_{1t} = \text{logit}^{-1}(X_t \hat{\theta}) \quad (14)$$

346 If a test result was available, the prediction must combine information  
 347 from the test as well as previous information. The way to estimate this pre-  
 348 dicted probability from  $p(\tilde{S}_t^{+*})$  and test results can be derived from Table 1.  
 349 The predicted probability of being status positive can be computed as:

$$p(\tilde{S}_t^+ | T_t^+, \tilde{S}_t^{+*}) = T_t^+ \cdot \frac{Se \cdot p(\tilde{S}_t^{+*})}{Se \cdot p(\tilde{S}_t^{+*}) + (1 - Sp)(1 - p(\tilde{S}_t^{+*}))} + (1 - T_t^+) \cdot \frac{(1 - Se) \cdot (1 - p(\tilde{S}_t^{+*}))}{(1 - Se) \cdot (1 - p(\tilde{S}_t^{+*})) + Sp \cdot (1 - p(\tilde{S}_t^{+*}))} \quad (15)$$

<sup>3</sup>Here  $\tilde{\tau}_{1t}$  is *predicted* from herd-month specific risk factors while  $\hat{\tau}_2$  is the same for all herds and *estimated* from historical data.

Table 1: Probability of test result by herd status. Cells on the first row are test positive herds with true positives on the left-hand side and false positives on the right-hand side. Cells on the second row are test negative herds with false negatives on the left-hand side and true negatives on the right-hand side.

		<i>Herdstatus<sub>t</sub></i>	
		+	-
<i>T<sub>est</sub><sub>t</sub></i>	+	$Se.p(S_t^+)$	$(1 - Sp)(1 - p(S_t^+))$
	-	$(1 - Se).p(S_t^+)$	$Sp.(1 - p(S_t^+))$

350 where  $T_t^+ = 1$  when the test at time  $t$  is positive,  $T_t^+ = 0$  when it is  
 351 negative

## 352 2.2 Application of the model to a control programme 353 for BVDV infection in cattle

### 354 2.2.1 Data

355 The model was evaluated on data collected for the surveillance of BVDV  
 356 infection in cattle in Loire-Atlantique, France. Data were available from  
 357 1687 dairy herds between the beginning of 2010 and the end of 2016. Under  
 358 the programme, each herd was tested twice a year with a bulk tank milk  
 359 antibody ELISA test. For each campaign of testing, tests were performed  
 360 for all the herds over a few weeks. Data on the number of cattle introduced  
 361 into each herd with the associated date of introduction were also available.  
 362 For the model evaluation, test data from the beginning of 2014 to the end  
 363 of 2016 were used. Risk factor data collected between 2010 and 2016 were  
 364 available to model (possibly lagged) associations between risk factors and  
 365 latent status.

### 366 2.2.2 Test results

367 Test results were reported as optical density ratios (ODR). In the Loire-  
 368 Atlantique CP, these ODRs are discretised into 3 categories using threshold

369 values of 35 and 60. ODR values below 35 are associated with low antibody  
370 levels and ODR values above 60 are associated with high antibody levels.  
371 Decision regarding which herds require further testing for the identification  
372 and removal of PI animals is complex and involves the combination of test  
373 categories on 3 consecutive tests, spanning a year.

374 In this work, the ODR values were discretised in order to convert them  
375 into either seropositive (antibodies detected) or seronegative (no antibodies  
376 detected) outcomes. The choice of the threshold to apply for the discreti-  
377 sation was based on the ODR distribution, which was clearly bimodal. For  
378 this purpose, the ODR distribution was modelled as a mixture of 2 normal  
379 distributions using the R `mixdist` package (Macdonald & Du, 2018). Assum-  
380 ing that one of the distributions was associated with seronegativity and the  
381 other one with seropositivity, the threshold that discriminated best between  
382 the 2 distributions was selected.

### 383 2.2.3 Selection of risk factors

384 A difficulty in the evaluation of putative risk factors was that Bayesian models  
385 usually take time to run, especially with large datasets as used here. It was  
386 therefore not possible to perform this selection with our Bayesian model.  
387 To circumvent this problem, logistic models as implemented in the R `glm`  
388 function (R Core Team, 2019) were used<sup>4</sup>. The outcome of these models was  
389 seroconversion defined as a binary event, and covariates of interest were risk  
390 factors for becoming status positive as defined through the  $\tau_1$  variable. All  
391 herds with 2 consecutive test results whose first result was negative (ODR  
392 below the chosen threshold) were capable of seroconverting. Of these herds,  
393 the ones that had a positive result (ODR above the chosen threshold) on  
394 the second test were considered as having seroconverted. The time of event  
395 (seroconversion or not) was considered the mid-point between the 2 tests.

396 Two types of risk factors of new infection were evaluated: infection through  
397 cattle introductions and infection through neighbourhood contacts (Qi *et al.*,  
398 2019). Cattle introduction variables were constructed from the number of an-  
399 imals introduced into a herd on a given date. In addition to the raw number of  
400 animals introduced, the natural logarithm of the number of animals (+1 be-  
401 cause  $\ln(0)$  is not defined) was also evaluated. This was to allow a decreasing  
402 effect of each animal as the number of animals introduced increased. Regard-

---

<sup>4</sup>The functions used to perform this evaluation are included in the [STOCfree package](#).

403 ing the neighbourhood risk, the test result data were used. For each testing  
404 campaign, the municipality-level prevalence of test positives (excluding the  
405 herd of interest) was calculated, and is subsequently termed 'local preva-  
406 lence'. It was anticipated that when local seroprevalence would increase, the  
407 probability of new infection in the herd of interest would increase as well.

408 For all candidate variables, a potential problem was delayed detection,  
409 which relates to the fact that a risk factor recorded at one point in time may  
410 be detected through testing much later, even if the test is sensitive. For ex-  
411 ample, if a trojan cow (a non-PI female carrying a PI calf) is introduced into  
412 a herd, the lactating herd will only seroconvert when the PI calf is born and  
413 has had contact with the lactating herd. Therefore, for each candidate vari-  
414 able, the data were aggregated between the beginning of an interval (labelled  
415 lag1, in months from the outcome measurement) and the end of this inter-  
416 val (labelled lag2, in months from the outcome measurement). Models with  
417 all possible combinations of time aggregation between lag1 and lag2 were  
418 run, with lag1 set to 0 and lag2 set to 24 months. The best variables and  
419 time aggregation interval were selected based on low AIC value, biological  
420 plausibility and suitability for the Bayesian model.

#### 421 2.2.4 Bayesian models

422 Four different Bayesian models were considered. For all models, historical  
423 data were used for parameter estimation and the probability of infection on  
424 the last month in the dataset was predicted.

425 **Model 1 - Perfect routine test:** in order to evaluate the monthly dy-  
426 namics of seropositivity and seronegativity, the Bayesian model was run  
427 without any risk factor and assuming that both test sensitivity and test  
428 specificity were close to 1. The prior distributions for sensitivity and speci-  
429 ficity were  $Se \sim Beta(10000, 1)$  (percentiles: 5 = 1, 50 = 1, 95 = 1) and  
430  $Sp \sim Beta(10000, 1)$ . Regarding infection dynamics, prior distributions were  
431 also specified for the prevalence of status positives (also test positives in this  
432 scenario) on the first testing time  $p(S_1^+) \sim Beta(1, 1)$  (uniform on 0-1), the  
433 probability of becoming status positive  $\tau_1 \sim Beta(1.5, 10)$  (percentiles: 5  
434 = 0.017, 50 = 0.109, 95 = 0.317), and the probability of remaining status  
435 positive  $\tau_2 \sim Beta(10, 1.5)$  (percentiles: 5 = 0.683, 50 = 0.891, 95 = 0.983).



436 **Model 2 - Perfect routine test and risk factors:** in order to quantify  
437 the association between risk factors and the probability of becoming status  
438 positive if the test were close to perfect, the Bayesian model was run with  
439 the risk factors identified as associated with seroconversion on the previous  
440 step and using the same priors for sensitivity, specificity and  $\tau_2$  as in Model  
441 1 ( $Se \sim Beta(10000, 1)$ ,  $Sp \sim Beta(10000, 1)$ ,  $\tau_2 \sim Beta(10, 1.5)$ ). The  
442 priors for risk factors were specified as normal distributions on the logit  
443 scale. The prior for the intercept was  $\theta_1 \sim \mathcal{N}(-3, 1)$  (on the probability  
444 scale - percentiles: 5 = 0.01, 50 = 0.047, 95 = 0.205). This represented  
445 the prior probability of a new infection in a herd purchasing no animal and  
446 with a local seroprevalence of 0. The priors for the other model coefficients  
447 were centred on 0 with a standard deviation of 2. On the logit scale, values  
448 of -4 (2 standard deviations in this case) correspond to probabilities close  
449 to 0 ( $\text{logit}(-4) = (0.018)$ ) and values of 4 to probabilities that are close to 1  
450 ( $\text{logit}(4) = (0.982)$ ).

451 **Model 3 - Imperfect routine test and risk factors:** the objective  
452 of this model was to incorporate the uncertainty associated with test re-  
453 sults in both parameter estimation and in the prediction of the probabili-  
454 ties of infection. The priors for test sensitivity and specificity were selected  
455 based on the ODR distributions for seronegatives and seropositives iden-  
456 tified by the mixture model. The following prior distributions were used:  
457  $Se \sim Beta(5000, 260)$  (percentiles: 5 = 0.946, 50 = 0.951, 95 = 0.955) and  
458  $Sp \sim Beta(5000, 260)$ . For the associations between risk factors and the  
459 probability of new infection, the same prior distributions as in Model 2 were  
460 used.

461 **Model 4 - Imperfect routine test, confirmatory testing and risk fac-**  
462 **tors:** the objective of this model was to assess the impact of confirmatory  
463 testing. The same prior distributions as in scenario 3 were used. In this  
464 case however, every time a positive test result was recorded, a new confir-  
465 matory test was randomly generated in the following month so that 85% of  
466 these tests were positive and 15% were negative. The confirmatory test was  
467 assumed to have both a sensitivity and a specificity close to 1.

468 For each model, 4 chains were run in parallel. The first 5 000 MCMC  
469 iterations were discarded (burn-in). The model was run for 5 000 more  
470 iterations of which 1 in 20 was stored for analysis. This yielded 1 000 draws

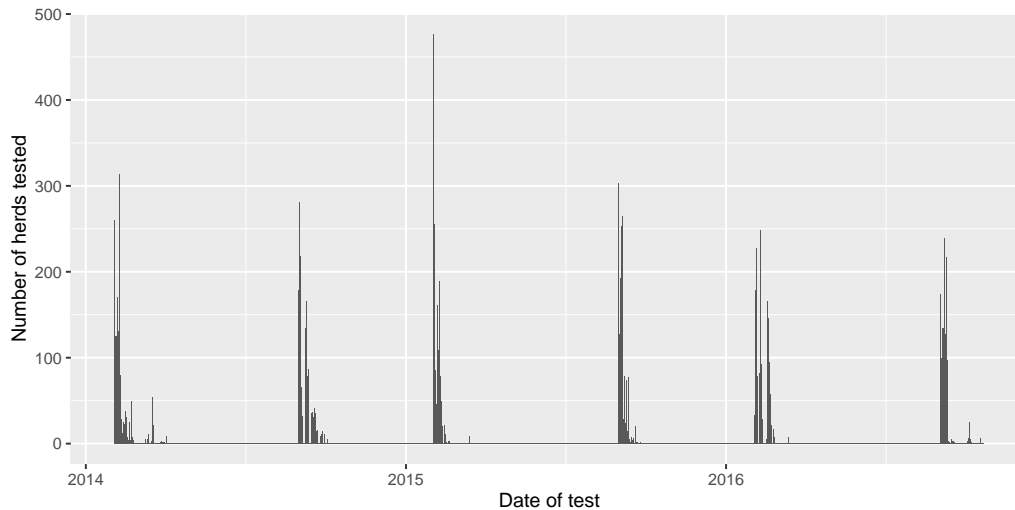


Figure 4: Distribution of the test dates between 2014 and 2017 in 1687 herds from Loire-Atlantique, France.

471 from the posterior distribution of each parameter. Convergence was assessed  
472 visually using traceplots. Each distribution was summarised with its median  
473 and 95% credibility interval.

## 474 3 Results

### 475 3.1 Test results

476 There were 9725 available test results from 1687 herds. Most herds were  
477 tested in February and September (See Figure 4). Two normal distributions  
478 were fit to the ODR data using the R `mixdist` package (Figure 5). The distri-  
479 bution for seronegatives had a mean and standard deviation of 7.1 and 16.3  
480 respectively. The distribution for seropositives had a mean and standard  
481 deviation of 57 and 13 respectively. There were 58.6% and 41.4% of obser-  
482 vations in the seronegative and seropositive distributions respectively. ODR  
483 values above 35 (21% of ODR values) were categorised as test positive and  
484 ODR values below 35 were categorised as test negative. The sensitivity and  
485 the specificity of the threshold value of 35 for the classification of test results  
486 with respect to seropositivity were estimated using the fitted distributions  
487 as the gold standard. These estimated sensitivity and specificity were 0.956

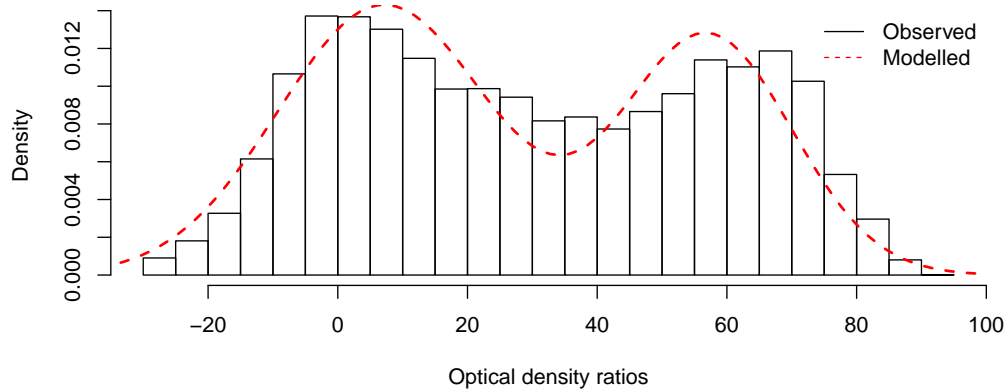


Figure 5: Distribution of the observed optical density ratios (histogramme) and fitted mixture of normal distributions (red dashed curves) for the bulk tank milk test results used in the analyses.

488 and 0.955 respectively. In the Bayesian models in which the latent status  
489 was seropositivity, the prior distributions for sensitivity and specificity were  
490 centred on these values.

### 491 3.2 Selection of risk factors

492 Risk factors related to animal introductions and seroprevalence were evalu-  
493 ated with logistic models. The model outcome was a seroconversion event.  
494 A first step of the analysis was, for each variable, to identify the time in-  
495 terval that was the most predictive of an observed seroconversion. Figure 6  
496 presents the AIC values associated with each possible interval for the vari-  
497 ables  $\ln(\text{Number of animals introduced} + 1)$  and local seroprevalence.

498 For the animal introduction variables, for the same time interval, the  
499 AICs of the models of the untransformed number of animals were higher  
500 than the ones for the log transformed values (not shown). It can also be  
501 noted that considering longer intervals (further away from the diagonal) was  
502 usually better than considering short intervals (close to the diagonal). It  
503 may be that some herds never buy any animal while, on average, herds that  
504 buy once have already done it in the past. In this case, it is possible that

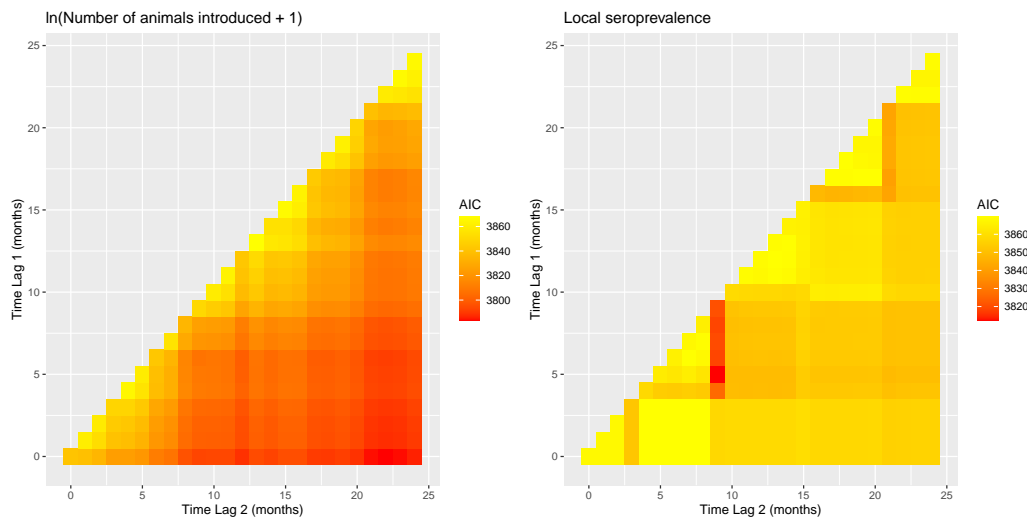


Figure 6: AIC values associated with logistic models of the association between 2 variables and the probability of seroconversion between 2 tests. The variable evaluated on the left-hand side panel is the sum of the log(number of animals introduced + 1) between lag1 and lag2. The variable evaluated on the right-hand side panel is the max of the local seroprevalence between lag1 and lag2.

505 the infection was introduced several times, while it is not possible to know  
506 which animal introduction was associated with herd seroconversion. This  
507 could explain the apparent cumulative effect of the number of introductions.  
508 The cells that are close to the diagonal are associated with short intervals.  
509 Considering one month intervals, the probability of infection was highest for  
510 introductions made 8 months from the month of seroconversion.

511 Local seroprevalence was evaluated from data collected in 2 different test-  
512 ing campaigns per year, as shown in Figure 4. For this reason, in the investi-  
513 gation of lagged relationships between local seroprevalence and the probabili-  
514 ty of seroconversion, the maximum local seroprevalence was computed, and  
515 not the sum as for the number of animals introduced. The strength of as-  
516 sociation between local seroprevalence and herd seroconversion was greatest  
517 for local seroprevalence 9 months prior to herd seroconversion.

518 A final multivariable logistic model with an animal introduction variable  
519 and a local seroprevalence variable was constructed. In the choice of the  
520 time intervals to include in this model, the following elements were consid-

Table 2: Results of the final logistic model of the probability of seroconversion between consecutive tests.

	lag1	lag2	Estimate	p-value
Intercept	-	-	-1.96	7.99e-306
ln(Number animals introduced +1)	8	8	0.38	5.70e-10
local seroprevalence	9	9	4.59	3.39e-13

521 ered. First, the Bayesian model runs with a monthly time step. Aggregating  
522 data over several months would result in including the same variable sev-  
523 eral times. Secondly, historical data may sometimes be limited. Having the  
524 smallest possible value for the end of the interval could be preferable. For  
525 this reason the variables considered for the final model were the natural log-  
526 arithm of the number of animals introduced 8 months prior to the month of  
527 seroconversion as well as the local seroprevalence 9 months prior to the month  
528 of seroconversion. The results of this model are presented in Table 2. All  
529 variables were highly significant. The model intercept was the probability of  
530 seroconversion in a herd introducing no animals and with local seroprevalence  
531 of 0 in each of the time intervals considered. The probability of seroconver-  
532 sion between 2 tests corresponding to this scenario was of 0.124. Buying 1,  
533 10 or 100 animals increased this estimated probability to 0.171, 0.866 and 1  
534 respectively. Buying no animals and observing a seroprevalence of 0.2 (pro-  
535 portion of seropositives in the dataset) was associated with a probability of  
536 seroconversion of 0.261.

### 537 3.3 Bayesian models

538 Running each of the 4 models for the 1687 herds with 3 years of data took on  
539 average 7 hours per model. In models 2 to 4, the candidate covariates were  
540 the natural logarithm of the number of animals introduced 8 months before  
541 status evaluation/prediction as well as the local seroprevalence 9 months  
542 prior. The 95% credibility interval for the estimated coefficient associated  
543 with local seroprevalence included 0. This variable was therefore removed  
544 from the models and only cattle introductions were considered.

Table 3: Median (2.5%, 97.5%) of the parameter posterior distributions used in the 4 Bayesian models evaluated. Model 1: Perfect routine test; Model 2: Perfect routine test and risk factors; Model 3: Imperfect routine test and risk factors; Model 4: Imperfect routine test, confirmatory testing and risk factors.

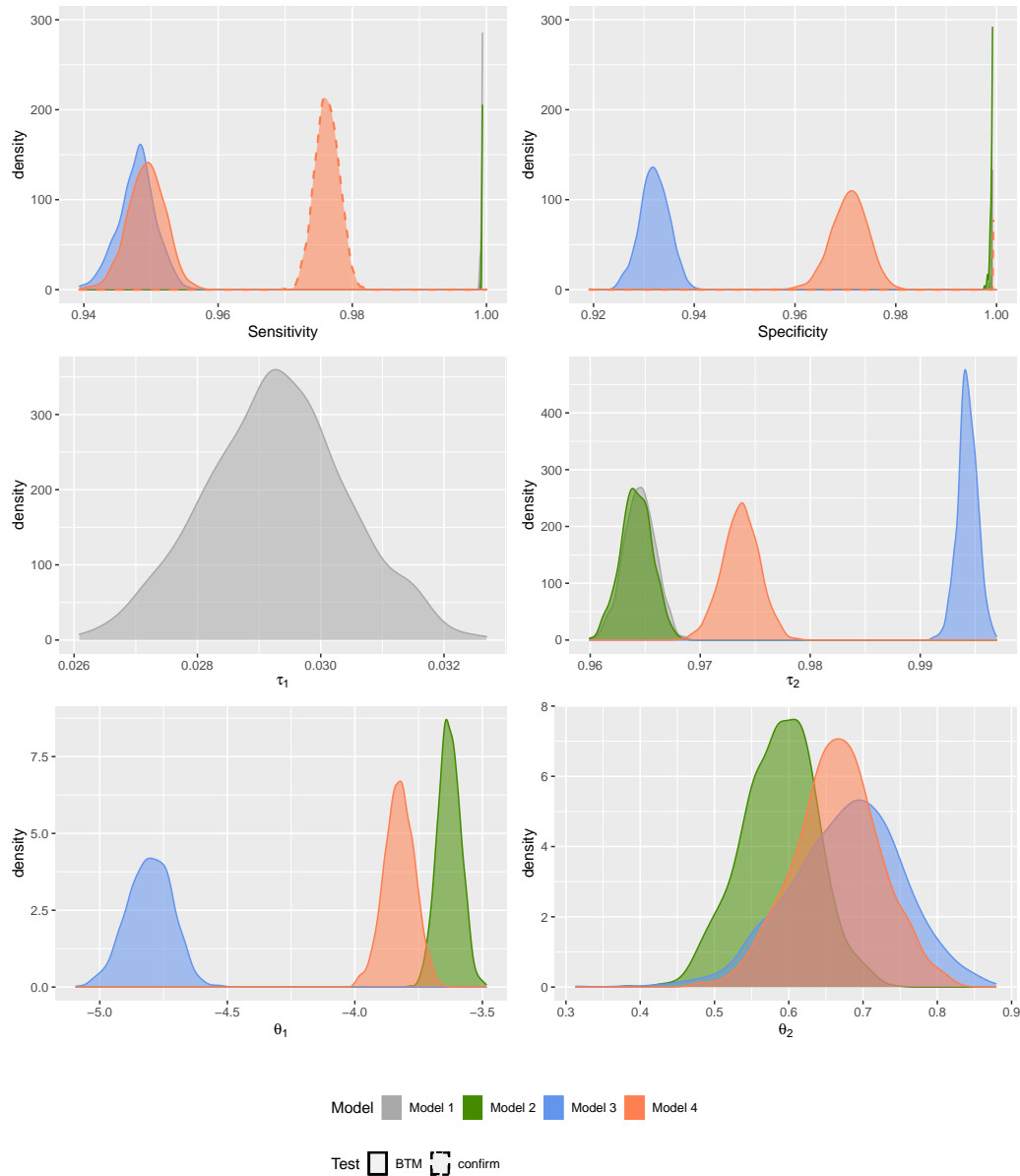
Parameter	Model 1	Model 2	Model 3	Model 4
Se BTM ODR	1 (0.999, 1)	1 (1, 1)	0.948 (0.942, 0.953)	0.949 (0.944, 0.955)
Se confirmatory	-	-	-	0.976 (0.973, 0.98)
Sp BTM ODR	1 (0.999, 1)	1 (0.999, 1)	0.932 (0.926, 0.938)	0.971 (0.964, 0.978)
Sp confirmatory	-	-	-	1 (1, 1)
$\tau_1$	0.029 (0.027, 0.032)	-	-	-
$\tau_2$	0.965 (0.962, 0.967)	0.964 (0.961, 0.967)	0.994 (0.993, 0.996)	0.974 (0.97, 0.977)
$\theta_1$ (Intercept)	-	-3.631 (-3.718, -3.545)	-4.803 (-4.985, -4.646)	-3.825 (-3.94, -3.711)
$\theta_2$	-	0.589 (0.482, 0.684)	0.682 (0.522, 0.813)	0.665 (0.547, 0.776)

### 545 3.3.1 Model parameters

546 Figure 7 and Table 3 show the distributions of model parameters for the 4  
 547 models. Figure 8 shows the predicted probability of becoming status positive  
 548 as a function of the number of animals introduced 9 months before status  
 549 evaluation.

550 In Models 1 and 2, the prior distributions put on sensitivity and speci-  
 551 ficity were very close to 1. With these models, the latent status corresponded  
 552 to the test result. In effect, they modelled the monthly probabilities of transi-  
 553 tion between BTM test negative and BTM test positive. In this case, the  
 554 median (percentile 2.5 - percentile 97.5) probability of becoming status posi-  
 555 tive between consecutive months was 0.029 (0.027 - 0.032). This represents  
 556 a probability of becoming status positive over a 12 month period of 0.298  
 557 (0.280 - 0.323). For status positive herds, the monthly probability of remain-  
 558 ing positive was of 0.965 which represents a probability of still being status  
 559 positive 12 months later of 0.652 (0.628-0.669). In model 2, a risk factor was  
 560 incorporated into the estimation. The model intercept was much lower than  
 561 the estimate from the logistic model estimated in the variable selection step.  
 562 This was due to the different time steps considered (1 month vs. half a year).  
 563 On the other hand, the estimate for the log number of animals introduced  
 564 was higher.

565 In model 3, the prior distributions for test sensitivity and specificity were  
 566 centred on 0.95 based on the mixture of 2 normal distributions for seroneg-



**Figure 7:** Parameters posterior distributions for the 4 Bayesian models. Model 1: Perfect routine test; Model 2: Perfect routine test and risk factors; Model 3: Imperfect routine test and risk factors; Model 4: Imperfect routine test, confirmatory testing and risk factors. Sensitivities and specificities close to 1 are not shown to facilitate reading. The dashed lines correspond to the distributions of the confirmatory tests. Parameters related to status dynamics are  $\tau_1$  (probability of becoming status positive between consecutive months) and  $\tau_2$  (probability of remaining status positive).  $\tau_1$  was only estimated for the model without risk factors (model 1). The parameters for the association between risk factors and the probability of becoming status positive are  $\theta_1$  and  $\theta_2$ .  $\theta_1$  is the intercept of the logistic model and  $\theta_2$  is the coefficient associated with the log of the number of animals introduced 8 months before status evaluation/prediction.

567 atives and seropositives that described best the BTM ODR data (see Sec-  
568 tion 3.1). With this model, the latent status corresponded to seropositivity.  
569 This assumption allowed the effect of having an imperfect test on the estima-  
570 tion of the different model parameters to be investigated. In this scenario,  
571 the posterior distribution for sensitivity was close to the prior, but the poste-  
572 rior for the specificity was slightly lower. On the other hand, the distribution  
573 for  $\tau_2$  was higher than when the test was considered perfect. This implies  
574 that the model identified some test positives as false positives, but that the  
575 ones that retained a positive status remained positive for longer. Compared  
576 to Model 2, the probability of becoming status positive was lower in herds  
577 buying no animals (model intercept), and tended to increase more rapidly  
578 with the number of animals introduced ( $\theta_2$ ), although for 100 animals intro-  
579 duced, the probability of becoming status positive was still lower than with  
580 the other models (Figure 8). Because of the imperfect sensitivity of routine  
581 testing, some herds that were seronegative at a test while seropositive at the  
582 previous or following tests were classified as false negative by the model and  
583 thereby were not included in the estimation of  $\tau_1$ , which may have decreased  
584 the estimated strength of association between cattle introduction and new  
585 infection. However, the estimates produced by this should be more accurate.

586 In model 4, confirmatory testing was added, with a testing procedure as-  
587 summed to have perfect sensitivity and specificity for the detection of farms  
588 with infected animals. This resulted in several differences with model 3,  
589 which illustrate the interplay between data and prior information. The added  
590 confirmatory negative results often contradicted the data because, they were  
591 generally followed by a positive routine test. This had the following conse-  
592 quences. The posterior distribution for the sensitivity of confirmatory testing  
593 was lower than its prior distribution, indicating that herds negative to con-  
594 firmatory testing were classified as false negatives more often than suggested  
595 by the priors. The fact that the estimated value for the specificity of BTM  
596 testing was higher than in Model 3 shows that herds positive to routine test-  
597 ing were considered to be true positives slightly more often. The fact that  
598 the estimated value for  $\tau_2$  was lower than in Model 3 shows that status posi-  
599 tive herds tended to clear infection more quickly, which allowed a more rapid  
600 status change between routine and confirmatory testing. Because Model 4  
601 resulted in more frequent changes in status, the coefficients for the associ-  
602 ation between cattle introduction and new infections (Figure 8) were closer  
603 between Model 4 and Model 2 than between Model 4 and Model 3.



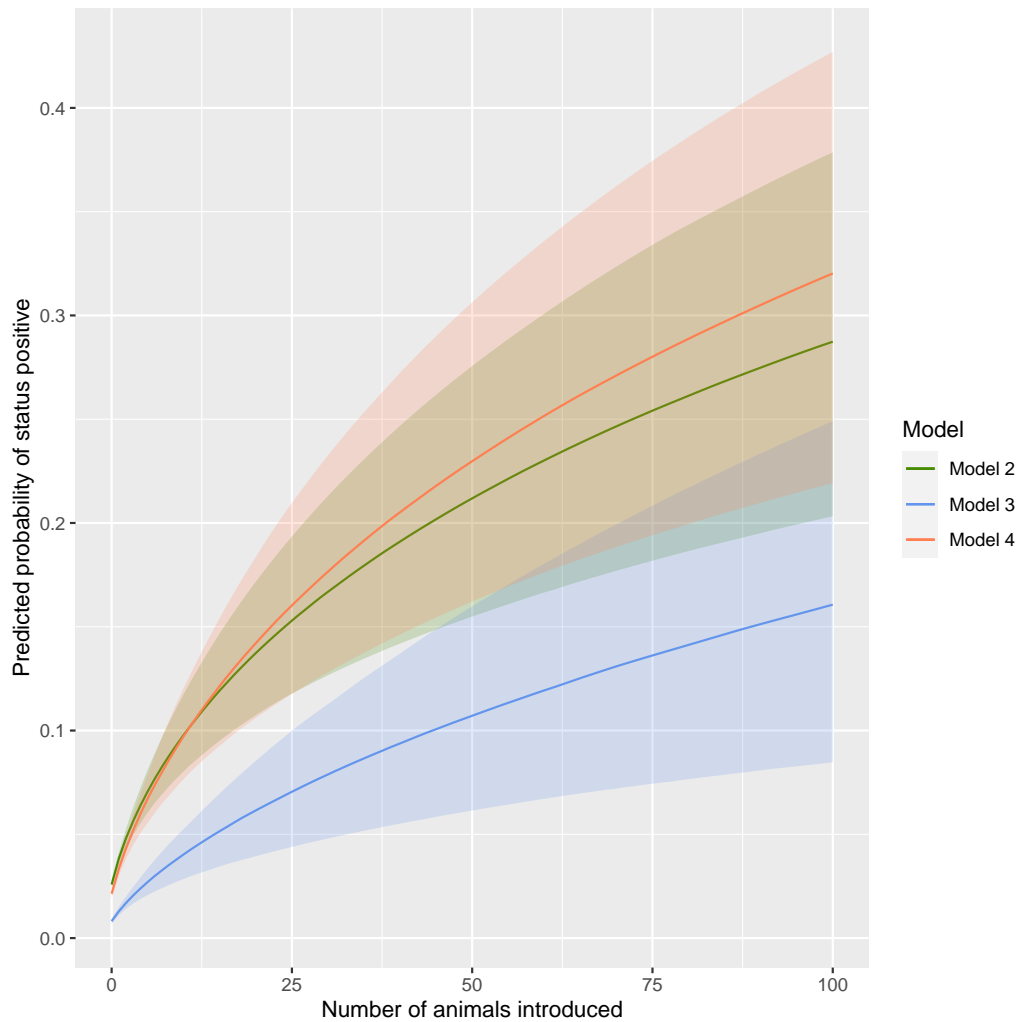


Figure 8: Predicted probability of new infection as a function of the number of animals introduced 8 months before the month of interest for the Bayesian models 2 to 4. Model 2: Perfect routine test and risk factors; Model 3: Imperfect routine test and risk factors; Model 4: Imperfect routine test, confirmatory testing and risk factors. The lines represent the median predicted values. The shaded areas represent the 95% credibility intervals.

### 604 3.3.2 Predicted probabilities of infection

605 Figure 9 shows the distributions of herd-level probabilities of infection pre-  
606 dicted by the 4 Bayesian models. These probability distributions are bimodal  
607 for all models. The left-hand side corresponds to herds that were predicted  
608 status negative on the month before the month of prediction. These are  
609 associated to becoming status positive, i.e.  $\tau_1$ . The right-hand side of the  
610 distributions corresponds to herds that were predicted status positive on the  
611 month before the month of prediction. These are associated to remaining  
612 status positive, i.e.  $\tau_2$ . For models 3 and 4, which incorporate both risk  
613 factors and test uncertainty, the modes are closer to 0 and 1 than for the  
614 other 2 models. For Model 4, there is a third mode between 0.4 and 0.5.  
615 This mode was associated with confirmatory testing.

616 Figure 10 shows the distributions of the predicted probability of being  
617 status positive for 4 herds. It can be seen that herds that were consistently  
618 negative (positive) to the test had extremely low (high) probabilities of being  
619 status positive. Accounting for the number of animals introduced increased  
620 the probability of infection in the herds that were test negative.

## 621 4 Discussion

622 This article describes a statistical framework for the prediction of an infection  
623 related status from longitudinal data generated by CPs against infectious  
624 diseases of farm animals. The statistical model developed estimates a herd  
625 level probability of being *latent status* positive on a specific month, based  
626 on input data that can vary in terms of the types of test used, frequency  
627 of testing and risk factor data. This is achieved by modelling the latent  
628 status with the same discrete time step, regardless of the frequency with  
629 which input data are available, and by modelling changes in the latent status  
630 between consecutive time steps. This model therefore fulfils one of our main  
631 objectives which was to be able to integrate heterogeneous information into  
632 the estimation. However, in order to be able to compare the output of this  
633 model run on data from different CPs, the definition of the latent status  
634 should be the same.

635 In this model, the latent status is mostly defined by the prior distribu-  
636 tions put on the different model parameters. In setting the prior distributions  
637 there are two issues: setting the distribution's central value (mean, median

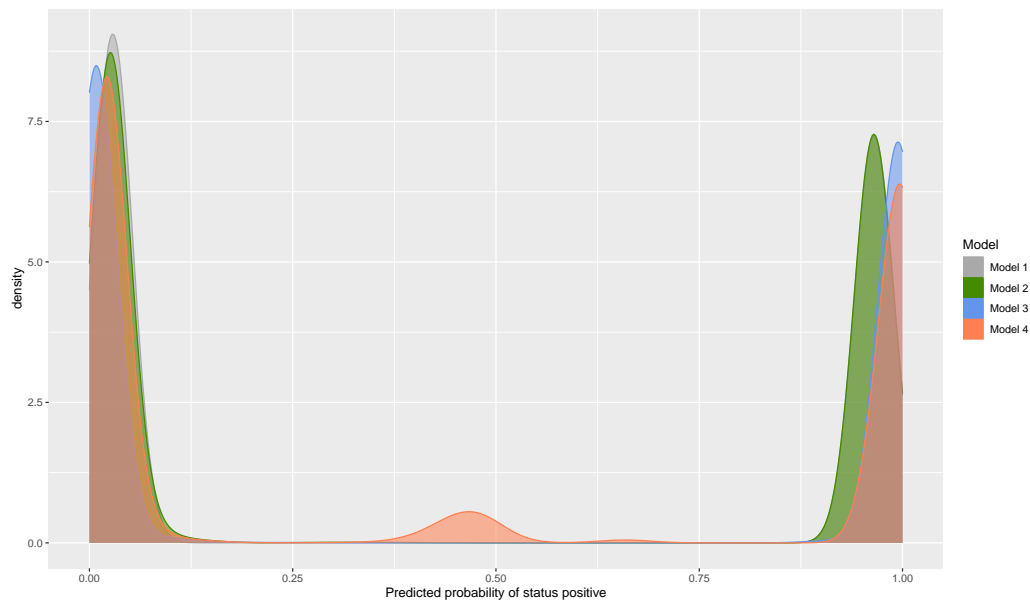


Figure 9: Distributions of the predicted probabilities of being status positive for all herds with the 4 Bayesian models evaluated. Model 1: Perfect routine test; Model 2: Perfect routine test and risk factors; Model 3: Imperfect routine test and risk factors; Model 4: Imperfect routine test, confirmatory testing and risk factors.

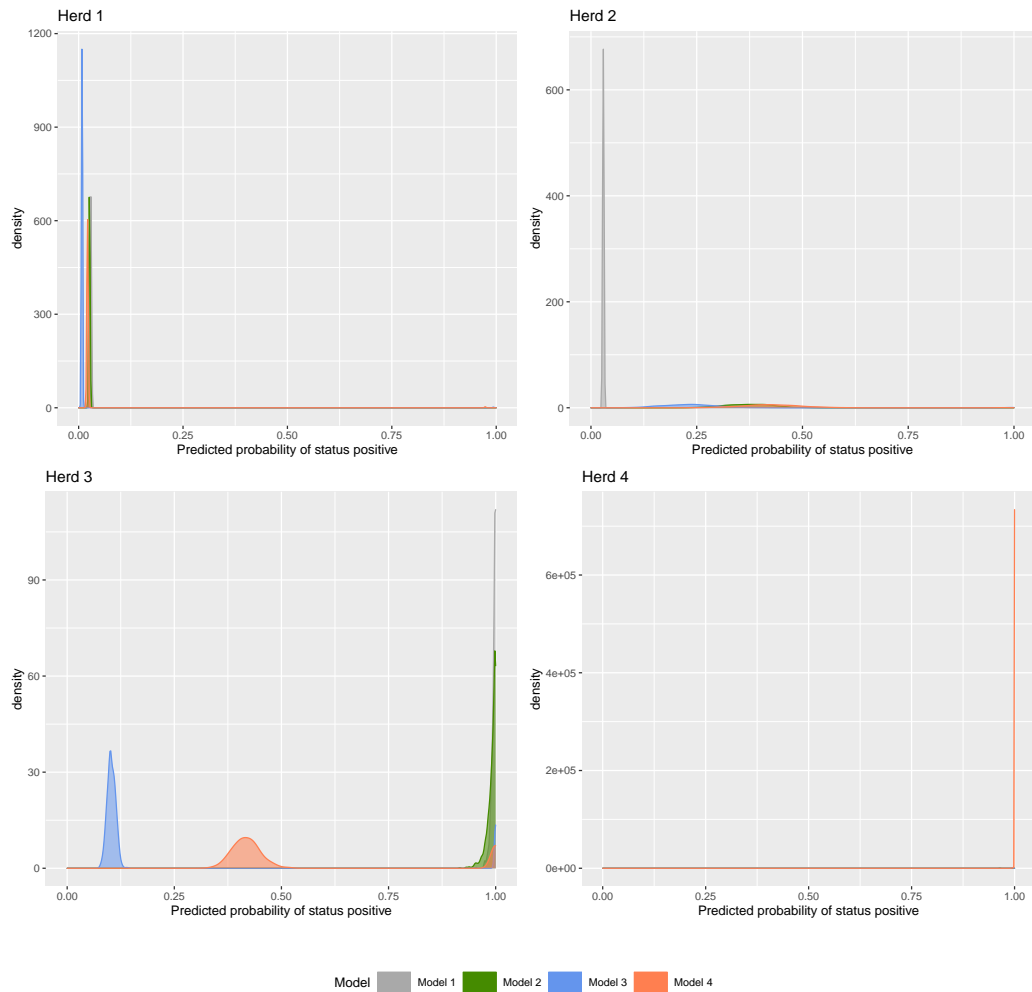


Figure 10: Distribution of predicted probabilities of being status positive on the month of prediction for 4 herds with the 4 models compared. Model 1: Perfect routine test; Model 2: Perfect routine test and risk factors; Model 3: Imperfect routine test and risk factors; Model 4: Imperfect routine test, confirmatory testing and risk factors. Herd 1 was test negative for 6 consecutive tests, introduced no animal. Herd 2 was test negative for 6 consecutive tests, introduced animals regularly (196 associated with the month of prediction). Herd 3 was test negative on the first 5 tests and test positive on the month of prediction, introduced animals regularly (3 introductions associated with the month of prediction). Herd 4 was test negative on the first 2 tests and test positive on the last 4 tests, introduced animals regularly.

638 ... ) and setting the distribution width. Choosing the wrong central value,  
639 i.e. the prior distribution does not include the true parameter value, can lead  
640 to systematic error (bias) or absence of convergence. This problem will be  
641 more important as prior distributions become narrower. Setting prior distri-  
642 butions that are too wide can lead to a lack of convergence, when multiple  
643 combinations of parameter values are compatible with the data. This was a  
644 problem in initial modelling of the BVDV data (not shown). Putting narrow  
645 prior distributions on test sensitivity and test specificity allowed the model  
646 to converge. These narrow distributions imply very strong hypotheses on  
647 test characteristics.

648 The definition of prior distributions for test characteristics that reflect  
649 the latent status of interest is challenging (Duncan *et al.*, 2016). This was  
650 apparent in the application to infection by the BVDV we presented. For  
651 the trade of animals from herds that are free from infection by the BVDV,  
652 the latent status of interest was the *presence of at least one PI animal in*  
653 *the herd*. The test data available to estimate the probability of this event  
654 were measures of bulk tank milk antibody levels which were used to define  
655 seropositivity as a binary event. Although milk antibody level is associated  
656 with the herd prevalence of antibody positive cows (Beaudeau *et al.*, 2001),  
657 seropositive cows can remain long after all the PIs have been removed from  
658 a herd. Furthermore, vaccination induces an antibody response which may  
659 result in vaccinated herds being positive to serological testing regardless of  
660 PI animal presence (Raue *et al.*, 2011; Booth *et al.*, 2013). Therefore, the  
661 specificity of BTM seropositivity, i.e. the probability for herds with no PI  
662 animals to be test negative, is less than 1. More importantly, this specificity  
663 depends on the context; i.e. on the CP. PI animals can be identified and  
664 removed more or less quickly depending on the CP, the proportion of herds  
665 vaccinating and the reasons for starting vaccination can differ between CPs.  
666 Test sensitivity can also be imperfect. Continuing with the example of bulk  
667 tank milk testing, contacts between PI animals present on the farm and the  
668 lactating herd may be infrequent, which would decrease sensitivity. In this  
669 case, the sensitivity of the testing procedure is the sensitivity of the test  
670 for the detection of seroconversion in a group of animals multiplied by the  
671 probability that the tested group has seroconverted if there is a PI animal in  
672 the herd. The probability of contact between PI animals and the lactating  
673 herd depends on how herds are organised, which could vary between CPs.  
674 This problem is alleviated when newborn calves are tested because the group  
675 of animals tested is the group in which the infectious animals are most likely

676 to be present. Furthermore, with BTM testing, the contribution of each  
677 seropositive cow to the BTM decreases as herd size increases which can result  
678 in differences in BTM test sensitivity associated with different herd sizes  
679 between CPs.

680 The effects of using different prior distributions for test characteristics  
681 on latent status definition, parameter estimation and probability prediction  
682 were evaluated. In models 1 and 2, the dichotomised BTM antibody test  
683 results were modelled assuming perfect sensitivity and perfect specificity.  
684 With these assumptions, the latent status was the dichotomised test results.  
685 In Model 3, the BTM test was assumed to have both a sensitivity and a speci-  
686 ficity concentrated around 95%, based on the normal distributions associated  
687 with seronegativity and seropositivity identified by a mixture model. The la-  
688 tent status in Model 3 can therefore be described as *seropositivity*. Because  
689 overall the probability of changing status was small, assuming an imperfect  
690 sensitivity lead to isolated negative test results in sequences of mostly posi-  
691 tive test results to be considered false negatives, as shown by the increase  
692 in the estimated value for  $\tau_2$  between Model 2 and Model 3. This illustrates  
693 that in addition to test characteristics, status dynamics will determine the  
694 latent status within herds. Model 4 was constructed to evaluate the impact  
695 of incorporating confirmatory testing into the model. In CPs, herds that test  
696 positive are usually re-tested in order to rule out a false positive test, and  
697 to identify infected animals if needed. The testing procedure used in con-  
698 firmatory testing usually has a high sensitivity and a higher specificity than  
699 routine testing in relation to the gold standard. When incorporated into the  
700 model, this high quality information, in conjunction with wider prior distri-  
701 butions on routine testing specificity, should allow the posterior distribution  
702 of the specificity of routine testing to be revised towards the gold standard.  
703 Indeed, if a confirmatory test comes back negative, then the corresponding  
704 latent status will become negative with high probability. Given the low prob-  
705 ability of becoming status negative between consecutive months, the latent  
706 status on the month of routine testing has an increased probability of be-  
707 ing negative, leading to a decrease in the specificity of routine testing. This  
708 could not be adequately demonstrated in Model 4, because simulating test  
709 results at random was often not consistent with patterns of test results in  
710 individual herds. However, this confirmed the importance of status dynamics  
711 in estimating the latent status.

712 Status dynamics contributed to the definition of the latent status in sev-  
713 eral ways. Negative test results interspersed with sequences of positive test

714 results will be classified as latent status positive (i.e. as false negatives) more  
715 often as test sensitivity decreases and  $\tau_2$  increases. Positive test results in-  
716 terspersed with sequences of negative test results will be classified as latent  
717 status negative (i.e. as false positives) with increased frequency as test speci-  
718 ficity and  $\tau_1$  each decrease. With a perfect test (sensitivity and specificity  
719 equal to 1), the model can learn the values of  $\tau_1$  and  $\tau_2$  from the data, and  
720 the prior distributions put on these parameters can be uninformative. With  
721 decreasing values for test sensitivity and specificity, the information provided  
722 through the prior distributions put on  $\tau_1$  and  $\tau_2$  becomes increasingly impor-  
723 tant. The informative value of  $\tau_1$  and  $\tau_2$  will increase as the probability of  
724 transition between latent status negative and latent status positive decrease,  
725 i.e. when  $\tau_1$  is small and  $\tau_2$  is high.

726 When data on risk factors of new infection are available, the  $\tau_1$  parameter  
727 is modelled as a function of these risk factors using logistic regression. In such  
728 a case, prior distributions are put on the parameters of the logistic regression  
729 and not on the the  $\tau_1$  parameter. In the application that we presented, we  
730 used a prior distribution corresponding to a low probability of new infection  
731 in the reference category (intercept: herds which introduced no animals) and  
732 we centred the prior distribution for the association with cattle introductions  
733 on a hypothesis of no association (mean = 0 on the logit scale). This allowed  
734 the model to estimate the association between the risk factor and the latent  
735 status from historical data and to use the estimated association to predict  
736 probabilities of being latent status positive on the month of prediction. As  
737 expected, the prior distributions put on test characteristics had an impact  
738 on the parameter estimates. In Model 3, the model intercept was lower and  
739 the estimated association between becoming latent status positive and cattle  
740 introduction was higher than in the other models. The most likely explana-  
741 tion for this is that Model 3 allowed the highest level of discrepancy between  
742 dichotomised test result and latent status, while assuming a low probability  
743 of changing status between months. This resulted in negative test results  
744 in herds that were regularly positive to be classified as latent status positive  
745 (false negatives, associated with lower test sensitivity, see Table 3) thereby re-  
746 moving opportunities for new infections in herds that were regularly positive  
747 while also buying animals. This would imply that the estimated association  
748 from model 3 is more closely associated with new infections than estimates  
749 from the other models because herds that are regularly test positive have  
750 less weight in the estimation. It would also have been possible to base the  
751 prior distributions for the model coefficients on published literature. Unfor-

752 tunately, estimates of the strengths of association between risk factors and  
753 the probability of new infection are not readily available from the published  
754 literature or are hard to compare between studies (van Roon *et al.*, 2020a).  
755 However, estimates from the literature could allow the prior distributions to  
756 be bounded within reasonable ranges.

757 Because the model takes a lot of time to run, the variables included in  
758 the logistic regression were first identified with logistic models estimated by  
759 maximum likelihood. This confirmed the importance of animal introduction  
760 and neighbourhood contacts in new infections (Qi *et al.*, 2019). However, in  
761 the Bayesian models, the 95% credibility for the association between local  
762 seroprevalence and new infection included 0 and this variable was therefore  
763 not included. The reason for this was not elucidated in this work. Other risk  
764 factors such as herd size, participation in shows or markets, the practice of  
765 common grazing have shown a consistent association with the probability of  
766 new infection by the BVDV (van Roon *et al.*, 2020a). These variables were  
767 not included in our model because the corresponding data were not available.  
768 One advantage of our approach is the possibility to choose candidate risk  
769 factors to include in the prediction of infection based on the data available in  
770 a given CP. The associations between the selected putative risk factors and  
771 the probability of new infection can be estimated from these data.

772 Given the reasonably good performance of tests for the detection of BVDV  
773 infection, the main advantage of incorporating these risk factors was not to  
774 complement the test results on a month a test was performed, but rather to  
775 enhance the timeliness of detection. Risk factors that are associated with  
776 new infection will increase the predicted probability of infection regardless  
777 of the availability of a test result. Therefore, when testing is not frequent,  
778 infected herds could be detected more quickly if risk factors of infection are  
779 recorded frequently. If the available data on risk factors of new infection  
780 captured all the possible routes of new infection, it would be possible to  
781 perform tests more frequently in herds that have a higher probability of  
782 infection as predicted by our model. In other words, our model could be  
783 used for risk-based surveillance (Cameron, 2012).

784 In the CP from which the current data were used, herds are tested twice  
785 a year. This could lead to a long delay between the birth of PI calves and  
786 their detection through bulk tank milk testing. We addressed this problem  
787 of *delayed detection* by proposing a method for the investigation of lagged  
788 relationships between risk factor occurrence and new infections, and by in-  
789 cluding lagged risk factor occurrences in the prediction of the probability of



790 infection. In our dataset, herds purchasing cattle were more likely to have  
791 seroconverted 8 months after the introduction. In the Bayesian model, cattle  
792 introduction was modelled as affecting the probability of becoming status  
793 positive 8 months after the introduction. It can be argued that infection is  
794 present but not detected during this period, as the expression *delayed detec-*  
795 *tion* suggests, and that the probability of infection should increase as soon  
796 as risk factor occurrence is recorded. Modelling this phenomenon would be  
797 possible by decreasing the test sensitivity for a period corresponding to the  
798 lag used in the current version of the model. This would imply that for this  
799 duration, any negative BTM test result would not provide any information  
800 about the true status regarding infection and that the herd would have an  
801 increased predicted probability of infection. This could be incorporated in  
802 future versions of the model.

803 There are several questions related to this modelling framework that  
804 would require further work. The model outputs are distributions of herd  
805 level probabilities of infection. Defining herds that are free from infection  
806 from these distributions will require decision rules to be developed based on  
807 distribution summaries (likely a percentile) and cut-off values. It would also  
808 be possible to model the probability of remaining infected between consecu-  
809 tive tests ( $\tau_2$ ) as a function of the control measures put in place in infected  
810 herds. Another area that requires further investigations is the evaluation  
811 of the modelling framework against a simulated gold standard to determine  
812 whether it provides an added value compared to simpler methods. The avail-  
813 ability of the model code as a Github repository allows interested users to  
814 improve or suggest improvements to our modelling framework.

## 815 Acknowledgments

816 This work is part of the STOC free project that was awarded a grant by  
817 the European Food Safety Authority (EFSA) and was co-financed by public  
818 organizations in the countries participating in the study.

819 We thank Groupement de Défense Sanitaire de Loire-Atlantique (GDS-  
820 44) for providing the data.

## 821 Conflict of interest disclosure

822 The authors of this article declare that they have no financial conflict of  
823 interest with the content of this article.

## 824 References

- 825 Beaudeau, F., Belloc, C., Seegers, H., Assié, S., Pourquier, P., & Joly, A.  
826 2001. Informative value of an indirect enzyme-linked immunosorbent as-  
827 say (ELISA) for the detection of bovine viral diarrhoea virus (BVDV)  
828 antibodies in milk. *Journal of veterinary medicine. B, Infectious diseases*  
829 *and veterinary public health*, **48**(Nov.), 705–712.
- 830 Booth, R. E., Cranwell, M. P., & Brownlie, J. 2013. Monitoring the bulk milk  
831 antibody response to BVDV: the effects of vaccination and herd infection  
832 status. *The Veterinary record*, **172**(Apr.), 449.
- 833 Cameron, A. R. 2012. The consequences of risk-based surveillance: Develop-  
834 ing output-based standards for surveillance to demonstrate freedom from  
835 disease. *Preventive veterinary medicine*, **105**(Aug.), 280–286.
- 836 Collins, J., & Huynh, M. 2014. Estimation of diagnostic test accuracy without  
837 full verification: a review of latent class methods. *Statistics in medicine*,  
838 **33**(Oct.), 4141–4169.
- 839 Duncan, A. J., Gunn, G. J., & Humphry, R. W. 2016. Difficulties arising  
840 from the variety of testing schemes used for bovine viral diarrhoea virus  
841 (BVDV). *The Veterinary record*, **178**(Mar.), 292.
- 842 Fernandes, L. G., Denwood, M. J., de Sousa Américo Batista Santos, C.,  
843 Alves, C. J., Pituco, E. M., de Campos Nogueira Romaldini, A. H., De Ste-  
844 fano, E., Nielsen, S. S., & Santos de Azevedo, S. 2019. Bayesian estimation  
845 of herd-level prevalence and risk factors associated with BoHV-1 infec-  
846 tion in cattle herds in the State of Paraíba, Brazil. *Preventive veterinary*  
847 *medicine*, **169**(Aug.), 104705.
- 848 Hui, S. L., & Walter, S. D. 1980. Estimating the error rates of diagnostic  
849 tests. *Biometrics*, **36**(Mar.), 167–171.

- 850 Johnson, W. O., Gardner, I. A., Metoyer, C. N., & Branscum, A. J. 2009.  
851 On the interpretation of test sensitivity in the two-test two-population  
852 problem: assumptions matter. *Preventive veterinary medicine*, **91**(Oct.),  
853 116–121.
- 854 Le Strat, Y., & Carrat, F. 1999. Monitoring epidemiologic surveillance data  
855 using hidden Markov models. *Statistics in medicine*, **18**(Dec.), 3463–3478.
- 856 Macdonald, P., & Du, J. 2018. *mixdist: Finite Mixture Distribution Models*.  
857 R package version 0.5-5.
- 858 Martin, P. A. J., Cameron, A. R., & Greiner, M. 2007. Demonstrating free-  
859 dom from disease using multiple complex data sources 1: a new method-  
860 ology based on scenario trees. *Preventive veterinary medicine*, **79**(May),  
861 71–97.
- 862 Norström, M., Jonsson, M. E., Åkerstedt, J., Whist, A. C., Kristoffersen,  
863 A. B., Sviland, S., Hopp, P., & Wahlström, H. 2014. Estimation of the  
864 probability of freedom from bovine virus diarrhoea virus in Norway using  
865 scenario tree modelling. *Preventive veterinary medicine*, **116**(Sept.), 37–  
866 46.
- 867 Nusinovici, S., Madouasse, A., Hoch, T., Guatteo, R., & Beaudeau, F. 2015.  
868 Evaluation of Two PCR Tests for *Coxiella burnetii* Detection in Dairy  
869 Cattle Farms Using Latent Class Analysis. *PloS one*, **10**, e0144608.
- 870 Plummer, M. 2017. *JAGS version 4.3.0 user manual*.
- 871 Qi, L., Beaunée, G., Arnoux, S., Dutta, B. L., Joly, A., Vergu, E., &  
872 Ezanno, P. 2019. Neighbourhood contacts and trade movements drive  
873 the regional spread of bovine viral diarrhoea virus (BVDV). *Veterinary*  
874 *research*, **50**(Apr.), 30.
- 875 R Core Team. 2019. *R: A Language and Environment for Statistical Com-*  
876 *puting*. R Foundation for Statistical Computing, Vienna, Austria.
- 877 Rabiner, L. R. 1989. A tutorial on hidden Markov models and selected  
878 applications in speech recognition. *Proceedings of the IEEE*, **77**(2), 257–  
879 286.

- 880 Raue, R., Harmeyer, S. S., & Nanjiani, I. A. 2011. Antibody responses to  
881 inactivated vaccines and natural infection in cattle using bovine viral di-  
882 arrhoea virus ELISA kits: assessment of potential to differentiate infected  
883 and vaccinated animals. *Veterinary journal (London, England : 1997)*,  
884 **187**(Mar.), 330–334.
- 885 Toft, N., Jørgensen, E., & Højsgaard, S. 2005. Diagnosing diagnostic  
886 tests: evaluating the assumptions underlying the estimation of sensitiv-  
887 ity and specificity in the absence of a gold standard. *Preventive veterinary  
888 medicine*, **68**(Apr.), 19–33.
- 889 van Roon, A., Mercat, M., van Schaik, G., Nielen, M., Graham, D., More,  
890 S., Guelbenzu, M., Fourichon, C., Madouasse, A., & Santman-Berends, I.  
891 2020a. Quantification of risk factors for BVDV in cattle herds: a systematic  
892 search and meta-analysis. *Journal of dairy science (In press)*.
- 893 van Roon, A. M., Santman-Berends, I. M. G. A., Graham, D., More, S. J.,  
894 Nielen, M., van Duijn, L., Mercat, M., Fourichon, C., Madouasse, A., Geth-  
895 mann, J., Sauter-Louis, C., Frössling, J., Lindberg, A., Correia-Gomes, C.,  
896 Gunn, G. J., Henry, M. K., & van Schaik, G. 2020b. A description and  
897 qualitative comparison of the elements of heterogeneous bovine viral di-  
898 arrhoea control programs that influence confidence of freedom. *Journal of  
899 dairy science*, **103**(May), 4654–4671.
- 900 Whittington, R., Donat, K., Weber, M. F., Kelton, D., Nielsen, S. S., Eisen-  
901 berg, S., Arrigoni, N., Juste, R., Sáez, J. L., Dhand, N., Santi, A., Michel,  
902 A., Barkema, H., Kralik, P., Kostoulas, P., Citer, L., Griffin, F., Barwell,  
903 R., Moreira, M. A. S., Slana, I., Koehler, H., Singh, S. V., Yoo, H. S.,  
904 Chávez-Gris, G., Goodridge, A., Ocepek, M., Garrido, J., Stevenson, K.,  
905 Collins, M., Alonso, B., Cirone, K., Paolicchi, F., Gavey, L., Rahman,  
906 M. T., de Marchin, E., Van Praet, W., Bauman, C., Fecteau, G., McKenna,  
907 S., Salgado, M., Fernández-Silva, J., Dziedzinska, R., Echeverría, G., Sep-  
908 pänen, J., Thibault, V., Fridriksdottir, V., Derakhshandeh, A., Haghkhah,  
909 M., Ruocco, L., Kawaji, S., Momotani, E., Heuer, C., Norton, S., Cad-  
910 mus, S., Agdestein, A., Kampen, A., Szteyn, J., Frössling, J., Schwan,  
911 E., Caldow, G., Strain, S., Carter, M., Wells, S., Munyeme, M., Wolf, R.,  
912 Gurung, R., Verdugo, C., Fourichon, C., Yamamoto, T., Thapaliya, S.,  
913 Di Labio, E., Ekgatat, M., Gil, A., Alesandre, A. N., Piaggio, J., Suanes,

- 914 A., & de Waard, J. H. 2019. Control of paratuberculosis: who, why and  
915 how. A review of 48 countries. *BMC veterinary research*, **15**(June), 198.
- 916 Zucchini, W., MacDonald, I. L., & Langrock, R. 2017. *Hidden Markov models*  
917 *for time series: an introduction using R*. Chapman and Hall/CRC.