



# Prior Precision, Prior Accuracy, and the Estimation of Disease Prevalence Using Imperfect Diagnostic Tests

Jenni L. McDonald and Dave James Hodgson\*

Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn, United Kingdom

## OPEN ACCESS

### Edited by:

Julian Drewe,  
Royal Veterinary College,  
United Kingdom

### Reviewed by:

Muhammad Zubair Shabbir,  
University of Veterinary and Animal  
Sciences, Pakistan  
Flavie Vial,  
Animal and Plant Health Agency,  
United Kingdom

### \*Correspondence:

Dave James Hodgson  
d.j.hodgson@exeter.ac.uk

### Specialty section:

This article was submitted to  
Veterinary Infectious Diseases,  
a section of the journal  
Frontiers in Veterinary Science

**Received:** 30 January 2018

**Accepted:** 03 April 2018

**Published:** 11 May 2018

### Citation:

McDonald JL and Hodgson DJ  
(2018) Prior Precision, Prior  
Accuracy, and the Estimation of  
Disease Prevalence Using Imperfect  
Diagnostic Tests.  
*Front. Vet. Sci.* 5:83.  
doi: 10.3389/fvets.2018.00083

Estimates of disease prevalence in any host population are complicated by uncertainty in the outcome of diagnostic tests on individuals. In the absence of gold standard diagnostics (tests that give neither false positives nor false negatives), Bayesian latent class inference can be applied to batteries of diagnostic tests, providing posterior estimates of the sensitivity and specificity of each test, alongside posterior estimates of disease prevalence. Here we explore the influence of precision and accuracy of prior information on the precision and accuracy of posterior estimates of these key parameters. Our simulations use three diagnostic tests, yielding eight possible diagnostic outcomes per individual. Seven degrees of freedom allow the estimation of seven parameters: sensitivity and specificity of each test, and disease prevalence. We show that prior precision begets posterior precision but only when priors are accurate. We also show that analyses without gold standard can use imprecise priors as long as they are initialised with accuracy. Imprecise priors risk the divergence of MCMC chains towards inaccurate posterior estimates, if inaccurate initial values are used. We note that inaccurate priors can yield inaccurate and imprecise inference. Bounded priors should certainly not be used unless their accuracy is well established. Inaccurate estimates of sensitivity or specificity can yield wildly inaccurate estimates of disease prevalence. Our analyses are motivated by studies of bovine tuberculosis in a wild badger population.

**Keywords:** diagnostics, Bayesian inference, sensitivity, specificity, prevalence, bovine tuberculosis, accuracy, precision

## INTRODUCTION

Uncertainty lies at the heart of real-world epidemiology. While hosts might be truly infected or uninfected, and diseased or not, our observation of these states suffers from imperfect detection of hosts, infection and disease. Ecological studies tend to deal with imperfect host detection using capture-mark-recapture methodologies [e.g., (1)], with limited consideration of biases in diagnoses themselves [but see (2)]. However, imperfect pathogen detection is a common occurrence when sampling live populations, with studies often drawing conclusions from the results of one or more tests, none of which are 100% accurate [e.g., (3)]. This is important because methods for the accurate detection of disease are pivotal to surveillance programmes that focus on the spatial and temporal spread of pathogens within and between populations, with infection prevalence often the primary parameter of interest (4–7).

For many wildlife diseases, post-mortem analysis provides gold standard diagnosis, however this is not an option for most ecological studies of wild animals, where corpses are hard to find and where an understanding of natural disease dynamics is the primary goal. Given the rarity of gold standard diagnostics for live animals, the development of statistical approaches for the evaluation of imperfect diagnostic tests has been an active field of research applied to human [e.g., (8)] and veterinary medicine [e.g., (9)]. These approaches can account for misclassification of both test positive individuals and test negative individuals. Specifically, they quantify test sensitivity, which measures the probability of a positive test outcome caused by the individual having the disease (the probability of a true positive outcome), and test specificity, which measures the probability of a negative test caused by the individual not having the disease (the probability of a true negative outcome).

Accounting for test sensitivity and specificity is vital to the accurate estimation of disease prevalence in a host population. Bayesian latent class analysis evaluates the performance of diagnostic tests in the absence of a reference test (10, 11) and consequently provides suitably adjusted estimates of prevalence. The development of Bayesian approaches for diagnostic test evaluation now provides a means to simultaneously estimate the performance of multiple tests in the light of the others and provide accurate estimates of disease prevalence that accounts for diagnostic uncertainty (12). A desirable aspect of Bayesian analysis is the pairing of data with prior information to estimate parameters that may have been previously unidentifiable, offering practical advantages over frequentist approaches (13). Despite these obvious benefits, often little is known about test performance in the field and the determination of an informative prior can be a challenging and subjective process. Although we know the choice of a prior contributes to the posterior distribution (12), it remains unclear how such prior sensitivity affects our inference regarding disease prevalence.

Here we explore how the accuracy and precision of prior information influence conclusions regarding disease prevalence. Our objectives are threefold. First, using simulated data for three diagnostic tests based on a previous analysis of bovine tuberculosis in a wild population of European badgers (9), we ask: (1) how does parameter identifiability compare between models that use precise (informative) and imprecise (vague or uninformative) priors? (2) How do inaccurate priors for test sensitivity and specificity

influence conclusions regarding overall prevalence? (3) How accurate are prevalence estimates that rely on the assumption that specificity and sensitivity are both perfect (i.e., both = 1)? From an applied perspective, disentangling disease processes from test performance will aid in forecasting long-term dynamics and in developing control strategies. This is especially important as diagnostic uncertainty has the potential to hinder eradication efforts by masking or exaggerating observed disease patterns. However, without robust statistical approaches it is impossible to understand how such bias may influence prevalence estimates.

## MATERIALS AND METHODS

### Simulated Data

We simulated results of three independent diagnostic tests for 875 individuals, using the prior modes of prevalence, specificity and sensitivity taken from Drewe et al. (9) (Table 1). The observations lead to a cross-classification table for the joint test results  $y = (y_{111}, y_{121}, y_{112}, y_{122}, y_{211}, y_{221}, y_{212}, y_{222})$  where  $y_{111}$  is number of sampled individuals that tested positive for all three diagnostic tests and  $y_{222}$  is number of sampled individuals that tested negative for all three diagnostic tests.

### Assessment of Prevalence Estimation

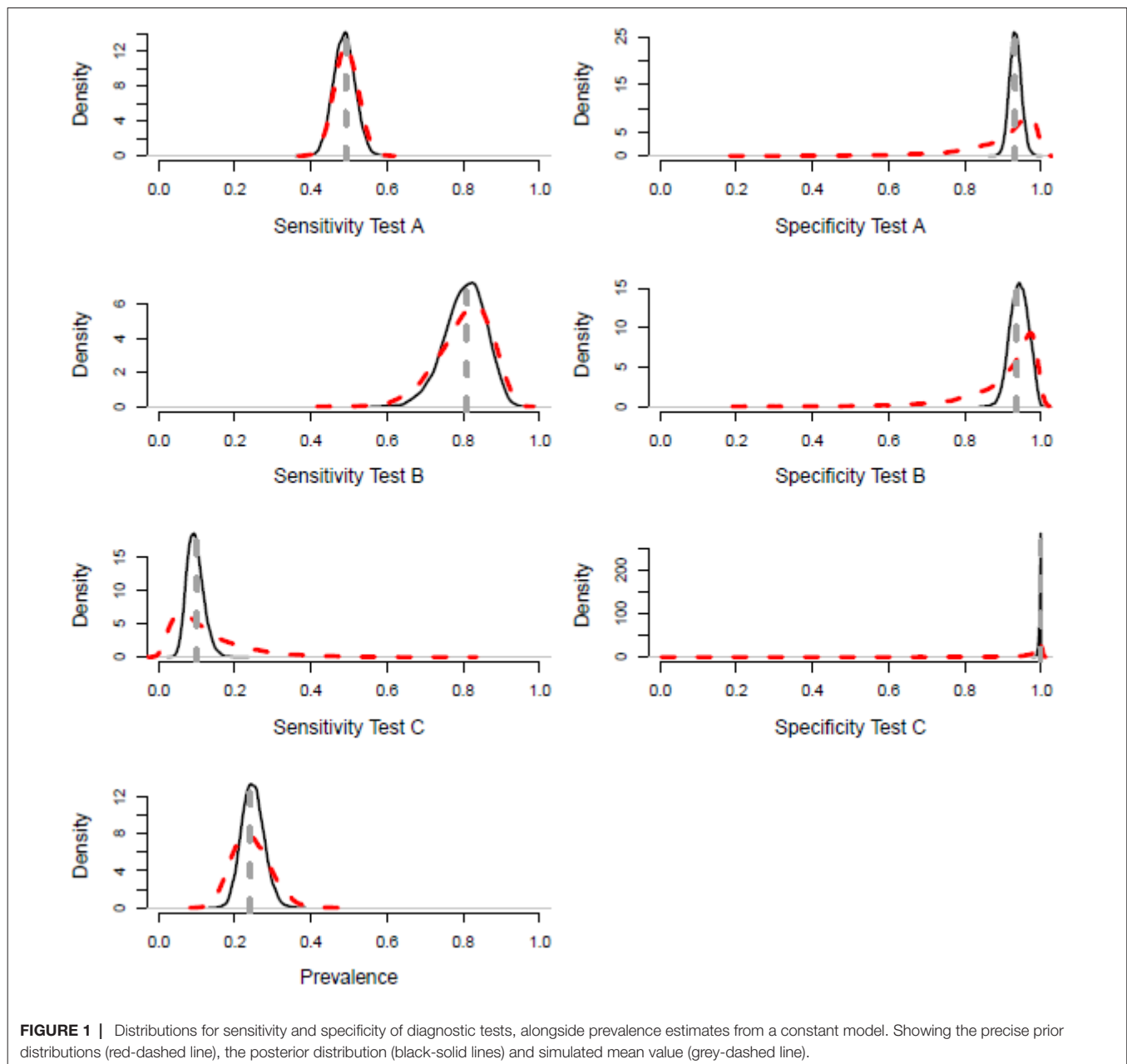
The parameters for the model include the three sensitivities ( $Se$ ), three specificities ( $Sp$ ) and prevalence ( $\pi$ ). Biological independence between tests was assumed. Multinomial cell probabilities for the population are given by:

$$\begin{aligned}
 y &\sim \text{multinomial}(n, p_{jjj}), j \in \{1, 2\} \\
 p_{111} &= \pi Se_A Se_B Se_C + (1 - \pi)(1 - Sp_A)(1 - Sp_B)(1 - Sp_C) \\
 p_{121} &= \pi Se_A(1 - Se_B)Se_C + (1 - \pi)(1 - Sp_A)Sp_B(1 - Sp_C) \\
 p_{112} &= \pi Se_A Se_B(1 - Se_C) + (1 - \pi)(1 - Sp_A)(1 - Sp_B)Sp_C \\
 p_{122} &= \pi Se_A(1 - Se_B)(1 - Se_C) + (1 - \pi)(1 - Sp_A)Sp_BSp_C \\
 p_{211} &= \pi Se_A Se_B(1 - Se_C) + (1 - \pi)Sp_A Sp_B(1 - Sp_C) \\
 p_{221} &= \pi(1 - Se_A)(1 - Se_B)Se_C + (1 - \pi)Sp_A Sp_B(1 - Sp_C) \\
 p_{212} &= \pi(1 - Se_A)Se_B(1 - Se_C) + (1 - \pi)Sp_A(1 - Sp_B)Sp_C \\
 p_{222} &= \pi(1 - Se_A)(1 - Se_B)(1 - Se_C) + (1 - \pi)Sp_A Sp_B Sp_C
 \end{aligned}$$

TABLE 1 | Simulated values taken from Drewe et al. (9) along with prior distributions used for precise, vague and badly stated priors.

Test	Parameter	Simulated values	Prior distributions on the logit scale Normal ( $\mu, \sigma$ )					
			Precise	Vague	Error	Error	Error	Error
Test A	Sensitivity	0.492	-0.032, 0.125	0, 1.65	<b>-0.847, 0.162</b>	<b>0.809, 0.139</b>	-0.032, 0.125	-0.032, 0.125
	Specificity	0.931	2.602, 1.073	0, 1.65	2.602, 1.073	2.602, 1.073	2.602, 1.073	2.602, 1.073
Test B	Sensitivity	0.809	1.443, 0.443	0, 1.65	1.443, 0.443	1.443, 0.443	1.443, 0.443	1.443, 0.443
	Specificity	0.936	2.683, 1.117	0, 1.65	2.683, 1.117	2.683, 1.117	<b>1.025, 0.271</b>	<b>6.907, 1.179</b>
Test C	Sensitivity	0.100	-2.197, 0.856	0, 1.65	-2.197, 0.856	-2.197, 0.856	-2.197, 0.856	-2.197, 0.856
	Specificity	0.999	4.595, 2.129	0, 1.65	4.595, 2.129	4.595, 2.129	4.595, 2.129	4.595, 2.129
	Prevalence	0.24	-1.153, 0.272	0, 1.65	-1.153, 0.272	-1.153, 0.272	-1.153, 0.272	-1.153, 0.272

Badly stated priors are precise but inaccurate for one of the sensitivities or specificities of the tests. These inaccuracies are emboldened in the body of the table. "Precise" priors are chosen to be precise on the probability scale, which means that canonical links between mean and variance can make them look imprecise on the logit scale when the mean is near zero or one. We provide the back-transformation of these priors onto the probability scale as red dashed lines in Figures 1 and 2.

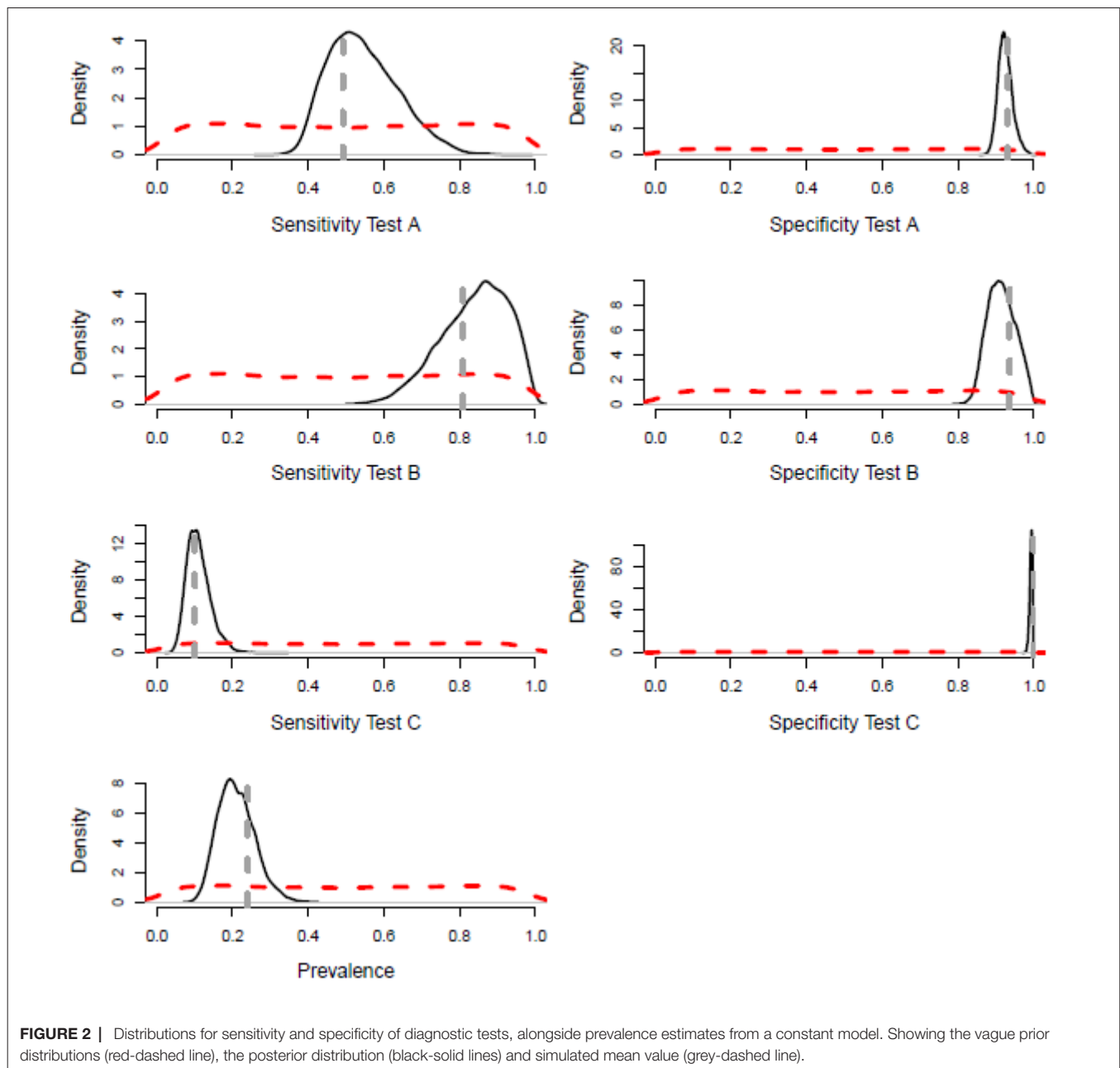


where  $j = 1$  describes a positive test outcome, and  $j = 2$  describes a negative test outcome for each of the three tests A, B and C. Prevalence  $\pi$  was set at 0.24.

To ensure parameters were bounded between 0 and 1 we modelled the logit of sensitivities, specificities and prevalence using Normal prior distributions with mean  $\mu$  and SD  $\sigma$ . We compared precise priors (small  $\sigma$ , implying strong prior knowledge) with vague priors (large  $\sigma$ , implying weak prior knowledge). Inaccurate but precise priors (implying strong prior belief in a wrong parameter value) were also incorporated to explore the influence of misinformed beliefs on inference of disease prevalence. Priors for the different modelling scenarios are shown in **Table 1**.

## Study System

Our simulations of test specificity and disease prevalence are based on the long term study of natural epidemiology of *M. bovis* among wild European badgers in Woodchester Park, Gloucestershire, UK. Here, badgers are regularly live trapped, sampled using diagnostic tests, and released (14). Three diagnostic tests were used routinely during the period assessed by Drewe et al. (9) to assess the infection status of each trapped individual. Blood samples are taken to test for antibodies to *M. bovis* using Stat-Pak (simulated as Test A), and further used to test for a cell-mediated response to *M. bovis* using interferon-gamma (IFN $\gamma$ ; Test B). Samples of faeces, urine, tracheal

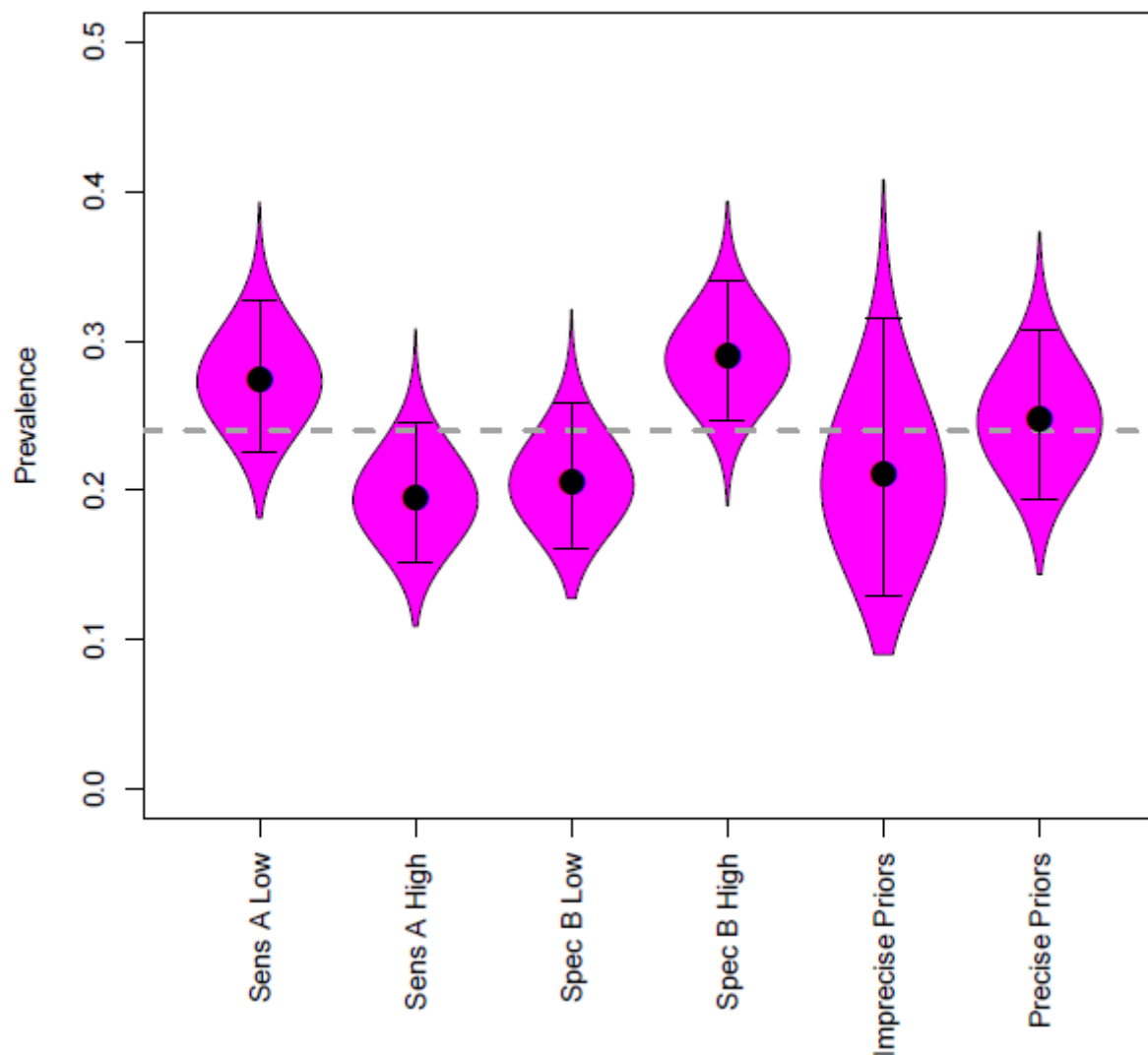


aspirate, oesophageal aspirate and swabs from bite wounds (where present) are collected for mycobacterial culture (Test C). Estimates of specificity and sensitivity of each of the diagnostic tests, used here to inform our simulations, are drawn from Drewe et al. (9), although these were subsequently updated by Buzdugan et al. (15). Given this description of the study system, it is difficult to justify the assumption of independence of test outcomes made by our simulations; tests A & B are applied to the same blood samples; the test sensitivities and specificities depend on different stages of disease progression in infected individuals. This, combined with the importance of the badger-TB system, is why we emphasise that this analysis is motivated by, but not definitive for, the prevalence of bovine tuberculosis

in badgers. A definitive analysis would have to account for non-independence of test outcomes, and for longitudinal patterns of disease progression within host individuals.

## Model Fitting

We fit all models using Bayesian methods and estimated the posterior distributions for all parameters using MCMC implemented in winBUGS (16) with the R2Winbugs package (17) in R (18). Convergence was assessed both visually ensuring mixing of the chains and formally within the model calculating the potential scale reduction factor ( $\hat{R}$ ). When  $\hat{R}$  is close to one we can be confident that convergence has been reached



**FIGURE 3 |** The effect of incorrectly specifying priors on disease prevalence estimates, compared to mean prevalence (grey dashed line). Along with prevalence obtained by uninformative and informative priors for all parameters.

(19). Consequently, posterior distributions were computed after a burn in of 5,000, followed by 50,000 iterations with a thinning interval of 10 iterations. WinBUGS code is provided as supplementary material.

## RESULTS

### Simulation Summary

By simulating data, we determine whether prevalence is estimated accurately (i.e., the true value should lie within 95% bounds of credibility of the posterior estimate) and precisely (i.e., the 95% bounds of credibility are usefully tight around the posterior estimate), using (1) raw test outcomes assuming perfect sensitivity and specificity, (2) models with incorrect,

precise priors, (3) models with accurate, precise priors and (4) models with imprecise (vague) priors.

### Simulated Population

A constant Bayesian model (i.e., assuming no variation in epidemiological or diagnostic parameters through time; (9, 11) with precise priors successfully estimated the diagnostic performance values and disease prevalence of the simulated population (Figure 1). The model also performed accurately with imprecise priors, although the posterior distributions were less precise, accounting for the additional uncertainty (Figures 2 and 3, Table 2). However, we note that with imprecise priors, the chains required realistic initial values to ensure convergence. Parallel chains occurred in models that excluded informative

priors and included randomly assigned initial values, indicating more than one area of high posterior probability. However, with realistic initial starting values for the chains all parameters were identifiable, despite vague priors (Figures 2 and 3).

Precise priors can aid identifiability, however with often limited information regarding the performance of diagnostic tests we explore how the accuracy of the prior information impacts conclusions regarding disease prevalence using four scenarios; (1) prior specifies lower test A sensitivity ( $\mu = 0.292$ ); (2) prior specifies higher test A sensitivity ( $\mu = 0.692$ ); (3) prior specifies lower test B specificity ( $\mu = 0.736$ ); (4) prior specifies higher test B specificity ( $\mu = 0.999$ ).

Incorrect specification of priors leads to inaccurate prevalence estimates (Figure 3). Prior information that reduces the sensitivity of a test leads to assumptions of more false negatives and consequently increased prevalence estimates (Figure 3). In contrast, prior information that increases the sensitivity of a test leads to assumptions of fewer false negatives and decreased prevalence estimates (Figure 3).

Incorrect priors on specificity values changes assumptions regarding false positives. Reduced test specificity, caused by badly specified priors, infers higher numbers of false-positives and therefore reduces estimates of disease prevalence (Figure 3). The flipside of this is that increased test specificity assumes lower rates of false positives and increases estimates of prevalence (Figure 3).

Raw values from diagnostic tests with low sensitivity produced inaccurate estimates of prevalence when perfect diagnoses were assumed (Table 2). We find that test B, which has high relative sensitivity and specificity, and therefore the lowest number of false negatives and low numbers of false positives, naturally provided the most accurate estimate of prevalence (Table 2).

## DISCUSSION

### Methodological Guidance

*Raw diagnostic outcomes should not be used to infer prevalence.* In the absence of knowledge regarding sensitivity and specificity of

the underlying tests, using the raw test results can provide highly inaccurate estimates of prevalence.

*To be certain of convergence run multiple chains.* From the simulations we have ascertained that parameters are identifiable despite vague priors. However, realistic starting values are required as chains can get stuck in different regions, indicating a bottleneck between two regions of high probability, or multiple posterior modes. The occurrence of multiple posterior modes would not be identified without running multiple chains. If realistic initial values are unknown, then inference of the “true” state should be made with caution.

*Beware precise priors.* Inaccurate, precise priors can bias parameter estimates and can lead to inaccurate conclusions regarding prevalence. This will be particularly true if priors are bounded above and/or below, for example if using Uniform prior distributions. Diagnosis of such inaccuracy might be possible if posterior distributions are clustered at either bound, however we recommend the avoidance of bounded priors unless prior knowledge is strong and accurate.

## General Discussion

Disease prevalence is fundamental to our understanding of wildlife epidemiology and is often the focal parameter when it comes to deriving management recommendations. However, the accuracy of diagnostics used to live sample wild populations is often uncertain and has the potential to alter conclusions regarding the prevalence of disease. Indeed, we have shown the importance of accounting for bias in determining disease parameters within wild populations, with raw data usually providing a poor representation of the true prevalence. Using a simulated population, we demonstrated how Bayesian latent class analysis can provide accurate estimates of test performance and infection prevalence in the absence of a reference test. However, this method is not without its caveats as poorly specified priors can heavily influence the posterior, altering conclusions regarding prevalence of infection.

Prevalence only provides a reliable surveillance indicator if issues of diagnostic uncertainty are accounted for. As our study shows, tests that have low sensitivities can vastly underestimate prevalence, due to the inclusion of high numbers of false negative individuals. Further, inaccuracies in prior information on specificity can yield inaccurate inference of credible ranges of prevalence, due to poor inference of false positive diagnoses. In the absence of knowledge regarding specificity and sensitivity, some ecological studies attempt to minimise the risk of misclassification by focusing on individuals that are diseased rather than infected, because disease tends to be accompanied by visible symptoms (2, 20). While this approach minimises false positives, it accepts that false negatives can occur and risks the underestimation of true prevalence. We also note that estimates of prevalence based on raw outcomes of diagnostic tests can also be useful for longitudinal studies of “relative” prevalence, or for comparisons among populations [e.g., (6)].

Latent class analysis provides a solution, and has been used extensively to estimate both sensitivities and specificities of diagnostic tests in the absence of a reference test across a range

**TABLE 2 |** Prevalence estimates derived from raw outcomes of each test (assuming perfect sensitivity and specificity), and inferred using latent class analysis, along with 95% credible intervals.

Method to estimate prevalence	Prevalence
<i>Assume perfect diagnostics</i>	
Test A	0.170
Test B	0.242
Test C	0.025
<i>Modelling diagnostic uncertainty</i>	
Precise priors	Mean (95% Credible Interval)
Vague priors	0.248 (0.194, 0.308)
	0.209 (0.129, 0.311)
<i>Incorrect prior specification</i>	
Reduced sensitivity test A	0.274 (0.226, 0.327)
Increased sensitivity test A	0.196 (0.152, 0.249)
Reduced specificity test B	0.206 (0.161, 0.259)
Increased specificity test B	0.289 (0.247, 0.338)

*The true value of prevalence in our simulation is 0.24.*

of diseases (9–11, 21). A prerequisite of Bayesian analysis is the combination of information from data and prior information. Commonly, when there is a lack of information surrounding parameters, vague priors are used to ensure the posterior distributions are driven by the data alone. This is advisable given we find posteriors to be sensitive to the selection of priors, similar to findings from previous studies (12). However, when a model is not identifiable, for example when there are more parameters than degrees of freedom, constraints or/and informative prior distributions are required to obtain a computational solution (8). When required for identifiability issues, informative priors should be used cautiously and only on parameters with a strong knowledge base. In these scenarios multiple chains are recommended to check for prior sensitivity. Additionally, in our simulated example (7 parameters, 7 degrees of freedom), we find parameters are identifiable but suffer from convergence issues if initial values are randomly assigned. This indicates more than one area of high posterior probability. Using precise initial values enables convergence in one area of parameter space, but accuracy can only be assumed if the initial values are themselves accurate. However, when there is a complete lack of knowledge then application of different starting values to explore differing areas of high posterior probability will be required.

Heterogeneity in test performance across tests and populations are commonly incorporated within the Bayesian approach presented here (11). The modelling we have performed did not consider variation in test performance and prevalence of infection through time. We recommend further development of latent class analysis to incorporate variation or trends in sensitivity, specificity and prevalence through time. Similar to the benefits of extending data input across populations (11), an advantage of incorporating time-varying data on diagnostic outcomes is an increase in the degrees of freedom provided by decomposition of diagnostic outcomes into timesteps. The cost of such a decomposition will be a reduction in the sample size of individuals diagnosed per timestep. As well as variation through space and time, diagnostic outcomes will vary according to the circumstances of the individual being tested. Detecting the clinical status of the host may depend on a range of additional factors such as their age, sex, coinfections, disease severity or stress. For example, pathogen detection may vary as a function of pathogen load, with seropositivity associated with advanced stages of disease (22). Reconciling links between both individual and population level test performance in an integrated framework is an important area of future research.

The choice of latent class model, for the estimation of diagnostic test performance and infection prevalence, will always depend on the purpose of the study. Latent class analysis, or “diagnosis without gold standard” relies on the use of multiple diagnostic tests to infer sensitivity, specificity and prevalence. If the performance of diagnostic tests is the reason for study, then it makes sense to study multiple populations that vary in prevalence, or a single population that varies through time, with one reliable test to use as a standard. If tests cannot be assumed to be independent, then covariance must be modelled

and more tests or replicate populations are required to cope with the demand on degrees of freedom. Mostly, in studies of wild host populations, there is no gold standard and little knowledge of test dependencies, and the simple approach we expound here is a good starting point. We have only simulated a single population to make our point about precision and accuracy, but the supplementary code (Data Sheet S1) we provide allows interested readers to explore various scenarios. A benefit of BUGS code is that it can be used as a building block for more complex modelling scenarios.

## CONCLUSION

Understanding links between diagnostic uncertainty and prevalence provides a key to explaining and predicting the population dynamics of infected hosts, and ultimately informs the development, and tests the efficacy, of management for disease control. We demonstrate the utility of Bayesian latent class analysis, developed to assess diagnostic sensitivity and specificity in the absence of gold standard tests. Analysis revealed complexities underpinning misclassification bias, including inaccuracy and imprecision of priors, which fundamentally influence our understanding of disease dynamics within wildlife-host systems. Increasingly epidemiological data is available at population and temporal scales necessary to estimate diagnostic parameters. We therefore recommend further model development and ultimately application of this approach to surveys in other populations and species.

## AUTHOR CONTRIBUTIONS

JM coded the models and drafted the paper. DH motivated the research, re-analysed the models and write the final draft.

## FUNDING

This research was supported by NERC grant NE/L007770/1 and the University of Exeter.

## ACKNOWLEDGEMENT

The authors thank Dez Delahay, Freya Smith and Julian Drewe for discussion of diagnostic modelling for bovine tuberculosis, and two anonymous referees for their review of this work.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fvets.2018.00083/full#supplementary-material>

## REFERENCES

- Graham J, Smith GC, Delahay RJ, Bailey T, McDonald RA, Hodgson D. Multi-state modelling reveals sex-dependent transmission, progression and severity of tuberculosis in wild badgers. *Epidemiol Infect* (2013) 141(7):1429–36. doi: 10.1017/S0950268812003019
- Lachish S, Gopalaswamy AM, Knowles SCL, Sheldon BC. Site-occupancy modelling as a novel framework for assessing test sensitivity and estimating wildlife disease prevalence from imperfect diagnostic tests. *Methods Ecol Evol* (2012) 3(2):339–48. doi: 10.1111/j.2041-210X.2011.00156.x
- Greer AL, Collins JP. Sensitivity of a diagnostic test for amphibian Ranavirus varies with sampling protocol. *J Wildl Dis* (2007) 43(3):525–32. doi: 10.7589/0090-3558-43.3.525
- Donnelly CA, Hone J. Is There an Association between Levels of Bovine Tuberculosis in Cattle Herds and Badgers? *Stat Commun Infect Dis* (2010) 2(1). doi: 10.2202/1948-4690.1000
- Boadella M, Gortazar C, Acevedo P, Carta T, Martín-Hernando MP, de La Fuente J, et al. Six recommendations for improving monitoring of diseases shared with wildlife: examples regarding mycobacterial infections in Spain. *Eur J Wildl Res* (2011) 57(4):697–706. doi: 10.1007/s10344-011-0550-x
- Delahay RJ, Walker N, Smith GC, Smith GS, Wilkinson D, Clifton-Hadley RS, et al. Long-term temporal trends and estimated transmission rates for *Mycobacterium bovis* infection in an undisturbed high-density badger (*Meles meles*) population. *Epidemiol Infect* (2013) 141(7):1445–56. doi: 10.1017/S0950268813000721
- Wobeser GA. *Essentials of disease in wild animals*. John Wiley & Sons (2013).
- Gonçalves L, Subtil A, de Oliveira MR, do Rosário V, Lee PW, Shiao MF. Bayesian Latent Class Models in malaria diagnosis. *PLoS ONE* (2012) 7(7):e40633. doi: 10.1371/journal.pone.0040633
- Drewe JA, Tomlinson AJ, Walker NJ, Delahay RJ. Diagnostic accuracy and optimal use of three tests for tuberculosis in live badgers. *PLoS ONE* (2010) 5(6):e11196. doi: 10.1371/journal.pone.0011196
- Enøe C, Georgiadis MP, Johnson WO. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Prev Vet Med* (2000) 45(1-2):61–81. doi: 10.1016/S0167-5877(00)00117-3
- Branscum AJ, Gardner IA, Johnson WO. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Prev Vet Med* (2005) 68(2-4):145–63. doi: 10.1016/j.prevetmed.2004.12.005
- Álvarez J, Perez A, Bezos J, Marqués S, Grau A, Saez JL, et al. Evaluation of the sensitivity and specificity of bovine tuberculosis diagnostic tests in naturally infected cattle herds using a Bayesian approach. *Vet Microbiol* (2012) 155(1):38–43. doi: 10.1016/j.vetmic.2011.07.034
- Dunson DB. Commentary: practical advantages of Bayesian analysis of epidemiologic data. *Am J Epidemiol* (2001) 153(12):1222–6. doi: 10.1093/aje/153.12.1222
- McDonald JL, Robertson A, Silk MJ. Wildlife disease ecology from the individual to the population: Insights from a long-term study of a naturally infected European badger population. *J Anim Ecol* (2018) 87(1):101–112:101–12. doi: 10.1111/1365-2656.12743
- Buzdugan SN, Vergne T, Grosbois V, Delahay RJ, Drewe JA. Inference of the infection status of individuals using longitudinal testing data from cryptic populations: Towards a probabilistic approach to diagnosis. *Sci Rep* (2017) 7(1):1111. doi: 10.1038/s41598-017-00806-4
- Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* (2000) 10(4):325–37. doi: 10.1023/A:1008929526011
- Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. *J Stat Softw* (2005) 12:1–16.
- R development core team. *R: A language and environment for statistical computing [Internet]*. Vienna, Austria: (2017).
- Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* (1998) 7:434–55.
- McDonald JL, Bailey T, Delahay RJ, McDonald RA, Smith GC, Hodgson DJ. Demographic buffering and compensatory recruitment promotes the persistence of disease in a wildlife population. *Ecol Lett* (2016) 19(4):443–9. doi: 10.1111/ele.12578
- Rahman AK, Saegerman C, Berkvens D, Fretin D, Gani MO, Ershaduzzaman M, et al. Bayesian estimation of true prevalence, sensitivity and specificity of indirect ELISA, Rose Bengal Test and Slow Agglutination Test for the diagnosis of brucellosis in sheep and goats in Bangladesh. *Prev Vet Med* (2013) 110(2):242–52. doi: 10.1016/j.prevetmed.2012.11.029
- Chambers MA, Waterhouse S, Lyashchenko K, Delahay R, Sayers R, Hewinson RG. Performance of TB immunodiagnostic tests in Eurasian badgers (*Meles meles*) of different ages and the influence of duration of infection on serological sensitivity. *BMC Vet Res* (2009) 5:42. doi: 10.1186/1746-6148-5-42

**Conflict of interest statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Hodgson and McDonald. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.