Bayesian latent class estimation of sensitivity and specificity parameters of diagnostic tests for bovine tuberculosis in chronically infected herds in Northern Ireland

Title: Bayesian latent class estimation of sensitivity and specificity parameters of diagnostic tests for bovine tuberculosis in chronically infected herds in Northern Ireland

Authors: A. Lahuerta-Marin, M.G. Milne, J. McNair, R.A. Skuce, S.H. McBride, F.D. Menzies, S.J.W. McDowell, A.W. Byrne, I.G. Handel, B.M. de C. Bronsvoort

PII: S1090-0233(18)30137-0
DOI: https://doi.org/10.1016/j.tvjl.2018.04.019
Reference: YTVJL 5156

To appear in:

Accepted date: 26-4-2018

Please cite this article as: A.Lahuerta-Marin, M.G.Milne, J.McNair, R.A.Skuce, S.H.McBride, F.D.Menzies, S.J.W.McDowell, A.W.Byrne, I.G.Handel, B.M.de C.Bronsvoort, Bayesian latent class estimation of sensitivity and specificity parameters of diagnostic tests for bovine tuberculosis in chronically infected herds in Northern Ireland (2010), https://doi.org/10.1016/j.tvjl.2018.04.019

Original Article

Bayesian latent class estimation of sensitivity and specificity parameters of diagnostic tests for bovine tuberculosis in chronically infected herds in Northern Ireland

A. Lahuerta-Marin [a,*], M.G. Milne [a], J. McNair [a], R. A. Skuce [a], S.H. McBride [a], F.D. Menzies [b], S.J.W. McDowell [a], A.W. Byrne [a, d], I.G. Handel [c], B.M. de C. Bronsvoort [c]

[a] *Agri-Food and Biosciences Institute (AFBI), Veterinary Sciences Division, Stoney Road, Belfast BT4 3SD, United Kingdom*
[b] *Veterinary Epidemiology Unit, Department of Agriculture, Environment and Rural Affairs (DAERA), Dundonald House, Belfast, BT4 3SB, United Kingdom*
[c] *The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Edinburgh EH25 9RG, United Kingdom*
[d] *School of Biological Sciences, Queen's University Belfast, Belfast BT7 1NN, United Kingdom*

* Corresponding author. Tel.: +44 2890525783.
E-mail address: angela.lahuerta-marin@afbini.gov.uk (A. Lahuerta-Marin).

Highlights

- Performance characteristics of bovine tuberculosis (bTB) tests were determined in chronically infected herds.
- Bayesian latent class analysis was used to estimate sensitivities and specificities.
- The single intradermal cervical comparative tuberculin (SICCT) test may have a lower sensitivity than reported previously.
- Combinations of SICCT and interferon-γ tests, and meat inspection data, affected estimates of sensitivity and specificity.
- The ability of SICCT and interferon-γ tests to detect infected animals was lower in dairy herds than beef herds.

Abstract

In the European Union, the recommended ante-mortem diagnostic methods for bovine tuberculosis (bTB) include the single intradermal cervical comparative tuberculin (SICCT) test and

the interferon-gamma (IFN-γ) test as an ancillary test. The SICCT test has a moderate sensitivity (Se) and high specificity (Sp), while the IFN-γ test has good Se, but a lower Sp than the SICCT test. A retrospective Bayesian latent class analysis was conducted on 71,185 cattle from 806 herds chronically infected with bTB distributed across Northern Ireland (NI) to estimate the Se and Sp of the common ante-mortem tests and meat inspection. Analyses were also performed on data stratified by farming type and herd location to explore possible differences in test performance given the heterogeneity in the population. The mean estimates in chronically infected herds were: (1) 'standard' SICCT: Se 40.5-57.7%, Sp 96.3-99.7%; (2) 'severe' SICCT: Se 49.0%-60.6%, Sp 94.4-99.4%; (3) IFN-γ(bovine–avian) using a NI optical density (OD) cut-off difference of 0.05: IFN-γ$(B–A)_{NI}$: Se 85.8-93.0%, Sp 75.6-96.2%; (4) IFN-γ(bovine–avian) using a standard 'commercial' OD cut-off difference of 0.1: IFN-γ$(B–A)_{0.1}$: Se 83.1-92.1%, Sp 83.1-97.3%; and (5) meat inspection: Se 49.0-57.1% Se, Sp 99.1-100%. Se estimates were lower in cattle from dairy farms than from beef farms. There were no notable differences in estimates by location of herds. Certain population characteristics, such as production type, might influence the ability of bTB tests to disclose truly infected cases.

Keywords: Bovine tuberculosis; Bayesian latent class analysis; Sensitivity; Specificity; Single intradermal cervical comparative tuberculin test; Interferon-γ test

Introduction

Bovine tuberculosis (bTB) is a chronic infectious disease caused by *Mycobacterium bovis* and cattle are considered to be the main reservoir of infection (Allen et al., 2010). Bovine tuberculosis is a serious animal health problem that leads to economic and international trade restrictions for those countries that are not officially tuberculosis free (OTF) (Zinsstag, 2006; Good and Duignan, 2011). bTB has important economic consequences in Northern Ireland (NI), where cattle are an important

part of the local economy and bTB is considered to be endemic[1]. Despite an ongoing eradication scheme, costing ~£30 million per annum[2], bTB has proved difficult to eliminate, perhaps because of the complex and multifactorial nature of the disease (Humblet et al., 2009). Some herds in NI are also 'chronically infected', and experience prolonged breakdowns or recurrent bTB infection (Doyle et al., 2016).

The eradication programme in NI is based on a 'test and-cull' strategy. Two of the recommended ante-mortem screening tests for bTB include the single intradermal comparative cervical tuberculin (SICCT) test and the interferon gamma (IFN-γ) test, which is approved as an ancillary test as specified in European Union (EU) Council Directive EC/1226/2002 amending Annex B to Directive 64/432/EEC[3]. Both tests have well documented limitations in terms of performance characteristics (de la Rua-Domenech et al., 2006). The SICCT test can have a 'standard' or 'severe' interpretation. A standard interpretation is read when there is a difference of greater than 4 mm between the thickness at the site of injection of the bovine antigen and the site of injection of the avian antigen. A severe interpretation is one in which the thickness at the bovine site is greater than the avian site by 2 mm. The standard interpretation has a good specificity (Sp > 99%), but previous studies have suggested that it can have relatively poor sensitivity (Se 51-85%) (Nunez-Garcia et al., 2017). This makes the SICCT test a good initial herd screening test, but with serious limitations if applied to chronically infected herds, because of the high probability of leaving undetected infection (Lahuerta-Marin et al., 2015, 2016).

Since 2004, the IFN-γ test has also been used in NI as a voluntary ancillary test in parallel with the SICCT test in high risk and chronically infected herds, with the main aim of identifying

[1] See: https://www.daera-ni.gov.uk/publications/tuberculosis-disease-statistics-northern-ireland-2014 (accessed 8 October 2016).
[2] £1 = approx. US$1.41, €1.15 at 5 April 2018.
[3] See: https://publications.europa.eu/en/publication-detail/-/publication/b4b312f4-ba77-4cc7-99cd-5fc19c85d2cd/language-en (accessed 8 October 2016).

additional truly infected animals. Previous research indicates that the IFN-γ test using a standard 'commercial' optical density (OD) cut-off difference of 0.1 has a higher Se than the SICCT test (88-94%), but a lower Sp (85-98%) (de la Rua-Domenech et al., 2006; Clegg et al., 2011; Downs, 2011). In NI, a more stringent IFN-γ cut-off difference (OD 0.05) is applied than the commercial cut-off, which increases the Se of the test, but reduces the Sp. The application of both SICCT and IFN-γ tests together should improve the prospect of clearing bTB from herds where control has been problematic.

An estimate of the performance characteristics of each test is necessary in order to maximise the success in detecting infection, whilst minimising the number of false positive animals. This optimal scenario is difficult to achieve for bTB, where screening ante-mortem tests are imperfect, and there is no gold standard test, which makes it difficult to determine the true disease status of tested animals (Dohoo et al., 2009). To tackle the lack of a gold standard test, Bayesian latent class analysis has been used to estimate the performance characteristics of tests for bTB (Clegg et al., 2011; Alvarez et al., 2012; EFSA, 2012; Karolemeas et al., 2012; Bermingham et al., 2015; Nunez-Garcia et al., 2017). However, previous attempts to derive estimates for Se and Sp in Northern Ireland using Bayesian latent class analysis (EFSA, 2012; Bermingham et al., 2015) have left some gaps in our understanding, namely: (1) the derivation of SICCT test parameters under both severe and standard interpretation; (2) the derivation of IFN-γ parameters under both standard commercial (cut-off OD difference 0.1) and NI (cut-off OD difference 0.05) interpretation; and (3) differences in test performance between beef and dairy cattle.

The aim of this study was to undertake new Bayesian latent class analysis to estimate the Se and Sp of each of the standard and severe SICCT tests, the IFN-γ(bovine–avian) test using a NI OD cut-off difference of 0.05, i.e. IFN-γ$(B–A)_{NI}$, the IFN-γ(B–A) using a standard 'commercial' OD cut-off difference of 0.1, i.e. IFN-γ$(B–A)_{0.1}$, the IFN-γ test using the 6 kDa early secretory antigenic

target, i.e. IFN-γ(ESAT6), and detection of visible lesions consistent with bTB via meat inspection, on a robust dataset of animals from chronically infected herds tested across NI.

Materials and methods

Study population

The cohort of herds of cattle selected to participate in this retrospective study were based on their herd bTB histories. Only herds with chronic or recurrent infection, or very large recent breakdowns, were eligible for this cohort as part of the NI IFN-γ scheme from 2004 to 2010. The herd selection criteria for the IFN-γ scheme were applied as described by Lahuerta-Marin et al. (2015, 2016); ethical approval was not required for this study. Herd participation within the NI IFN-γ scheme was voluntary and the population under test was a convenience series; thus, the study population was not representative of the entire cattle population in NI. A total of 71,185 animals belonging to 806 cattle herds were included in the analysis; these animals were treated as a single population. The data set were curated and validated, and missing data were excluded. The analysis was carried out following the Standards for the Reporting of Diagnostic Accuracy Studies that Use Bayesian Latent Class Models (STARD-BLCM) guidelines (see Appendix: Supplementary material 1)[4].

Diagnostic testing

The SICCT test was performed according to requirements specified in EU Council Directive 64/432/EEC[5]. SICCT tests were performed by Department of Agriculture, Environment and Rural Affairs (DAERA) Veterinary Officers (VOs) and designated Local Veterinary Inspectors (LVIs) on behalf of DAERA within the participating Divisional Veterinary Office (DVO) regions. For IFN-γ testing, blood was collected during day 1 of the SICCT test before tuberculin injections by the DAERA VO, and samples were sent to AFBI for independent analysis within 8 h of collection. IFN-γ tests for purified protein derivatives (PPDs) for bovine–avian (B–A) and ESAT6 antigens were

[4] See: http://www.equator-network.org/reporting-guidelines/stard-blcm/ (accessed 8 October 2016).
[5] See: https://www.daera-ni.gov.uk/articles/bovine-tuberculosis-tb-legislation (accessed 8 October 2016).

performed, as described by Welsh et al. (2002). For IFN-γ(A–B), two cut-off differences were used to interpret the results: (1) NI cut-off, with a net difference between the bovine PPD (PPDB) and avian PPD (PPDA) ODs (PPDB–PPDA) ≥ 0.05, if PPDB ≥ 0.1 OD; and (2) the commercially recommended cut-off, with a net difference PPDB–PPDA ≥ 0.1 OD. For IFN-γ(ESAT6), a net OD cut-off difference of 0.05 was used. Meat inspection was carried out as described in Lahuerta-Marin et al. (2016), whereby all animals sent to slaughter were assessed by meat inspectors for lesions consistent with bTB.

Bayesian latent class analysis

Hui and Walter (1980) published a Bayesian latent class model to evaluate diagnostic tests in the absence of a 'gold standard' test; some of the assumptions of the approach under this two-test, two-population latent class model are: (1) when multiple populations are being compared, each population prevalence should be different; (2) the Se and Sp of the test are the same across test populations; and (3) the tests are conditionally independent. Whilst the two-test, two-population approach is the 'classical' latent class model, the framework can be extended to any scenario whereby $S \geq R/(2^{R-1} - 1)$, where S = number of populations and R = number of tests. Thus, a three-test, one population study system as presented here should be sufficient (Toft et al., 2005). The selection of a single study population, as opposed to two or more populations, reflects the limitations that come from identifying sub-populations which differ in prevalence, but not in how the each sub-population reacts to the test (Toft et al., 2005).

The analysis was performed as described by Hui and Walter (1980) and Bronsvoort et al. (2009), using three tests and one population. This model assumes that, for the population under study, the counts of the different combinations of test results for the three tests follow a multinomial distribution (i.e. + + +; + + –; – + +; – + –; + – +; + – –; – –+; – – –). In the model, seven parameters in total must be estimated, comprising the prevalence of the disease in the population and both the Se

and Sp for each of the three tests. The data provide seven degrees of freedom; therefore the number of degrees of freedom are equal to the number of parameters to be estimated when conditional independence is assumed (Jones et al., 2010; Lewis and Torgerson, 2012). However, both the SICCT and IFN-γ tests are based on cell-mediated immunity and, therefore, are not independent. The implications of conditional dependence between tests were explored by running separate models accounting for model co-variance (see Appendix: Supplementary material 2: Part 1); however these models may be considered non-identifiable, requiring nine parameters to be estimated with only seven degrees of freedom (Toft et al., 2007; Jones et al., 2010). Where model selection took place, lower deviance information criterion (DIC) values were used to identify better fitting models.

Our model compared the distribution of results from various combinations of tests, i.e. SICCT standard, SICCT severe, IFN-γ$(B–A)_{NI}$, IFN-γ$(B–A)_{0.1}$, IFN-γ(ESAT6) and meat inspection, to derive estimates for the Se and Sp of each test (Toft et al., 2005). Where meat inspection data were available (n = 49,540), the first and second tests were the SICCT test and one of the IFN-γ tests, respectively, while the third test was the meat inspection data. However, parameter estimates for the whole population were derived using only SICCT, IFN-γ(B–A) and IFN-γ(ESAT6) tests, since meat inspection data were not available for the whole population. The latent state being estimated was the 'true' infection status of animals, which was defined as an animal infected with *M. bovis*; we made no inference as to whether the hosts were infectious, latently infected or had active infection.

All models were implemented in JAGS within the R statistical software environment, using packages *rjags* and *runjags*. Estimates of the prevalence were allowed to range from 0.1 to 0.3, following a uniform distribution. Parameter values for the Se and Sp of each test were estimated using flat, vague priors (*beta (1,1)*), although the impact of changing the priors was also explored (see Appendix: Supplementary material 2: Part 2). Three Monte-Carlo Markov chains (MCMCs) were each run for a total of 50,000 iterations, with the first 20,000 iterations discarded as a 'burn-in' and

the subsequent 30,000 iterations retained for posterior inference, with 'thinning' performed every 10 iterations. The estimates and the Bayesian credibility intervals were reported; the credibility interval represents the limits in which the parameter estimate falls, with 95% credibility. The outputs, including the MCMC trace-plots, posterior density distribution plots and cross-correlation plots were assessed further to ensure that autocorrelation was low. Chain convergence after the initial burn-in was assessed by visual inspection of the MCMCs and via diagnostics proposed by Brooks and Gelman (1998), whereby convergence is indicated with a value close to one. An example of the typical script used in this analysis is included in Supplementary material 3 (see Appendix).

Following the analysis to determine estimates of Se and Sp for the whole test population, the data set was split into individual groups and the same modelling approach that was applied to the whole test population was used to derive estimates for Se and Sp for each group independently. A main assumption of the classic Hui-Walter latent class model, with two tests and two populations, is that disease prevalence differs across populations, but that the test Se and Sp are the same. In practice, such a population split is difficult to derive. Therefore, at no point was a multi-population parameter estimation approach taken; every analysis was conducted using one group with three tests. The whole test population was firstly split by production type (i.e. dairy-beef) (Lahuerta-Marin et al., 2016), then the whole test population was then split on the basis of location in NI: (1) North (Coleraine, Ballymena, Larne and Londonderry); (2) South-West (Omagh, Enniskillen and Dungannon); and (3) South-East (Armagh, Newry and Newtownards).

Results

Descriptive results

A total of 71,185 animals from 806 herds were tested for bTB using the SICCT and IFN-γ(B–A) tests from 2004 to 2010; 60,594 animals were tested using the IFN-γ(ESAT6) test from 2005 to 2010, since IFN-γ(ESAT6) data were not available for 2004. Overall, 1.2% of animals tested were

SICCT reactors when the 'standard' interpretation was used; this proportion increased to 1.5% if the 'severe' interpretation was used. The percentages of animals positive in the IFN-γ tests were 6.5% in the IFN-γ$(B–A)_{NI}$ and IFN-γ(ESAT6) tests, and 4.7% in the IFN-γ$(B–A)_{0.1}$ test (Table 1).

Although valid meat inspection results were available for 49,540 animals in the initial cohort, 7840 animals were slaughtered within the first 2 months after ante-mortem testing. To avoid classification bias, these animals were excluded from the meat inspection data set; thus, 41,700 meat inspection results were included in the models for estimation of meat inspection performance. The distribution of results by visible lesion (VL) found at meat inspection is shown in Table 2. Of cattle included in the study, 51% were from dairy units and 49% were from beef herds. The distribution of animals tested by location was 43% in the South-East, 37% in the South-West and 19% in the North (total n = 71,185). The South-East region and animals from dairy herds accounted for the highest proportions of positive test results (Table 3).

Models

Estimates of Se ranged from 40.5% for the SICCT test with standard interpretation to 93.0% for the IFN-γ$(B–A)_{NI}$ test, without inclusion of meat inspection data (Table 4). The estimate for Se using the SICCT test with standard interpretation was 57.7% when the data set from meat inspection was used in the model. Estimates of Se were 93.0% in the IFN-γ$(B–A)_{NI}$ test and 92.1% in the IFN-γ $(B–A)_{0.1}$ test, without inclusion of meat inspection data (Table 4). The IFN-γ$(B–A)_{NI}$ test had a Se of 85.8%.

Sp estimates for the SICCT test with standard interpretation were 99.7% without inclusion of meat inspection data and 97.5% when meat inspection data were included. Estimates of Sp were 99.4% in the severe SICCT test, 95.8% in the IFN-γ$(B–A)_{NI}$ test, 97.0% in the IFN-γ$(B–A)_{0.1}$ test and

95.2% in the IFN-γ(ESAT6) test, without inclusion of meat inspection data. The estimated Sp of the IFN-γ$(B–A)_{NI}$ test was 83.0% when calculated using meat inspection data.

These estimates were derived using a model which assumed conditional independence. Whilst conditional dependence was considered (see Appendix: Supplementary material 1; Part 1), the parameter estimates between these models and the models which did not account for conditional dependence were minimal. Furthermore, the DIC values between models which accounted for conditional dependence and those that did not were small (< 3), indicating that the inclusion of a covariate term did not improve the fit of the model. The diagnostics for the models which included conditional dependence exhibited chain separation, high values for the potential scale reduction factor and higher levels of correlation between parameters than those derived in models which assumed conditional independence. Given that models which accounted for conditional dependence exhibited lower DIC values and a decline in model fit, whilst not markedly changing parameter estimates, the decision was made to proceed with those models with conditional independence. Although the models with conditional dependence were not identifiable, this was considered acceptable, since the simpler model (without priors) could deliver the same information regarding parameter estimates as the models with priors included, without compromising the model fit.

Production type

Se estimates for the SICCT standard test, the IFN-γ(ESAT6) test and meat inspection were 39.3%, 55.7% and 50.7% in dairy animals and 41.4%, 76.1% and 73.1% in beef animals, respectively, whereas Sp values for these tests were similar between dairy and beef animals (Table 5). Se estimates using the IFN-γ$(B–A)_{NI}$ test were 93.9% in dairy cattle and 92.6% in beef cattle without inclusion of meat inspection data, compared to 87.6% in dairy cattle and 89.8% in beef cattle when meat inspection data were included (Table 5).

Location of herd

Test performances among differing regions in NI are presented in Table 6. Se estimates for meat inspection were 55.7% in the North, 63.8% in the South-West and 62.9% in the South-East. Se estimates for the IFN-γ(ESAT6) test were 70.3% in the South-West, 65.4% in the North and 62.1% in the South-East without using meat inspection data, while Se estimates using the IFN-γ$(B–A)_{NI}$ test were 96.7% in the North, 92.7% in the South-East and 92.8% in the South-West. A similar pattern was observed if meat inspection data were used within the model. Sp estimates by test were similar among different regions.

Discussion

Bayesian latent class analysis is an accepted methodology to estimate the performance parameters of diagnostic tests in the absence of a 'gold standard'. This method has been used widely for assessing the performance of diagnostic tests in animal populations (Bronsvoort et al., 2009; Clegg et al., 2011). Our Se estimates for the SICCT test (40.5-57.7% Se at standard interpretation and 49.0-60.6% Se at severe interpretation), were lower than in previous studies performed with diagnostic data from NI. For example, the European Food Standards Authority (EFSA) Scientific Expert Opinion on the use of a IFN-γ test for diagnosis of bTB declared that the Se for the SICCT standard interpretation was 52.5-60.1% (EFSA, 2012); while Bermingham et al. (2015) estimated a 56% Se using the SICCT standard interpretation and 71% Se using the SICCT severe interpretation. On the other hand, Pollock et al. (2003) estimated a Se of 86% for NI, but with only 22 animals tested in the study.

Previous estimates derived from latent class models on data from the Republic of Ireland estimated a 56-67% Se for the SICCT test with standard interpretation and 65-77% Se for the SICCT test with severe interpretation (Clegg et al., 2011). In England, Se estimates for the SICCT test were 36-66 with standard interpretation and 48-72% with severe interpretation, depending on the model of

infection transmission (Conlan et al., 2012). However, it should be noted that the very low SICCT test Se values reported here are obtained when two IFN-γ tests are used. When a test combination of the IFN-γ test and meat inspection is used alongside the SICCT test, the Se estimates are higher, being 53.1%-57.7% at standard interpretation and 60.6% at severe interpretation.

Since the low values from the SICCT test observed here may be related to bias introduced from constructing models using two IFN-γ tests, i.e. IFN-γ(ESAT6) and IFN-γ(B–A), covariance between the two IFN-γ tests and the SICCT test was investigated. Our results indicated that model fit was not improved by the inclusion of covariance terms, while parameter values remained largely unaffected. Therefore, it is difficult to comment on whether fitting a model with two IFN-γ tests has contributed towards the low SICCT test Se values. The observed lack of covariance between the SICCT and IFN-γ tests has been reported previously from Germany and NI (Bermingham et al., 2015; Pucken et al., 2017). However, the models in these examples were identifiable when covariance was included, whereas the models presented here are not. This exemplifies the limitation of a single population latent class analysis with three non-gold standard tests, since a multi-population latent class analysis may have facilitated the addition of covariance terms to the model whilst remaining identifiable. Such an analysis was not feasible in the present study, since it was not possible to split the populations into groups which differed in prevalence but not in other test characteristics.

The Sp values of the SICCT test were also lower than expected, being 96.3-99.7% with standard interpretation and 94.4-99.4% with severe interpretation. The EFSA report also found that the Sp values for NI were < 99%, with a range of 96.4-97.5%; this may be related to the NI dataset, which included a ‘large proportion of positive IFN-γ test results in which both skin test and meat inspection results were negative’ (EFSA, 2012). The reasons for this are not clear, but may be related to a characteristic of the test population; the herds included in this study are chronically infected

herds, which have experienced prolonged, recurrent or severe breakdowns. Farmers may be choosing to cull IFN-γ positive animals in order to clear infection from the herd (Goodchild et al., 2015).

The Se estimates for the performance of the IFN-γ(B–A) test at different cut-offs were similar, with 92% Se at 0.1 OD cut-off difference and 93% Se using the current NI OD cut-off difference of 0.5. However, Sp at an OD cut-off difference of 0.1 was 97.3%, compared with the Sp at the current NI OD cut-off difference (95.8%). This information should be taken into consideration for possible changes to the current eradication programme in NI, specifically when both the SICCT and IFN-γ(B–A) tests are used. In a previous study using diagnostic data from NI, the IFN-γ(B–A) test had an estimated Se of 88.5-93.6%, but a lower estimated Sp (66.1-69.1%) than our estimates (EFSA, 2012).

Previous research suggests that dairy and beef cattle populations exhibit differences in their response to the SICCT test. One study found a higher risk of animals being misdiagnosed as false negatives in dairy herds compared to beef herds (Lahuerta-Marin et al., 2016). In the RI, dairy herds are also a higher risk group for bTB breakdowns (Doyle et al., 2014), a higher risk for residual infection (Lahuerta-Marin et al., 2016) and have an increased risk of disclosing a lesion at routine slaughter (Olea-Popelka et al., 2008). Under-performance in ante-mortem and post-mortem tests in dairy herds has been reported previously (O'Hagan et al., 2015; Lahuerta-Marin et al., 2016). Given these known differences, we explored heterogeneity between the beef and dairy cattle population test characteristics. In doing so, we also acknowledge that the results of this analysis, along with the analysis exploring regional differences in the SICCT test, should be interpreted with caution, since selection bias may exist between the groups. Furthermore, whilst it was not possible to conduct a multi-population latent class analysis, it is useful from a practical perspective to be aware of any heterogeneity within the cattle population.

Our estimates showed lower Se estimates for all tests in animals from dairy farms compared to beef farms, which is in line with current understanding. Whether this is a real observation or an artefact of the modelling process remains to be seen. However, lower Se of diagnostic tests in dairy herds has been reported previously (O'Hagan et al., 2015; Lahuerta-Marin et al., 2016) and may be explained in part because dairy cows generally live longer and remain longer within the herd than beef animals. Therefore, dairy animals can be increasingly exposed to bTB over time (Cagiola et al., 2004). Dairy breeds might have different levels of genetic resistance to *M. bovis* (Allen et al., 2010). Finally, cows in dairy herds may be under physiological stress due to pregnancy, which can affect the performance of bTB tests that rely on an immunological reaction (Buddle et al., 1994; Wood et al., 1991). Our results indicate that test characteristics differ only slightly with geography, with a marginally lower SICCT standard Se in the North region compared to the South-West South-East.

These results provide a useful estimate of the test characteristics of the SICCT standard, SICCT severe, IFN-γ$(B–A)_{NI}$, IFN-γ$(B–A)_{0.1}$ and IFN-γ(ESAT6) tests, and meat inspection, in chronically infected herds in NI. However, caution should be exercised when interpreting the scope of these test performance estimates, since the cattle included in the IFN-γ scheme are from chronically infected herds and have entered the scheme voluntarily, and thus may not be representative of the entire cattle population in NI. Additionally, only one population was used to derive the estimates as part of the latent class analysis. Using this single population of herds presents a technical limitation to latent class analysis, since estimates of all potential parameters that could be included in a model (potentially 13, comprising three Se and three Sp values, prevalence and six pairwise correlations) requires more degrees of freedom than provided by the data. Furthermore, the use of priors and covariance structures were thoroughly considered as part of the modelling process. However, simpler models without covariance and priors resulted in models with better fit, without compromising on parameter estimates. Thus, the use of this simpler models was considered acceptable in this circumstance.

This may be improved by exploring different covariance structures or by using a test whose specificity can be fixed at 1, thus reducing the number of parameters to be estimated by the model. We recommend that veterinary practitioners be mindful that the Se of the SICCT test may be lower than previously advised. We are also aware that a previous Bayesian latent class analysis to estimate some test performances was performed by EFSA (2012). However, there are significant differences with the analysis performed in this study that justifies this communication, and can be useful for policy making in NI. These two main differences are: (1) the modelling by EFSA only uses distributions of the tests with one part of the study, namely animals with meat inspection findings, whilst our results cover live and post-mortem groups of animals; and (2) the test performance estimates performed by EFSA do not include the IFN-γ(ESAT6) test, SICCT severe interpretation or the IFN-γ(B–A) test with the current NI cut-off OD difference of 0.1.

Conclusions

Our results show that the SICCT test, although a very good initial herd screening test, has a limited Se in herds with a history of persistent bTB problems. Thus, it is recommended for use together with an ancillary test, such as the IFN-γ(B–A) test, which increases the test Se. Furthermore, in this population it would be recommended to evaluate the IFN-γ(B–A) test at the commercial cut-off OD difference instead of the test at the current NI cut-off OD difference in terms of cost-effectiveness (marginal decrease in Se, for a much improved Sp). The ability of the ante-mortem tests to detect infected animals appeared to be lower in dairy herds than beef herds. The results from the present study, along with other recent research findings, suggest that dairy herds represent a sub-population of at-risk animals in terms of disease detection. Thus, it might be more difficult to clear infection from this type of herd with the currently available tests. These findings should be taken into account when reviewing surveillance and eradication programmes, especially in countries or regions where bTB is endemic, such as NI.

Conflict of interest statement

None of the authors has any financial or personal relationships that could inappropriately influence or bias the content of the paper.

Acknowledgements

This study is part of a larger project on the evaluation of the performance characteristics of the test in chronic bTB herds in NI from 2004 to 2010. It was financed by DAERA (E&I grant code 11/03/10). The authors would like to thank Dr Mairead Bermingham for useful comments and advice with the Bayesian models.

References

Allen, A.R., Minozzi, G., Glass, E.J., Skuce, R.A., McDowell, S.W.J., Woolliams, J.A., Bishop, S.C., 2010. Bovine tuberculosis: The genetic basis of host susceptibility. Proceedings of the Royal Society of London B: Biological Sciences 277, 2737-2745.

Alvarez, J., Perez, A., Bezos, J., Marques, S., Grau, A., Saez, J. L., Minguez, O., de Juan, L., Dominguez, L., 2012. Evaluation of the sensitivity and specificity of bovine tuberculosis diagnostic tests in naturally infected cattle herds using a Bayesian approach. Veterinary Microbiology 155, 38-43.

Bermingham, M.L., Handel, I.G., Glass, E.J., Woolliams, J.A., de Clare Bronsvoort, B.M., McBride, S.H., Skuce, R.A., Allen, A.R., McDowell, S.W.J., Bishop, S.C., 2015. Hui and Walter's latent-class model extended to estimate diagnostic test properties from surveillance data: a latent model for latent data. Scientific Reports 5, 11861.

Bronsvoort, B.M., Koterwas, B., Land, F., Handel, I.G., Tucker, J., Morgan, K.L., Tanya, V.N., Abdoel, T.H., Smits, H.L., 2009. Comparison of a flow assay for brucellosis antibodies with the reference cELISA test in West African *Bos indicus*. Public Library of Science One 4, e5221.

Brooks S.P., Gelman A., 1998. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7, 434-455.

Buddle, B.M., Aldwell, F.E., Pfeffer, A., de Lisle, G.W., Corner, L.A., 1994. Experimental *Mycobacterium bovis* infection of cattle: Effect of dose of *M. bovis* and pregnancy on immune responses and distribution of lesions. New Zealand Veterinary Journal 42, 167-172.

Cagiola, M., Feliziani, F., Severi, G., Pasquali, P., Rutili, D., 2004. Analysis of possible factors affecting the specificity of the gamma interferon test in tuberculosis-free cattle herds. Clinical and Diagnostic Laboratory Immunology 11, 952-956.

Clegg, T.A., Duignan, A., Whelan, C., Gormley, E., Good, M., Clarke, J., Toft, N., More, S.J., 2011. Using latent class analysis to estimate the test characteristics of the γ-interferon test, the single intradermal comparative tuberculin test and a multiplex immunoassay under Irish conditions. Veterinary Microbiology 151, 68-76.

Conlan, A.J., McKinley, T.J., Karolemeas, K., Pollock, E.B., Goodchild, A.V., Mitchell, A.P., Birch, C.P., Clifton-Hadley, R.S., Wood, J.L., 2012. Estimating the hidden burden of bovine tuberculosis in Great Britain. PLoS Computational Biology 8, e1002730.

de la Rua-Domenech, R., Goodchild, A.T., Vordermeier, H.M., Hewinson, R.G., Christiansen, K.H., Clifton-Hadley, R.S., 2006. Ante mortem diagnosis of tuberculosis in cattle: A review of the tuberculin tests, gamma-interferon assay and other ancillary diagnostic techniques. Research in Veterinary Science 81, 190-210.

Dohoo, I., Martin, W., Stryhn, H., 2009. Veterinary Epidemiologic Research, 2nd Edition. VER, Charlottetown, Prince Edward Island, Canada, pp. 91-127.

Downs, S.H., Parry, J., Nunez-Garcia, J., Abernethy, D.A., Broughan, J.M., Cameron, A.R. , Cook, A.J., de la Rua Domensch, R., Goodchild, A.V., Greiner, M., et al., 2011. Meta-analysis of diagnostic test performance and modelling of testing strategies for control of bovine tuberculosis in GB. Proceedings of the Society for Veterinary Epidemiology and Preventative Medicine (SVEPM), Leipzig, Germany, pp. 139-153.

Doyle, L.P., Gordon, A.W., Abernethy, D.A., Stevens, K., 2014. Bovine tuberculosis in Northern Ireland: Risk factors associated with time from post-outbreak test to subsequent herd breakdown. Preventive Veterinary Medicine 116, 47-55.

Doyle, L.P., Courcier, E.A., Gordon, A.W., O'Hagan, M.J.H., Menzies, F.D., 2016. Bovine tuberculosis in Northern Ireland: Risk factors associated with duration and recurrence of chronic herd breakdowns. Preventive Veterinary Medicine 131, 1-7.

EFSA (European Food Standards Authority), 2012. Scientific opinion on the use of a gamma interferon test for the diagnosis of bovine tuberculosis. EFSA Journal 10, 2975.

Good, M., Duignan, A., 2011. Perspectives on the history of bovine TB and the role of tuberculin in bovine TB eradication. Veterinary Medicine International 2011, 410470.

Goodchild, A.V., Downs, S.H., Upton, P., Wood, J.L., de la Rua-Domenech, R., 2015. Specificity of the comparative skin test for bovine tuberculosis in Great Britain. Veterinary Record 177, 258.

Hui, S.L., Walter, S.D., 1980. Estimating the error rates of diagnostic tests. Biometrics 36, 167-171.

Humblet, M.F., Boschiroli, M.L., Saegerman, C., 2009. Classification of worldwide bovine tuberculosis risk factors in cattle: A stratified approach. Veterinary Research 40, 50.

Jones, G., Johnson, W.O., Hanson, T.E., Christensen, R., 2010. Identifiability of models for multiple diagnostic testing in the absence of a gold standard. Biometrics 66, 855-863.

Karolemeas, K., de la Rua-Domenech, R., Cooper, R., Goodchild, A.V., Clifton-Hadley, R.S., Conlan, A.J., Mitchell, A.P., Hewinson, R.G., Donnelly, C.A., Wood, J.L., et al., 2012. Estimation of the relative sensitivity of the comparative tuberculin skin test in tuberculous cattle herds subjected to depopulation. Public Library of Science One 7, e43217.

Lahuerta-Marin, A., Gallagher, M., McBride, S., Skuce, R., Menzies, F., McNair, J., McDowell, S. W., Byrne, A.W., 2015. Should they stay, or should they go? Relative future risk of bovine tuberculosis for interferon-gamma test-positive cattle left on farms. Veterinary Research 46, 90.

Lahuerta-Marin, A., McNair, J., Skuce, R., McBride, S., Allen, M., Strain, S.A., Menzies, F.D., McDowell, S.J., Byrne, A.W., 2016. Risk factors for failure to detect bovine tuberculosis in cattle from infected herds across Northern Ireland (2004-2010). Research in Veterinary Science 107, 233-239.

Lewis, F.I., Torgerson, P.R., 2012. A tutorial in estimating the prevalence of disease in humans and animals in the absence of a gold standard diagnostic. Emerging Themes in Epidemiology 9, 9.

O'Hagan, M.J., Courcier, E.A., Drewe, J.A., Gordon, A.W., McNair, J., Abernethy, D.A., 2015. Risk factors for visible lesions or positive laboratory tests in bovine tuberculosis reactor cattle in Northern Ireland. Preventive Veterinary Medicine 120, 283-290.

Olea-Popelka, F.J., Costello, E., White, P., McGrath, G., Collins, J.D., O'Keeffe, J., Kelton, D.F., Berke, O., More, S., Martin, S.W., 2008. Risk factors for disclosure of additional tuberculous cattle in attested-clear herds that had one animal with a confirmed lesion of tuberculosis at slaughter during 2003 in Ireland. Preventive Veterinary Medicine 85, 81-91.

Pollock, J.M., McNair, J., Bassett, H., Cassidy, J. P., Costello, E., Aggerbeck, H., Rosenkrands, I., Andersen, P., 2003. Specific delayed-type hypersensitivity responses to ESAT-6 identify tuberculosis-infected cattle. Journal of Clinical Microbiology 41, 1856-1860.

Pucken, V.B., Knubben-Schweizer, G., Dopfer, D., Groll, A., Hafner-Marx, A., Hormansdorfer, S., Sauter-Louis, C., Straubinger, R.K., Zimmermann, P., Hartnack, S., 2017. Evaluating diagnostic tests for bovine tuberculosis in the southern part of Germany: A latent class analysis. PloS One 12, e0179847.

Toft, N., Akerstedt, J., Tharaldsen, J., Hopp, P., 2007. Evaluation of three serological tests for diagnosis of maedi-visna virus infection using latent class analysis. Veterinary Microbiology 120, 77-86.

Toft, N., Jorgensen, E., Hojsgaard, S., 2005. Diagnosing diagnostic tests: Evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. Preventive Veterinary Medicine 68, 19-33.

Welsh, M.D., Kennedy, H.E., Smyth, A.J., Girvin, R.M., Andersen, P., Pollock, J.M., 2002. Responses of bovine $WC1^{+}$ γδ T cells to protein and nonprotein antigens of *Mycobacterium bovis*. Infection and Immunity 70, 6114-6120.

Wood, P.R., Corner, L.A., Rothel, J.S., Baldock, C., Jones, S.L., Cousins, D.B., McCormick, B.S., Francis, B.R., Creeper, J., Tweddle, N.E., 1991. Field comparison of the interferon-gamma

assay and the intradermal tuberculin test for the diagnosis of bovine tuberculosis. Australian Veterinary Journal 68, 286-290.

Zinsstag, J., Schelling, E., Roth, F., Kazwala, R., 2006. Economics of bovine tuberculosis. In: Thoen, C.O., Steele, J.H., Gilsdorf, M.J. (Eds). *Mycobacterium bovis* Infection in Animals and Humans. Blackwell, Boston, Massachusetts, USA, pp. 68-83.

Figures and table

Table 1
Distribution of results for five bovine tuberculosis ante-mortem tests in Northern Ireland (NI).

	Number (%) of positive tests	Number (%) of negative tests	Total
SICCT standard	1047 (1.5)	70,138 (98.1)	71,185
SICCT severe	1603 (2.3)	69,582(97.7)	71,185
IFN-γ$(B–A)_{NI}$	4606 (6.5)	66,567 (93.5)	71,173
IFN-γ$(B–A)_{0.1}$	3352 (4.7)	67,821 (95.3)	71,173
IFN-γ(ESAT6)	3919 (6.5)	56,675 (93.5)	60,594

SICCT, single intradermal comparative cervical tuberculin test using standard or severe interpretations; IFN-γ$(B–A)_{NI}$, interferon γ (bovine–avian) test using the NI optical density (OD) cut-off difference of 0.05; IFN-γ$(B–A)_{0.1}$, IFN-γ(B–A) using a commercial OD cut-off difference of 0.1; IFN-γ(ESAT6), IFN-γ using the ESAT6 antigen.

Table 2
Distribution of diagnostic test crude sensitivity (Se) and specificity (Sp) stratified by the meat inspection results in Northern Ireland (NI).

		Positive (VL)	Negative (NVL)	Total	Se (95% CI)	Sp (95% CI)
SICCT standard	Positive	531	515	1046	0.59 (0.56-0.63)	0.93 (0.92-0.94)
	Negative	362	6432	6794		
SICCT severe	Positive	626	837	1463	0.70 (0.67-0.73)	0.88 (0.87-0.89)
	Negative	267	6110	6377		
IFN-γ$(B–A)_{NI}$	Positive	789	2653	3442	0.88 (0.86-0.90)	0.62 (0.61-0.63)
	Negative	104	4293	4397		
IFN-γ$(B–A)_{0.1}$	Positive	732	1835	2567	0.82 (0.79-0.84)	0.74 (0.73-0.75)
	Negative	161	5111	5272		
IFN-γ(ESAT6)	Positive	481	724	1205	0.60 (0.56-0.63)	0.88 (0.87-0.89)
	Negative	326	5183	5509		

VL, visible lesion at meat inspection; NVL, no visible lesion at meat inspection; 95% CI, 95% confidence interval; SICCT, single intradermal comparative cervical tuberculin test using standard or severe interpretations; IFN-γ$(B–A)_{NI}$, interferon γ (bovine–avian) test using the NI optical density (OD) cut-off difference of 0.05; IFN-γ$(B–A)_{0.1}$, IFN-γ(B–A) using a commercial OD cut-off difference of 0.1; IFN-γ(ESAT6), IFN-γ using the ESAT6 antigen.

Table 3
Number of animals positive to different diagnostic tests for bovine tuberculosis in Northern Ireland (NI) stratified by production type and location.

		SICCT		IFN-γ(B–A)		IFN-γ(ESAT6)	Meat inspection
Variable	*n*	Standard	Severe	NI cut-off	0.1 cut-off	NI cut-off	VL
Production type							
Dairy	36,095	544	897	2728	1983	1964	765
Beef	35,090	503	706	1878	1369	1955	689
Location							
North	13,548	137	258	862	619	618	180
South-West	26,572	372	607	1799	1292	1518	475
South-East	31,065	538	758	1945	1441	1783	799

VL, visible lesion; SICCT, single intradermal comparative cervical tuberculin test using standard or severe interpretations; IFN-γ$(B–A)_{NI}$, interferon γ (bovine–avian) test using the NI optical density (OD) cut-off difference of 0.05 or the commercial OD cut-off difference of 0.1; IFN-γ(ESAT6), IFN-γ using the ESAT6 antigen.

Table 4
Bayesian latent class estimates for sensitivity and specificity for SICCT, IFN-γ(B–A)$_{NI}$, IFN-γ(B–A)$_{0.1}$, IFN-γ(ESAT6) and meat inspection in Northern Ireland (NI) based on various combinations of tests and cut-off differences fitted with the whole cohort of tested animals (n = 71,185).

			Credibility interval (%)	
Parameter	Test	Mean (%)	2.5	97.5
Sensitivity	SICCT standard	40.5	37.4	43.7
	IFN-γ(B–A)$_{NI}$	93.0	90.2	95.5
	IFN-γ(ESAT6)	65.4	61.7	69.0
Specificity	SICCT standard	99.7	99.6	99.7
	IFN-γ(B–A)$_{NI}$	95.8	95.6	96.0
	IFN-γ(ESAT6)	95.3	95.1	95.5
Sensitivity	SICCT severe	49.0	45.8	52.2
	IFN-γ(B–A)$_{NI}$	92.0	89.2	94.6
	IFN-γ(ESAT6)	57.1	53.9	60.3
Specificity	SICCT severe	99.4	99.3	99.4
	IFN-γ(B–A)$_{NI}$	96.2	96.0	96.4
	IFN-γ(ESAT6)	95.3	95.1	95.5
Sensitivity	SICCT standard	44.3	41.1	47.6
	IFN-γ(B–A)$_{0.1}$	92.1	89.1	94.7
	IFN-γ(ESAT6)	68.9	65.3	72.5
Specificity	SICCT standard	99.6	99.6	99.7
	IFN-γ(B–A)$_{0.1}$	97.3	97.1	97.4
	IFN-γ(ESAT6)	95.2	95.0	95.3
Sensitivity	SICCT standard	57.7	52.7	62.7
	IFN-γ(B–A)$_{NI}$	85.8	81.2	89.8
	Meat inspection	55.5	50.7	60.1
Specificity	SICCT standard	97.5	96.9	98.0
	IFN-γ(B–A)$_{NI}$	83.0	81.8	84.2
	Meat inspection	99.1	98.4	99.8
Sensitivity	SICCT severe	60.6	54.5	65.5
	IFN-γ(B–A)$_{NI}$	87.7	85.4	89.8
	Meat inspection	49.0	42.8	54.2
Specificity	SICCT severe	94.4	93.7	95.1
	IFN-γ(B–A)$_{NI}$	75.6	74.0	77.8
	Meat inspection	99.9	99.8	100
Sensitivity	SICCT Stand.	53.1	48.1	57.8
	IFN-γ(B–A)$_{0.1}$	83.1	80.3	86.0
	Meat inspection	57.1	51.1	62.7
Specificity	SICCT standard	96.3	95.7	96.8
	IFN-γ(B–A)$_{0.1}$	83.1	82.1	85.1
	Meat inspection	99.8	99.4	100

SICCT, single intradermal comparative cervical tuberculin test using standard or severe interpretations; IFN-γ(B–A)$_{NI}$, interferon γ (bovine–avian) test using the NI optical density (OD) cut-off difference of 0.05; IFN-γ(B–A)$_{0.1}$, IFN-γ(B–A) using a commercial OD cut-off difference of 0.1; IFN-γ(ESAT6), IFN-γ using the ESAT6 antigen; 95% Bayesian credibility interval, representing the limits in which the parameter estimate falls with 95% credibility.

Table 5
Bayesian latent class estimates for sensitivity and specificity for SICCT, IFN-γ(B–A)$_{NI}$ and IFN-γ(B–A)$_{0.1}$, IFN-γ(ESAT6) and meat inspection stratified by production type in Northern Ireland (NI).

		Beef			Dairy		
		Mean (%)	Credibility interval (%)		Mean (%)	Credibility interval (%)	
			2.5	97.5		2.5	97.5
Model not including meat inspection data							
Sensitivity	SICCT standard	41.4	37.3	45.6	39.3	34.6	44.0
	IFN-γ(B–A)$_{NI}$	92.6	88.9	95.7	93.9	89.7	97.4
	IFN-γ(ESAT6)	76.1	71.3	80.6	55.7	50.4	61.0
Specificity	SICCT standard	99.8	99.6	99.8	99.6	99.5	99.7
	IFN-γ(B–A)$_{NI}$	96.8	96.6	97.1	94.8	94.4	95.1
	IFN-γ(ESAT6)	95.5	95.2	95.7	95.1	94.8	95.4
Model including meat inspection data							
Sensitivity	SICCT standard	63.8	59.0	68.7	48.8	43.2	54.6
	IFN-γ(B–A)$_{NI}$	89.8	86.0	93.0	87.6	82.8	91.9
	Meat inspection	73.1	68.4	77.6	50.7	45.2	56.1
Specificity	SICCT standard	99.6	99.5	99.7	99.3	99.1	99.4
	IFN-γ(B–A)$_{NI}$	95.2	94.9	95.5	91.5	91.0	92.1
	Meat inspection	98.0	97.8	98.2	97.4	97.1	97.6

SICCT, single intradermal comparative cervical tuberculin test using standard or severe interpretations; IFN-γ(B–A)$_{NI}$, interferon γ (bovine–avian) test using the NI optical density (OD) cut-off difference of 0.05; IFN-γ(B–A)$_{0.1}$, IFN-γ(B–A) using a commercial OD cut-off difference of 0.1; IFN-γ(ESAT6), IFN-γ using the ESAT6 antigen; 95% Bayesian credibility interval, representing the limits in which the parameter estimate falls with 95% credibility.

Table 6
Bayesian latent class estimates for sensitivity and specificity for SICCT, IFN-γ(B–A)$_{NI}$ and IFN-γ(B–A)$_{0.1}$, IFN-γ(ESAT6) and meat inspection stratified by location in Northern Ireland (NI).

		North			South-West			South-East		
		Mean (%)	Credibility interval (%)		Mean (%)	Credibility interval (%)		Mean (%)	Credibility interval (%)	
			2.5	97.5		2.5	97.5		2.5	97.5
Model not including meat inspection data										
Sensitivity	SICCT standard	39.7	31.8	48.0	41.5	36.3	46.7	40.0	35.7	44.5
	IFN-γ(B–A)$_{NI}$	96.7	91.4	99.7	92.8	88.2	96.6	92.7	88.5	96.3
	IFN-γ(ESAT6)	65.4	55.6	74.6	70.3	64.2	76.0	62.1	57.0	67.2
Specificity	SICCT standard.	99.8	99.7	99.9	99.6	99.5	99.7	99.5	99.5	99.6
	IFN-γ(B–A)$_{NI}$	95.5	95.1	96.0	95.2	94.8	95.5	96.4	96.1	96.7
	IFN-γ(ESAT6)	95.9	95.5	96.2	94.9	94.6	95.2	95.3	95.1	95.6
Model including meat inspection data										
Sensitivity	SICCT standard	57.1	47.0	66.9	57.9	51.6	64.2	55.4	50.3	60.6
	IFN-γ(B–A)$_{NI}$	94.1	88.7	98.0	89.8	85.0	93.9	87.2	82.8	91.1
	Meat inspection	55.7	46.1	65.0	63.8	57.6	69.9	62.9	57.8	67.9
Specificity	SICCT standard	99.7	99.5	99.8	99.4	99.2	99.5	99.3	99.2	99.5
	IFN-γ(B–A)$_{NI}$	93.0	92.4	93.6	93.0	92.5	93.4	94.5	94.2	94.9
	Meat inspection	99.2	98.9	99.4	98.2	97.9	98.4	96.9	96.6	97.1

SICCT, single intradermal comparative cervical tuberculin test using standard or severe interpretations; IFN-γ(B–A)$_{NI}$, interferon γ (bovine–avian) test using the Northern Irish (NI) optical density (OD) cut-off difference of 0.05; IFN-γ(B–A)$_{0.1}$, IFN-γ(B–A) using a commercial OD cut-off difference of 0.1; IFN-γ(ESAT6), IFN-γ using the ESAT6 antigen; 95% Bayesian credibility interval, representing the limits in which the parameter estimate falls with 95% credibility.