Assessment of the sensitivity of the gamma-interferon test and the single intradermal comparative cervical test for the diagnosis of bovine tuberculosis under field conditions

A. PRAUD[1]*, M. L. BOSCHIROLI[2], L. MEYER[3], B. GARIN-BASTUJI[2]
AND B. DUFOUR[1]

[1] *Université Paris-Est, Alfort National Veterinary School, Epidemiology and Infectious Diseases Unit, ENVA USC ANSES, Maisons-Alfort, F-94704, France*
[2] *ANSES, Animal Health Laboratory, National Tuberculosis Reference Laboratory, Bacterial Zoonoses Unit, Maisons-Alfort, F-94706, France*
[3] *INSERM, Centre for Research in Epidemiology and Population Health (CESP), U1018, Faculté de Médecine Paris-Sud, Le Kremlin-Bicêtre; AP-HP, Hopital Bicêtre, Epidemiology and Public Health Service, Université Paris-Sud, Le Kremlin-Bicêtre, 94276, France*

SUMMARY

In some French départements, the eradication of bovine tuberculosis is incomplete and usual skin tests [single intradermal tuberculin test (SIT) and single intradermal comparative cervical test (SICCT)] have poor specificity due to cross-reactions with non-pathogenic mycobacteria, causing economic losses. In Côte d'Or (Burgundy, France), an experimental serial testing scheme based on the combination of SICCT and gamma-interferon (IFN-γ) tests has been initiated in order to shorten the interval between suspicion and its invalidation in herds with false-positive results to skin tests. Our aim was to assess the scheme's sensitivity and to compare it to the sensitivity of the screening scheme recommended by the European Commission. Our study included 1768 animals from Côte d'Or. The sensitivities of both schemes were estimated using a Bayesian approach. The individual sensitivity of the IFN-γ test [88·1%, 95% credibility interval (CrI) 72·8–97·5] was not significantly different from individual SICCT sensitivity (80·3%, 95% CrI 61·6–98·0) and individual SIT sensitivity (84·2%, 95% CrI 59·0–98·2). The individual specificity of the IFN-γ test was 62·3% (95% CrI 60·2–64·5). No significant difference could be demonstrated between the sensitivities of the serial testing scheme used in Côte d'Or (73·1%, 95% CrI 41·1–100) and the European Union serial testing scheme (70·1%, 95% CrI 31·5–100·0).

Key words: Bayesian analysis, tuberculosis (TB), screening programme.

INTRODUCTION

Bovine tuberculosis (bTB) is a zoonotic infectious disease mainly due to *Mycobacterium bovis*. In cattle, its clinical expression is discrete and frequently unnoticeable until several years of evolution.

Although France has been considered as officially free from bTB since December 2000, the eradication of the disease is still incomplete in a few French départements like Côte d'Or (Burgundy, France) [1].

In France, the detection of bTB is mainly based on meat inspection at the abattoir and skin test performed on cattle aged >6 weeks, at different

* Author for correspondence: Dr A. Praud, Infectious Diseases Unit/EPIMAI Laboratory, Alfort National Veterinary School, 7 av. du Général de Gaulle, 94704 Maisons-Alfort cedex, France. (Email: apraud@vet-alfort.fr)
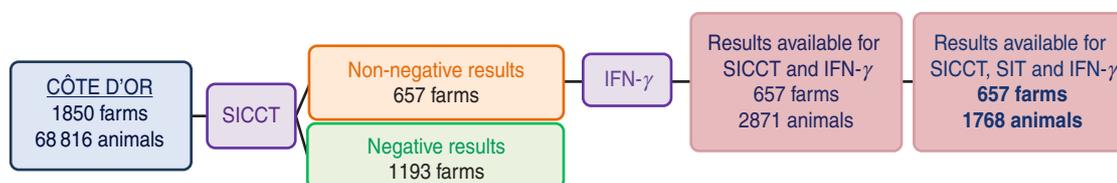
Fig. 1 [*colour online*]. Sampling scheme. SIT, Single intradermal tuberculin test; SICCT, single intradermal cervical comparative test; IFN-γ, gamma-interferon test.

frequencies depending on the region's epidemiological status with regard bTB. Two types of skin tests are used: the single intradermal tuberculin test (SIT) and the single intradermal comparative cervical test (SICCT). Both tests have imperfect sensitivity and specificity in addition to technical drawbacks, even though SICCT is a little more specific than SIT [2]. The biological bases of skin tests are detailed in [3]. European directive CE/64/432 recommends the use of SIT as a first attempt for screening in herds. When positive or doubtful results to SIT occur in a herd, animals with non-negative results may be slaughtered in order to confirm the infection. The only opportunity to invalidate the suspicion is to repeat the skin test (using an SICCT) and to obtain only negative results in the herd. This second skin test must be performed at least 6 weeks after the SIT because of a desensitization phenomenon. During this period, trade of cattle and products from the herd is forbidden.

Another test is available for the diagnosis of bTB: the gamma-interferon (IFN-γ) test. IFN-γ is a cytokine produced by T-lymphocytes of bTB-infected animals after stimulation of *M. bovis* antigens [4]. The IFN-γ test has been used in Australia since the 1980s [5] and in Europe since the 2000s [6]. According to the literature, this test is more sensitive than SICCT [7], and often as specific as it; however, some studies show lower specificities of IFN-γ test compared to SICCT [6]. The IFN-γ test is based on the same biological principle as skin tests [3] but it detects bTB in the early stages of infection, and can be performed immediately after a skin test without desensitization [8].

In some French areas like Côte d'Or, the specificity of skin tests is known to be particularly poor, because of cross-reactions due to non-pathogenic mycobacteria. In this département, SICCT is used as a first attempt instead of SIT, in order to improve the specificity of the screening, but false-positive results occur frequently, hampering trade and causing subsequent economic losses [1].

An experimental screening scheme using serial testing combining SICCT and IFN-γ tests has been initiated in Côte d'Or since 2008, to detect herds with false-positive results to SICCT and to shorten the interval between suspicion and its invalidation in these herds. This use of the IFN-γ test (i.e. serial testing with SICCT) is not included in the European directive CE/64/432, raising issues about the certification of the status of animals exported by France with regard to bTB. The aim of our study was to assess the sensitivity of the screening scheme used in Côte d'Or (serial application of SICCT and IFN-γ tests; SICCT-IFN-γ) and to compare it to the sensitivity of the screening scheme recommended by the European Commission (serial application of SIT and SICCT; SIT-SICCT).

METHODS

Sample and tests

The data presented in this study is part of annual testing data for 1850 farms from Côte d'Or, based on individual skin tests performed on 68816 heads of cattle aged >12 months. The current study includes 2871 animals, sampled from cattle raised on 657 farms in which one or more animals tested positive or doubtful in the annual SICCT testing regimen between 2009 and 2012 (i.e. 35·5% of the 1850 Côte d'Or farms). Animals were included in the study if they had a positive or a doubtful result to SICCT ('reactors'), or if they had close contacts with reactors (2871 animals, i.e. 4·2% of the 68816 tested animals). The sampling scheme is displayed in Figure 1. Thus, 88·8% (2549/2871) of the animals had positive (or doubtful) SICCT results and 11·2% (322/2871) had negative SICCT results. They were included because they presented a higher risk of bTB than the other animals in the herd.

Each animal has been included only once in the sample (in order to satisfy independence). This point

Table 1. *Interpretation of the results of the single intradermal cervical comparative test (SICCT)*

Skin-fold thickening at the point of bovine PPD injection (DB* = B3−B0)	Difference of skin-fold thickening at the points of bovine PPD injection (DB* = B0−B3) and avian PPD injection (DA† = A3−A0)	Result of SICCT
DB >2 mm	DB−DA > 4 mm	Positive
	1 mm ⩽ DB−DA ⩽ 4 mm	Doubtful
	DB−DA < 1 mm	Negative
DB ⩽ 2 mm		Negative

PPD, Purified protein derivative.
* DB, Skin-fold thickening at the point of bovine PPD injection.
† DA, Skin-fold thickening at the point of avian PPD injection.

has been strictly verified using the national identification number given to each animal and the date the test was performed.

Most (591) of these 657 farms kept their bTB-free status after further investigation and 66 were found to be infected (according to Directive 64/432/EEC and Arrêté Ministériel du 15/09/03 modifié).

In Côte d'Or, each animal aged >12 months was annually subjected to SICCT, using standard methods. Skin tests were performed by sanitary veterinarians ('vétérinaires sanitaires', i.e. private veterinarians in charge of official veterinary controls and application of health policy) between 2009 and 2012. Bovine and avian purified protein derivative (bovine and avian PPD, 0·1 ml) were injected intradermally and separately in the mid-cervical region, after the measurement of skin-fold thicknesses at day 0 (thickness at the point of bovine PPD injection: B0; thickness at the point of avian PPD injection: A0). In France, bovine PPD is used at a concentration of 20 000 international units (IU)/ml and avian PPD at a concentration of 25 000 IU/ml. The swelling induced by the delayed-type hypersensitivity reaction was evaluated 72 h after the injection, with the measurement of the skin-fold thicknesses at day 3 (B3 and A3, respectively), and interpreted according to Table 1.

Qualitative skin test results (i.e. 'positive', 'doubtful', 'negative') were available for 2871 animals but detailed quantitative results (i.e. skin-fold thickening measured in millimetres) were registered by veterinarians and therefore available in the database for only 1768 of the 2871 animals. For this last case, SIT results were deduced from SICCT results interpreting only the DB value. When DB was ⩽ 2 mm, the SIT result was negative; when DB was between 2 and 4 mm, the SIT result was doubtful and when DB was >4 mm,

the SIT result was positive. The results for SICCT were thus available for 2871 animals and results for SIT for 1768 animals.

In farms where non-negative results (i.e. positive or doubtful results) to SICCT were observed, an IFN-γ test was performed between 72 h and 5 days after the tuberculin injections, on a batch of cattle including animals showing non-negative results by SICCT, in order to confirm or invalidate the suspicion. The IFN-γ tests were performed at the Veterinary Laboratory of Côte d'Or. The samples were identified by a code number, transported to the laboratory by veterinarians within 8 h and tests were performed blind to the results of the other tests. The data was collected between 2009 and 2012 by veterinarians and the Veterinary Laboratory of Côte d'Or, and gathered by local veterinary officers (Direction Départementale de Protection des Populations de Côte d'Or; DDPP21).

The principle of the IFN-γ test is based on the fact that when sensitized T-lymphocytes from blood of an infected animal are exposed to different antigens (avian and bovine PPD and recombinant antigens), the cells will release the cytokine IFN-γ [3]. The first step of the test is culture of whole blood in the presence or absence of different antigens and the separation of plasma from these cultures after incubation at 37 °C. The second step is measurement of IFN-γ from plasma, using an ELISA. The levels of IFN-γ released from blood cultures stimulated with different antigens are compared [9]. Two kinds of antigens were used: bovine and avian PPD commercialized with the Bovigam® kit (Bovigam®; Prionics AG, Switzerland) and specific antigen ESAT-6 (SSI, Denmark), in order to improve the specificity [10]. The Bovigam test was performed according to the manufacturer's recommendations and as described previously [11]. Optical densities (ODs) were transformed into

Table 2. *Interpretation of the results of the gamma-interferon (IFN-γ) test performed with the Bovigam® kit and ESAT-6 specific antigens*

Bovigam result	ESAT-6 result	IFN-γ test result
Positive	Positive	Positive
Positive	Negative	
Negative	Positive	Divergent
Negative	Negative	Negative

percentage values by comparing the test-sample OD to the positive control OD, both ODs being corrected by subtracting the negative control OD [11]. Results ⩾0·04 were considered positive and <0·04 considered negative under this cut-off. The Bovigam and ESAT-6 results were interpreted jointly using the rules presented in Table 2. When animals had a negative result by IFN-γ test, the 'officially bTB free' status of their herd was maintained. Otherwise, the herd was considered as 'suspected of bTB' and further investigation was carried out (slaughter of animals with non-negative results, search for lesions and PCR or isolation of *M. bovis* on lesions or lymph nodes).

The skin and IFN-γ tests were performed as part of the usual official screening scheme on farms from Côte d'Or. The EU recommendations about screening procedures and EU ethical guidelines and animal welfare regulations were strictly adhered to.

Statistical methods

Since the infectious status of each animal regarding bTB was unknown, each test's sensitivity and specificity were estimated using a latent class analysis implemented through a Bayesian approach using Markov Chain Monte Carlo (MCMC) algorithms. The Bayesian approach has often been used to estimate characteristics of tests in the absence of a gold standard, whether in veterinary or human medicine [12–15]. The main advantage of Bayesian models for latent class analysis is the combination of prior information (previous studies or expert opinion) with evidence from new data (likelihood).

The model used was a generalization of the Bayesian independence model to allow for dependent test outcomes. It has been applied in previous studies [12, 15–18] and used to estimate the covariance of the sensitivities (γSe) and specificities (γSp) of each pair of tests (see online Appendix). Two tests are considered independent as to whether or not an animal is diseased if the probability of positive (or negative) outcome of the first test is the same, whatever the outcome, for the other test [19]. Skin and IFN-γ tests share common steps in their triggering pathways [3]. Consequently, they are expected to be conditionally dependent [16].

The prior distributions of sensitivities, specificities and prevalence were modelled as beta distributions [19]. To construct a beta prior distribution, the most probable value of the parameter (or 'best guess'; θ_0) and a 'lower limit' (θ_L; i.e. a value for which the experimenter is 95% certain that the parameter will be larger) were determined. Many different studies were performed to estimate sensitivity (Se) and specificity (Sp) of bTB screening tests (using different sampling schemes, laboratory and statistical methods). For this reason, diffuse prior distributions combining literature information (meta-analysis of previous studies [6] using the same methods for the performance of SIT [5, 10, 20–22], SICCT [23–25] and IFN-γ tests [5, 21–23, 26, 27]) and experts' advice were introduced in the model. Table 3 provides the parameters of the beta (a,b) distribution for each parameter estimated by the model.

Six sensitivity covariances (one for each pair of tests) and six specificity covariances were required for the model. No information was available on covariances: they were modelled as uniform distributions between a lower limit and an upper limit defined as follows:

$$\max(-(1-Se_1)*(1-Se_2), -Se_1*Se_2 \leqslant \gamma Se$$
$$\leqslant \min(Se_1*(1-Se_2), Se_2*(1-Se_1)),$$
$$\max(-(1-Sp_1)*(1-Sp_2), -Sp_1*Sp_2 \leqslant \gamma Sp$$
$$\leqslant \min(Sp_1*(1-Sp_2), Sp_2*(1-Sp_1)).$$

The statistical analysis was performed with the WinBUGS program [28]. The MCMC algorithm convergence was assessed by checking the stabilization of the plots of iterate parameter value, after a given number of samples and by running multiple chains from dispersed starting values. Early samples (1000 out of 50 000) were discarded as a 'burn-in' period. A sensitivity analysis was also performed by making the prior distributions more diffuse in order to check that the parameter estimates were little affected by these variations [19]. The posterior inferences were expressed as means and 95% credibility intervals (CrIs) [12, 19].

After estimating the sensitivities and specificities of the two tests, they were evaluated when used in

Table 3. *Parameters of the Beta(a,b) distributions used as priors in the latent class model implemented through a Bayesian approach*

Test	Parameter	Best guess (θ_0)	Lower limit (95%) (θ_L)	Parameters of prior beta (a,b) distributions	
				a	b
SIT	Se	84%	63%	9·4	1·79
	Sp	97%	76%	1·59	0·05
SICCT	Se	80%	52%	5·73	1·43
	Sp	99%	80%	0·10	0·001
IFN-γ	Se	88%	75%	21·11	2·88
	Sp	97%	85%	6·87	0·21

SIT, Single intradermal tuberculin test; SICCT, single intradermal cervical comparative test; IFN-γ, gamma-interferon test; Se, sensitivity; Sp, specificity.

Table 4. *Number and proportion of concordant results for each pair of tests in the studied sample*

Tests	Number of concordant results (proportion)
SIT/SICCT	1581/1768 (89·4%)
SIT/IFN-γ	685/1768 (38·7%)
SICCT/IFN-γ	716/1768 (40·5%)

SIT, Single intradermal tuberculin test, SICCT, single intradermal cervical comparative test; IFN-γ, gamma-interferon test.

Table 5. *Cross-results to single intradermal cervical comparative test (SICCT) and gamma-interferon (IFN-γ) test in the studied sample (doubtful and divergent results were considered as positive)*

SICCT result	IFN-γ result	No. of animals (proportion)
Positive	Positive	1046 (36·4%)
Positive	Negative	1503 (52·4%)
Negative	Positive	139 (4·8%)
Negative	Negative	183 (6·4%)
Total		2871 (100%)

series. Under this screening scheme, the overall result was positive when the results were positive (or doubtful /divergent) with both tests. The conditional dependence between results (γSe) estimated through the Bayesian approach was also taken into account [16]. The sensitivity of the serial application of SIT and SICCT ('EU protocol') was defined as:

$$Se_{EU\ protocol} = Se_{SIT} * Se_{ICCT} + \gamma Se_{SIT/SICCT},$$

where Se is sensitivity and γ is covariance. The sensitivity of the serial application of SICCT and IFN-γ tests ('Côte d'Or protocol') was defined as:

$$Se_{Côte\ d'Or\ protocol} = Se_{SICCT} * Se_{IFN-\gamma} + \gamma Se_{SICCT/IFN-\gamma}$$

The specificity of the serial application of the two tests was also defined as:

$$SpSerial = 1 - ((1 - Sp_{test1}) * (1 - Sp_{test2})) - \gamma Sp.$$

RESULTS

Qualitative results for SICCT and IFN-γ tests were available for 2871 animals and results for SIT were available for 1768/2871 animals in the database. Results for SIT and IFN-γ tests were concordant for 38·7% of the animals (685/1768) whereas results for SICCT and IFN-γ tests were concordant for only 40·5% of the animals (716/1768) (Table 4). For the 2871 animals for which SICCT and IFN-γ test results were available, 1881 (65·5%) obtained concordant results with both IFN-γ tests (Bovigam kit and recombinant antigen ESAT-6) and 990 (34·5%) obtained a divergent result to the IFN-γ test. Most (793) of these 990 animals had a positive result with Bovigam and a negative result with ESAT-6 and 197 animals had a negative result with both Bovigam and ESAT-6. As explained above, doubtful results to SICCT and divergent results to IFN-γ test were considered as positive. Tables 5 and 6 give the result

Table 6. *Cross-results to single intradermal tuberculin test (SIT), single intradermal cervical comparative test (SICCT) and gamma-interferon (IFN-γ) test in the studied sample (doubtful results were considered as positive)*

SIT result	SICCT result	IFN-γ result	No. of animals (proportion)
Positive	Positive	Positive	546 (30·9%)
Positive	Positive	Negative	953 (53·9%)
Positive	Negative	Positive	78 (4·4%)
Positive	Negative	Negative	109 (6·2%)
Negative	Negative	Positive	21 (1·2%)
Negative	Negative	Negative	61 (3·4%)
Total			1768 (100%)

Table 7. *Sensitivity and specificity of each test and their covariance estimated through a Bayesian model for latent class analysis*

Test	Parameter	Mean	95% credibility interval
SIT	Se	0·842	0·590 to 0·982
	Sp	0·104	0·089 to 0·119
SICCT	Se	0·803	0·616 to 0·980
	Sp	0·083	0·070 to 0·095
IFN-γ	Se	0·881	0·728 to 0·975
	Sp	0·623	0·602 to 0·645
SIT/SICCT	γSe	0·025	−0·048 to 0·122
	γSp	0·037	0·030 to 0·046
SICCT/IFN-γ	γSe	0·023	−0·038 to 0·105
	γSp	−0·001	−0·008 to 0·005
SIT/IFN-γ	γSe	0·026	−0·025 to 0·105
	γSp	0·006	0·0003 to 0·012

SIT, Single intradermal tuberculin test, SICCT, single intradermal cervical comparative test; IFN-γ, gamma-interferon test; Se, sensitivity; Sp, specificity; γSe, sensitivity covariance; γSp, specificity covariance.

profiles of the studied animals to the skin and IFN-γ tests.

The sensitivities and specificities of the tests and their covariances were then estimated using a latent class analysis implemented through a Bayesian approach (Table 7). The average individual sensitivity of the IFN-γ test ($Se_{IFN-γ} = 0.881$, 95% CrI 0·728–0·975) was slightly higher than the average individual sensitivity of SIT ($Se_{SIT} = 0.842$, 95% CrI 0·590–0·982) and SICCT ($Se_{SICCT} = 0.803$, 95% CrI 0·616–0·980) but there was no significant difference between the sensitivities of the tests ($P > 0.05$). The sensitivity covariance of SIT and SICCT was $γSe_{SIT/SICCT} = 0.025$ (95% CrI −0·048 to 0·122). The sensitivity covariance of SICCT and IFN-γ tests was $γSe_{SICCT/IFN-γ} = 0.023$ (95% CrI −0·038 to 0·105). Most animals included in the protocol had positive or doubtful results by skin test: for this reason, specificities estimated for SIT and SICCT were considerably low and were thus not biologically interpretable. The individual specificity of the IFN-γ test was 0·623 (95% CrI 0·602–0·645). The individual prevalence of bTB in the studied sample was 2·0% (95% CrI 1·2–3·2).

In Côte d'Or, SICCT and IFN-γ tests were combined in series to improve the specificity of the screening, whereas the European Commission recommends the use of a serial testing scheme with SIT and SICCT, raising issues about a potential defect of sensitivity of the French protocol. The estimation of the sensitivity of each sequence was based on the individual test sensitivities described above. The diagnostic sensitivity of the Côte d'Or protocol (i.e. serial application of SICCT and IFN-γ tests) was 73·1% (95% CrI 41·1–100·0) and the diagnostic sensitivity of the EU protocol (i.e. serial application of SIT and SICCT) was 70·1% (95% CrI 31·5–100·0). No significant difference between the sensitivities of the protocols could be shown ($P = 0.638$).

DISCUSSION

The aim of our study was to assess and to compare the sensitivity of the screening scheme used in Côte d'Or (serial testing using SICCT and IFN-γ tests) and the screening scheme recommended by European Commission (serial testing scheme using SIT and SICCT).

The estimation of the accuracy of a test should ideally be derived from testing a statistically relevant panel of animals. The history of these animals and their infection status should be known and the panel should be representative of the region where the test is to be used [29]. The selection of an appropriate gold-standard-infected population (i.e. culture-positive animals) is extremely cumbersome, requires a significant budget and could lead to an overestimation of the diagnostic sensitivity of the tests under study, since culture-positive animals are probably suffering from an advanced stage of infection. The technical and financial constraints highlighted above justify the statistical approach (i.e. latent class modelling) used to assess the sensitivities of tests and

protocols. Latent class analysis implemented through a Bayesian approach is frequently used in evaluation of diagnostic tests, in veterinary or human medicine [14]. It offers some advantages over frequentist methods, e.g. the possibility of taking into account prior information about test accuracy or prevalence in the form of expert opinion or literature data. These priors should be generated independently of the study but collected in a similar context [19]. Another advantage of the Bayesian method is that it provides true probability intervals (credibility interval contain the true parameter with 95% certainty), whereas a 95% frequentist confidence interval is considered to contain the true parameter value 95% of the time [19]. The width of the credibility intervals also depends on the adaptation of the prior distributions to the data. The prior distributions used in this study might not be ideal, but reflects the central tendency and variation of the accuracy of the respective tests in previous studies regarding detection of *M. bovis* in cattle.

The tests were considered as conditionally dependent. Taking into account this dependence was essential; otherwise the estimates of characteristics might have been biased due to an underestimation of the classification errors [30]. Posterior estimates of covariances confirmed the hypothesis of conditional dependence between tests, with low values of covariance.

The study was based on a representative sample of the bovine population submitted to IFN-γ testing in Côte d'Or, i.e. animals from herds where non-negative results to SICCT were observed.

This sampling scheme engenders some biases. For example, no information was available on animals with false-negative results to SICCT: these animals could have been detected by the IFN-γ test since this test is known to detect bTB infection earlier than skin tests, leading to an underestimation of the sensitivity of the IFN-γ test [10, 31, 32]. Furthermore, specificity estimates for SIT and SICCT are artificially low and not biologically interpretable.

The individual sensitivity of the IFN-γ test (88·1%, 95% CrI 72·8–97·5) was not significantly different from individual SICCT sensitivity (80·3%, 95% CrI 61·6–98·0) and individual SIT sensitivity (84·2%, 95% CrI 59·0–98·2). The sensitivity of SIT has been deduced from bovine PPD results of SICCT: this might engender some biases (especially an overestimation of the sensitivity of SIT since in SICCT the skinfold is systematically measured with a cutimeter whereas the interpretation of SIT results is usually subjective). This overestimation does not seem to

raise a major issue since this bias leads to an overestimation of the sensitivity of the EU protocol and does not impact the estimation of the sensitivity of the SICCT-IFN-γ test protocol.

In previous studies, the IFN-γ test was described as at least as sensitive as skin tests or more sensitive than these tests since the IFN-γ test allows detection of the infection earlier than skin tests [10, 31, 32]. According to Dean *et al.* [33], the period between the infection and the possibility of detection does not depend on the infecting dose. Two studies using the IFN-γ test with a Bayesian approach have recently been conducted in Ireland [34] and Spain [35]. The sensitivity of the IFN-γ test was 89·3% (95% CrI 77·5–97·2) in Spain and 64·1% (95% CrI 60·8–67·8) in Ireland. The results found in Spain are similar to ours.

As explained above, most animals included in the protocol had positive or doubtful results by skin test: for this reason, specificities estimated for SIT and SICCT were considerably low and were thus not interpretable. The individual specificity of the IFN-γ test was 0·623 (95% CrI 0·602–0·645). The discordance between test results could be due to the lack of specificity of skin tests in Côte d'Or.

A high rate of animals tested positive by SICCT and IFN-γ tests (36·4%) but the infection could be confirmed (according to Directive 64/432/EEC and Arrêté Ministériel du 15/09/03) in only 66 (10·0%) out of 657 farms. This can probably be related to the poor predictive value of positive results to screening tests since SICCT and IFN-γ tests lack specificity (particularly in areas where cross-reactions to mycobacteria are frequent, such as Côte d'Or) and the prevalence of bTB is low. Furthermore, it should be borne in mind that in field, tests results are interpreted at the herd level. Herd sensitivity is higher than individual sensitivity (it increases with the number of bTB-infected animals) and herd specificity is considerably lower than individual specificity (it decreases very quickly with the number of bTB-free animals). At the herd level, the sensitivities of the schemes are thus higher than at the individual level.

The assessment of the characteristics of the IFN-γ test varies greatly according to the methods used in performing the test (antigens, cut-offs, interpretation criteria, etc.) and the epidemiological context. For instance, using specific antigens like ESAT-6 or CFP-10 improves the specificity of the IFN-γ test [3]. Furthermore, in most European countries, the IFN-γ test is used in parallel with skin tests in herds infected with bTB in order to improve disease eradication,

whereas in Côte d'Or it is used in a serial testing scheme with skin tests to speed up the release of herds with false-positive results by SICCT.

Moreover, the IFN-γ test is easier to perform than skin tests, since animals need to be restrained only once, to draw blood samples. This is particularly important in Côte d'Or, where herds are large and mainly comprised of beef cows that spend most of their time grazing and are not used to be handled. Finally, the IFN-γ test can be standardized: its result does not depend on the technician.

The main drawbacks of this test are its cost (40–60 euros per test per animal) and technical constraints: the delay between blood sampling and analysis at the laboratory should not exceed 8 h [36]. In Côte d'Or, this condition was satisfied thanks to a specific system of sample collection and the immediate analysis of the samples by the veterinary laboratory of Côte d'Or.

The imperfect diagnostic performance of both tests could justify their combined use [37]. With a parallel testing scheme, the overall result is positive when at least one test gives a positive result, whereas with a serial testing scheme, the overall result is positive when the results are positive by both tests.

The parallel combination increases the sensitivity of the screening. Accordingly, it could be recommended in the context of suspected infection with the aim of accelerating the eradication of disease. In this context, many sera may have to be tested in a flock and the rapidity of the process influences the success of eradication operations. In France, skin and IFN-γ tests are used in parallel in two areas (Côte d'Or and Dordogne) in infected herds and in herds epidemiologically linked to an outbreak of bTB. This kind of testing scheme is recommended by European Directive CE/64/432.

A serial testing scheme increases the diagnostic specificity of screening and could therefore be recommended in areas where prevalence of bTB is low but cross-reactions are frequent, in order to avoid false-positive results. An experimental use of serial SICCT and IFN-γ tests has been initiated in France in Côte d'Or but is not allowed under European Directive CE/64/432. The aim of our study was thus to evaluate the sensitivity of this protocol. Subsequently, we assessed the sensitivities of serial testing schemes including, respectively, SICCT and IFN-γ tests (Côte d'Or protocol) and SIT and SICCT (EU protocol). The individual sensitivity of the Côte d'Or protocol was 73·1% (95% CrI 41·1–100·0) and the individual sensitivity of the EU protocol was 70·1% (95% CrI 31·5–100·0). Our study could not demonstrate the existence of a difference between the sensitivities of these protocols ($P = 0.638$). The risk of importing infected cattle considered bTB free from Burgundy did not seem higher than from areas following the CE/64/432 Directive recommendations. Nevertheless, it is essential to take into account the width of the credibility intervals: this could be due to lack of informative data. Our study suggests it will be of interest to use the IFN-γ test in combination with SIT or SICCT in serial testing in order to improve the specificity of the screening scheme; however, considering the width of the credibility intervals and the biases due to the field protocol, a new study on larger samples of animals is currently being conducted in Côte d'Or in order to verify and refine these results.

To conclude, our study is an original survey aiming at estimating the sensitivity of the IFN-γ test in serial combination with SICCT for the diagnosis of bTB, in a French area with low bTB prevalence but numerous false-positive results to skin tests. The estimation of the sensitivity of the IFN-γ test showed this test to be as sensitive as SICCT and SIT. Our results are consistent with previous studies, even if the epidemiological context of bTB in Côte d'Or is very specific. When studying IFN-γ and skin tests in series, no significant difference of sensitivity could be shown between the Côte d'Or protocol (serial testing scheme with SICCT and IFN-γ tests) and the EU protocol (serial testing scheme with SIT and SICCT). Nevertheless, the credibility intervals of the estimated parameters were very wide and the protocol engenders some biases; therefore caution should be exercised when interpreting the results.

SUPPLEMENTARY MATERIAL

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0950268814000338.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Fediaevsky A, *et al.*** Bovine tuberculosis in France in 2011: the number of outbreaks keeps decreasing [in French]. *Bulletin Epidémiologique Santé Animale – Alimentation* 2012; **54**: 4–12.

2. **Alvarez J, *et al.*** Effect of paratuberculosis on the diagnosis of bovine tuberculosis in a cattle herd with a mixed infection using interferon-gamma detection assay. *Veterinary Microbiology* 2009; **135**: 389–393.

3. **Pai M, Riley LW, Colford JM.** Interferon-gamma assays in the immunodiagnosis of tuberculosis: a systematic review. *Lancet Infectious Diseases* 2004; **4**: 761–776.

4. **Pollock JM, Neill SD.** *Mycobacterium bovis* infection and tuberculosis in cattle. *Veterinary Journal* 2002; **163**: 115–127.

5. **Wood PR, *et al.*** Field comparison of the interferon-gamma assay and intradermal tuberculin test for the diagnosis of bovine tuberculosis. *Australian Veterinary Journal* 1991; **68**: 286–290.

6. **De la Rua-Domenech R, *et al.*** 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* 2006; **83**: 190–210.

7. **Whipple DL, *et al.*** Comparison of the sensitivity of the caudal fold skin test and a commercial gamma-interferon assay for diagnosis of bovine tuberculosis. *American Journal of Veterinary Research* 1995; **56**: 415–419.

8. **Ryan TJ, Buddle BM, De Lisle GW.** An evaluation of the gamma interferon test for detecting bovine tuberculosis in cattle 8 to 28 days after tuberculin skin testing. *Research in Veterinary Science* 2000; **69**: 57–61.

9. **Buddle BM, Livingstone PG, de Lisle GW.** Advances in ante-mortem diagnosis of tuberculosis in cattle. *New Zealand Veterinary Journal* 2009; **57**: 173–180.

10. **Pollock JM, *et al.*** Specific delayed-type hypersensitivity responses to ESAT-6 identify tuberculosis-infected cattle. *Journal of Clinical Microbiology* 2003; **41**: 1856–1860.

11. **Faye S, *et al.*** Determination of decisional cut-off values for the optimal diagnosis of bovine tuberculosis with a modified IFNγ assay (Bovigam®) in a low prevalence area in France. *Veterinary Microbiology* 2011, **151**: 60–67.

12. **Branscum AJ, Gardner IA, Johnson WO.** Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. *Preventive Veterinary Medicine* 2005; **68**: 145–163.

13. **Rutjes AWS, *et al.*** Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technology Assessment* 2007; **11**: 1–67.

14. **Meyer N, Vinzio S, Goichot B.** Bayesian statistics: an approach adapted to clinical studies [in French]. *La Revue de Médecine Interne* 2009; **30**: 242–249.

15. **Praud A, *et al.*** Assessment of diagnostic sensitivity and specificity of an indirect ELISA kit for the diagnosis of *Brucella ovis* infection in rams. *BMC Veterinary Research* 2013; **8**: 68.

16. **Gardner IA, *et al.*** Conditional dependence between tests affects the diagnosis and surveillance of animal disease. *Preventive Veterinary Medicine* 2000; **45**: 107–122.

17. **Dendukuri N, Joseph L.** Bayesian approaches to modelling the conditional dependence between multiple diagnostic tests. *Biometrics* 2001; **57**: 158–167.

18. **Georgiadis MP, *et al.*** Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Journal of Applied Statistics* 2003; **52**: 63–78.

19. **Enøe C, Georgiadis MP, Johnson WO.** Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease is unknown. *Preventive Veterinary Medicine* 2000; **45**: 61–81.

20. **Domingo ML, *et al.*** Comparison of the efficacy of skin test and gamma-interferon test for the diagnosis of bovine tuberculosis in the field. *Medicina Veterinaria* 1995; **12**: 307–317.

21. **Gonzalez Llamazares OR, *et al.*** Field evaluation of the single intradermal cervical tuberculin test and the interferon-gamma assay for detection and eradication of bovine tuberculosis in Spain. *Veterinary Microbiology* 1999; **70**: 55–6.

22. **Cagiola M, *et al.*** Analysis of possible factors affecting the specificity of the gamma interferon test in tuberculosis-free cattle herds. *Clinical and Diagnostic Laboratory Immunology* 2004; **11**: 652–656.

23. **Neill SD, *et al.*** The diagnostic of bovine tuberculosis by blood testing. In: *Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine*, 1994. Queen's University, Belfast, pp. 1–8.

24. **Doherty ML, *et al.*** Effect of a recent injection of purified protein derivative on diagnostic tests for tuberculosis in cattle infected with *Mycobacterium bovis*. *Research in Veterinary Science* 1995; **58**: 217–221.

25. **Costello E, *et al.*** Performance of the single intradermal comparative tuberculin test in identifying cattle with tuberculosis lesions in Irish herds. *Veterinary Record* 1997; **141**: 22–224.

26. **Buddle BM, *et al.*** Use of ESAT-6 in the interferon-gamma test for diagnosis of bovine tuberculosis following skin testing. *Veterinary Microbiology* 2001; **80**: 37–46.

27. **Vordermeier HM, *et al.*** Use of synthetic peptides derived from the antigens ESAT-6 and CFP-10 for differential diagnosis of bovine tuberculosis in cattle. *Clinical and Diagnostic Laboratory Immunology* 2001; **8**: 571–578.

28. **Lunn DJ, *et al.*** WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 2000; **10**: 325–337.

29. **OIE (World Organisation for Animal Health).** Principles and methods of validation of diagnostic assays for infectious diseases. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. Paris: OIE, 2011.

30. **Vacek PM.** The effect of conditional dependence on the evaluation of diagnostic tests. *Biometrics* 1985; **41**: 959–968.

31. **Neill SD, Bryson DB, Pollock JM.** Pathogenesis of tuberculosis in cattle. *Tuberculosis* 2001; **81**: 9–86.

32. **Wood PR, Jones SL.** Bovigam: an in vitro cellular diagnostic test for bovine tuberculosis. *Tuberculosis* 2001, **81**: 147–155.

33. **Dean GS, et al.** Minimum infective dose of *Mycobacterium bovis* in cattle. *Infection and Immunity* 2005; **73**: 6467–6471.

34. **Clegg T, et al.** Using latent class analysis to estimate the test characteristics of the gamma-interferon test, the single intradermal comparative tuberculin test and a multiplex immunoassay under Irish conditions. *Veterinary Microbiology* 2011; **151**: 68–76.

35. **Alvarez J, et al.** Evaluation of the sensitivity and specificity of bovine tuberculosis diagnostic tests in naturally infected cattle herds using a Bayesian approach. *Veterinary Microbiology* 2012; **155**: 38–43.

36. **Gormley E, et al.** The effect of the tuberculin test and the consequences of a delay in blood culture on the sensitivity of a gamma-interferon assay for the detection of *Mycobacterium bovis* infection in cattle. *Veterinary Immunology and Immunopathology* 2004; **102**: 413–420.

37. **Weinstein S, Obuchowski NA, Lieber ML.** Clinical evaluation of diagnostic tests. *American Journal of Roentgenology* 2005; **184**: 14.