The effect of alternative testing strategies and bio-exclusion practices on Johne's disease risk in test-negative herds

S. J. More,^{*1} E. S. G. Sergeant,[†] S. Strain,[‡] W. Cashman,[§] Kevin Kenny,[#] and D. Graham^{||}

^{*}Centre for Veterinary Epidemiology and Risk Analysis, UCD School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

[†]AusVet Animal Health Services Pty. Ltd., PO Box 2321, Orange NSW 2800, Australia

[‡]Veterinary Sciences Division, Agri-Food and Biosciences Institute, Belfast BT4 3SD, United Kingdom

[§]Riverstown Cross, Glanmire, Co. Cork, Ireland

[#]Central Veterinary Research Laboratory, Backweston, Cellbridge, Co. Kildare, Ireland

^{||}Animal Health Ireland, Main Street, Carrick-on-Shannon, Co. Leitrim, Ireland

ABSTRACT

Herd classification is a key component of national Johne's disease (JD) control programs. Herds are categorized on the basis of test results, and separate sub-programs are followed for test-positive and test-negative herds. However, a test-negative herd result does not necessarily equate to JD freedom for reasons relating to disease pathogenesis and available diagnostic tests. Thus, in several countries, JD control programs define test-negative herds as having a "low risk" of infection below a specified prevalence. However, the approach is qualitative, and little quantitative work is available on herd-level estimates of probability of freedom in test-negative herds. This paper examines the effect over time of alternative testing strategies and bio-exclusion practices on JD risk in test-negative herds. A simulation model was developed in the programming language R. Key model inputs included sensitivity and specificity estimates for 3 individual animal diagnostic tests (serum ELISA, milk ELISA, and fecal culture), design prevalence, testing options, and testing costs. Key model outputs included the probability that infection will be detected if present at the design prevalence or greater (herd sensitivity; *SeH*), the probability that infection in the herd is either absent or at very low prevalence (i.e., less than the design prevalence; *ProbF*), the probability of an uninfected herd producing a false-positive result [$P(\text{False+})$], and mean testing cost (*HerdCost*) for different testing strategies. The output *ProbF* can be updated periodically, incorporating data from additional herd testing and information on cattle purchases, and could form the basis for an output-based approach to herd classification. A high *ProbF* is very difficult to achieve, reflecting the low sensitivity of the evaluated tests. Moreover, *ProbF* is greatly affected by any risk

of introduction of infection, decreasing in herds with poor bio-exclusion practices despite ongoing negative test results. The value of $P(\text{False+})$ was substantial when tests with imperfect specificity were used. Testing strategies can substantially influence testing costs but with little effect on test performance. This study illustrates an output-based approach to herd classification, with potential for national and field applications.

Key words: Johne's disease, testing strategy, bio-exclusion, probability of freedom

INTRODUCTION

Herd classification is a key component of national Johne's disease (JD) control programs, enabling commerce to continue while protecting herds with low infection risk. Herds are categorized on the basis of test results, and separate sub-programs are followed for test-positive and test-negative herds (Tavornpanich et al., 2012). In herds with test-negative results, herd owners seek to minimize the risk of introduction of *Mycobacterium avium* ssp. *paratuberculosis* (Map, the causative agent of JD) through attention to bio-exclusion and to build confidence of JD freedom. However, JD freedom among test-negative herds is very difficult to prove because of the prolonged incubation period and the poor individual animal sensitivity (*Se*) of available diagnostic tests, especially during the early stages of the disease. Animals most commonly acquire Map infection during the early stages of life (Whittington and Windsor, 2009; Windsor and Whittington, 2010; Lombard, 2011). Time between onset of infection to detectable fecal shedding is variable but usually >2 yr, and animals can be infectious, and thus shedding Map, for a variable period before progressing to clinical disease (Sweeney, 2011). Therefore, animals can be infected, and infectious, for extended periods before developing clinical signs (Nielsen and Toft, 2008).

Several countries have developed programs for test-negative herds, cognizant of the above-mentioned con-

Received July 6, 2012.

Accepted November 2, 2012.

¹Corresponding author: simon.more@ucd.ie

cerns, to enable cattle herd owners to objectively and transparently demonstrate a specified low risk of Map infection, including the Market Assurance Program in Australia (**CattleMAP**; Animal Health Australia, 2008) and the Voluntary Herd Status Program in the United States (United States Department of Agriculture, 2010). In CattleMAP, movement to higher herd assurance (classification) levels [monitored negative (MN)1 to MN2 to MN3] relies on ongoing test-negative results at, generally biennial, intervals while complying with defined testing regimens and rigorous herd biosecurity requirements. Detailed program rules are in place (Animal Health Australia, 2008).

As yet, limited quantitative work has been done on herd-level estimates of probability of freedom in test-negative herds. Sergeant et al. (2008) investigated the probability of low JD prevalence in Danish dairy herds, but only at a single point in time. The rules in the above-mentioned Australian and US programs are essentially based on qualitative assessments, and there is currently no understanding of the relative difference in probability of freedom between herds at different assurance (classification) levels, such as MN1 versus MN2. A quantitative assessment would be useful, both to provide numeric estimates of probability of freedom and to identify the factors that are most influential in increasing, or decreasing, these estimates over time. Rapid progress has been made in quantitative methods to substantiate freedom from infection, most recently using scenario trees (Martin et al., 2007a). To date, these methods have primarily been used at a regional or national level to substantiate freedom from a range of infectious diseases (Martin et al., 2007b; More et al., 2009; Schuppers et al., 2010). This paper seeks to extend this methodology to examine the effect of alternative testing strategies and bio-exclusion practices in test-negative herds on probability of JD freedom over time. The work has been developed in an Irish context, where a national JD control program is currently being designed (More et al., 2011).

MATERIALS AND METHODS

The Model

A stochastic simulation model was developed in the R Programming environment (R v 2.12.1; R Project for Statistical Computing, Vienna, Austria), with a web-based interface for ease of user input. The model simulates testing of cattle herds for the presence of Map infection and allows the evaluation and comparison of a variety of testing strategies and alternative tests. Key model inputs are entered as probability distributions to reflect uncertainty about their true values, and the

model was run for 1,000 iterations for all simulations, so that outputs are also presented as probability distributions. The model web interface is available at http://epitools.ausvet.com.au/content.php?page=AH1_JD.

The model has 2 primary outputs. The first is an estimate of herd sensitivity (**SeH**) of a given testing strategy; that is, the probability of detecting infection in a herd if it is present at a true prevalence exceeding a specified threshold value or design prevalence. The second is the level of confidence that the true prevalence is less than the specified design prevalence value, assuming negative testing results (**ProbF**), equivalent to the negative predictive value of the herd test; **ProbF** can be estimated over multiple time periods assuming periodic (e.g., annual) testing and allowing for the probability of introduction during the intervening period, based on the numbers of animals introduced and the likelihood of infection in introduced animals. If a positive result on a definitive test such as fecal culture is recorded at any time, the herd is considered infected and calculation of **SeH** and **ProbF** is unnecessary.

Modeling was conducted using a range of small to moderate herd sizes, based on Irish data (Supplemental Table S1; <http://www.journalofdairyscience.org/>). As a consequence, **SeH** was estimated using an approximation of the hypergeometric probability distribution, as follows:

$$SeH = 1 - (1 - Se \times n/N)^d,$$

assuming individual animal specificity (**Sp**) = 100% (after follow-up testing with a highly specific, confirmatory test such as fecal culture), and where **Se** = individual animal sensitivity, n = sample size, N = population size (MacDiarmid, 1988), and d (the assumed number of infected animals in the herd) = $P^* \times N$ rounded up to next integer, where P^* is the specified design prevalence. For testing strategies in which multiple tests were used and interpreted in series (e.g., serum ELISA followed by fecal culture), $Se = Se_1 \times Se_2 \dots \times Se_n$. For testing options in which $Sp < 100\%$ (serum or milk ELISA without follow-up), **SeH** was estimated using the modified hypergeometric probability distribution, as described by Cameron and Baldock (1998).

Confidence that the true prevalence is less than P^* was estimated as follows:

$$ProbF = NPV = [1 - (1 - PriorF)] / [1 - (1 - PriorF) \times SeH],$$

where NPV = negative predictive value and **PriorF** is the level of confidence that the prevalence is less than P^* before the testing was undertaken (Martin et al.,

2007a). Note that NPV is generally applied at the level of the animal, whereas $PriorF$, as applied here, is a herd-level measure. $PriorF$ was estimated as follows:

$$PriorF_t = 1 - [(1 - ProbF_{t-1}) + PIntro_t - (1 - ProbF_{t-1}) \times PIntro_t],$$

where t is the time period (yr) and $PIntro$ is the probability of introduction during the previous time period (Martin et al., 2007a). For a single herd test or the first in a series of annual tests, $ProbF_{t-1}$ is estimated by the complement of the estimated herd prevalence for the herd type (beef or dairy), whereas for subsequent time periods, $ProbF_{t-1}$ is the estimate from the previous period.

For any time period, the probability of introduction is estimated using a standard binomial probability calculation (Dohoo et al., 2009):

$$PIntro = 1 - (1 - P_{beef})^{n_{beef}} \times (1 - P_{dairy})^{n_{dairy}},$$

where P_{beef} and P_{dairy} are the probabilities of infection in the source animals for beef and dairy, respectively, and n_{beef} and n_{dairy} are the numbers of beef and dairy animals, respectively, introduced during the time period.

The probability of an uninfected herd producing a false-positive result (one or more positive reactors), $P(False+)$, was calculated as follows:

$$P(False+) = 1 - SpH = 1 - Sp^n,$$

where SpH is herd specificity and Sp is individual animal specificity (Martin et al., 1992).

The final output of the model is the mean testing cost for a given testing scenario in an uninfected herd, **HerdCost**, based on sample collection and testing costs, number of animals tested, estimated test specificity, and the estimated cost of follow-up testing on positive reactors:

$$\begin{aligned} HerdCost = & SamplingCost + n \times TestCost_1 \\ & + P(Test+) \times (SamplingCost + n \\ & \times ProportionPositive \times TestCost_2), \end{aligned}$$

where $SamplingCost$ is the cost of a farm visit to collect samples, $TestCost_1$ and $TestCost_2$ are the individual sample costs (laboratory cost, sample collection cost, and consumables) of the screening and confirmatory tests, respectively, $P(Test+)$ is the probability that one or more animals will give a positive test result [= $P(False+)$ for an uninfected herd], and $ProportionPositive$ is the expected mean number of positive results.

Model Inputs

General. Input values for the model included herd and sample size for the tested herd, specified design prevalence, individual animal test sensitivity in the sampled age cohort, testing costs, the numbers of beef or dairy animals introduced during the previous period and the corresponding risk of infection in those animals, and the prior confidence of freedom at the start of the period (Table 1). Individual animal test sensitivities were input as Pert probability distributions because the Pert distribution is an appropriate distribution for expressing uncertainty about parameters defined by a most likely value and a range, which was the format of values elicited from the literature review. Beta distributions were used for probabilities of infection in introduced animals because these values were reported as proportions with 95% confidence limits, which can be easily expressed as beta distributions. All other inputs were fixed values. Model input values were determined, through consensus among all study authors, after considering relevant published data and any additional considerations relevant to Ireland.

Herd Size. We calculated the herd size distribution for Irish cattle herds in 2009, by enterprise type (Supplemental Table S1; <http://www.journalofdairyscience.org/>). Using data from the national Animal Health Computer System (AHCS, Department of Agriculture, Food and the Marine, Dublin, Ireland), we calculated the average number of cattle greater than 6 wk of age at each full herd-level tuberculosis summary test during 2009, after excluding dealers, factory agents, and feedlot herds. Partial herd test records were deleted and tests with zero animals were ignored (no stock on day of test). Enterprise type was categorized from AHCS herd test summary records for tuberculosis. Data processing was carried out using SAS software, version 9.1.3 (SAS Institute Inc., Cary, NC).

Infection Probabilities. The probabilities of infection in source animals for beef and dairy, P_{beef} and P_{dairy} , respectively, were based on results from a JD survey in Ireland that was conducted in 2005 (Good et al., 2009).

Testing Costs and Operating Test Characteristics. Current indicative costs associated with JD testing in Ireland are given in Supplemental Table S3 (<http://www.journalofdairyscience.org/>).

Nielsen and Toft (2008) present the most recently published review of the operating characteristics of diagnostic tests for JD. We critically reviewed these summary estimates, supplemented by subsequent peer-reviewed publications reporting estimates of absolute JD test Se and Sp (Supplemental Table S2; <http://www.journalofdairyscience.org/>). We only considered individual animal JD diagnostic tests currently avail-

Table 1. A summary of model inputs

Input parameter ¹	Distribution	Minimum	Most likely	Maximum
Serum ELISA sensitivity	Pert	0.07	0.15	0.22
Milk ELISA sensitivity	Pert	0.06	0.13	0.19
Fecal culture sensitivity (screening)	Pert	0.16	0.23	0.3
Fecal culture sensitivity (follow-up)	Pert	0.60	0.65	0.70
Value				
Serum ELISA specificity	Fixed value		0.985	
Milk ELISA specificity	Fixed value		0.985	
Fecal culture specificity	Fixed value		1	
Prior <i>ProbF</i> (Dairy)	Fixed value		0.794	
Prior <i>ProbF</i> (Beef)	Fixed value		0.925	
Design prevalence	Fixed value		5% or 1 animal	
Mode 95th percentile				
<i>PIntro</i> (Dairy)	Beta		0.027	0.031
<i>PIntro</i> (Beef)	Beta		0.031	0.035

¹*ProbF* = level of confidence that the true prevalence is less than the specified design prevalence value, assuming negative testing results; *PIntro* = probability of introduction during the previous time period.

able and used in Ireland: serum ELISA, milk ELISA, and fecal culture incorporating a confirmatory PCR test of all mycobacterial growth positive cultures. We focused on estimates for infected (as opposed to infectious or affected) cattle, following the terminology outlined by Nielsen and Toft (2008), to provide conservative Se estimates.

An Se value of 0.15 (most likely, range: 0.07–0.22) was used for individual serum ELISA, being the summary measure for infected animals from the review by Nielsen and Toft (2008). van Weering et al. (2007) reported a relative Se of 0.87 for individual milk ELISA compared with individual serum ELISA, when using one of the most commonly used kits and manufacturer's cut-off. Therefore, for individual milk ELISA, an Se value of 0.13 ($= 0.15 \times 0.87$, most likely, range: 0.06–0.19) was used. Other recent reports for individual milk ELISA Se were noted but not considered further, including 0.39 from Nielsen and Toft (2008). This value was derived from a single study using an in-house ELISA kit that used an antigen derived from *Mycobacterium avium avium* and was tested in herds preselected to have a high proportion of milk ELISA-positive animals, whereas Collins et al. (2005) found milk and serum ELISA to have similar Se. An Se of 0.23 (most likely, range: 0.16–0.30) was used for individual fecal culture at the initial screening test, based on the work of McKenna et al. (2005). This estimate was based on liquid culture using the TREK system (TREK Diagnostics System Inc., Cleveland, OH), which is the only system currently used for Map culture in Ireland. Sensitivity estimates using other culture methods were not considered relevant. Based on work by Nielsen et al. (2002), the Se of individual fecal culture as a confirmatory

test (on cattle with positive ELISA readings) was 0.65 (most likely, range 0.60–0.70).

The Sp of individual serum ELISA and individual milk ELISA were each considered 0.985 (most likely, range: 0.95–1.0). This is lower than that suggested by several authors (Supplemental Table 2; <http://www.journalofdairyscience.org/>) to account for the potential effect of repeated skin testing for bovine tuberculosis (Lilenbaum et al., 2009). In Ireland, cattle are tested at least once yearly using the single intradermal comparative tuberculin test, as part of the national control program, but more frequently if in a high-risk herd or area (More and Good, 2006). Further, the recommended cut-off for the commercial Pourquier ELISA kit (Institut Pourquier, Montpellier, France) has changed over time, with the Vidal-Diez et al. (2009) study being based on a higher Pourquier cut-off value. The range was higher than the 0.85 to 1.0 reported by Nielsen and Toft (2008); however, the 0.85 value was derived from a diagnostic kit not used in Ireland. van Weering et al. (2007) found a negligible difference in Sp using individual serum or milk ELISA. An Sp of 1.0 (most likely, range: 0.996–1.0) was used for individual fecal culture, with the range being based on the work of Vidal-Diez et al. (2009). Although a perfect Sp does not account for passive transit in a contaminated environment, such as noninfected cattle ingesting the organism from the environment and passively shedding it (Sweeney et al., 1992), any recovery of Map as a result of such “pass-through” still implies infection in the herd.

Design Prevalence. Two estimates of design prevalence were considered during modeling: a single infected animal and 5% of animals infected. A single infected animal is biologically plausible in an Irish context, given

the relatively small herd sizes and current knowledge of individual animal seroprevalence (Good et al., 2009).

Model Evaluation Options

The model was used to calculate SeH , $ProbF$, $P(False+)$, and $HerdCost$ in a dairy herd at a single point in time, given negative results at a single test (or testing combination), based on differences in (1) test type (serum ELISA; fecal culture; serum ELISA and, if positive, confirmatory fecal culture; milk ELISA and, if positive, confirmatory fecal culture); (2) herd size (100 and 50 cows); (3) number of cows >2 yr old sampled (for 100-cow herd: 100, 50, 25, and 10; for 50-cow herd: 50, 25, and 10 cows); and (4) design prevalence (1 animal, 5% of animals).

These model outputs were also calculated for a dairy herd over a 5-yr period, assuming each of the following: 100 dairy cows, annual testing with 100% sampling of animals >2 yr old, and design prevalence of 1 animal. The model was run using 2 test types (serum ELISA and, if positive, confirmatory fecal culture; milk ELISA and, if positive, confirmatory fecal culture) and 4 scenarios of animal introductions: (1) closed herd; (2) single beef bull purchased annually, normal risk source (the farm of origin was selected without regard to JD risk; therefore, P_{beef} and P_{dairy} are equivalent to animal-level JD prevalence in the Irish national herd); (3) 5% of dairy herd replaced annually, low-risk (25% of normal) source (the farm of origin was considered at low JD risk; P_{beef} and P_{dairy} are equivalent to 25% of animal-level JD prevalence in the Irish national herd); and (4) 5% of dairy herd replaced annually, normal risk source.

Further, these scenarios were evaluated based on 2 alternative assumptions: a zero risk of introduction before yr 1, or current practices considered reflective of practices during the 10 yr preceding the yr 1 test. To illustrate, applying this latter assumption to the second scenario of animal introductions, it was assumed that a single beef bull had been purchased from a normal risk source each year during both the 10 yr preceding and the 5 yr under investigation.

RESULTS

Table 2 presents SeH , $ProbF$, $P(False+)$, and $HerdCost$ in a dairy herd at a single point in time, given negative results at a single test (or testing combination), based on differences in test type(s), herd size, number sampled, and design prevalence. The following results were found.

In regard to test type, $P(False+)$ was high when serum ELISA (or milk ELISA) was used without con-

firmatory fecal culture. In addition, $ProbF$ was slightly higher when fecal culture (alone) was used compared with either serum ELISA and fecal culture or milk ELISA and fecal culture (which were very similar); however, the associated test costs were much greater. When tests were used in series (serum ELISA and fecal culture; milk ELISA and fecal culture), $HerdCost$ varied slightly, depending on whether the herd was uninfected or infected at the specified design prevalence.

Herd sensitivity (SeH) increased with herd size (100 vs. 50 dairy cows), although the difference was greater when a design prevalence of 5% was used rather than 1 animal. The value of $ProbF$ increased only slightly with increasing herd size, independent of the design prevalence selected, whereas $P(False+)$ increased with herd size independent of design prevalence.

An increase in number sampled had a greater influence on SeH than on $ProbF$. Decreasing the proportion of the herd tested did not markedly influence $ProbF$ in a negative herd because of the large influence of the prior value, but it did markedly reduce the likelihood of detecting an infected herd due to a decrease in SeH .

Herd sensitivity and $ProbF$ (albeit to a lesser extent) were lower when design prevalence was set to 1 animal rather than 5% of animals.

The impact of bio-exclusion practices, specifically the number and source of introduced animals, on confidence of JD freedom in herds testing negative (using milk ELISA, with fecal culture as a confirmatory test) at each annual test during a 5-yr period of observation is illustrated in Figures 1 and 2. Similar results were found when testing was conducted using serum ELISA with fecal culture as a confirmatory test (Supplemental Figures S1 and S2; <http://www.journalofdairyscience.org/>). Bio-exclusion practices were extremely influential, both during the 5-yr period of observation (Figure 1) and in the preceding 10 yr (Figure 2). The confidence of freedom decreased during the observation period as the probability of introduction of infection through purchase increased (such as when 5% of the dairy herd was replaced annually, without regard to infection risk in the source herd), despite negative test results throughout. Infection risk in the source herd—normal or low (25% of normal) risk—had a substantial effect on confidence of freedom during the 5 yr of observation. The 2 testing methods (serum ELISA and fecal culture vs. milk ELISA and fecal culture) gave similar confidence of freedom, although testing costs were lower with milk ELISA and fecal culture.

DISCUSSION

This paper presents quantitative estimates of confidence of JD freedom over time in dairy herds, using

Table 2. Model outputs for a dairy herd, given negative results at a single test (or testing combination) and differences in test type(s), herd size, number sampled, and design prevalence¹

Test type and herd size (animals >2 yr)	Number sampled	Design prevalence (DP)	SeH^2	$ProbF^3$	$P(False+)^4$	$HerdCost^5$ (€)	
						In uninfected herds	In herds infected at DP
Individual serum ELISA							
50	50	1 animal	0.68	0.93	0.64	425	425
50	25	1 animal	0.44	0.88	0.40	237.50	237.50
50	10	1 animal	0.21	0.83	0.18	125	125
50	50	5%	0.76	0.94	0.64	425	425
50	25	5%	0.51	0.89	0.40	237.50	237.50
50	10	5%	0.25	0.84	0.18	125	125
100	100	1 animal	0.89	0.97	0.87	800	800
100	50	1 animal	0.66	0.92	0.64	425	425
100	25	1 animal	0.42	0.87	0.40	237.50	237.50
100	10	1 animal	0.19	0.83	0.18	125	125
100	100	5%	0.93	0.98	0.87	800	800
100	50	5%	0.74	0.94	0.64	425	425
100	25	5%	0.49	0.89	0.40	237.50	237.50
100	10	5%	0.24	0.84	0.18	125	125
Individual fecal culture							
50	50	1 animal	0.23	0.86	0	1,925	1,925
50	25	1 animal	0.12	0.82	0	987.50	987.50
50	10	1 animal	0.05	0.81	0	425	425
50	50	5%	0.54	0.90	0	1,925	1,925
50	25	5%	0.31	0.85	0	987.50	987.50
50	10	5%	0.13	0.82	0	425	425
100	100	1 animal	0.23	0.84	0	3,800	3,800
100	50	1 animal	0.12	0.82	0	1,925	1,925
100	25	1 animal	0.06	0.81	0	987.50	987.50
100	10	1 animal	0.02	0.80	0	425	425
100	100	5%	0.73	0.94	0	3,800	3,800
100	50	5%	0.46	0.88	0	1,925	1,925
100	25	5%	0.26	0.84	0	987.50	987.50
100	10	5%	0.11	0.82	0	425	425
Individual serum ELISA and fecal culture							
50	50	1 animal	0.10	0.82	0	480.64	487.82
50	25	1 animal	0.05	0.81	0	264.76	272.08
50	10	1 animal	0.02	0.80	0	135.52	142.66
50	50	5%	0.26	0.84	0	480.64	498.43
50	25	5%	0.14	0.82	0	264.76	273.94
50	10	5%	0.06	0.81	0	135.52	139.05
100	100	1 animal	0.10	0.82	0	908.42	914.69
100	50	1 animal	0.05	0.81	0	480.64	487.76
100	25	1 animal	0.02	0.80	0	264.76	272.15
100	10	1 animal	0.01	0.80	0	135.52	142.64
100	100	5%	0.40	0.87	0	908.42	938.87
100	50	5%	0.22	0.84	0	480.64	498.39
100	25	5%	0.11	0.82	0	264.76	273.99
100	10	5%	0.05	0.81	0	135.52	139.06
Individual milk ELISA and fecal culture							
50	50	1 animal	0.08	0.81	0	305.64	311.80
50	25	1 animal	0.04	0.81	0	152.26	158.45
50	10	1 animal	0.02	0.80	0	60.52	66.66
50	50	5%	0.23	0.84	0	305.64	320.86
50	25	5%	0.12	0.82	0	152.26	160.12
50	10	5%	0.06	0.81	0	60.52	63.53
100	100	1 animal	0.09	0.81	0	608.42	613.80
100	50	1 animal	0.04	0.81	0	305.64	311.84
100	25	1 animal	0.02	0.80	0	152.26	158.55
100	10	1 animal	0.01	0.80	0	60.52	66.60
100	100	5%	0.35	0.86	0	608.42	634.96

Continued

Table 2 (Continued). Model outputs for a dairy herd, given negative results at a single test (or testing combination) and differences in test type(s), herd size, number sampled, and design prevalence¹

Test type and herd size (animals >2 yr)	Number sampled	Design prevalence (DP)	SeH^2	$ProbF^3$	$P(False+)^4$	$HerdCost^5$ (€)	
						In uninfected herds	In herds infected at DP
100	50	5%	0.19	0.83	0	305.64	320.91
100	25	5%	0.10	0.82	0	152.26	160.08
100	10	5%	0.04	0.81	0	60.52	63.51

¹In all cases, the probability of introduction before the test was assumed to be zero. Results from a Johne's disease (JD) survey in Ireland suggest a dairy herd prevalence of 20.6% (Good et al., 2009). Therefore, in these calculations, 0.79 was used as the pretest probability of herd JD freedom.

² SeH = probability of detecting infection if present at a prevalence exceeding a specified threshold value or design prevalence.

³ $ProbF$ = level of confidence that the true prevalence is less than the specified design prevalence value, assuming negative testing results.

⁴ $P(False+)$ = proportion of uninfected herds producing a false-positive result (one or more positive reactors).

⁵ $HerdCost$ = mean testing cost for a given testing scenario in either an uninfected herd or a herd infected at the specified design prevalence, based on sample collection and testing costs (Supplemental Table S3; <http://www.journalofdairyscience.org/>), number of animals tested, estimated test specificity, and the estimated cost of follow-up testing on positive reactors.

Figure 1. Confidence of freedom from infection in a dairy herd over 5 yr of annual testing given 4 scenarios of animal introductions, assuming negative testing results at each annual test and a zero risk of introduction before yr 1. The following were assumed: a herd of 100 dairy cows, 100% sampling of all animals >2 yr of age, a design prevalence of 1 animal, and a pretest probability of freedom of 0.79. Each animal was tested using individual milk ELISA and, if positive, confirmatory individual fecal culture. In each scenario, a herd sensitivity (SeH) of 0.08 was achieved, and the probability of a herd false positive was zero. Annual testing costs are presented in Table 2.

Figure 2. Confidence of freedom from infection over 5 yr of annual testing given 4 scenarios of animal introductions, assuming negative testing results at each annual test. Current practices were assumed reflective of practices during the 10 yr preceding the test in yr 1, leading to different pretest probabilities of freedom (at yr 0) in each scenario. The following were assumed: a herd of 100 dairy cows, 100% sampling of all animals >2 yr of age, and a design prevalence of 1 animal. The pretest probability of freedom for the 4 scenarios (in the order presented in the figure legend, starting with "Closed dairy herd") was 0.80, 0.58, 0.57, and 0.20. Each animal was tested using individual milk ELISA and, if positive, confirmatory individual fecal culture. In each scenario, a herd sensitivity (SeH) of 0.08 was achieved, and the probability of a herd false positive was zero. Annual testing costs are presented in Table 2.

a mathematical model, given alternative testing strategies and bio-exclusion practices. There were several key findings. Most notably, a high degree of confidence of freedom from infection (*ProbF*) is very difficult to achieve, reflecting the low *Se* of evaluated tests. As a consequence, only a modest annual increase in *ProbF* was achieved in closed dairy herds with each year of negative test results (Figures 1 and 2), from a pretest probability of disease of 0.79, equivalent to 1 minus the dairy herd JD prevalence in Ireland (Good et al., 2009). The value of *ProbF* was greatly affected by any risk of introduction of infection; as highlighted in Figures 1 and 2, *ProbF* decreased in herds with poor bio-exclusion practices, despite ongoing negative test results. Other variables, including herd size, test type, number sampled, and design prevalence, had limited effect on *ProbF*. Herd sensitivity (*SeH*) was greatly influenced by design prevalence (being substantially greater for a design prevalence of 5% compared with 1 animal), number sampled (increasing with increasing number sampled), and test type (increasing with increased test *Se*; Table 2). As expected, *P(False+)* is substantial when tests with imperfect *Sp* are used; it also increased with increasing herd size (Table 2). Testing strategies can substantially influence testing costs, but with little effect on test performance, as measured using *SeH* and *ProbF*. These modeling results have direct policy relevance and will inform the design and implementation of a voluntary national control program in Ireland.

In current national programs, confidence of JD freedom is assumed to increase as herds move between herd assurance (classification) levels (in CattleMAP, for example, MN2 is considered as less risk than MN1). The model results are in agreement with this principle, but suggest only marginal improvement in confidence of freedom from the starting point (the probability of freedom before testing) and between herd levels, directly because of poor test sensitivity. As illustrated in Figure 1, confidence of freedom in a closed dairy herd with annual testing increases slowly, given model assumptions, from 0.79 before testing, to 0.82 and 0.87 after the first and fifth years of testing. Current programs are based on input-based standards, which are prescriptive, focusing on actions that need to be taken (Cameron, 2012). The use of *ProbF* as outlined in this paper (or *Pr_{Low}* in an earlier paper; Sergeant et al., 2008) illustrates an output-based approach to herd classification (Cameron, 2012), and it has potential for national and field application.

The current results highlight the critical role of effective bio-exclusion practices in JD control and require that the herd owner maintains a closed herd or sources low-risk animals (e.g., from herds that have been subject to a similar testing regimen), or both, if

confidence of JD freedom is to be maintained. Otherwise, confidence of JD freedom is very substantially and adversely affected, despite ongoing test-negative results, as illustrated in Figures 1 and 2. These results highlight the critical role of on-farm risk assessment and management plans in the management of JD risk (Garry, 2011; Roussel, 2011), with particular emphasis on bio-exclusion practices in test-negative herds. The risks associated with introduced cattle are very well recognized in both the Australian and US programs. In CattleMAP, for example, the general rule is that cattle can only be introduced onto a CattleMAP property if they originate from a herd already in CattleMAP that has the same or higher status. Bio-exclusion is a substantial challenge in countries such as Ireland, where between-farm cattle movements are common (Ashe et al., 2009). The model focuses on the infection risks associated with the introduction of cattle. We acknowledge that other bio-exclusion risks may be important on some farms, including the introduction of slurry (Whittington et al., 2004; Lombard et al., 2006) or colostrum (Garry, 2011), which will be addressed through a veterinary on-farm risk assessment.

The factors affecting *SeH* in the current study, including design prevalence, number sampled, and test type, are in agreement with theoretical principles (Christensen and Gardner, 2000) and with the results of Tavoranpanich et al. (2008). The latter authors found that *SeH* was strongly associated with within-herd prevalence, the number of Map organisms shed in feces by infected cows, and the number of samples tested. Herd sensitivity is determined primarily by herd and sample size, the performance of the tests used, and the design prevalence specified. Moreover, *ProbF* is derived from *SeH* and is therefore influenced by the same factors, but is also influenced by the prior value used and the probability of introducing infection in the preceding time period. Thus, *ProbF* has a distinct advantage over *SeH* for herd classification in that *ProbF* can be updated regularly (e.g., annually) after taking account of new data from periodic testing and from any intervening risk of introduction of infection. To take advantage of the ability to update *ProbF* over time, an appropriate time interval between testing events must be established. This interval needs to be sufficiently long to meet the assumption of independence between tests, but at the same time short enough to be practical for implementation in the program and to provide ongoing assurance of herd classification. In the current model, we considered that a 12-mo interval between tests was most appropriate, although a shorter interval, perhaps 6 mo, might be possible and affordable with milk-based testing. Shorter intervals, such as 3 mo based on the milk ELISA, are likely to compromise the independence

of test events, whereas longer intervals are too long to maintain assurance without ongoing evidence of a herd's low infection risk.

The tests selected for the current model were based on individual animal tests currently available and used in Ireland. As illustrated previously, estimates of test operating characteristics were informed by the work of Nielsen and Toft (2008), supplemented by subsequent peer-reviewed publications reporting estimates of absolute JD test Se and Sp. A range of other sample types (bulk tank milk, environmental samples, pooled fecal samples) and test methods (PCR, phage assays) are described in the literature (van Weering et al., 2007; Donaghy et al., 2008; Tavoranpanich et al., 2008; Carvalho et al., 2009; Foddai et al., 2010) and could be incorporated into the current model. The use of herd-based tests (where appropriate) has the potential to reduce overall testing costs with little if any negative effect on *SeH* (Raizman et al., 2004; Berghaus et al., 2006; Lombard et al., 2006). The model also used a single estimate of Se for each test, regardless of age. This is a conservative approach, with the effect of reducing demonstrated gains in *ProbF* over time. The model could be further developed to incorporate age-related Se estimates, when available, in association with herd-specific, age-related animal data. Similarly, by varying the serological cut-off, it would be possible to vary the test Se and Sp. In this way, cattle with high titers might be expected to have a higher likelihood of being truly infected and not require confirmation by culture, thereby reducing the overall test-associated costs. Although the current literature does not currently support such a modification of the model, it could be incorporated as data become available. The current model is particularly useful in evaluating the effect of different testing strategies over time, through comparison of confidence achieved (*ProbF*) and cost (*HerdCost*).

CONCLUSIONS

For a herd to maintain reasonable *ProbF*, it is important to maintain bio-exclusion practices and minimize high-risk purchases. This study emphasized the critical role of on-farm risk assessment and associated management plans in JD control efforts. Confidence of freedom from JD was increased only marginally with repeated negative herd tests. The paper provides a novel approach to categorization of test-negative herds, offering purchasers a means to evaluate and compare risks that is not available in other programs. The current paper specifically focuses on testing issues relating to test-negative herds. Subsequently, we will examine testing

issues in test-positive herds, evaluating the cost and effectiveness of different testing strategies.

ACKNOWLEDGMENTS

We thank Isabella Higgins (UCD Centre for Veterinary Epidemiology and Risk Analysis, University College Dublin, Dublin, Ireland) for completing the herd size calculations. Funding support from Animal Health Ireland (Carrick-on-Shannon, Co. Leitrim, Ireland) is gratefully acknowledged. The authors have no conflicts of interest to declare.

REFERENCES

- Animal Health Australia. 2008. CattleMAP: Rules and guidelines of the Australian Johne's disease Market Assurance Program for cattle. Accessed Jun. 17, 2012. www.animalhealthaustralia.com.au/wp-content/uploads/2011/04/CattleMAP-manual.pdf.
- Ashe, S., S. J. More, J. O'Keeffe, P. White, G. McGrath, and I. Aznar. 2009. Survival and dispersal of a defined cohort of Irish cattle. *Ir. Vet. J.* 62:44–49.
- Berghaus, R. D., T. B. Farver, R. J. Anderson, C. C. Jaravata, and I. A. Gardner. 2006. Environmental sampling for detection of *Mycobacterium avium* ssp. *paratuberculosis* on large California dairies. *J. Dairy Sci.* 89:963–970.
- Cameron, A. R. 2012. The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. *Prev. Vet. Med.* 105:280–286.
- Cameron, A. R., and F. C. Baldock. 1998. Two-stage sampling in surveys to substantiate freedom from disease. *Prev. Vet. Med.* 34:19–30.
- Carvalho, I. A., A. Silva, V. E. B. Campos, and M. A. S. Moreira. 2009. Short communication: detection of *Mycobacterium avium* subspecies *paratuberculosis* by polymerase chain reaction in bovine milk in Brazil. *J. Dairy Sci.* 92:5408–5410.
- Christensen, J., and I. A. Gardner. 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. *Prev. Vet. Med.* 45:83–106.
- Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. *Clin. Diagn. Lab. Immunol.* 12:685–692.
- Dohoo, I., W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER Inc., Charlottetown, Prince Edward Island, Canada.
- Donaghy, J. A., M. T. Rowe, J. L. W. Rademaker, P. Hammer, L. Herman, V. De Jonghe, B. Blanchard, K. Duhem, and E. Vindel. 2008. An inter-laboratory ring trial for the detection and isolation of *Mycobacterium avium* ssp. *paratuberculosis* from raw milk artificially contaminated with naturally infected faeces. *Food Microbiol.* 25:128–135.
- Foddai, A., C. T. Elliott, and I. R. Grant. 2010. Rapid assessment of the viability of *Mycobacterium avium* ssp. *paratuberculosis* cells after heat treatment, using an optimized phage amplification assay. *Appl. Environ. Microbiol.* 76:1777–1782.
- Garry, F. 2011. Control of paratuberculosis in dairy herds. *Vet. Clin. North Am. Food Anim. Pract.* 27:599–607.
- Good, M., T. Clegg, H. Sheridan, D. Yearsely, T. O'Brien, J. Egan, and P. Mullooney. 2009. Prevalence and distribution of paratuberculosis (Johne's disease) in cattle herds in Ireland. *Ir. Vet. J.* 62:597–606.
- Lilenbaum, W., C. Marassi, R. Varges, L. Madieros, W. Oeleman, and L. Fonesca. 2009. Occurrence of false-positive results in three paratuberculosis ELISAs performed in a tuberculous herd. *Vet. Res. Commun.* 33:693–699.

- Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. *Vet. Clin. North Am. Food Anim. Pract.* 27:525–535.
- Lombard, J. E., B. A. Wagner, R. L. Smith, B. J. McCluskey, B. N. Harris, J. B. Payeur, F. B. Garry, and M. D. Salman. 2006. Evaluation of environmental sampling and culture to determine *Mycobacterium avium* subspecies *paratuberculosis* distribution and herd infection status on US dairy operations. *J. Dairy Sci.* 89:4163–4171.
- MacDiarmid, S. C. 1988. Future options for brucellosis surveillance in New Zealand beef herds. *N. Z. Vet. J.* 36:39–42.
- Martin, P. A. J., A. R. Cameron, K. Barfod, E. S. G. Sergeant, and M. Greiner. 2007b. Demonstrating freedom from disease using multiple complex data sources: 2: Case study—Classical swine fever in Denmark. *Prev. Vet. Med.* 79:98–115.
- Martin, P. A. J., A. R. Cameron, and M. Greiner. 2007a. Demonstrating freedom from disease using multiple complex data sources 1: A new methodology based on scenario trees. *Prev. Vet. Med.* 79:71–97.
- Martin, S. W., M. Shoukri, and M. A. Thorburn. 1992. Evaluating the health status of herds based on tests applied to individuals. *Prev. Vet. Med.* 14:33–43.
- McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for *Mycobacterium avium* ssp. *paratuberculosis* using tissue and fecal culture as comparison standards. *Vet. Microbiol.* 110:105–111.
- More, S. J., A. R. Cameron, M. Greiner, R. S. Clifton-Hadley, S. Correia Rodeia, D. Bakker, M. D. Salman, J. M. Sharp, F. De Massis, A. Aranaz, M. B. Boniotti, A. Gaffuri, P. Have, D. Verloo, M. Woodford, and M. Wierup. 2009. Defining output-based standards to achieve and maintain TB freedom in farmed deer, with reference to member states of the European Union. *Prev. Vet. Med.* 90:254–267.
- More, S. J., M. L. Doherty, J. O'Flaherty, L. Downey, K. McKenzie, and C. Devitt. 2011. Animal Health Ireland: Providing national leadership and coordination of non-regulatory animal health issues in Ireland. *Rev. Sci. Tech. Off. Int. Epizoot.* 30:715–723.
- More, S. J., and M. Good. 2006. The tuberculosis eradication programme in Ireland: A review of scientific and policy advances since 1988. *Vet. Microbiol.* 112:239–251.
- Nielsen, S. S., C. Grønbaek, J. F. Agger, and H. Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture diagnosis of paratuberculosis. *Prev. Vet. Med.* 53:191–204.
- Nielsen, S. S., and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon- γ assay and faecal culture techniques. *Vet. Microbiol.* 129:217–235.
- Raizman, E. A., S. J. Wells, S. M. Godden, R. F. Bey, M. J. Oakes, D. C. Bentley, and K. E. Olsen. 2004. The distribution of *Mycobacterium avium* ssp. *paratuberculosis* in the environment surrounding Minnesota dairy farms. *J. Dairy Sci.* 87:2959–2966.
- Roussel, A. J. 2011. Control of paratuberculosis in beef cattle. *Vet. Clin. North Am. Food Anim. Pract.* 27:593–598.
- Schuppers, M. E., C. F. Frey, B. Gottstein, K. D. C. Stärk, U. Kihm, and G. Regula. 2010. Comparing the demonstration of freedom from *Trichinella* infection of domestic pigs by traditional and risk-based surveillance. *Epidemiol. Infect.* 138:1242–1251.
- Sergeant, E. S. G., S. S. Nielsen, and N. Toft. 2008. Evaluation of test-strategies for estimating probability of low prevalence of paratuberculosis in Danish dairy herds. *Prev. Vet. Med.* 85:92–106.
- Sweeney, R. W. 2011. Pathogenesis of paratuberculosis. *Vet. Clin. North Am. Food Anim. Pract.* 27:537–546.
- Sweeney, R. W., R. H. Whitlock, A. N. Hamir, A. E. Rosenberger, and S. A. Herr. 1992. Isolation of *Mycobacterium paratuberculosis* after oral inoculation in infected cattle. *Am. J. Vet. Res.* 53:1312–1314.
- Tavornpanich, S., C. A. Muñoz-Zanzi, S. J. Wells, E. A. Raizman, T. E. Carpenter, W. O. Johnson, and I. A. Gardner. 2008. Simulation model for evaluation of testing strategies for detection of paratuberculosis in midwestern US dairy herds. *Prev. Vet. Med.* 83:65–82.
- Tavornpanich, S., S. J. Wells, C. P. Fossler, A. J. Roussel, and I. A. Gardner. 2012. Evaluation of an alternative method of herd classification for infection with paratuberculosis in cattle herds in the United States. *Am. J. Vet. Res.* 73:248–256.
- United States Department of Agriculture. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. Effective September 1, 2010. Animal and Plant Health Inspection Service, United States Department of Agriculture. APHIS 91–45–016. Accessed Jun. 17, 2012. [http://www.johnesdisease.org/Uniform Program Standards for the Voluntary Bovine National Johne's Disease Program.pdf](http://www.johnesdisease.org/Uniform%20Program%20Standards%20for%20the%20Voluntary%20Bovine%20National%20Johne's%20Disease%20Program.pdf).
- van Weering, H., G. Van Schaik, A. Van der Meulen, M. Waal, P. Franken, and K. Van Maanen. 2007. Diagnostic performance of the Pourquier ELISA for detection of antibodies against *Mycobacterium avium* subspecies *paratuberculosis* in individual milk and bulk milk samples of dairy herds. *Vet. Microbiol.* 125:49–58.
- Vidal-Diez, A., M. E. Arnold, R. Sayers, I. Gardner, and A. J. C. Cook. 2009. Use of Bayesian models to predict the prevalence of UK dairy herds infected with *Mycobacterium avium* subsp. *paratuberculosis* (Map). Pages 75–112 in SB4022: An Integrated Strategy to Determine the Herd Level Prevalence of Johne's Disease in the UK Dairy Herd. Final report, November 2009. Accessed Jun. 17, 2012. <http://archive.defra.gov.uk/foodfarm/farmanimal/diseases/atoz/documents/johnes-report0911.pdf>.
- Whittington, R. J., D. J. Marshall, P. J. Nicholls, I. B. Marsh, and L. A. Reddacliff. 2004. Survival and dormancy of *Mycobacterium avium* ssp. *paratuberculosis* in the environment. *Appl. Environ. Microbiol.* 70:2989–3004.
- Whittington, R. J., and P. A. Windsor. 2009. In utero infection of cattle with *Mycobacterium avium* ssp. *paratuberculosis*: A critical review and meta-analysis. *Vet. J.* 179:60–69.
- Windsor, P. A., and R. J. Whittington. 2010. Evidence for age susceptibility of cattle to Johne's disease. *Vet. J.* 184:37–44.