Evaluating assessments of TB freedom in possums: How close are we?

OSPRI

R-10769

Evaluating assessments of TB freedom in possums: How close are we?

Maria Cecilia Latham, Graham Nugent

Landcare Research

https://doi.org/10.7931/dl1t7035.2

Prepared for:

OSPRI

15 Willeston Street
PO Box 3412
Wellington 6140
New Zealand

September 2016

Landcare Research, Gerald Street, PO Box 69040, Lincoln 7640, New Zealand,
Ph +64 3 321 9999, Fax +64 3 321 9998, www.landcareresearch.co.nz

Reviewed by:	Approved for release by:
Bruce Warburton and Andrew Gormley Research Leaders Landcare Research	Dan Tompkins Portfolio Leader – Managing Invasives Landcare Research

Landcare Research Contract Report: LC2683

Disclaimer

While every effort has been made to ensure the accuracy of the information provided in this report, no warranty or representation is provided regarding the accuracy of such information, and Landcare Research does not accept liability for any losses or damage arising directly or indirectly from reliance on the information.

Contents

Summary

Project and Client

- Landcare Research, Lincoln, was contracted by OSPRI to evaluate the accuracy of subjective Proof of Freedom probabilities (iP_{free}) estimated by Area Disease Managers (ADMs) in 2015. The work was undertaken between February and September 2016.

Objective

- To determine whether subjective iP_{free} estimates provided by ADMs in 2015 are accurate, by:
- Conducting a meta-analysis comparing the probabilities of TB freedom assigned to individual Vector Control Zones (VCZs) by ADMs in 2015 against the surveillance outcomes of possum surveys conducted within those VCZs for the period 2011–2015.

Methods

- In early 2015, all ADMs were asked to subjectively estimate what they thought was the then current 'interim' probability that TB had been eliminated (iP_{free}) for all VCZs within their VRA.
- Estimates of iP_{free} and possum survey data were collated for VCZs where TB had been recorded historically. The survey data were used within the POF utility to obtain surveillance sensitivity estimates ($SS_{1/N}$) under two scenarios: including or excluding possum non-habitat. We then used conventional probability theory to calculate the probability of detection if TB had been present in possums in 2 ($SS_{2/N}$), 3 ($SS_{3/N}$), etc., up to 40 cells.
- Surveillance sensitivity estimates were then combined with the iP_{free} estimates to calculate the number of TB+ve possums expected to be recorded across surveys based on a Poisson distribution.
- We summarized the number of expected detections according to four groups (low, medium-low, medium-high, and high) based on the iP_{free} value assigned to each VCZ. Additionally, we summarized the expected number of detections across surveys based on the historical source of TB (whether in possums, wildlife but not possums, and livestock).

Results

- Across all 496 surveys analysed, the expected number of TB+ve possums (assuming a Poisson distribution) was 58 and 67 without and with habitat clipping respectively. For the much smaller number of 213 surveys with at least 400 data points (chew cards, possum traps, and possum captures), the respective numbers were only slightly lower (54 and 60). In contrast to these predictions, the 496 surveys resulted in the detection of just 3 TB+ve possums.

- From the 133 surveys in VCZs with medium-high iP_{free} (0.7–0.95), the Poisson model we used suggested that these surveys would have detected at least 13 TB+ve possums (or more precisely, detected TB+ve possums in 13 1-ha cells), but none were actually detected.
- From surveys in VCZs with medium-low iP_{free} (0.4–0.7), 15 detections were expected from a total of 51 surveys, but none was observed during possum surveys.
- All TB+ve possums detected during the period 2011–2015 were detected in VCZs with low iP_{free} (<0.4), but still this number was lower than what was expected (32).

Conclusions

- Far fewer TB+ve possums are being detected during possum surveys than expected, given the ADMs subjective assessments of TB freedom in possums. This implies the assessments made by ADMs are conservatively low.
- Given the low number of TB+ve possums actually detected during surveys, it is likely that some of this conservatism derives from the ongoing detection of TB in other hosts. However, we infer the bulk of the effect is likely to be due simply to the adoption of a precautionary approach that is resulting in considerable over-expenditure on control and/or surveillance.
- An alternative explanation for the higher number of TB+ve possums predicted compared with what is actually detected is that surveillance sensitivity is being overestimated in the POF utility. This may arise from some of the parameters being biased high. However, comparisons of field-based estimates with those derived from the POF utility (as part of another OSPRI-funded project) point to no significant bias in $SS_{1/N}$, particularly for mid-range estimates.
- A high-level implication of the downward bias in the ADMs' assessment toward TB freedom is that much more of New Zealand is already free of TB in possums than they currently believe. From a management perspective, the implication is that the transition from the control/eradication phase to the surveillance/freedom phase should be made earlier and/or the amount of surveillance imposed to declare a VCZ free of TB could be reduced.

Recommendations

- OSPRI should consider:
  - adopting a more formalised, semi-quantitative approach to estimating iPfree that is aimed at reducing the downward bias documented here
  - routinely updating iP_{free} estimates and using them more consistently as a key decision support tool in deciding where, when, and how much control and/or surveillance should be conducted.

1 Introduction

Landcare Research, Lincoln, was contracted by OSPRI to evaluate the accuracy of subjective Proof of Freedom probabilities (iP_{free}) estimated by Area Disease Managers (ADMs) in 2015. The work was undertaken between February and September 2016.

2 Background

Approximately 40% of New Zealand (10.4 million hectares) has historically been designated as being in a Vector Risk Area (VRA), i.e. areas potentially containing wildlife infected with bovine tuberculosis (TB), an infectious disease caused by the bacterium *Mycobacterium bovis*. The primary maintenance wildlife host for TB in New Zealand is the brushtail possum (*Trichosurus vulpecula*), which is able to maintain the disease independently and is now the main source of infection of TB in livestock (both cattle and deer; Livingstone et al. 2015). Many other wildlife species (particularly ferrets, pigs, wild deer) also contract TB through spillover from possums but, with the exception of ferrets in the few places where their densities are high, these are rarely considered capable of maintaining the disease.

As of 1 July 2016, OSPRI had reduced the total area designated as VRA to 8.2 million ha. This represented better than expected progress toward the target of TB eradication from 2.5 million ha by 2026 that had been set in 2011 under the 3rd National Pest Management Plan for TB (NPMP). Accordingly, the 4th NPMP adopted an ambitious new goal of achieving TB 'freedom' from possums everywhere by 2040, and complete biological eradication from all of New Zealand by 2055.

'TB freedom' and 'eradication' are terms of convenience, both representing the probability that TB is absent from a given possum population (P_{free}). Eradication is used to indicate near certainty of TB absence (i.e. $P_{free} > 0.999$), whereas TB freedom is used to indicate some lesser level of confidence (e.g. 0.95). TB freedom is a stepping stone on the path to eradication.

The strategic approach to TB eradication has (at least until now) consisted of three phases. The first 'control' phase aims to break the TB cycle in possums, and centres on the use of intensive lethal population control to reduce the densities of possums to well below the threshold at which TB infection is able to persist by intra-species transmission. This typically requires high intensity initial knockdown that reduces a target possum population by >90%, followed by subsequent maintenance control spread over 10–15 years (Nugent et al. 2015). Over the control phase, the 'interim' probability that TB has been eliminated (iP_{free}) increases (i.e. rises from near zero to a much higher level). Once iP_{free} is assessed (usually subjectively) as having reached a specified level (traditionally ~0.8), management emphasis shifts from breaking the TB cycle to 'proving' that TB has been eliminated. In this 'surveillance' phase, surveys are conducted to quantitatively assess the likelihood of TB infection still persisting in possums. If no TB is found during these surveys, the control history is subsequently modelled using the Spatial Possum Model (Ramsey & Efford 2009) to estimate the <u>prior</u> probability that TB has been eliminated from possums in the area, whereas the surveillance data are modelled in a Bayesian updating framework (the 'POF utility'; Anderson 2011; Anderson et al. 2013) to quantitatively assess the <u>posterior</u> P_{free}.

Once a specified ‘stopping’ level for P_{free} has been attained (currently set at 0.95; AHB 2012), a decision is usually made to revoke the VRA status for the particular area and declare it free of TB in possums. At that point, all or most active management ceases (and the funding is redirected to other areas still classed as VRA), and the area enters an ‘assurance’ phase during which low-intensity and/or passive or incidental surveillance and the passage of time without any detection of TB result in further increases in P_{free} towards 0.999.

In this strategic control–surveillance–assurance framework, the two key decision points for management planning and resource allocation occur at the transition between phases. The processes and protocols for the transition between surveillance to freedom and assurance have been well defined (Anderson 2011; Anderson et al. 2015). In contrast, the transition between control and surveillance has been based largely on qualitative rules of thumb based primarily on the duration of the control phase.

In early 2015, all Area Disease Managers (ADMs) were asked to estimate subjectively what they thought was the then current iP_{free} for all Vector Control Zones (VCZs) within their VRA. Their estimates tended to be low ($iP_{free} = 0$), based on strong evidence of continued presence of TB in possums, or high ($iP_{free} > 0.95$), based on actual calculations undertaken for the revocation process.

These iP_{free} estimates have quickly become a powerful tool fboth or assessing overall progress toward TB freedom nation-wide, and also for ranking VCZs in terms of determining operational priorities for the allocation of future possum control and surveillance effort. Importantly, the iP_{free} estimates are likely to be used increasingly as ‘starting rules’ to determine when to start surveillance and the minimum amount of surveillance required to enable declarations of TB freedom at the lowest possible cost. However, the accuracy of these subjective estimates has not been assessed. In particular, if the iP_{free} estimates are biased substantially low, total expenditure on control and surveillance is likely to end up being much higher than is actually necessary to achieve the goal of TB eradication.

This study therefore aimed to assess the accuracy of the ADM’s iP_{free} estimates by comparing them *en masse* against actual surveillance outcomes from the numerous possum surveys conducted by OSPRI in recent years (2011–2015). For VCZs with similar iP_{free}, we aimed to assess whether the number of times TB was detected in possums was consistent with the expected number of positive detections assuming the iP_{free} estimates were accurate. The primary focus was on VCZs that were either at the end of the control phase or already in the surveillance phase, i.e. VCZs where a crucial management decision needs to be made: whether or not to start surveillance or how much more surveillance is required to achieve TB freedom.

3 Objective

To determine whether subjective iP_{free} estimates provided by ADMs in 2015 are accurate, by:

- conducting a meta-analysis comparing the probabilities of TB freedom assigned to individual VCZs by ADMs in 2015 against the surveillance outcomes of possum surveys conducted within those VCZs for the period 2011–2015.

4 Methods

4.1 Data sources

We chose to limit the analysis to the period of the 3rd NPMP (2011–2015) because the majority of surveys conducted within that period would have been designed specifically to estimate P_{free}, whereas earlier surveys were usually not.

We obtained the 2015 iP_{free} estimates for each of the 788 VCZs in New Zealand from a database compiled by Landcare Research (from information supplied by OSPRI) as part of a project aimed at modelling alternative national strategies for TB management (Howard et al. 2015). To maximise the number of possum surveys we could include in our analyses, we used the 2015 iP_{free} estimates to generate estimates for the years 2011–2014 simply by reducing the 2015 iP_{free} by 0.05 for each preceding year (with zero as the minimum possible value). This backwards projection is considered to be very conservative because it assumes that 19 years of control is required to achieve TB freedom in previously infected areas, which is longer than usually predicted from simulation of typical possum control histories (e.g. Barron et al. 2013). In this way, any downward bias in the 2015 iP_{free} estimates would have been reduced for the previous years.

Data from possum surveys were obtained through close collaboration with Mark Neill and Ben Ainsworth (OSPRI). We identified 434 VCZs where TB infection was believed or confirmed to have been present historically in wildlife, either through confirmed detection in wildlife or through inference from livestock surveillance data. The remaining VCZs were excluded from the analysis. From the sample of 434 VCZs with TB recorded historically, 204 VCZs had TB recorded only in livestock, 81 had TB recorded only in wildlife, and 149 had TB recorded in both livestock and wildlife.

We focussed exclusively on possum surveillance, because this is the metric that is calculated by the POF utility, and because in most VCZs possums are regarded as the only true long-term maintenance host. We therefore excluded sentinel surveillance data (pigs, deer, and ferrets). Of the 434 VCZs with confirmed or inferred TB in wildlife, 199 had no possum surveillance over the period 2011–2015 (40 were not surveyed at all, 75 were surveyed only for sentinel species, and 84 were surveyed for possums only before 2011).

In the 235 VCZs remaining, a total of 496 ‘surveys’ were conducted across the 5-year period, but this number is inflated by the accidental spillover of surveying devices between VCZs – sometimes a few devices deployed in a survey were accidentally placed in a neighbouring VCZ. Using a cut-off of 400 surveillance devices within a survey (possums themselves, traps, or chewcards), the number of VCZs with surveys of a reasonable sample size reduces to 136, and the number of surveys included in our analyses reduces to 213.

The survey data did not include any so-called ad hoc ‘surveys’ involving TB+ve possums detected incidentally during other activities such as commercial fur trapping. They also did not include recent (2015–2016) surveys of areas never subject to possum control that were aimed at confirming TB presence rather than quantifying iP_{free}.

Table 1 provides a summary of the number of devices (either possums trapped, or, mostly, chew cards, or traps deployed) per VCZ per year used for possum surveillance sensitivity estimates. In total, 1,009,480 data points were included in the meta-analysis.

Table 1: Total number, range and average number of devices used during possum surveys (with ≥400 devices deployed) across 136 VCZs for each year during the period 2011–2015. In a few cases, possums themselves were the survey 'device'

Variable	Year				
	2011	2012	2013	2014	2015
Total number of devices	183 145	252 040	190 704	256 063	127 528
Range of number devices per VCZ	476–10 359	444–14 757	611–22 754	1067–18 824	463–14 403
Mean number of devices per VCZ	3330	4272	5297	6566	5314
Number of VCZs with surveys	55	59	36	39	24

4.2 Analysis

For each eligible survey, we first used the POF utility to obtain a surveillance sensitivity ($SS_{1/N}$) estimate for that survey. The $SS_{1/N}$ estimates are the estimated probability that TB would have been detected if it had been present in possums in a single 1-ha cell in an area of N ha (i.e. in surveillance jargon, we estimated the SS for a design prevalence of 1/N). From that we used conventional probability theory to calculate the probability of detection if TB had been present in 2 cells ($SS_{2/N}$), 3 cells ($SS_{3/N}$), and so on (up to 40 cells). For each plausible number of infected cells, $SS_{n/N}$ represents the probability that TB would have been detected during the surveillance survey conducted within each VCZ if it had been present in possums in *n* cells.

For each survey we analysed two scenarios: (1) all 1-ha cells included, regardless of possum carrying capacity (i.e. no habitat clipping), and (2) all 1-ha cells with a zero possum carrying capacity excluded (i.e. with habitat clipping that excludes all non-habitat). We expected that scenario (2) would provide higher estimates of surveillance sensitivity and thus higher expected number of TB+ve possum detections. All other parameters in the POF utility were set to the default values (Anderson 2011). Further, for records that had no trap-night information, we assumed that possum traps were set for 4 nights, and that chew cards were set for 7 nights, with traps set at positive detection locations for 12 trap-nights.

We then used the iP_{free} estimates to predict the expected number of TB+ve possums present in the VCZs surveyed. To do this, the subjective iP_{free} were first converted to a predicted probability distribution for the number of TB+ve possums in each VCZ assuming a Poisson distribution. This is considered to be conservative as it implies a random spatial distribution of infected possums, whereas the true distribution of any residual infection is likely to be highly clumped, with multiple infected possums occupying a single cell (although this will depend on the possum home range size relative to the cell size).

Finally, we multiplied the two probabilities above (i.e. probability that there were x TB+ve possums × probability of detecting n cells with TB+ve possums) and summed them across all plausible number of infected possums (up to 40) to determine for each survey, how many cells with TB+ve possums we would have expected to find during that survey. These values were then summed across all VCZs where surveys had occurred to determine the number of times we would have expected a positive survey (i.e. detection of TB in possums). This number was then compared against the observed number of TB+ve possums detected using $\chi 2$ tests. This comparison was conducted separately for surveys in VCZs with low (<0.4), medium-low (0.4–0.7), medium-high (0.7–0.95) and high (>0.95) iPfree estimates, with our primary focus being on the two medium groupings, i.e. those VCZs in the range in which the subjective estimates are most likely to be crucially important in guiding management decisions on when to initiate surveillance and/or how much more surveillance is required to be able to declare TB freedom.

For the subset of VCZs with medium-high iPfree, results are also summarised according to the source of TB infection for (a) VCZs with TB confirmed in possums, (b) VCZs with TB confirmed in wildlife but not definitely in possums, and (c) inferred but not confirmed in wildlife from cattle breakdowns deduced as being from a wildlife source.

5 Results

The 2015 iP_{free} estimates for all VCZs are shown in Figure 1. In the North Island, areas with low to medium-low iP_{free} estimates are centred around the Kaimanawa Mountains and the Kaweka Ranges, and also in the Tararua Ranges. In the South Island, low iP_{free} have been estimated for the West Coast and the Kaikoura Ranges, and also around Middlemarch and the Rock and Pillar Range.

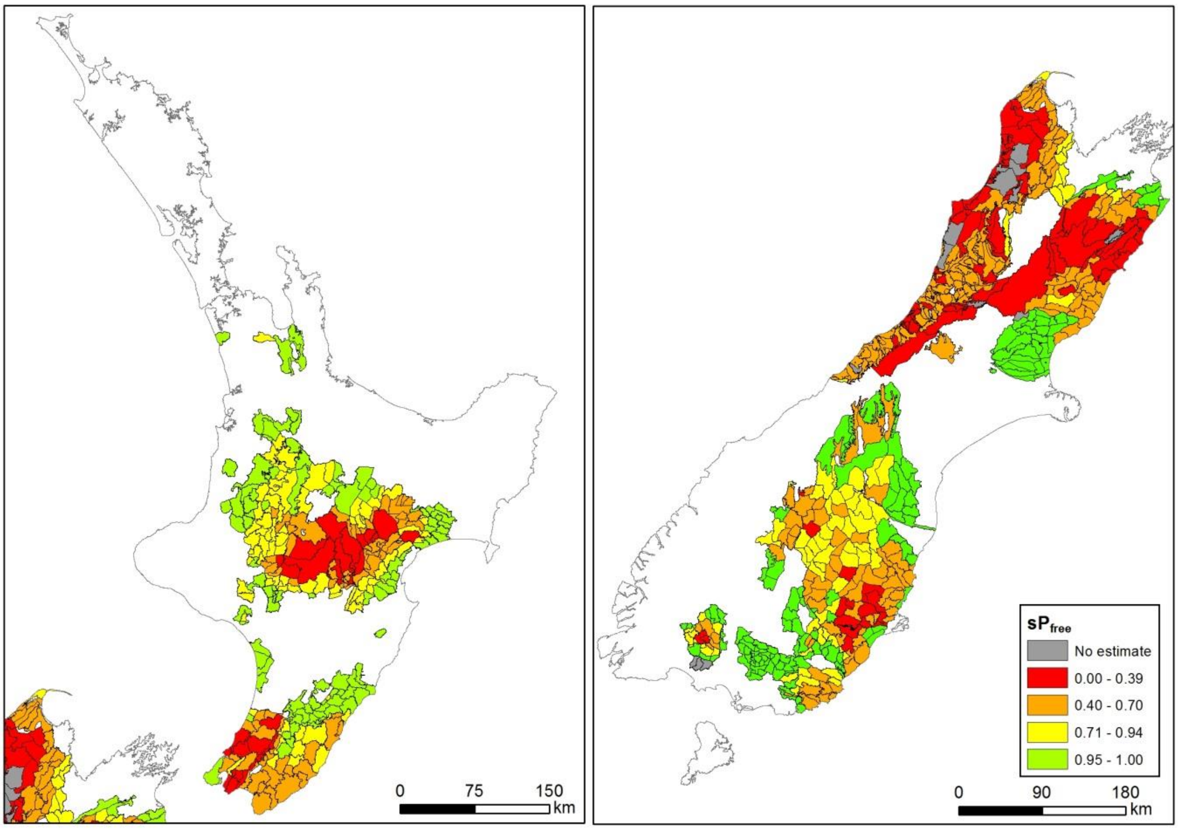

Figure 1 Probability of Freedom from TB in possums subjectively estimated by Area Disease Managers (iP_{free}) for each Vector Control Zone in 2015.

Table 2 Average iP_{free} at the time of survey, surveillance sensitivity ($SS_{1/N}$), and number of cells with TB+ve possums expected to be found over (a) all VCZs surveyed and (b) VCZs surveyed with more than 400 devices during each year for the period 2011–2015

Year	Number of surveys	Average iP_{free}	Without habitat clipping			With habitat clipping		
			Average $SS_{1/N}$	Maximum $SS_{1/N}$	Expected no. TB+ve possums	Average $SS_{1/N}$	Maximum $SS_{1/N}$	Expected no. TB+ve possums
a) All surveys								
2011	122	0.60	0.14	0.83	19	0.17	0.91	22
2012	129	0.66	0.18	0.80	13	0.22	0.83	15
2013	102	0.67	0.15	0.86	10	0.18	0.90	12
2014	82	0.72	0.23	0.87	4	0.29	0.94	5
2015	61	0.68	0.13	0.80	12	0.15	0.90	13
b) Surveys with >400 data points in a given year								
2011	55	0.66	0.30	0.83	18	0.36	0.91	20
2012	59	0.73	0.39	0.80	12	0.47	0.83	14
2013	36	0.72	0.43	0.87	8	0.50	0.90	8
2014	39	0.84	0.49	0.87	4	0.60	0.94	5
2015	24	0.75	0.32	0.81	12	0.39	0.90	13

In line with the huge variation in surveillance effort between surveys (Table 1), surveillance sensitivity estimates (SS1/N) ranged between 0 for surveys with very few data points up to 0.87 without habitat clipping and up to 0.94 when non-habitat was excluded (Table 2). The iPfree estimates were broadly similar each year, tending to be on average lower in the earlier years, which is an artefact of the backwards projection of iPfree for the early years.

Across all 496 surveys, the predicted or expected number of cells with TB+ve possums (assuming a Poisson distribution) was 58 and 67 without and with habitat clipping respectively (Table 2). For the much smaller number of larger surveys with at least 400 data points, the respective numbers were only slightly lower (54 and 60). In contrast to these predictions, the 496 surveys resulted in the detection of just 3 TB+ve possums with two detected in 2011 (Hatepe and Karamea Ground VCZs) and one in 2012 (Rolleston Range VCZ). Two of these detections were in surveys of areas in which possum control had either not been previously implemented (Mt Algidus outbreak, Rolleston Range VCZ) or only recently (and poorly) implemented (Hatepe VCZ; Nugent et al. 2015). The third, at Karamea, was a detection in a farmland area subject to sustained control but which is believed to suffer from ongoing immigration of TB+ve possums from neighbouring forest (Warburton et al. 2012).

There were no TB+ve possums detected in the 2013–2015 periods, and, other than at Karamea, there were no detections of TB+ve possums during the whole 2011–2015 in any area subject to sustained possum control.

Table 3 Average iP_{free} at the time of survey, surveillance sensitivity ($SS_{1/N}$, with habitat clipping), and number of cells with TB+ve possums expected to be found over 213 major (>400 monitoring devices) surveys conducted during the period 2011–2015. VCZs were separated into four groups according to the iP_{free} estimate from the year in which surveys were conducted

iP_{free} group	Number of surveys	Average iP_{free}	Average $SS_{1/N}$	Expected no. TB+ve possums	Observed no. TB+ve possums
Low (<0.4)	13	0.14	0.34	32	2
Medium-low (0.4–0.7)	51	0.58	0.41	15	0
Medium-high (0.7–0.95)	133	0.82	0.49	13	0
High (>0.95)	16	0.95	0.52	0.4	0

Separating the major surveys (>400 data points) in four iPfree groups indicates that the majority (62%) of such surveys were, as expected, conducted in VCZs where the iPfree was close to but still below the current surveillance stopping rule of Pfree = 0.95 (Table 3). On average, the SS1/N for these surveys was 0.49, i.e. on average, one in two such surveys would have detected TB even if TB+ve possums were present in only a single cell. However, the high iPfree values (0.7–0.95) indicate ADMs' belief that the likelihood of TB being present in possums is small. Nonetheless, the Poisson model we used suggests these surveys should have detected at least 13 TB+ve possums (or more precisely, detected TB+ve possums in 13 1-ha cells), although none were actually detected. To test the statistical significance of that difference, we assumed that 13 represented the total number of infected cells from a Poisson distribution with a mean iPfree of 0.82 (Table 3) – this equates to an expectation of 11.8 single cell detections, 1.1 2-cell detections, and a negligible number of 3+ cell detections. Based on that, we expected ~12 surveys to result in positive detections in one or more cells. The difference from the observed outcomes (i.e. no TB+ve possums) is significant ($\chi 2 = 12.2$, $P < 0.001$). A mean iPfree > 0.9 (resulting in an expectation of 5 positive surveys) would be required to make the difference statistically non-significant.

Likewise, for the group of surveys with medium-low iPfree (0.4–0.7), 15 detections were expected in total, but none was observed during possum surveys. Using the same approach as above, the expected detection of 15 infected cells equates to an expectation of about 13 positive surveys, given a Poisson distribution based on the mean iPfree of 0.58. The difference from the observed outcome (no TB+ve surveys) is significant ($\chi 2 = 16.1$, $P < 0.001$). A mean iPfree of almost 0.8 (resulting in an expectation of 5–6 positive surveys) would be required to make the difference statistically non-significant. It is woth noting that many of these VCZs had TB recorded recently in livestock and sentinels.

A few surveys were conducted in areas already considered likely to be free (iPfree ≥ 0.95; Table 3). The expected numbers of detections of cells with TB+ve possums were very low (0.4), partly because of the small number of surveys but largely because of the high iPfree. No TB was detected in sentinels or livestock at the time of these surveys (or since) in these VCZs.

There were also a few surveys conducted in areas with a low iPfree (Table 3). Despite the small number of surveys, the expected number of detections was high, and in line with that TB was detected during 2 surveys, which was still much smaller than the number of TB+ve possums expected.

Table 4 Average iP_{free} at the time of survey, surveillance sensitivity ($SS_{1/N}$, with habitat clipping), and number of cells with TB+ve possums expected to be found over those VCZs with medium-high iP_{free} (0.7–0.95) that were surveyed during the period 2011–2015. VCZs were separated according to whether TB had been detected and confirmed historically in possums themselves, or only in other wildlife sentinels (with or without livestock), or only in livestock

TB confirmed	Number of VCZs	Number of surveys	Average iP_{free}	Average $SS_{1/N}$	Expected no. TB+ve possums
Possums	3	3	0.83	0.70	0.4
Wildlife	37	65	0.82	0.50	7
Livestock	42	65	0.82	0.48	6

For those 133 surveys conducted within VCZs with medium-high iP_{free}, we further separated them into groups based on the animal where TB had been detected historically (Table 4). From these, only 3 surveys were conducted within VCZs that had TB confirmed in possums at some point in the past (Table 4). The number of TB+ve possums that were expected to be detected across these surveys was very low (0.4), and indeed no TB+ve possums were detected. The expected number of detections of TB+ve possums for VCZs where TB was detected in wildlife but not confirmed in possums was 7 (equating to ~6 positive surveys), whereas in VCZs with TB detected in livestock but not confirmed in wildlife this number was 6 (equating to ~5 positive surveys).

6 Conclusions

Far fewer TB+ve possums are being detected during possum surveys than expected, given the ADMs subjective assessments of TB freedom in possums. This suggests that those assessments (and our backwards extrapolation from them) are conservatively low. The implication is that much more of New Zealand is already free of TB in possums than is currently believed by the key disease managers. Under such circumstances, more surveillance effort is being spent in declaring areas free of TB than is actually needed. However, there are a range of possible reasons for the apparent conservatism in the ADMs-iPfree estimates, and also some methodological limitations inherent in our approach.

Beginning with the latter, our approach is an approximation that equates the number of infected cells (specified as the spatially explicit design prevalence) with the number of infected possums present. However, an infected cell could, of course, potentially contain multiple possums. Our approximation will be least reliable for high design prevalences, but we consider it to be acceptable for the range of moderate to high iPfree estimates that is our primary focus, i.e. VCZs that are close to or already in the surveillance phase. For these, the Poisson distribution predicts that for the vast majority of instances in which TB is actually present it will be present in only one or two cells. Consider, for example, 100 surveys/VCZs with an iPfree of 0.9. The Poisson expectation/prediction is that only 10 of these will contain infected possums and that 9 of those 10 will contain only one infected cell, with only a 1% chance that 2 or more cells are infected.

Another possible bias is the simplistic backwards projection of iPfree estimates from 2015 to 2011. However, as noted in the methods, we consider our extrapolations are highly likely to have resulted in overstated iPfree estimates for the early years, given that the extrapolations effectively assume that almost 20 years of sustained control is required to achieve TB freedom. If, therefore, the backwards extrapolated values are an overstatement, the predicted number of positive detections that we estimated would be biased low. Applying a less conservative approach to the backwards projection of iPfree would have resulted in even higher predicted number of positive detections.

An alternative explanation for the higher number of TB+ve possums predicted, compared with what is actually detected, is that surveillance sensitivity is being overestimated in the POF utility. This may arise from some of the parameters being biased high. For example, the sensitivity of the laboratory test used to detect TB in possums may not be as high as is assumed (0.95), or the probability of capturing a possum at a positive chew card detection may be overestimated, or the estimates of g0 and/or sigma may be too high. With regards to this last point, home range use (sigma) and detection probability (g0) are known to vary between possums in farmland vs forest vs those in the high country (Yockney et al. 2013). Given that our analyses straddled VCZs that included all these habitat types, we would have had to use different g0 and sigma parameters for each VCZ to account for this. However, these habitat-specific estimates are simply not available. Further, although overestimation of surveillance sensitivity might be an issue when the SS1/N estimates are very high (Nugent et al. 2014), comparisons of field-derived estimates with those derived from the POF utility point to no significant biased in SS1/N for mid-range estimates (P. Sweetapple, Landcare Research, unpubl. data). Finally, it is also worth noting that we did not include surveillance sensitivity derived from TB-negative sentinel surveys in our analyses. This means the SS1/N we included in our calculations are actually underestimated, which would quite likely offset the potential overestimation of SS1/N described above.

The conservatism in the iPfree estimates may partly reflect simple pessimism by some or all ADMs. However, it may also have an epidemiological basis – the iPfree estimates are likely to reflect their belief that all wildlife in a VCZ are free of TB, not just possums, so will be lower where vector-induced TB is believed to have occurred in livestock or where TB has been detected in sentinels. The conundrum is that failure to detect TB in such sentinels is evidence of TB absence in possum, whereas detection of TB in sentinels does not necessarily confirm ongoing TB presence in possums. This can be because deer, in particular, can remain alive in an infected state for many years without transmitting TB, but eventually transmit it to scavengers such as pigs and ferrets, and also possibly possums (Barron et al. 2013). It can also be because ferrets may be able to cycle TB intra-specifically for some years (or even indefinitely) after the usual source of infection (TB+ve possums) has been eliminated.

To check for this, we examined the subset of VCZs with moderate-high iPfree estimates (0.7–0.95) in which there had been detection of vector-induced TB in livestock or in sentinels. More than 85% of these VCZs have had no TB detected in sentinels since 2009, the remaining VCZs had detection of TB in pigs and ferrets in 2012-2013 but none in 2014–2015. Likewise, 95% of these VCZs have had no TB detected in livestock since 2009, with only four infected herds detected during the period 2011–2013, but none detected in the last two years (2014–2015). So even though ADMs appear to have assigned moderate-high iPfree estimates to VCZs where detections of non-possum TB infection occurred historically, most of these would have occurred at least 6 or more years ago.

Following this line of thought, we also examined the larger group of 184 surveys conducted within VCZs with medium iPfree (0.4–0.95), where all detections of TB since (and including) 2011 were in sentinels or livestock but none in possums. For this group, we expected that ADMs would have substantially reduced iPfree for those VCZs with recent infection (2011-onwards), but there was no strong indication of that (mean iPfree = 0.70, c.f. to a mean iPfree = 0.78 for VCZs with TB detected in livestock and/or sentinels pre-2011). Further, although the proportion of expected positive surveys was slightly larger for those VCZs with recent TB detections (8/54, c.f. 16/130 for VCZs with TB detected pre-2011), it was not statistically significant ($\chi 2 = 0.05$, $P = 0.82$). This indicates that either ADMs are downgrading iPfree estimates to very low levels (<0.4) wherever TB has been recently detected or they are largely basing the iPfree estimates on the robustness of the possum control history (or both).

Thus, in summary, we conclude that as a general rule ADMs tend to be more pessimistic about progress toward TB freedom in possums than they should be. Part of the pessimism is attributable to ongoing detection of TB in other hosts, where conservatism is likely to be appropriate. However, we infer the bulk of the effect is likely to be due simply to the adoption of a precautionary approach – better to underestimate progress, and do more control than the minimum necessary, than to risk being wrong and stopping control too soon.

The high-level implication of the downward bias in iPfree is, first, that much more of New Zealand is already free of TB in possums than is currently believed by the key disease managers. From a management perspective, the implication is that the transition from the control/eradication phase to the surveillance/freedom phase should be made earlier and/or the amount of surveillance imposed to declare a VCZ free of TB could be reduced. The latter is supported by those VCZs that have been declared free of TB since 2011: for the majority of these, the posterior P_{free} calculated quantitatively was found to be well above the stopping rule of 0.95 – usually >0.97 and sometimes approaching 0.99 (G. Nugent, Landcare Research, unpubl. data). This indicates that substantially more than the minimum necessary surveillance has been conducted in such VCZs.

If, as we infer, the downward bias in iPfree estimation is resulting in over-expenditure on control and/or surveillance, the question that arises is how to minimise this problem. One solution is simply to lower the 'surveillance start' and 'stopping' rules from the current levels of 0.8 and 0.95 respectively to (say) 0.65 and 0.90 respectively. Another is to simply make ADMs aware of the bias and encourage them to try to be more aggressively optimistic in assessing progress. A third option is to develop a set of guidelines for assessing iPfree that includes consideration of factors such as:

(i) the evidence that TB was actually established (or well established) in the possum population

(ii) the duration and intensity of the possum control programme imposed to break the TB cycle

(iii) passive surveillance data from livestock and other sources indicating TB presence or absence

(iv) the likelihood of immigration by TB+ve possums from neighbouring VCZs.

A set of such guidelines are currently being developed as part of the broader development of technical guidelines for the 4th NPMP (G. Nugent, pers. comm.). An extended version of this could involve using the Spatial Possum Model (Ramsey & Efford 2009) to quantify iPfree annually based on the known or inferred history of possum control, as is currently done as part of the 'declaration of freedom' process. However, this would only cover point (ii) in the list above, implying that some subjective assessment by ADMs would still be required.

7 Recommendations

OSPRI should consider:

- adopting a more formalised approach to estimating iP_{free} that is aimed at reducing the downward bias documented here;
- routinely updating iP_{free} estimates and using them more consistently as a key decision support tool in deciding where, when, and how much control and/or surveillance should be conducted.

8 Acknowledgements

Funding for this work was provided by OSPRI. We thank Dean Anderson and Andrew Gormley (Landcare Research) for helpful discussions on the POF utility. We also thank Paul Livingstone, Ben Ainsworth, and Mark Neill (OSPRI).

9 References

AHB 2012. Animal Health Board Annual Report 2011/2012. http://www.tbfree.org.nz/Portals/0/Interactive%20e-book%20research%20report.pdf. Animal Health Board, Wellington, New Zealand.

Anderson DP, Ramsey DSL, Nugent G, Bosson M, Livingstone P, Martin PAJ, Sergeant E, Gormley AM, Warburton B 2013. A novel approach to assess the probability of disease eradication from a wild-animal-reservoir host. Epidemiology and Infection 141: 1509–1521.

Anderson DP, Ramsey DSL, de Lisle GW, Bosson M, Cross ML, Nugent G 2015. Development of integrated surveillance systems for the management of tuberculosis in New Zealand wildlife. New Zealand Veterinary Journal 63(S2): 89–97.

Anderson DP 2011. Development of a Proof of Freedom utility. Landcare Research Contract Report LC218 for the Animal Health Board. 30 p.

Barron MC, Nugent G, Cross ML 2013. Importance and mitigation of the risk of spillback transmission of *Mycobacterium bovis* infection for eradication of bovine tuberculosis from wildlife in New Zealand. Epidemiology and Infection 141: 1394–1406.

Howard S, Gormley A, Nugent G, Braaksma N 2015. Developing a national TB freedom possum model. Landcare Research Contract Report LC2409 for the Animal Health Board. 45 p.

Livingstone PG, Hancox N, Nugent G, Mackereth G, Hutchings SA 2015. Development of the New Zealand strategy for local eradication of tuberculosis from wildlife and livestock. New Zealand Veterinary Journal 63(S2): 4–18.

Nugent G, Buddle BM, Knowles GJE 2015. Epidemiology and control of *Mycobacterium bovis* infection in brushtail possums (*Trichosurus vulpecula*), the primary wildlife host of bovine tuberculosis in New Zealand. New Zealand Veterinary Journal 63(S2): 28–41.

Nugent G, Whitford J, Latham C, Morris G 2014. Rapid declaration of TB freedom in possums in Hokonui Hills. Landcare Research Contract Report LC2076 for the Animal Health Board. 45 p.

Ramsey DSL, Efford M 2010. Management of bovine tuberculosis in brushtail possums in New Zealand: predictions from a spatially explicit, individual-based model. Journal of Applied Ecology 47: 911–919.

Warburton B, Anderson D, Nugent G, Whitford, Neill M 2012. Identification of risk factors associated with new and persistent infection in cattle herds at Karamea. Landcare Research Contract Report LC1143 for the Animal Health Board. 84 p.

Yockney IJ, Nugent G, Latham MC, Perry M, Cross ML, Byrom AE 2013. Comparison of ranging behaviour in a multi-species complex of free-ranging hosts of bovine tuberculosis in relation to their use as disease sentinels. Epidemiology and Infection 141: 1407–16.