Sampling protocol to determine sensitivity and specificity of Kauri dieback testing

PREPARED FOR MPI

BY EMILIE VALLEE, GEOFF JONES AND NAOMI COGGER

MASSEY UNIVERSITY

AUGUST 2019

Objectives

To design a sampling protocol to collect the data needed to assess the diagnostic sensitivity and specificity of aerial inspection and laboratory test of soil sampling conducted in the same way as ground-truthing to diagnose kauri dieback. Knowing the sensitivity and specificity will inform the Kauri dieback surveillance program, in particular in the design of a protocol to classify an area as free of the pathogen. However, the design of a survey to detect freedom from disease is outside of the scope of this project.

Definitions

Gold standard test: a test with a sensitivity and specificity of 100%, that is a test that produces no false-positive or false-negative test results.

Sensitivity: refers here to the diagnostic sensitivity, the proportion of infected units that will test positive

Specificity: refers here to the diagnostic specificity, the proportion of healthy units that will test negative

Prevalence: the proportion of units in a site that truly are infected

Testing in series: Using two diagnostic tests on the same unit, with a unit being considered positive when both tests produce a positive result. For all other combinations of the unit would be classified negative.

Test accuracy: The ability of a diagnostic test to correctly discriminate between disease and non-disease states of the unit tested.

We also distinguish in the text the presence of *P. agathidicida* in the soil, the infection by *P. agathidicida* and the kauri dieback disease itself, which is the expression of symptoms.

Current methods for detecting Kauri dieback

Kauri dieback is a complex disease with slow evolution that will eventually result in the death of a kauri tree. There is a long non-clinical, infected stage (long incubation period). It is caused by the fungus-like pathogen *Phytophthora agathidicida*, although other *Phytophthora* species are likely involved (Waipara et al., 2013). It is transmitted by movements of soils, on the shoes or gears of visitors to kauri forests, by vehicles, during heavy rain or floods or by animals such as wild pigs. The pathogen can survive for years in the soil. It damages the roots of kauri and impairs the tree's water and nutrient intake (Figure 1). The clinical disease has been reported since 2006 although it had likely been present for years, and the pathogen formally identified in 2015. The clinical signs of the disease can be observed either from the ground (bleeding gum, yellowing of leaves, possibly canopy thinning) and when flying over the canopy (thinning canopy, dead branches or the tree itself).

Currently, the two methods are used in combination to detect and monitor Kauri dieback; an aerial survey of the canopy, followed by laboratory testing to ground-truth the results of the inspection. It has been suggested that aerial surveys are a cost- and time-effective tool to be used in combination with other tests, and not useful for early detection, suggesting imperfect sensitivity given our definition of true positive, since *P. agathidicida* will be present in the soil before the apparition of clinical signs visible by aerial inspection (Figure 1).

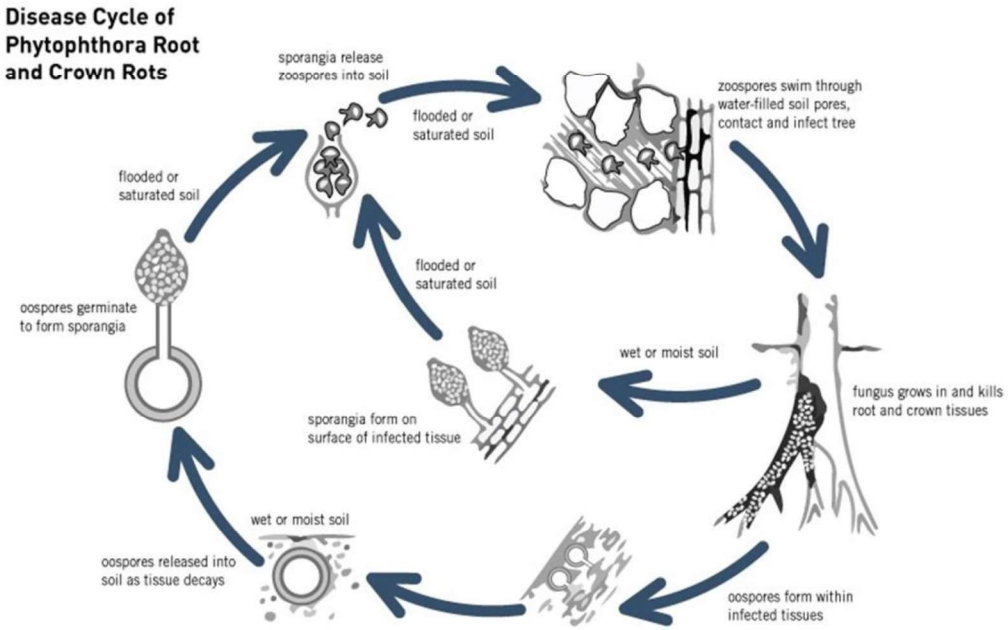

Figure 1: Disease cycle of Phytophora root and crown rots (taken from https://www.kauridieback.co.nz/how-does-it-spread/)

Aerial inspection

Aerial surveillance has been described in detail by Jamieson *et al.* (2014). Briefly, the approach combines visual inspection when flying over the site in a helicopter or aircraft with high-resolution photography of kauri canopy taken during the flight with GPS data. The specificity of aerial inspection is expected to be imperfect because trees can have similar symptoms to kauri dieback as a result of environmental stress (drought), age, presence of other *Phytophthora* pathogens, or damage from lightning or fall of the neighbouring tree. In a study by Jamieson et al. (2014) in the Hunua Ranges none of the trees identified

during the aerial survey as having symptoms compatible with *P. agathidicida* infection had soil samples that tested positive for *P. agathidicida*.

Laboratory testing to ground truth (GT) aerial surveillance

Laboratory testing, including recommendations on soil sampling, are described in detail in Beever et al. (2010) and Dick & Bellgard (2010). Briefly, soil samples are taken from eight locations around a tree: four are "cardinal points" taken one and two meters from the base of the tree; one sample aligned with a lesion if present. The samples are then pooled to produce a final soil volume of approximately 1000 cm^3. The laboratory testing can distinguish between *P. agathidicida* and other *Phytophthora* species, which suggests a high specificity. That is, the test is unlikely to classify a sample as positive when the pathogen is not present. However, the results of a study by Singh et al. (2017) suggest the test is likely to have imperfect sensitivity that is the ability to detect the pathogen when present. Specifically, of the 44 samples taken from a tree with visible signs of disease, in the Waitākere Ranges, only 27 tested positive. By comparison, a PCR test on the same 44 samples was positive for 41 samples.

Approaches to evaluating diagnostic test accuracy

To date, the measures used to evaluate testing for Kauri dieback have focused on accuracy that is the proportion of units correctly classified as diseased. The problem with using accuracy is that the results are affected by the prevalence in the population studied. Other methods of assessing test performance involve assessing the diagnostic sensitivity and specificity. In this context, sensitivity refers to the proportion of truly positive units that test positive, while diagnostic specificity refers to the proportion of truly disease-free units that test negative. Sensitivity and specificity are selected as their values are independent of the prevalence of the disease in the population. Sensitivity and specificity provide important information about test performance but are not affected by differences in the occurrence of disease, and as such will not vary between sites. Further, knowledge of test sensitivity and specificity, rather than accuracy, is essential when designing a surveillance system to monitor the spread of disease and develop a protocol for a survey to demonstrate freedom-of-disease in an area.

Traditionally, calculation of sensitivity and specificity required comparison of tests results to a gold standard, so that true disease status of the units is known. Table 1 depicts how the results of a test against a gold standard would typically be displayed. Using the format presented in Table 1, we would calculate the sensitivity and specificity of a test as follows:

$$\text{Sensitivity} = a/(a+c)$$
$$\text{Specificity} = d/(b+d)$$

Table 1: Classification of data used to evaluate a test when a gold-standard test is available.

	Gold standard disease present	**Gold standard disease absent**	
Test positive	True positive (TP): a	False positive (FP) b	Total test positive: a+b
Test negative	False negative: c	True negative (TN): d	Total test negative: c+d
	Total diseased: a+c	Total population: b+d	Total tested: a+b+c+d

While conceptually appealing, we can rarely evaluate a test against a true gold standard. Therefore, alternative methods have been developed to assess diagnostic test performance in the absence of a gold standard. In this report we propose estimating the sensitivity and specificity of the two methods currently used to detect Kauri dieback, namely aerial inspection and laboratory testing conducted currently for ground-truthing, using the two-population model described by Hui and Walter (1980). Briefly, the two population model estimates the sensitivity and specificity without a gold standard by testing two different populations using both tests. The following section of the report describes the approach taken to determine the number of samples at each site that should be tested to allow sensitivity and specificity to be estimated. The two-population method for evaluating tests in the absence of a gold-standard is used in animal health, and the methodology has been recognised by the World Animal Health Organisation (OIE) as appropriate.

Simulation to determine the number of samples required to evaluate tests

Assumptions

Currently, surveillance operations involve aerial surveillance to identify trees with clinical signs. For any kauri tree with clinical signs coherent with dieback, a process of 'ground-truthing' is then undertaken. This process involves laboratory testing of soil collected from near the tree to determine if *P. agathidicida* is present. The two tests are interpreted in series: the tree is considered positive if it has clinical signs consistent with Kauri dieback and the *P. agathidicida* must be isolated from the soil.

In this report, we are determining the number of units required to estimate the sensitivity and specificity of the aerial inspection and the whole process from sampling soil around a tree at a distance not exceeding three times the drip line to laboratory testing. In the case of laboratory testing, we are evaluating the whole procedure. That is the study is concerned not just with the laboratory component but includes the number of soil samples collected at each site, location of sampling relative to the tree, sample volume, storage, transport mode and duration, storage in the laboratory. Therefore, the process of collecting and testing soil samples must be standardised. Further, to evaluate the tests, however, we require a single definition of the true positive status. For this study, the definition of a truly positive is:

> *Presence of P. agathidicida in the soil around the tree at a distance not exceeding three times the drip line, with or without signs of Kauri dieback.*

The process of determining the number of samples required probability distributions for the parameters we wanted to estimate; namely the sensitivity and specificity of the aerial inspection and laboratory testing for ground truth. All recommendations assume the prior distributions for the 1) sensitivity and specificity obtained from experts during the expert elicitation process and 2) the prevalence of *P. agathidicida* at each site. Should the true values used differ substantially, then the uncertainty around the estimated for sensitivity and specificity of the two tests may be greater than predicted. Also, we have assumed:

1. That the analytical sensitivity and specificity for each test has been deemed sufficient and that the repeatability is considered acceptable.
2. For each test, the sensitivity and specificity do not change from site to site or from tree to tree.
3. The two tests are conditionally independent; this means that for a given disease status, the result of one test does not depend on the result of the other test. We believe that this is a valid assumption because one test is based on the visual observation of clinical signs and the other one on the detection of *P. agathidicida* in the surroundings of the tree (Gardner et al. 2000).
4. The study will evaluate the entire procedure, including the pre-laboratory stage, so the process must be the same for all units.

Determining prior distributions

Data upon which to base the prior distributions were lacking. Therefore, a process of expert elicitation was undertaken to construct appropriate distributions for the sensitivity and specificity of aerial inspection and laboratory testing of soil collected from around Kauri trees. The method used to elicit expert opinion is described in Hemming et al. (2018). Briefly, the experts listed in Appendix 1 were asked to complete an online questionnaire. The questionnaire asked them to consider each test separately and

provide the minimum, maximum and most likely values for the sensitivity and specificity along with their confidence their answer encompassed the truth (see Appendix 2 for details). Round 1 answers were shared with experts at a workshop on 28th May 2019 and participants allowed to discuss the responses and ask questions. The online questionnaire was sent out for a second time, along with the results of the first round and a summary of the discussion at the workshop. After the second round, the answers to each question were averaged and scaled to determine the 80% confidence interval. Figure 2 depicts the beta distributions created by transforming the minimum, maximum and most likely values using the 'prevalence' package (Version 0.4.0) in the statistical package R.

During the first round of expert elicitation experts were also asked about possible locations, with the following criteria: 1) different prevalence between the two sites, 2) not completely free of *P. agathidicida* and not fully infected, 3) roughly homogenous distribution of infected trees within the site, 4) not currently a controlled area, and 5) not treated with phosphite. The research team, in discussion with MPI, decided on the two sites to be used in the evaluation of the diagnostic tests. The prior distribution of the prevalence of Kauri dieback was then determined using the results from an aerial survey conducted in 2018.

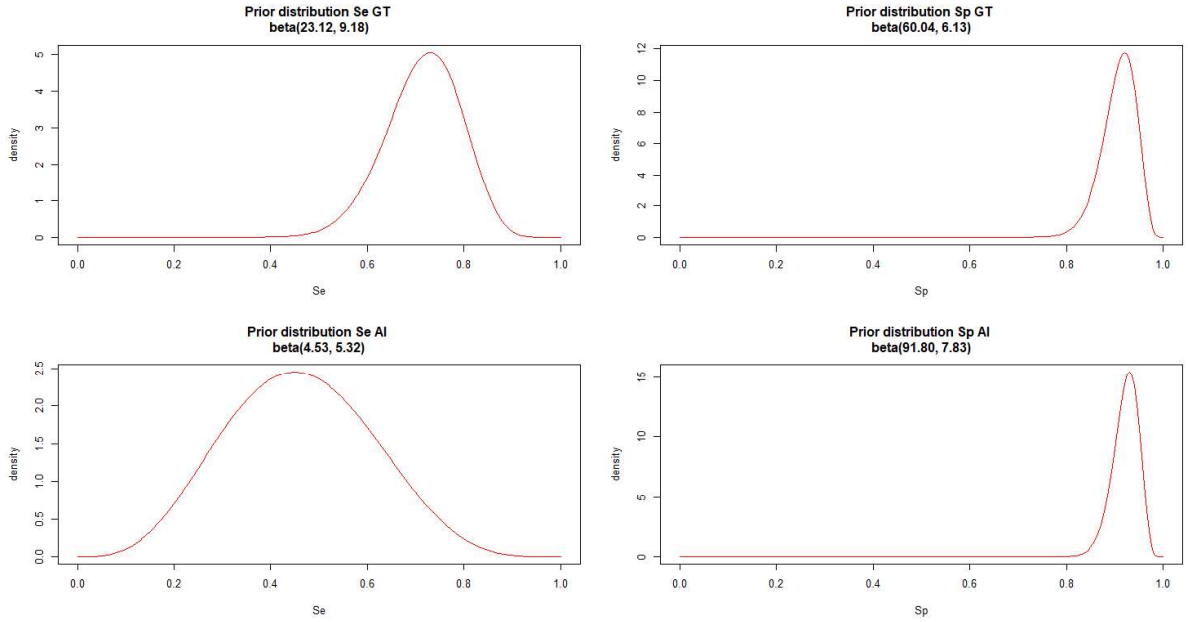

Figure 2: Prior probability distributions, based on expert opinion, for sensitivity (Se) and specificity (Sp) of aerial inspection (AI) and laboratory testing of soil samples taken from the soil around kauri, at a distance not exceeding three times the drip line (GT).

Numbers of samples

We ran 12 separate simulations to determine the effect on precision on the sensitivity and specificity of each test when the number of samples taken from each site varied from between 50 to 1000. For each simulation, 1000 iterations were drawn from the prior distribution (see Figure 2) and Bayesian latent class analysis was then used to determine the 95% credible intervals for the test parameters for each iteration. Therefore for each of the 12 simulations, there were 1000 values for the credible intervals for sensitivity and specificity of the aerial inspection and laboratory testing of soil taken around kauri trees. For each simulation, there were 1000 values for the credible intervals for each of the measures of test performance. To summarise the results from the 12 simulations simulation the mean credible interval

and half-width of the interval were calculated. Credible intervals measure uncertainty around a parameter: the wider the interval, the higher the uncertainty and the lower the precision.

The results showed that for a given number of samples, the precision was better for the sensitivity than the specificity of laboratory testing of soil samples taken from around kauri trees (Table 2). Increasing the number of samples improved the precision of the estimates in all cases (Figure 3). However, the improvements in precision slowed when 400 samples were taken from each site (i.e. 800 from both sites) and became negligible from 600 per site onwards (total 1200 trees). Further, simulations were run to determine if varying the proportion of 1,200 tests table from each site could improve precision. Results suggest that take more samples from the high prevalence site could result in better precision for the test estimates than taking an equal number from each site.

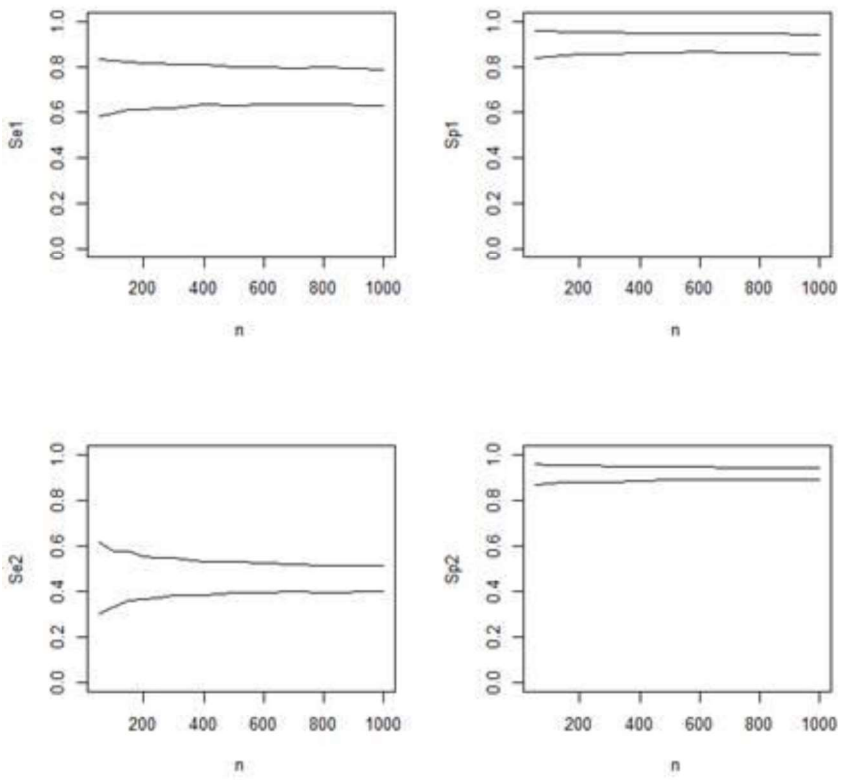

Figure 3: Effect of increasing number of samples tested (n) at each site at each site on the precision, measured as 95% credible interval over 1000 simulations, on the sensitivity and specificity of aerial inspection (Se2, Sp2) and laboratory testing of soil samples taken from the soil around kauri, at a distance not exceeding three times the drip line (Se1, Sp1).

Table 2: Effect of increasing number of samples tested at each site from 50 to 1000 at each site on the precision, measured as credible interval half-width average over 1000 simulations, on the sensitivity and specificity of aerial inspection and laboratory testing of soil samples taken from the soil around kauri, at a distance not exceeding three times the drip line.

Number of samples	Laboratory testing		Aerial inspection	
	Sensitivity	Specificity	Sensitivity	Specificity
50	0.127	0.06	0.158	0.046
100	0.117	0.056	0.125	0.042
150	0.107	0.053	0.107	0.039
200	0.103	0.05	0.097	0.037
250	0.099	0.049	0.089	0.035
300	0.096	0.047	0.083	0.034
400	0.089	0.044	0.075	0.031
500	0.084	0.043	0.069	0.029
600	0.083	0.043	0.064	0.028
700	0.08	0.041	0.061	0.027
800	0.078	0.04	0.058	0.026
1000	0.08	0.043	0.056	0.026

Table 3: Effect of altering the proportion of 1200 samples taken at the high and low prevalence site on the precision, measured as credible interval half-width average over 1000 iteration, on the sensitivity and specificity of aerial inspection and laboratory testing of soil samples taken from the soil around kauri, at a distance not exceeding three times the drip line.

Number samples		Laboratory testing		Aerial inspection	
High prevalence site	Low prevalence site	Sensitivity	Specificity	Sensitivity	Specificity
200	1000	0.101	0.04	0.091	0.025
300	900	0.094	0.04	0.08	0.026
400	800	0.091	0.041	0.074	0.026
500	700	0.087	0.042	0.069	0.027
600	600	0.081	0.041	0.063	0.028
700	500	0.079	0.043	0.061	0.029
800	400	0.079	0.045	0.059	0.031
900	300	0.075	0.046	0.056	0.033
1000	200	0.074	0.05	0.055	0.036

Recommended sampling protocol

Sampling sites

Discussions with experts identified two sites in Pukekaroro that could be used to determine the sensitivity and specificity of the two tests using the latent class approach. One site had a prevalence of symptomatic trees of 61% out of 1513 trees (Figure 3A), the other 17% out of 1648 trees (Figure 3B) in a January 2018 aerial inspection round. *P. agathidicida* presence is confirmed in the high prevalence site, and both sites are contiguous and of easy access.

Selection of sampling units

The unit chosen for this study is a kauri tree and the surrounding soil to a distance of no more than three times the dripline for each tree. At each site, the sample units need to be chosen randomly, with no regard for the disease status of the unit, and each unit must be tested by both aerial surveillance and laboratory testing of soil samples. The order in which the aerial surveillance and laboratory testing of soil samples are tests are done as long as the timeframe is sufficiently short that it is reasonable to assume that the true positive status hasn't changed.

Number of samples

Across the two sites, we recommend 1,200 sampling units be selected at random, with no regard to whether signs consistent with Kauri dieback disease are present. Should the number of sampling units be deemed too expensive, it should be possible to gain reasonable precision with 800 samples. Regardless of the total number of samples, the proportion of samples from the higher prevalence site should be greater. For example, if 1,200 trees are sampled, then 800 to 900 should come from the high prevalence site.

Note: The final precision will depend on the actual data collected. The simulations assumed the data were compatible with the prior probability distributions, but the real data may be different, which may ultimately increase the uncertainty compared to the simulations.

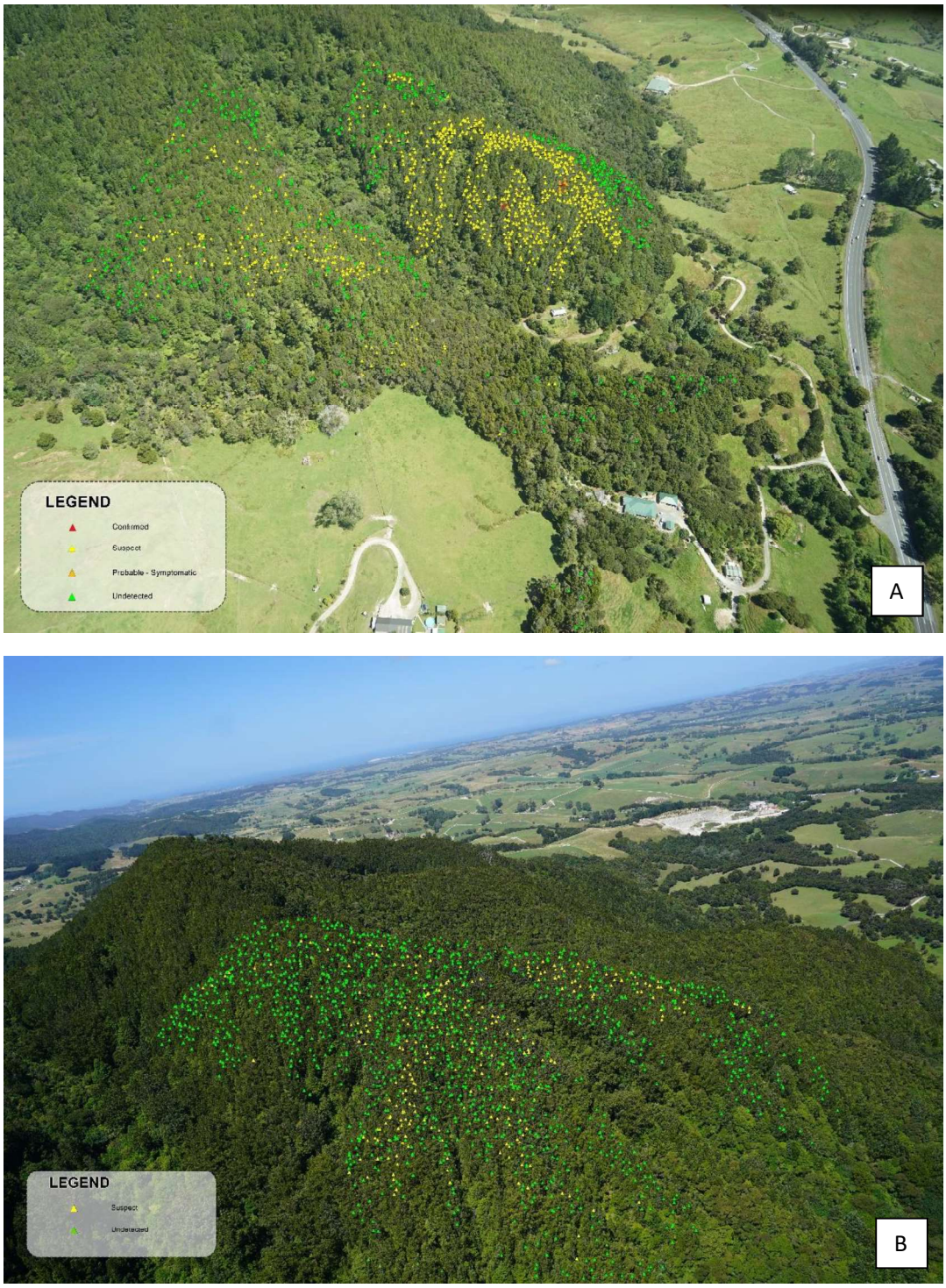

Figure 3: High (A) and low (B) prevalence sites (credit A. Macdonald, Biospatial)

Testing protocol for aerial inspection

For each of the sampling units, aerial inspection should be at a similar time of day, close in calendar time and under same weather conditions. The rationale for this is that these factors can affect the visibility and hence, the test performance. Also, when conducting the aerial inspection, the proforma in Appendix 3 should be completed for every run.

Testing protocol for aerial inspection laboratory testing of soil samples

In the case of laboratory testing of soil samples, we are evaluating the whole procedure. That is the study is concerned not just with the laboratory component but includes the number of soil samples collected at each site, location of sampling relative to the tree, sample volume, storage, transport mode and duration, storage in the laboratory. Therefore, the same process must be used for all sampling units. Given that the goal of the exercise is to create values that can be used to design surveillance system, it would be advisable that whatever method used for soil sampling in this evaluation is the same as what will be used for future surveillance activities. Put another way, the estimates obtained in the project will only apply to the process used here and a change to the protocol this exercise would have to be repeated. Also, when collecting the sample for each sampling unit, the proforma in the Best Practice Guideline for soil survey methodology for *P. agathnida* (Beauchamp, 2016).

Conclusions and recommendations

Kauri dieback is currently detected using aerial inspection followed by laboratory testing of soil samples collected to ground-truth the aerial inspection. Work has been undertaken to determine the accuracy of the tests. However, test accuracy varies depending on the prevalence of the diseases in the population tested. Therefore, medical and human health professionals also consider diagnostic sensitivity and specificity when evaluating tests. A Bayesian latent class analysis method should be used to determine the sensitivity and specificity because there is no gold-standard test for Kauri dieback. This report describes the method used to estimate the number of sampling units required, at two sites to evaluate the diagnostic sensitivity and specificity of the two tests. The sampling unit is a kauri tree and the surrounding soil to a distance of no more than three times the dripline for each tree. We recommend:

1. Random selection of at least 800 sampling units, but ideally 1 200, across both sites. It is important that the sampling units selected without regard for whether signs of Kauri dieback are present on the trees.
2. Favourably sample from the high prevalence site at a ratio of between 2:1 or 3:1.
3. Each sampling unit must be tested using both aerial surveillance and laboratory testing of soil samples.
4. The order in which aerial surveillance and laboratory testing of soil samples are performed does not matter and can in whatever order is most convenient. The only requirement is that time frame between test is sufficiently short allow an assumption that the actual disease status has not changed between testing.
5. The method used to collect, store and test soil samples must be consistent between samples as the analysis will evaluate the process as a whole.

In conclusion, the recommended number of sampling units was based on the prior probability distributions, constructed using expert opinion. Should the actual data have a different distribution, the uncertainty for the sensitivity and specificity of the two tests may be higher than predicted.

References

Beauchamp, T (2016). Soil survey methodology for Phytophthora agathidicida. Report prepared for Kauri Dieback Program, Ministry of Primary Industries.

Beever, R. E., Bellgard, S. E., Dick, M. A., Horner, I. J., & Ramsfield, T. D. (2010). *Detection of Phytophthora taxon Agathis (PTA): Final report*.

Dick, M. A., & Bellgard, S. E. (2010). *Preliminary survey for Phytophthora taxon Agathis*.

Gardner, I. A., Stryhn, H., Lind, P. D., & Collins, M. T. (2000). Conditional dependence affects the diagnosis and surveillance of animal diseases. *Preventive Veterinary Medicine, 45*, 107-122.

Hemming, V., Burgman, M. A., Hanea, A. M., McBride, M. F., & Wintle, B. C. (2018). A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution, 9*, 169-180. doi:10.1111/2041-210X.12857

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

Jamieson, A., Bassett, I. E., Hill, L. M. W., Davis, A., Waipara, N. W., Hough, E. G., & Horner, I. J. (2014). Aerial surveillance to detect kauri dieback in New Zealand. *New Zealand Plant Protection, 67*, 60-65.

Singh, J., Curran-Cournane, F., Waipara, N. W., Schwendenmann, L., & Lear, G. (2017). *Comparison of methods used to detect the organism responsible for kauri dieback, Phytophthora agathidicida, from soil samples*. Auckland, New Zealand.

Waipara, N. W., Hill, S., Hill, L. M. W., Hough, E. G., & Horner, I. J. (2013). Surveillance methods to determine tree health, distribution of kauri dieback disease and associated pathogens. *New Zealand Plant Protection, 66*.

Appendices

Appendix 1: Experts consulted for prior probability distributions and affiliations

Name	Affiliation
Tony Beauchamp	Department Of Conservation
Stanley Bellgard	Manaaki Whenua – Landcare Research
Gavin Clapperton	Northland Regional Council
Ian Horner	Plant & Food Research
Andrew Macdonald	Biospatial
Kim Parker	Waikato Regional Council
Peter Scott	Plant & Food Research
Nari Williams	Scion

Appendix 2: Expert elicitation questionnaire (distributed online)

We are a team of researchers at Massey University working on evaluating 2 diagnostic or detection tests for Kauri dieback: **(1) aerial surveillance** and **(2) soil sampling**. To do this, we are using Bayesian methods, which require us to collect expert opinions on likely values of tests sensitivity and specificity, as well as disease prevalence. We ask you some questions around test performance in the questionnaire to collect your "best guess" of what you think are the most likely values. We will also try to capture your uncertainty about these values. The questions relate to scenarios needed for this study and may differ from the current use and purpose of the tests.
This is the second round of the expert elicitation procedure; please make sure you have read the summary of round 1 answers and the workshop discussion. You can use this information to change your answers from round 1 if you want to.

Aerial surveillance:
Aerial surveillance is currently used only to decide which trees to sample from, but for this study, we will be assuming it is as a test to detect "true positive" trees.

Soil sampling:
For this study, we will evaluate the 8 point sampling protocol, and assume the analysis is done tree by tree.

The unit of interest is a tree and the soil around it (up to 3 times the dripline). A "true positive" status is defined as the presence of P.A. in the soil around the tree up to 3 times the dripline.

Note: the questions below refer to trees in forests where P.a. has been detected, not to forests that are believed to be free of P.a.
If you have any question regarding this research, please contact Dr Naomi Cogger N.Cogger@massey.ac.nz

Out of 100 mature Kauri trees or rickers, **WITH P.a. actually, present around its roots***:*

Q1.
What is the minimum number of trees you expect to test **POSITIVE** by soil sampling, independently of whether the tree shows symptoms? (e.g. I think P.A. will be in at least 50 of the soil samples and the laboratory procedure will detect it at least 95% of the time if present in the sample, so my answer is 50*0.95 = 47.5)

Q2.
What is the maximum number of trees you expect to test **POSITIVE** by soil sampling, independently of whether the tree shows symptoms? (e.g. I think P.A. will be present in up to 80 of the soil samples and the laboratory procedure could possibly detect it all the time if present in the sample, so my answer is 80)

Q3.
What is the most likely number of trees you expect to test **POSITIVE** by soil sampling, independently of whether the tree shows symptoms?

Q4. How confident are you that the interval defined by the minimum and maximum numbers you entered before does contain the true value? (E.g. I think the minimum is 90 and the maximum 100, and I'm 95% sure the true value is between 90 and 100)

Out of 100 mature Kauri trees or rickers, **WITH P.a. actually, present around its roots**:

Q5.
What is the minimum number of trees you expect to test **POSITIVE** by aerial surveillance, considering that this would require a tree with the presence of P.a. to show symptoms of dieback and for these symptoms to be classified as positive by the aerial surveillance? (E.g. I think at least 30% of trees with P.A. in the soil show symptoms and at least 80% of those will be seen by aerial surveillance, so my answer is 30*0.8 = 24)

Q6.
What is the maximum number of trees you expect to test **POSITIVE** by aerial surveillance, considering that this would require a tree with presence of P.a. to show symptoms of dieback and for these symptoms to be classified as positive by the aerial surveillance?(E.g. I think that no more than 70 trees with P.A. in the soil show symptoms and that aerial surveillance could potentially identify all of them, so my answer is 70)

Q7.
What is the most likely number of trees you expect to test **POSITIVE** by aerial surveillance, considering that this would require a tree with the presence of P.a. to show symptoms of dieback and for these symptoms to be classified as positive by the aerial surveillance?

Q8. How confident are you that the interval defined by the minimum and maximum numbers you entered before does contain the true value? (E.g. I think the minimum is 90 and the maximum 100, and I'm 95% sure the true value is between 90 and 100)

Out of 100 mature Kauri trees or rickers, **WITHOUT P.a. around their roots***:*

Q9.
What is the minimum number of trees you expect to test **NEGATIVE** by soil sampling, independently of whether the tree shows symptoms? (E.g. I think there can be 1 false positive in the lab, so my answer is 99)

Q10.
What is the maximum number of trees you expect to test **NEGATIVE** by soil sampling, independently of whether the tree shows symptoms? (E.g. I think there cannot be any false positive, so my answer is 100)

Q11.
What is the most likely number of trees you expect to test **NEGATIVE** by soil sampling, independently of whether the tree shows symptoms?

Q12. How confident are you that the interval defined by the minimum and maximum numbers you entered before does contain the true value? (E.g. I think the minimum is 90 and the maximum 100, and I'm 95% sure the true value is between 90 and 100)

Out of 100 mature Kauri trees or rickers, **WITHOUT P.a. around their roots***:*

Q13.
What is the minimum number of trees you expect to test **NEGATIVE** by aerial surveillance? (e.g., if you think up to 20 of these 100 trees, will show similar symptoms due to drought, age or other causes, and the aerial surveillance cannot make the difference at all, you expect at least 80 to test negative)

Q14.
What is the maximum number of trees you expect to test **NEGATIVE** by aerial surveillance? (e.g. if you think a maximum of 10 trees will show similar symptoms due to drought, age or other causes, and the aerial surveillance can actually correctly identify the real cause of symptoms up to 60% of the time - i.e. classify as positive only 40% of these 10 trees - you expect 96 to test negative)

Q15.
What is the most likely number of trees you expect to test **NEGATIVE** by aerial surveillance?

Q16. How confident are you that the interval defined by the minimum and maximum numbers you entered before does contain the true value? (E.g. I think the minimum is 90 and the maximum 100, and I'm 95% sure the true value is between 90 and 100)

Appendix 3: Proforma (aerial inspection)

For each flight:

1. Date
2. Time of the day
3. Weather
   ☐ Clear sky

   ☐ Partially cloudy (<50%, the sun not hidden)

   ☐ Mostly cloudy (>50% or sun hidden)

   ☐ Fully cloudy

   ☐ Rainy
4. Wind speed
5. Site
6. Flight altitude
7. Helicopter/aircraft model
8. Camera model

For each tree:

9. Latitude
10. Longitude
11. Age

    ☐ Ricker

    ☐ Mature tree
12. Canopy diameter (m)
13. Canopy health score 1 2 3 4 5