Quantifying the hidden costs of imperfect detection for early detection surveillance.

Alexander J. Mastin*[1], Frank van den Bosch[2], Femke van den Berg[3], and Stephen R. Parnell[1]

[1]Ecosystems and Environment Research Centre, School of Environment and Life Sciences, University of Salford, Greater Manchester, M5 4WT, UK
[2]Computational and Systems Biology, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK
[3]Fera, National Agri-Food Innovation Campus, Sand Hutton, York, YO41 1LZ, UK

*Corresponding author. Email address: a.mastin@salford.ac.uk

Abstract

The global spread of pathogens poses an increasing threat to health, ecosystems, and agriculture worldwide. As early detection of new incursions is key to effective control, new diagnostic tests which can detect pathogen presence shortly after initial infection hold great potential for detection of infection in individual hosts. However, these tests may be too expensive to be implemented at the sampling intensities required for early detection of a new epidemic at the population level. To evaluate the trade-off between earlier and/or more reliable detection and higher deployment costs, we need to consider the impacts of test performance, test cost, and pathogen epidemiology. Regarding test performance, the period before new infections can be first detected and the probability of detecting them are of particular importance. We propose a generic framework which can be easily used to evaluate a variety of different detection methods and identify important characteristics of the pathogen and the detection method to consider when planning early detection surveillance. We demonstrate the application of our method using the plant pathogen *Phytophthora ramorum* in the UK, and find that visual inspection for this pathogen is a more cost effective strategy for early detection surveillance than an early detection diagnostic test.

Introduction

Increased trade, travel, transportation and tourism resulting from globalisation have facilitated the establishment of nonendemic pests (including animals, plants, and pathogens) in new areas [Brasier, 2008, Waage and Mumford, 2008, Anderson et al., 2004, Chapman et al., 2017]. Due to the considerable impacts these can have on human, animal, plant, and ecosystem health [Vitousek et al., 1997], it is of vital importance that new invasions are detected as early as possible, thereby allowing the implementation of control strategies to eliminate the pest before it becomes unmanageable [Cunniffe et al., 2016]. However, detecting pests present at a low level in the population can require considerable surveillance resources. This problem is further compounded when the pest is not easily detectable at an early stage in the establishment process. In particular, the inability of visual inspection to detect infection during the 'presymptomatic' period prior to the development of visible disease makes early detection more challenging [Leclerc et al., 2014], even when the probability of correctly detecting infected hosts (the 'diagnostic sensitivity') and the intensity of surveillance are high. Despite this, visual detection remains the cornerstone of early detection surveillance for emerging plant and animal pathogens. Indeed, within the UK, foot and mouth disease [Anderson, 2002], bluetongue [Landeg, 2007], chalara dieback [Woodward and Boa, 2013], and ramorum disease [Lane et al., 2003,Brasier et al., 2004,Webber et al., 2010] were all first found by visual detection.

Early detection surveillance schemes need to be biologically, statistically, and economically informed in order to be effective, yet many statistical approaches

fail to account for the dynamics of pathogen spread [Parnell et al., 2017]. Our previous work has shown that the proportion of infected hosts (the 'prevalence' of infection) at the time of first detection can be estimated by accounting for the exponential growth rate of the pathogen (r) as well as the rate of sampling [Parnell et al., 2012,Parnell et al., 2015]. We have also demonstrated how the prevalence at first detection is impacted when there is a time delay before infection is first detectable (which we term the 'detection lag'). Figure 1 shows the change in the 'apparent prevalence' (i.e. the proportion of detectable hosts) over time for two different detection methods with different detection lags. A detection lag shifts the growth curve to the right by λ days - meaning that the apparent prevalence for any given true prevalence (e.g. the prevalence at time T in Figure 1) will decrease as the detection lag is increased. Since this means that infection is harder to detect, the required sampling effort to detect infection at this point, and therefore the overall sampling cost, will increase. In response to issues such as this, there has been a particular focus in recent years on the development of new molecular diagnostic tests which can detect infection in the host at an early stage. These tests have been considered key to outbreak preparedness [Royal Society, 2002], but their superior test performance characteristics come with financial costs associated with test purchase or development.

The total cost of an early detection surveillance scheme (which is a key consideration due to the long durations over which these schemes must be maintained) is therefore impacted by both the costs of applying the detection method to individual hosts and the increased sampling effort required for detection methods with longer detection lags (such as visual inspection). The threshold at which the cost of a high sampling effort outweighs the cost of a more expensive test capable of earlier detection is influenced by how quickly the pathogen is expected to spread in the population. Despite the importance of this issue to the selection of appropriate detection methods for early detection surveillance, we currently have no method of quantifying this trade-off. In the current paper, we develop a novel, generic, method to address this deficiency and demonstrate its application by quantifying the costs of using molecular tests instead of visual inspection for detection of the oomycete *Phytophthora ramorum*) on rhododendron in the UK.

Methods

Developing a generic rule of thumb

Our previous work has demonstrated how to integrate epidemiological characteristics of a pathogen with both characteristics of the detection methods used and the statistical and financial considerations associated with sampling itself [Parnell et al., 2015, Chavez et al., 2016, Parnell et al., 2012, Mastin et al., 2017], to which the reader is directed for further information. Our current method rests upon the same assumptions as this earlier work - namely, that a surveillance

programme is already in place, in which N samples are collected every Δ days, using a detection method with a sensitivity of Se. Given that Δ is sufficiently small for us to approximate sampling as a continuous process, the rate of collection of test positive samples at time t (assuming a perfect diagnostic specificity) will be $\left(Se\left(\frac{N}{\Delta}\right)q_t\right)$, where q_t is the prevalence at this time. In the early stages of invasion, we can assume that the prevalence is growing exponentially at a rate of r new infections per infection per day, but in the presence of a detection lag (which we denote as λ) the 'detectable prevalence' will be lower than this. Our previous work has demonstrated that in these cases the prevalence at first detection follows an exponential distribution, from which we can estimate any desired cumulative percentile (q^x) as follows:

$$q^x = -\ln\left(1 - \left(\frac{x}{100}\right)\right)\left(\frac{re^{r\lambda}}{Se \cdot \frac{N}{\Delta}}\right) \tag{1}$$

Equation (1) is therefore useful for situations in which we may be interested in specifying an 'acceptable upper bound' of the prevalence at first detection (q^x), with $\left(\frac{x}{100}\right)$ being an acceptable probability of reaching this prevalence or less at the time of first detection. This is also the conceptual basis of many 'absence sampling' programmes - in which case the aim is to demonstrate (with some degree of confidence) that, if present, the prevalence is lower than a given threshold.

We can also reformulate Equation (1) to estimate the sampling effort required to be $x\%$ confident of detecting infection by some fixed prevalence when using a particular detection method. Multiplying this with the per-sample 'cost' of using the chosen detection method (c_{method}), which includes both the cost of visiting the host and using the method, will give estimates of the total (variable) cost (C_{assess}) of detecting infection by q^x using that method. This gives us $C_{assess} = \left(\frac{N_{method}}{\Delta_{method}}\right)c_{method}$, where $\left(\frac{N_{method}}{\Delta_{method}}\right)$ is the rate of sampling when using the detection method under consideration. Rearranging Equation (1) and substituting the new total cost formulation gives:

$$C_{assess} = -\ln\left(1 - \left(\frac{x}{100}\right)\right)\left(\frac{re^{r\lambda_{method}}}{q^x Se_{method}}\right)c_{method} \tag{2}$$

We can use Equation (2) to quantify the relative performance of two different detection methods (each of which may have different values of Se, λ, and c) by taking the ratio of the total costs. After simplification, we obtain the following:

$$\frac{C_{assess_1}}{C_{assess_2}} = \left(\frac{\left(\frac{c_{method_1}}{Se_1}\right)}{\left(\frac{c_{method_2}}{Se_2}\right)}\right)e^{-r(\lambda_2 - \lambda_1)} \tag{3}$$

As well as indicating the relative costs of the two detection methods for early detection surveillance, we can also consider Equation (3) as a threshold. At the equivalence point (where either test would result in the same sampling costs: i.e. $\left(\frac{C_{assess_1}}{C_{assess_2}}\right) = 1$), the following holds:

$$\left(\frac{\left(\frac{c_{method_1}}{Se_1}\right)}{\left(\frac{c_{method_2}}{Se_2}\right)}\right) = e^{r(\lambda_2-\lambda_1)} \quad (4)$$

The left side of Equation (4) is the relative expected sampling cost required for one positive detection when sampling from positive cases for method 1 compared to method 2, and the right side represents the relative increase in prevalence over the detection lag period of method 2 compared to that of method 1. If the term on the right of Equation (4) is greater than that on the left, the most economically viable option would be to select detection method 1. If the contrary is true, detection method 2 should be selected.

The *Phytophthora ramorum* case study

We demonstrate how to estimate the cost ratio in Equation (3) using data on *P. ramorum* in the UK, which typically affects woody ornamental shrubs (such as rhododendron) and larch trees - with the former playing a large role in spread and the latter being of particular economic and ecological importance. We have selected this pathogen because of the availability of data on its spread and detection rather than it being a prime candidate for early detection surveillance in the UK, where it is no longer considered eradicable [Potter and Urquhart, 2017] (although early detection and eradication in sub-regions is still relevant).

In response to the emergence of *P. ramorum* as an important plant pathogen in the UK, a surveillance strategy was instigated [Forestry Commission, 2018], conducted by trained inspection teams and based on the use of visual inspection and/or lateral flow devices (LFDs). LFDs are portable, easy-to-use, immunochromatographic tests which can be applied in the field, making them potentially useful for early detection surveillance. Although *Phytophthora* genus-specific LFDs are currently used for rapid confirmation of suspicious lesions detected by visual inspection, in the current study we consider their value as a replacement for visual detection (i.e. applied to randomly selected shrubs regardless of symptoms). We consider only surveillance of rhododendron, in which symptoms such as leaf necrosis are most apparent [Harris and Webber, 2016], and assume that that the diagnostic specificity for detection of *P. ramorum* will be perfect, since all suspected positive samples will undergo laboratory confirmation.

The parameter estimates used in the current model are shown in Table 1. We estimated the exponential growth rate of *P. ramorum* in rhododendron as the mean of the range of 0.001 to 0.005 shrubs per infected shrub per day reported in a recent paper [Chavez et al., 2016]. A study of natural transmission of *P. ramorum* in rhododendron found a high level of symptom expression after 14 days [Denman et al., 2008], which we took as a plausible upper bound for the presymptomatic period(and therefore the detection lag for visual detection). We estimated the detection lag of the LFD as three days, based upon a study of detection of *P. ramorum* on rhododendon leaves using PCR and culture [Beales,

2007], and a study of LFD detection of the pathogen *Botrytis cinera* [Tomlinson et al., 2010]. We used data from a proficiency test of 16 plant health inspectors for detection of ramorum and other *Phytopthora* diseases in rhododendron (Defra project PH0439: 'Improving tools and approaches for Plant Health Inspectorate activities – detection, surveillance and monitoring') to estimate the sensitivity of visual inspection. Since these individuals were not necessarily specialists on *P. ramorum*, we assumed that a surveillance program would use the top 10 performing inspectors, and so the six lowest performing inspectors were removed from further analysis. Using isolation as a gold standard, a total of 588 correct diagnoses of suspected ramorum disease were made from the 900 positive inspector-samples (accounting for each positive sample being inspected by multiple inspectors), giving an estimated sensitivity of 0.65. The same samples were tested with a commercially available LFD (*Phytophthora* spp. ALERT-LF$^{\mathrm{TM}}$; Neogen Corporation, UK), for which 39 of the 73 positive samples were correctly identified, giving a test sensitivity estimate of 0.53.

Method validation

Because of the difficulties in comparing the costs of detection by a specified exact prevalence in the presence of stochasticity, we evaluated the performance of our method by reformulating Equation (3) to relate to the ratio of prevalences at first detection, assuming a fixed total cost. This ratio can be shown to be mathematically equivalent to the cost ratio for detection by some fixed prevalence in Equation (3) by first reformulating Equation (2) to isolate q^x and then taking the ratio of these prevalences. For each detection method, we simulated deterministic logistic growth in the apparent prevalence of *P. ramorum* using the parameter estimates in Table 1 and starting from an apparent prevalence of 1e−8 (selected as an estimate of the rhododendron population of the UK and in order to reduce left censoring of low prevalences at first detection). Supplementary Figure 1 shows the initial simulated growth in the true and apparent prevalences. For each total cost, we estimated the sample size per visit (N) as $\left(\frac{C_{assess}\Delta}{c_{method}}\right)$, assuming a sampling interval (Δ) of 28 days. We then applied the binomial theorem (see [Parnell et al., 2015, Chavez et al., 2016, Parnell et al., 2012, Mastin et al., 2017]) to the predictions of the logistic growth model to estimate the probability of detection at each consecutive sampling point. For each total cost, we ran 100,000 realisations of a sequential sampling process, using a stochastic method (described in [Mastin et al., 2017]) to determine whether each sampling resulted in detection or not - at which point, the simulation was stopped and the prevalence recorded. We then estimated the 95th percentile of these prevalences at first detection for each test and each total cost (results shown in Supplementary Figure 2), as well as the ratio of these prevalences (see Supplementary Figure 3). In order to capture the effect of random error in this ratio, we also randomly paired each individual simulated prevalence at first LFD detection with that for visual inspection and estimated the ratio. The median and the 95% probability interval (2.5th-97.5th percentiles) of these estimates

for each total cost are shown in Supplementary Figure 4.

Results

Applying the estimates in Table 1 to Equation (3), we found that the cost of using an LFD for early detection surveillance was 1.9 times higher than using visual inspection. This result was confirmed using our Monte Carlo simulation model, which found that the relative prevalence at first detection when using the LFD was consistently 1.9 times higher than that when using visual inspection, over a range of total variable sampling costs (see Supplementary Figures 2 and 3). We found a similar pattern in the individual ratio estimates, with a median ratio of $\frac{1.9}{1}$ and a 95% probability interval of around $\left(\frac{1}{20.7}\right)$ to $\left(\frac{73.0}{1}\right)$ (see Supplementary Figure 4).

We also investigated the impact of parameter uncertainty on the optimal detection method for minimising total cost, as shown in Figure 2 and Supplementary Figure 5. Figure 2 shows the effect of varying those parameters impacting upon the apparent prevalence curve (i.e. detection lag and exponential growth rate) on the x-axis, and those parameters impacting upon the cost of detecting infections (i.e. test sensitivity and detection method costs) on the y-axis, using the formulation described in Equations (3) and (4) and in the Methods above. An alternate visualisation of the same results is shown in Supplementary Figure 5, which shows the effect of varying individual epidemiological or detection parameters. In both cases, the parameter ranges for which inspection based upon visual inspection would be economically preferable are unshaded, and those for which the LFD should be used are shaded. Current parameters are shown as dotted lines. Assuming other parameters are fixed, the frontier between these two planes is reached with an epidemic growth rate of around 0.06; a sensitivity ratio of 1.54; a detection lag difference of 206 days; or a cost ratio of 0.85 (Supplementary Figure 5).

Discussion

Recent developments in molecular biology, chemistry, and immunology have resulted in the development of a wide range of new diagnostic tests which can detect infection before the development of symptoms. This information is considered highly important for mounting an effective response to epidemics [Thompson et al., 2016], and therefore the potential for earlier detection has been heralded by some as the future of disease surveillance. However, these attributes come with a cost - in particular, the direct financial cost associated with their purchase - which may make them less cost effective over the large areas and long durations required for an effective early detection surveillance system. As a result, visual detection remains the mainstay of early detection surveillance for animal and plant pathogens.

When selecting a suitable detection method for early detection surveillance,

Par	Interpretation	Value
r	Epidemic growth rate	0.0033 hosts host^{-1} day^{-1}
Se_1	Sensitivity of LFD	0.53
Se_2	Sensitivity of visual inspection	0.65
λ_1	LFD detection lag	3 days
λ_2	Visual inspection detection lag	14 days
c_{test}	Cost of LFD use (visit + test)	£10 host^{-1} + £6 host^{-1}
c_{visual}	Cost of visual inspection (visit + inspection)	£10 host^{-1} + £0 host^{-1}

Table 1: Parameter values used for the *Phytophthora ramorum* case study

we are therefore faced with the challenge of weighing the benefits associated with the earlier and/or more reliable detection achievable with new molecular tests against the lower costs (and therefore higher achievable sampling rate) when using visual detection. In doing this, we must also account for the epidemiological characteristics of the pathogen, since the relative increase in required sampling effort (and therefore cost) for a given detection lag will be greater for faster spreading pathogens. Despite the central importance of this issue to the sustainability of a surveillance system, there has been little attempt to date to quantify the value of these attributes for early detection surveillance. Our method addresses this deficiency whilst also linking directly with methods used for declaring the absence of a pathogen from a population.

To summarise the basis of our method, we assume a pathogen invades a new population at some unknown point in time and starts to spread. Given we have a surveillance programme in place during this spread (collecting N samples every Δ days), there is an $x\%$ chance that the prevalence will be less than the output of Equation (1) at the time of first detection (assuming that our detection method has a detection lag of λ and a diagnostic sensitivity of Se - although further work is needed to identify how to incorporate a changing diagnostic sensitivity over the detection lag period and beyond [DiRenzo et al., 2018]). Equation (2) allows us to estimate the expected surveillance cost for detection by any specific prevalence using any specific detection method. This focus on a specified 'maximum acceptable' prevalence is the basis of most regulatory surveillance efforts for pathogens thought to be absent from an area of interest, with the threshold prevalence either prescribed by intergovernmental standard-setting organisations or determined by consideration of the impact of the pathogen and the availability of control measures. Given that an initial evaluation has been conducted and at least one detection method under consideration has been found to be economically viable for use in surveillance, we have developed a

method of comparing the total surveillance costs of different detection methods (see Equations (3) and (4)), which can be used to select a surveillance strategy that is cost effective and sustainable for the necessary long periods of time. We note that our method does not currently explicitly account for other surveillance aims [Häsler et al., 2011, Grosbois et al., 2015], such as prevalence monitoring or model parameterisation.

Using data obtained from the literature on the epidemiology of European *P. ramorum* strains in rhododendron and on the performance of different detection methods, and assuming random sampling of hosts regardless of their expression of symptoms, we find that the costs of early detection of this pathogen at any prevalence are lower for visual inspection than for a commercially available LFD. Figure 2 shows that this conclusion is relatively robust to changes in parameter values, unless there are considerable increases in the exponential growth rate; the relative sensitivity of the LFD; or the absolute difference in detection lags. These changes could occur with the evolution of new strains (with faster growth rates and/or longer presymptomatic periods), or through improvements in the sensitivity of the LFD (although a perfect LFD sensitivity would only just reach the frontier in Figure 2). Waiting for symptom expression before using the LFD, as is generally currently used in the field, would have constrained both the detection lag and the sensitivity of the LFD to be no greater than that for visual inspection and would therefore have resulted in a higher cost ratio.

Although we have used an example of a plant pathogen in the current report, our method can be applied to any emerging pathogen or parasite, given that sampling is an ongoing process with a reasonably short sampling interval and that the pathogen is not already established in the population. Our analysis (as demonstrated in Figure 2) identifies a number of pathogen and detection method characteristics which can increase the cost effectiveness of using a molecular detection method instead of visual detection for early detection surveillance. These are listed below, along with some examples of pathogens which may be worthy of such consideration:

(1) Fast-spreading pathogens (i.e. a high exponential growth rate), such as poliovirus, foot and mouth disease virus, or *Puccinia graminis f. sp. tritici.*

(2) Considerably earlier detection than visual inspection (as may be seen with a long presymptomatic period), such as with ebolavirus, *Leptospira interrogans*, or *Candidatus* Liberibacter spp.

(3) Higher test sensitivity than visual inspection (such as when clinical symptoms are not easily identified), for example, visceral leishmaniasis caused by *Leishmania* spp, *Mycobacterium bovis* (cervical skin test vs serological test), or cassava brown streak virus.

(4) Comparable (or lower) test cost to visual inspection, such as with *Plasmodium falciparum*, *Brucella abortus* (e.g. using the Rose Bengal test), or remote sensing for *Xylella fastidiosa* (where high coverage can be achieved at comparatively lower costs).

Exploring these other applications would be valuable, as would the application of our method to more realistic spread models and real-world data.

Acknowledgements

Part of this work was conducted under the Defra-Fera long term services agreement. We would like to thank our colleagues at Fera in providing data for the *P. ramorum* case study; Kelvin Hughes, Head of the APHA Inspectorate Programme, for his insightful comments; and two anonymous reviewers for helpful comments and suggestions.

Rothamsted Research receives support of the Biotechnological and Biological Research Council of the United Kingdom. SP received funding from the USDA APHIS Farm Bill.

Authors' contributions

FVDBosch and SP conceived of the study and developed the original methodology. FVDBerg identified the case study, procured data, and ran initial analyses. AM developed the final models, ran analyses, and drafted the manuscript. All authors contributed to the manuscript and gave final approval for publication.

References

[Anderson, 2002] Anderson, I. (2002). *Foot and mouth disease 2001 : lessons to be learned inquiry report*. Stationery Office, London.

[Anderson et al., 2004] Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., and Daszak, P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, 19(10):535–544.

[Beales, 2007] Beales, P. (2007). Detection of phytophthora ramorum in watercourses. defra project report ph0317.

[Brasier, 2008] Brasier, C. (2008). The biosecurity threat to the uk and global environment from international trade in plants. *Plant Pathology*, 57(5):792–808.

[Brasier et al., 2004] Brasier, C., Denman, S., Rose, J., Kirk, S., Hughes, K., Griffin, R., Lane, C., Inman, A., and Webber, J. (2004). First report of ramorum bleeding canker on quercus falcata, caused by phytophthora ramorum. *Plant Pathology*, 53(6):804–804.

[Chapman et al., 2017] Chapman, D., Purse, B. V., Roy, H. E., and Bullock, J. M. (2017). Global trade networks determine the distribution of invasive non-native species. *Global Ecology and Biogeography*, 26(8):907–917.

[Chavez et al., 2016] Chavez, V. A., Parnell, S., and Van den Bosch, F. (2016). Monitoring invasive pathogens in plant nurseries for early-detection and to

minimise the probability of escape. *Journal of theoretical biology*, 407:290–302.

[Cunniffe et al., 2016] Cunniffe, N. J., Cobb, R. C., Meentemeyer, R. K., Rizzo, D. M., and Gilligan, C. A. (2016). Modeling when, where, and how to manage a forest epidemic, motivated by sudden oak death in california. *Proceedings of the National Academy of Sciences*, 113(20):5640–5645.

[Denman et al., 2008] Denman, S., Moralejo, E., Kirk, S., Orton, E., and Whybrow, A. (2008). Sporulation of phytophthora ramorum and p. kernoviae on asymptomatic foliage and fruit. In *In: Frankel, Susan J.; Kliejunas, John T.; Palmieri, Katharine M., tech. coords. 2008. Proceedings of the sudden oak death third science symposium. Gen. Tech. Rep. PSW-GTR-214. Albany, CA: US Department of Agriculture, Forest Service, Pacific Southwest Research Station. pp. 201-207*, volume 214.

[DiRenzo et al., 2018] DiRenzo, G. V., Campbell Grant, E. H., Longo, A. V., Che-Castaldo, C., Zamudio, K. R., and Lips, K. R. (2018). Imperfect pathogen detection from non-invasive skin swabs biases disease inference. *Methods in Ecology and Evolution*, 9(2):380–389.

[Forestry Commission, 2018] Forestry Commission (2018). Phytophthora ramorum.

[Grosbois et al., 2015] Grosbois, V., Häsler, B., Peyre, M., Hiep, D. T., and Vergne, T. (2015). A rationale to unify measurements of effectiveness for animal health surveillance. *Preventive veterinary medicine*, 120(1):70–85.

[Harris and Webber, 2016] Harris, A. and Webber, J. (2016). Sporulation potential, symptom expression and detection of phytophthora ramorum on larch needles and other foliar hosts. *Plant Pathology*, 65(9):1441–1451.

[Häsler et al., 2011] Häsler, B., Howe, K. S., and Stärk, K. D. (2011). Conceptualising the technical relationship of animal disease surveillance to intervention and mitigation as a basis for economic analysis. *BMC health services research*, 11(1):225.

[Landeg, 2007] Landeg, F. (2007). Bluetongue outbreak in the uk. *The Veterinary record*, 161(15):534.

[Lane et al., 2003] Lane, C., Beales, P., Hughes, K., Griffin, R., Munro, D., Brasier, C., and Webber, J. (2003). First outbreak of phytophthora ramorum in england, on viburnum tinus. *Plant Pathology*, 52(3):414–414.

[Leclerc et al., 2014] Leclerc, M., Doré, T., Gilligan, C. A., Lucas, P., and Filipe, J. A. (2014). Estimating the delay between host infection and disease (incubation period) and assessing its significance to the epidemiology of plant diseases. *PloS one*, 9(1):e86568.

[Mastin et al., 2017] Mastin, A. J., van den Bosch, F., Gottwald, T. R., Alonso Chavez, V., and Parnell, S. R. (2017). A method of determining where to target surveillance efforts in heterogeneous epidemiological systems. *PLOS Computational Biology*, 13(8):1–23.

[Parnell et al., 2015] Parnell, S., Gottwald, T., Cunniffe, N., Chavez, V. A., and Van Den Bosch, F. (2015). Early detection surveillance for an emerging plant pathogen: a rule of thumb to predict prevalence at first discovery. 282(1814):20151478.

[Parnell et al., 2012] Parnell, S., Gottwald, T., Gilks, W., and Van den Bosch, F. (2012). Estimating the incidence of an epidemic when it is first discovered and the design of early detection monitoring. *Journal of theoretical biology*, 305:30–36.

[Parnell et al., 2017] Parnell, S., van den Bosch, F., Gottwald, T., and Gilligan, C. A. (2017). Surveillance to inform control of emerging plant diseases: an epidemiological perspective. *Annual review of phytopathology*, 55:591–610.

[Potter and Urquhart, 2017] Potter, C. and Urquhart, J. (2017). Tree disease and pest epidemics in the anthropocene: A review of the drivers, impacts and policy responses in the uk. *Forest Policy and Economics*, 79:61–68.

[Royal Society, 2002] Royal Society (2002). *Infectious diseases in livestock*. The Royal Society, London.

[Thompson et al., 2016] Thompson, R. N., Gilligan, C. A., and Cunniffe, N. J. (2016). Detecting presymptomatic infection is necessary to forecast major epidemics in the earliest stages of infectious disease outbreaks. *PLOS Computational Biology*, 12(4):1–18.

[Tomlinson et al., 2010] Tomlinson, J., Dickinson, M., and Boonham, N. (2010). Detection of botrytis cinerea by loop-mediated isothermal amplification. *Letters in applied microbiology*, 51(6):650–657.

[Vitousek et al., 1997] Vitousek, P. M., D'antonio, C. M., Loope, L. L., Rejmanek, M., and Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology*, pages 1–16.

[Waage and Mumford, 2008] Waage, J. and Mumford, J. (2008). Agricultural biosecurity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1492):863–876.

[Webber et al., 2010] Webber, J., Mullett, M., Brasier, C., et al. (2010). Dieback and mortality of plantation japanese larch (larix kaempferi) associated with infection by phytophthora ramorum. *New Disease Reports*, 22(19):2044–0588.

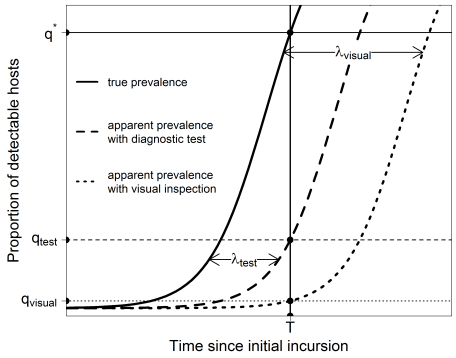


Figure 1: Effect of different detection lag periods on the apparent prevalence ('proportion of detectable hosts') at the time of first detection. Deterministic logistic growth in the true prevalence of infection (proportion of infected individuals) over time is shown in the solid line, and the 'apparent prevalences' for two detection methods (a diagnostic test and visual inspection) with different detection lag periods (λ) are shown as dashed lines. Assuming we are using visual inspection for early detection and we detect infection for the first time at time T, the apparent prevalence would be q_{visual}. However, due to the detection lag, the true prevalence is much higher - at q^*. In order to detect at a true prevalence equal to q_{visual}, the sampling effort (and therefore the cost) would have to be greatly increased. When using a diagnostic test with a shorter detection lag (λ_{test}), the apparent prevalence at time T (q_{test}) is higher, which can be achieved with a lower sampling effort.

[Woodward and Boa, 2013] Woodward, S. and Boa, E. (2013). Ash dieback in the uk: a wake-up call. *Molecular plant pathology*, 14(9):856–860.

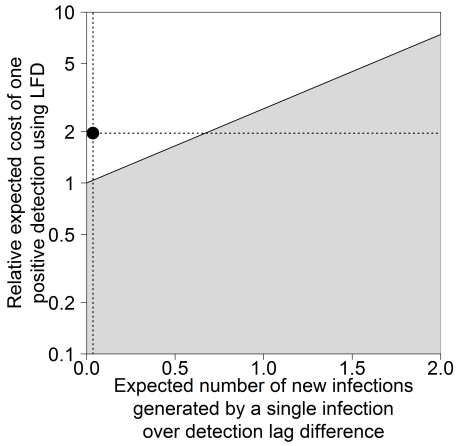
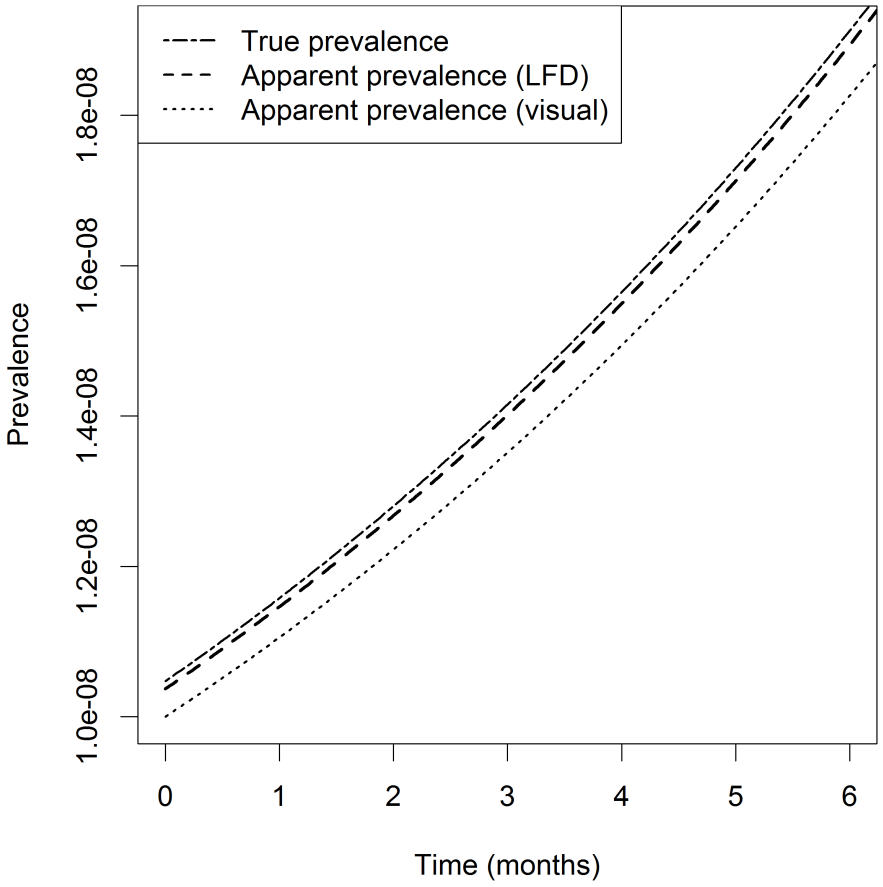
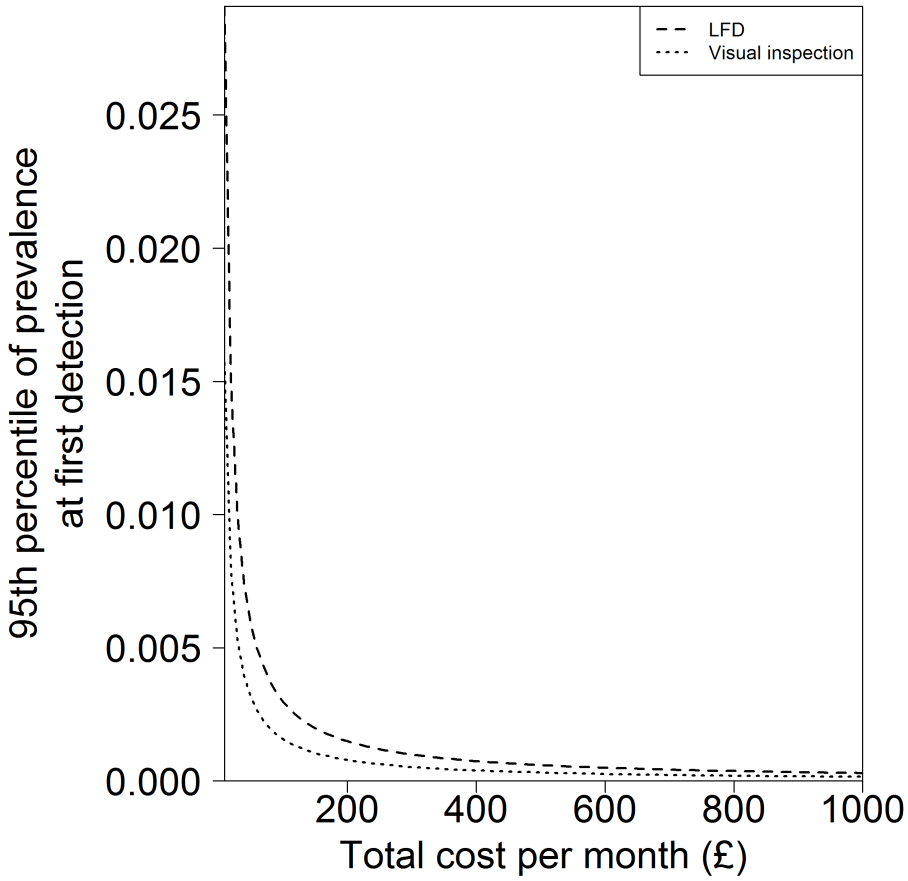


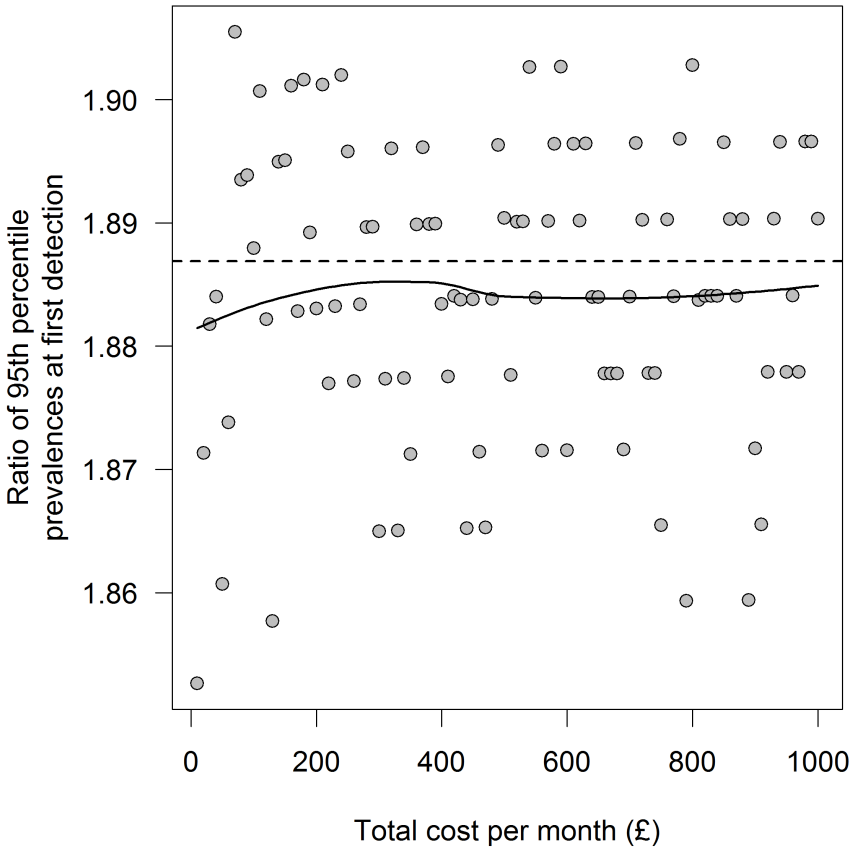
Figure 2: Effect of varying epidemiological and detection method parameters on the optimal detection strategy for early detection. We use the constructs in Equation (4) as a framework, so the x-axis represents the terms on the right side of this equation $\left(e^{r(\lambda_2-\lambda_1)}\right)$, and the y-axis represents those on the left $\left(\frac{\left(\frac{c_{method_1}}{Se_1}\right)}{\left(\frac{c_{method_2}}{Se_2}\right)}\right)$ (on a log scale, since these are ratio measurements). Higher values of r and/or a greater difference between the detection lag (assuming that the LFD lag is shorter than that for visual inspection) will be towards the right of the x-axis. On the y-axis, diagnostic methods with equal sensitivities and costs would be placed in the middle, with decreasing LFD sensitivity and/or higher costs moving towards the top of this axis and decreasing visual detection sensitivity and/or higher costs towards the bottom. The shaded area indicates parameter combinations giving a total cost ratio $\left(\frac{C_{assess_{test}}}{C_{assess_{visual}}}\right)$ of less than 1, indicating that using the LFD will minimise total costs. The unshaded area indicates where the total cost ratio is greater than 1 (where visual inspection will minimise total costs). The dotted horizontal and vertical lines indicate the values of the parameters used in the current analysis.



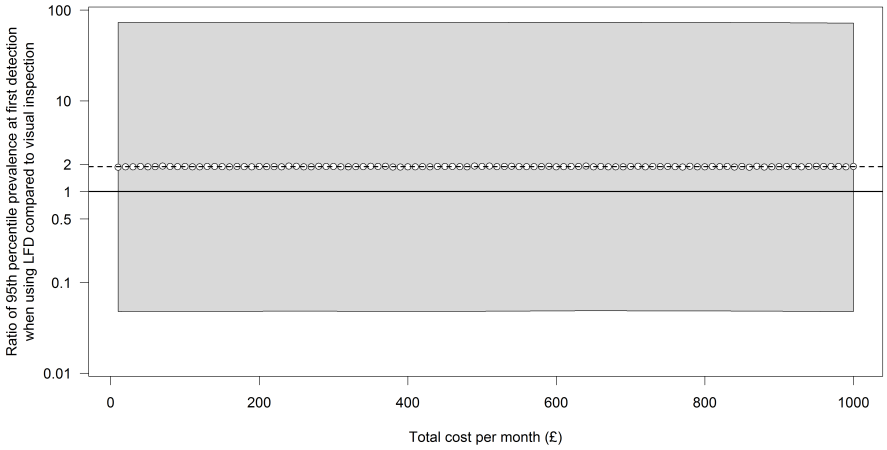
Supplementary Figure 1: Changes in the true and apparent prevalences over six months, as used in the simulation model.



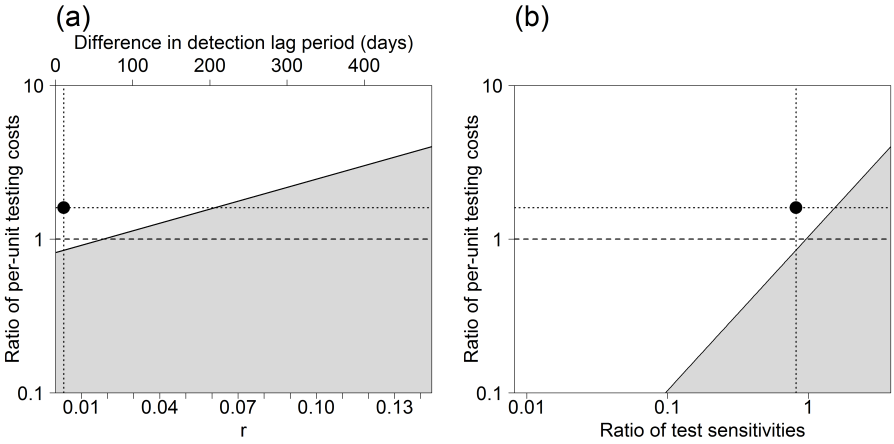
Supplementary Figure 2: 95th percentile of prevalence at first detection for LFD and visual inspection from Monte Carlo simulation $\left(P\left(Det_t\right) = 1 - \left(1 - Se \times q_t\right)^N\right)$ with 100,000 iterations, using the parameters in the main text for a range of total sampling costs between £10 and £1,000.



Supplementary Figure 3: Predicted ratio of 95th percentile prevalences at first detection when using LFD compared to using visual inspection from simulation for a range of sampling costs. Solid line estimated using locally weighted regression. The predicted ratio using the rule of thumb described in the text is shown as the dashed horizontal line.



Supplementary Figure 4: Individual-level estimates of the ratio of prevalences at first detection when using LFD compared to using visual inspection from simulation. White points show the median ratio for each total sampling cost, and the shaded area shows the range of 2.5th to 97.5th percentiles (estimated using locally weighted regression). The dashed horizontal line shows the prediction using the rule of thumb described in the text, and the solid horizontal line shows the equivalence point (a ratio of 1.0).



Supplementary Figure 5: Effect of varying epidemiological and detection method parameters on the optimal detection strategy for early detection. The x-axis represents the exponential growth rate, r, or the number of days earlier that the LFD can detect infection compared to visual inspection, $(\lambda_{test} - \lambda visual)$ (plot (a)); or the relative sensitivity of the LFD compared to visual inspection, on a log scale (plot (b)). The y-axis shows the relative per-sample cost of using the LFD compared to visual inspection, $\left(\frac{c_{test}}{c_{visual}}\right)$, on a log scale. The dashed horizontal line indicates a per-unit sampling cost ratio of 1 - whereby the cost of either sampling approaches is equal (and which also represents the minimum realistically achievable if the cost of visual inspection is zero, as we assume here). The shaded area indicates parameter combinations giving a total cost ratio $\left(\frac{C_{assess_{test}}}{C_{assess_{visual}}}\right)$ of less than 1 (i.e. using the LFD will minimise total costs); the unshaded area indicates where the total cost ratio is greater than 1 (i.e. visual inspection will minimise total costs); with the solid line representing the frontier between these two planes. The dotted horizontal and vertical lines indicate the values of the parameters used in the current analysis.