The value of volunteer surveillance for the early detection of biological invaders.

Frank van den Bosch[1*], Neil McRoberts[1], Yoann Bourhish[2], Stephen Parnell[3], Kirsty L. Hassall[2].

[1] Quantitative Biology & Epidemiology Group, Plant Pathology Department, University of California, Davis, One Shields Ave, Davis CA 95616, USA. (*Visiting Scholar)

[2] Net zero and resilient farming (YB), Applied and Research Statistics, Intelligent Data Ecosystems (KLH), Rothamsted Research, Harpenden AL5 2JQ, UK.

[3] Warwick Crop Centre, School of Life Sciences, University of Warwick, Wellesbourne Campus, Warwick, UK.

Suggested reviewers:

Hans Metz, j.a.j.**metz**@biology.**leidenuniv**.nl

Janis Dickinson, jld84@cornell.edu

John Coulston, jcoulston@fs.fed.us

Gemma Foster, gemma.foster@coventry.ac.uk

wally.**gilks**@maths.**leeds**.ac.uk

ABSTRACT.

Early detection of invaders requires finding small numbers of individuals across large landscapes. It has been argued that the only feasible way to achieve the sampling effort needed for early detection of an invader is to involve volunteer groups (citizen scientists, passive surveyors, etc.). A key concern is that volunteers may have a considerable false-positive and false-negative rate. The question then becomes whether verification of a report from a volunteer is worth the effort. This question is the topic of this paper. Since we are interested in early detection we calculate the Z% upper limit of the one sided confidence interval of the incidence (fraction infected) and use the term maximum expected plausible incidence for this.

We compare the maximum plausible incidence when the expert samples on his/her own, $\widetilde{q_E}$, and the maximum plausible incidence when the expert only verifies cases reported by the volunteer surveyor to be infected, $\widetilde{q_V}$. The maximum plausible incidences $\widetilde{q_E}$ and $\widetilde{q_V}$ are related as,

$$\widetilde{q_V} = \frac{\theta_{fp}}{1-\theta_{fn}}\,\widetilde{q_E}$$

where θ_{fp} and θ_{fn} are the false positive and false negative rate of the volunteer surveyor, respectively. We also show that the optimal monitoring programme consists of verifying only the cases reported by the volunteer surveyor if,

$$\frac{T_X}{T_N} < \frac{\theta_{fp}\ T_X}{1-\theta_{fn} T_N} < \frac{\theta_{fp}}{1-\theta_{fn}},$$

where T_N is the time needed for a sample taken by the expert and T_X is the time needed for an expert to verify a case reported by a volunteer surveyor.

Our results can be used to calculate the maximum plausible incidence of a plant disease based on reports of passive surveyors that have been verified by experts and data from experts sampling on their own. The results can also be used in the development phase of a surveillance project to assess whether including verifying passive surveyor reports is useful in the early detection of exotic invaders.

Keywords: Early detection, volunteers, surveillance, false positive, false negative, cost.

INTRODUCTION.

Early detection is a key requirement for successful eradication or containment of exotic invasive species (Ferguson et al., 2001). Early detection requires finding small numbers of individuals across large landscapes. The sampling effort and budget needed to achieve this are often well beyond the capacity of regulatory surveys. Volunteer data are frequently the first records of invading species. For example, Epanchin-Niell et al (2021) found that in the US around 25% of exotic invaders were detected by the general public and individual operators, in New Zealand the figure was even around 60%. In their review of existing and emerging tools for early detection of exotic invaders, Larson et al (2020) concluded: ''Programs that promote public participation in large-scale biodiversity identification and monitoring (such as iNaturalist and eBird) may be the best resources for early detection''. Dickinson et al (2010) argues that the only feasible way of achieving the sampling effort needed to meet the biosecurity objective of early detection is to involve volunteer groups in data collection.

A key concern with sightings of exotic invaders reported by volunteers is the quality of the data. It is to be expected that sightings by non-specialists have a considerable false-positive and false-negative rate. In their assessment of data from the iNaturalist platform, one of the most widely used citizen science platforms, Soroye et al (2022) found that poor data quality is one of the key risks in volunteer data gathering. A study on the ability of citizen scientists to identify bumblebee species, for example, showed that, depending on the observer, as few as 20% of the bumblebees were named correctly (Falk et al. 2019). For a range of amphibians it was found that the false-positive rate ranged from 0.01 to 0.09 (Cruickshank et al., 2019), although high, considerably better than the bumblebee recognition. Moreover, Given invasive pests are novel species, misidentification rates and are likely to be on the upper end of misidentification rates. This implies that a first report of an invader by a volunteer cannot be taken as conclusive proof that the invader has entered the area of interest. Verification of the sightings by an expert is essential, but comes at a cost.

The question then becomes whether verification of a report from a volunteer is worth the effort or if it is more effective when experts go directly into the field themselves to sample. What is the value of volunteer reporting for the early detection of an exotic invader if the volunteer is error-prone? That question

is the central topic of this paper. We will restrict our attention to pests and diseases of plants. In the paper we will use the terminology of an infectious plant disease, but the results hold for insect plant pests as well.

Detection surveys for invading plant pathogens proceeds in two stages:

1. Disease freedom. Surveillance is started when the pathogen is believed to not yet be present. This implies that for one or more surveillance rounds no detections are made. However, since sampling is a stochastic process it might be that the pathogen is present but missed by chance. The important question is thus, what could be the true incidence (fraction of plants infected), although still missed by chance, when no detections are made.

2. First detection. At some point in the sequence of surveillance rounds an infected plant will be found for the first time. This establishes that the invader has arrived. The question, then, is whether the surveyor found the very first case or that a considerable fraction of the plants are already infected.

3. We are thus concerned with situations in which the observations consist of cases of no detections and cases of the earliest detections of an infected tree or other plant; in other words, situations where the invasive species is not yet recognised to be invaded yet. These cases limit the contribution volunteers can make to this process. In the case of no detection or first detection it is incorrect to assume that a report by a volunteer of an infected host is an actual positive, as described above. An expert will always have to verify whether the report concerns a true positive. In that sense a volunteer can only provide information, that after the expert's verification, is redundant. From the moment the first true positive is established, so it is certain the species has invaded, further effort to delineate the outbreak and/or estimate incidence or density, information from the volunteer can be incorporated (including the appropriate methods to deal with probabilities on false positives and negatives in the estimates) without the need of verifying every single report . Moreover, for delineating an outbreak and estimating population densities volunteer reports from areas inaccessible to experts (e.g. private lands) or further afield than possible for experts to visit, are a valuable volunteer contribution.

Eradication and containment programs are very expensive and their total cost depends on the disease incidence at the start of the management programme. If initially too few resources are allocated to the eradication/containment programme the disease will escape control and the costs to get the outbreak eventually under control increase sharply (Cuthbert et al 2022). Therefore, it is of key importance to allocate enough resources when the invader is detected, implying we need to be sure that the actual incidence of the outbreak is smaller than our estimated incidence. More precisely, we are interested in the upper limit of the *Z*% one sided confidence interval of the incidence. Figure 1a illustrates this where the probability *P*, of incidence *q*, is plotted. Throughout the paper we will calculate such upper limits, $\tilde{q}$, of the incidence to be expected. We refer to this as the maximum plausible incidence. This upper limit is (figure 1a) calculated from,

$$\int_0^{\tilde{q}} P(q)\, dq = Z/100 \quad . \tag{1}$$

<figure 1 around here>

Several methods have been published about repeated sampling of populations to estimate incidence (Cameron & Baldock, 1998; Cannon 2022; Coulston et al., 2008). In these papers the disease incidence is assumed to be constant. In reality, for invading pathogens, the pathogen population and equivalently the population of infected hosts will often grow exponentially during the early period of invasion. Following the ideas developed by Metz (1983) several authors have studied the cases of disease freedom and first detection with exponential growth of the number of infected hosts (Bourhis et al., 2018; Parnell et al., 2015; Mastin et al., 2017; Bourhis et al., 2019). These authors studied the scenario in which an expert does multiple surveillance rounds, in which they assess several plants for the presence/absence of disease, and with a fixed time interval between surveillance rounds. From the data gathered, the maximum plausible incidence, $\tilde{q}$, (as defined above) is calculated.

The scenario we study in this paper is one where the expert verifies reports from the volunteer and we compare that with the scenario where experts sample for themselves without prior scouting by volunteers. We assume the expert can assess the infection status with certainty for example because they can bring samples into the laboratory and perform any diagnostics needed (also see the discussion for more details about this assumption). We derive expressions for the maximum plausible incidence, $\tilde{q}$, and compare

this maximum plausible incidence when the expert verifies volunteer reports, $\widetilde{q_V}$, with the scenario where the expert goes into the field and chooses their own hosts to assess for disease, $\widetilde{q_E}$. By this comparison, we will be able to quantify the value of volunteer reporting for the early detection of an invader. Our purpose is to derive general results which explain how these various quantities combine to determine the value of voluntary surveillance.

Our key aim is to derive simple explicit equations for the maximum plausible incidence. This will enable practitioners developing surveys to use our results without having to take recourse to extensive numerical computations for which they the need to involve a computer expert. We also aim at deriving simple equations measuring the value of volunteer reporting that, again, can directly be used by practitioners developing surveys. Therefore, we restrict, in this paper, our attention to a set of cases that does yield simple explicit expressions for the maximum plausible incidence. In the discussion, we will describe further extensions.

In the material and methods, we describe the model for sampling to establish disease freedom and first detection. These lead to the use of numerical procedures to calculate the maximum plausible incidence. To find simple explicit expressions we derive a series of approximations that yield explicit expressions and give insight into the value of volunteer surveillance for early detection. We will assess the accuracy of the approximations by comparing the maximum plausible incidence calculated from the full model and from these approximations.

Although we use 'the volunteer' and 'the expert' in the text there usually are more volunteers and experts involved. The key assumption here is that the inter-observer variation in detection skill is not taken into account (see discussion).

MATERIAL AND METHODS.

We use $\widetilde{q_E}$ and $\widetilde{q_V}$ to denote the upper limit of the confidence interval of q_0 for sampling by experts only and for verifying reports of volunteer surveyors, respectively. We use $\widetilde{q_{EV}}$ for surveys including both experts sampling on their own and validation of reports of volunteers. In the sections where approximations are

compared with exact solutions we will use $q_{\bullet exact}$ and $q_{\bullet approx}$, where · can be *E* or *V*, to denote the exact and the approximated upper limit, respectively.

Preamble:

1. *The probability for a volunteer surveyor to report a positive host:*

Disease incidence (the proportion of trees or plants, referred to generically as "hosts", in a survey area that are infected) is denoted by *q*. The probability that a volunteer surveyor observes an infected host to be uninfected, known as the false negative rate, is θ_{fn}. The probability that the volunteer surveyor observes an uninfected host to be infected, known as the false positive rate, is θ_{fp}. We denote the uninfected as 0 and the infected as 1. Table 1, the confusion matrix, summarises the probabilities.

<table 1 around here>

The probability for the volunteer surveyor to observe an infected, 1, host is,

$$p_1(disease = 1|obs = 1) = (1-q)\theta_{fp} + q(1-\theta_{fn}) \quad . \qquad (2)$$

The probability to observe an uninfected, 0, host is,

$$p_0 = (1-q)(1-\theta_{fp}) + q\,\theta_{fn} = 1 - p_1 \quad . \qquad (3)$$

After some rearrangement we see from (2) and (3) that the probability of a volunteer-reported positive detection being a false positive, k_1, is,

$$k_1 = \frac{(1-q)\theta_{fp}}{(1-q)\theta_{fp} + q(1-\theta_{fn})} \quad . \qquad (4)$$

2. *Multiple monitoring rounds.*

We assume that the epidemic is growing exponentially in time with rate *r*. This assumption is reasonable because we are only interested in small values of *q*. The incidence increases as $q(t) = q_{in}e^{rt}$, where q_{in} is the initial incidence. We want to estimate the incidence at the most recent monitoring round, q_0. At each previous monitoring round the incidence was smaller (figure 1b). We will number the monitoring rounds

starting with 0 for the most recent monitoring round. The time interval between two previous rounds i and i-1 is Δ. From the exponential growth we find that $Z_i = \left(e^{r\Delta}\right)^{-i} := \lambda^{-i}$, and thus $q_i = \lambda^{-i} q_0$. λ can be interpreted as the multiplication factor of the incidence in a Δ time step, figure 1b.

Disease freedom sampling.

1. *Regulatory survey only.*

The probability of species detection by experts from a regulatory agency is modelled to depend only on the prevalence of the pest and the number of hosts sampled. In a monitoring programme of K rounds (where the most recent round is round 0 and the first round is round K) the expert samples N_K, N_{K-1}, …., N_2, N_1, N_0 hosts. The expert concludes that none of these hosts are positive for the invasive species. We denote the number of true positives in monitoring round *i* by Y_{Ni}. When the incidence is q_i the probability of not finding any infected hosts in a sample of size N_i is $(1-q_i)^{Ni}$. Therefore, the probability of not finding any infected hosts in all *K* monitoring rounds is given by,

$$P(y_{Ni} = 0|q_0) = \prod_{i=0}^{K}(1-q_i)^{N_i}. \quad (5)$$

We will use Bayes' equation to calculate $P(q_0|y_{Ni}=0)$,

$$P(q_0|y_{Ni} = 0) = \frac{\tilde{P}(q_0)\,P(y_{Ni}=0|q_0)}{\int_0^1 \tilde{P}(q_0)\,P(y_{Ni} = 0|q_0)dq_0}. \quad (6)$$

We assume that there is no pre-existing knowledge of the incidence and thus the prior, $\tilde{P}(q_0)$, is taken as a uniform density between 0 and 1, also known as an uninformative prior (more details surrounding this choice of prior is given in the see the discussion for notes on the prior). This results in,

$$P(q_0|y_{Ni} = 0) = \frac{\prod_{i=0}^{K}(1-q_i)^{N_i}}{\int_0^1 \prod_{i=0}^{K}(1-q_i)^{N_i}dq_0}. \quad (7)$$

Using equations (1) and (7), and noting that $q_i=\lambda^{-i}q_0$, we can now numerically calculate the upper limit of the *Z*% confidence limit of q_0, $\widetilde{q_E}$. This $\widetilde{q_E}$is informally called, as discussed above, the maximum plausible incidence.

2. *Volunteer surveillance only.*

The volunteer surveyor reports x_K, x_{K-1}, …., x_2, x_1, x_0 , x_i>0, infected hosts. We denote the number of true positives in monitoring round *i* by y_{xi}. In the absence of disease all hosts reported by the volunteer surveyor are verified by the expert and found not infected, ${k_1}^{x_i}$. The probability of not finding any infected hosts in all *K* monitoring rounds is thus given by,

$$P(y_{xi} = 0|q_0) = \prod_{i=0}^{K} {k_1}^{x_i}. \quad (8)$$

Using Bayes' equation to calculate $P(q_0|y_{xi}=0)$ as above we find,

$$P(q_0|y_{xi} = 0) = \frac{P(y_{xi}=0|q_0)}{\int_0^1 P(y_{xi} = 0|q_0)dq_0} = \frac{\prod_{i=0}^{K} {k_1}^{x_i}}{\int_0^1 \prod_{i=0}^{K} {k_1}^{x_i} dq_0} . \quad (9)$$

Using equations (1) and (9), and noting that $q_i=\lambda^{-i}q_0$, we can numerically calculate the upper limit of the *Z*% confidence limit of q_0, $\widetilde{q_V}$.

3. *Combined volunteer surveillance and regulatory survey.*

In the situation where the incidence is very small, the volunteer surveyor reports x_K, x_{K-1}, …., x_2, x_1, x_0 hosts as infected and all of these are verified by the expert. On top of this the expert samples N_K, N_{K-1}, …., N_2, N_1, N_0 hosts themselves. In this case,

$$P(y_{xi} = y_{Ni} = 0|q_0) = \prod_{i=0}^{K} (1 - q_i)^{N_i} {k_1}^{x_i} . \quad (10)$$

and using Bayes' equation to calculate $P(q_0|y_i=0)$ as above we find,

$$P(q_0|y_{xi} = y_{Ni} = 0) = \frac{\prod_{i=0}^{K} (1-q_i)^{N_i} {k_1}^{x_i}}{\int_0^1 \prod_{i=0}^{K} (1-q_i)^{N_i} {k_1}^{x_i} dq_0} . \quad (11)$$

From which we can, numerically, calculate the upper limit of the *Z*% confidence limit of q_0, $\widetilde{q_{EV}}$.

First detection.

1. *Regulatory survey only.*

Following Parnell et al (2012) the expert samples N_K, N_{K-1}, …., N_2, N_1, N_0 hosts. In the survey rounds K to 1 none of the sampled hosts is infected, $(1-q_i)^{Ni}$, $i \in [K,\ldots,1]$. Only in the last round, round $i=0$, one or more sampled hosts turn out to be infected, $(1-(1-q_0)^{N0})$. We have,

$$P(y_{Ni=0}, y_{N0} \neq 0 | q_0) = \prod_{i=1}^{K}(1-q_i)^{N_i} \ (1-(1-q_i)^{N_0}) . \quad (12)$$

As in the disease freedom case we calculate $P(q_0 | \, y_{Ni=0}, y_{N0} \neq 0)$ using Bayes' equation with a uniform prior and find,

$$P(q_0 | \, y_{Ni=0}, y_{N0} \neq 0) = \frac{\prod_{i=1}^{K}(1-q_i)^{N_i} \ (1-(1-q_i)^{N_0})}{\int_0^1 \prod_{i=1}^{K}(1-q_i)^{N_i} \ (1-(1-q_i)^{N_0}) dq_0} . \quad (13)$$

Using equations (1) and (13) we can numerically calculate the upper limit of the Z% confidence limit of q_0 , $\widetilde{q_E}$.

2. *Volunteer surveillance only.*

The volunteer surveyor again reports x_K, x_{K-1}, …., x_2, x_1, x_0 cases. All reported cases in surveillance round K to 1 turn out to be not infected after the expert verifies the finds, $k_1^{x_i}$, $i=[K,1]$. In the surveillance round 0 one or more reported cases are confirmed to be infected after expert verification, $1-k_1^{x_0}$. We then get,

$$P(y_{i=0}, y_0 \neq 0 | q_0) = \prod_{i=1}^{K} {k_1}^{x_i} \ (1-{k_1}^{x_0}) . \quad (14)$$

As in the case of disease freedom we calculate $P(q_0 | \, y_{i=0}, y_0 \neq 0)$ using Bayes' equation with a uniform prior,

$$P(q_0 | \, y_{i=0}, y_0 \neq 0) = \frac{\prod_{i=1}^{K} {k_1}^{x_i} \ (1-{k_1}^{x_0})}{\int_0^1 \prod_{i=1}^{K} {k_1}^{x_i} \ (1-{k_1}^{x_0}) dq_0} . \quad (15)$$

From which we can numerically calculate the upper limit of the Z% confidence limit of q_0, $\widetilde{q_V}$.

3. *Combined volunteer surveillance and regulatory survey.*

The volunteer surveyor reports x_K, x_{K-1}, …., x_2, x_1, x_0 hosts as infected and all of these are verified by the expert. On top of this the expert samples N_K, N_{K-1}, …., N_2, N_1, N_0 hosts themself. In survey rounds K to 1 all hosts turn

out to be uninfected. In the most recent round, round 0, one or more hosts are found to be infected. We then have,

$$P(y_{xi} = y_{Ni} = 0,\ y_{x0} \neq 0 \text{ or } y_{N0} \neq 0 | q_0) = \prod_{i=1}^{K} (1-q_i)^{N_i} {k_1}^{x_i} \left(1 - (1-q_0)^{N_0} {k_1}^{x_0}\right) \quad (16)$$

Using Bayes' equation to calculate $P(q_0 | y_{xi} = y_{Ni} = 0,\ y_{x0} \neq 0 \text{ or } y_{N0} \neq 0)$ as above we find,

$$P(q_0 | y_{xi} = y_{Ni} = 0,\ y_{x0} \neq 0 \text{ or } y_{N0} \neq 0) = \frac{\prod_{i=0}^{K} (1-q_i)^{N_i} {k_1}^{x_i} \left(1 - (1-q_0)^{N_0} {k_1}^{x_0}\right)}{\int_0^1 \prod_{i=0}^{K} (1-q_i)^{N_i} {k_1}^{x_i} \left(1 - (1-q_0)^{N_0} {k_1}^{x_0}\right) dq_0} \quad (17)$$

Using equations (1) and (17) we can numerically calculate the upper limit of the Z% confidence limit of q_0, $\widetilde{q_{EV}}$.

Approximations.

Equations (7), (9), (11), (13), (15) and (17) can be approximated to give simple expressions for the *Z*% upper limit of the one sided confidence interval for q_0, the maximum plausible incidence. First, we write,

$${k_1}^{x_i} = \left(\frac{\left(1-\lambda^{-i} q_0\right)\theta_{fp}}{\left(1-\lambda^{-i} q_0\right)\theta_{fp} + \lambda^{-i} q_0 (1-\theta_{fn})}\right)^{x_i} = \left(1 - \frac{\lambda^{-i} q_0 (1-\theta_{fn})}{\left(1-\lambda^{-i} q_0\right)\theta_{fp} + \lambda^{-i} q_0 (1-\theta_{fn})}\right)^{x_i} \quad (18)$$

Since we are only interested in small values of q_0 we can write,

$$\frac{\lambda^{-i} q_0 (1-\theta_{fn})}{\left(1-\lambda^{-i} q_0\right)\theta_{fp} + \lambda^{-i} q_0 (1-\theta_{fn})} \approx \frac{(1-\theta_{fn})}{\theta_{fp}} \lambda^{-i} q_0 \quad (19)$$

and finally, using a Taylor expansion,

$$\left(1 - \frac{(1-\theta_{fn})}{\theta_{fp}} \lambda^{-i} q_0\right)^{x_i} \approx e^{-\frac{(1-\theta_{fn})}{\theta_{fp}} \lambda^{-i} x_i q_0} \quad (20)$$

Moreover since we are only interested in small values of q_0 we use, $\int_0^1 P(y_i = 0|q_0) dq_0 \approx \int_0^\infty P(y_i = 0|q_0) dq_0$ and $\int_0^1 P(y_{i=0}, y_0 \neq 0|q_0) dq_0 \approx \int_0^\infty P(y_{i=0}, y_0 \neq 0|q_0) dq_0$.

For the Disease freedom situation, equations (7), (9) and (11), we find that the probability distribution of the incidence, $P(q_0)$, is of exponential form,

$$P(q_0) = \alpha e^{-\alpha q_0} \quad (21)$$

And using equation (1) we get $\int_0^{\tilde{q}} \alpha e^{-\alpha q_0}\, dq_0 = -e^{-\alpha q_0} + 1$ and equating this with Z/100 we find the upper limit of the *Z*% confidence interval for q_0, is,

$$\widetilde{q_\bullet} < \frac{-\ln(1-\frac{Z}{100})}{\alpha} \widetilde{q_\bullet} < \frac{-\ln(1-\frac{Z}{100})}{\alpha} . \quad (22)$$

Where $\widetilde{q_\bullet}$ can be $\widetilde{q_E}$, $\widetilde{q_V}$ or $\widetilde{q_{EV}}$ depending on the case under consideration, and α is as defined in Table 2 where equation (20) is used throughout.

<Table 2 around here>

For the First Detection cases, equations (13), (15) and (17), we find that the probability distribution is a hypo-exponential density, of form,

$$P(q_0) = \frac{1}{\frac{1}{A}-\frac{1}{B}} (e^{-Aq_0} - e^{-Bq_0}). \quad (23)$$

Using equation (1) to calculate the upper limit $\tilde{q}$ does not give an explicit expression of $\tilde{q}$ in the model parameters and to obtain an approximation, we appeal to the law of large numbers and the z-score of the standard normal distribution to arrive at an approximation for $\tilde{q}$. The mean and variance of the hypo-exponential distribution are $E(q_0) = \frac{1}{A} + \frac{1}{B}$ and $r(q_0) = \frac{1}{A^2} + \frac{1}{B^2}$, respectively. Now assume that for a large number of samples, the hypo-exponential density can be approximated by a normal density. Then the z-score, $\tilde{Z}$, is,

$$\tilde{Z} = \frac{q_0 - \left(\frac{1}{A}+\frac{1}{B}\right)}{\sqrt{\frac{1}{A^2}+\frac{1}{B^2}}} . \quad (24)$$

Which for the 95% tail $\tilde{Z}$=1.64, for the 99% tail $\tilde{Z}$=2.33. Solving for $\tilde{q}$ we find,

$$\widetilde{q_\bullet} = \frac{1}{A} + \frac{1}{B} + \tilde{Z}\sqrt{\frac{1}{A^2} + \frac{1}{B^2}} . \quad (25)$$

Where $\widetilde{q_\bullet}$ can be $\widetilde{q_E}$, $\widetilde{q_V}$ or $\widetilde{q_{EV}}$ depending on the case under consideration, and A and B are defined in Table 2.

The accuracy of the approximations.

We will calculate for a range of epidemic and surveillance parameters the upper limit of the confidence interval for q_0, for the distributions (7), (9), (11), (13), (15) and (17), $\widetilde{q_{\bullet-exact}}$ and for their approximating distributions, $\widetilde{q_{\bullet-approx}}$, given in Table 2. The relative difference between the two tells us about the accuracy of the approximation. For this analysis we need realistic values of the epidemic growth rate of plant diseases. Table 3 summarises the growth rate of six tree diseases, some from natural systems and some from production orchard systems. The graphs to assess the accuracy of the approximations will be made for a pathogen with a large epidemic growth rate, citrus canker, and for one with a small epidemic growth rate, ash dieback.

<Table 3 around here>

Time budgets of the expert and volunteer surveillance.

Money and time are key constraints in monitoring programmes. It may take experts less time to sample a host themselves than to verify a report from a volunteer surveyor, for instance because of the time requirements to transfer the information from the volunteer surveyor to the expert and for the expert to verify that the validation survey is located correctly. In other cases it may take less time to verify a report, for example when sufficiently clear photographic material is available. In this case, however, the expert needs to trust the volunteer that the photo was taken where the volunteer says it was taken. Here, we assume the expert has in total T time units to do the work. To sample one host themselves an expert takes T_N time units, while to verify a reported plant it takes T_X time units. Then,

$$T = T_N N + T_X X \quad (26)$$

Which is the same as,

328 $$X = \frac{T}{T_X} - \frac{T_N}{T_X} N \quad . \qquad (27)$$

329 Now consider the probability that the incidence of the disease is smaller than a value *q**, which is given by,

330 $$P(q^*) = \int_0^{q^*} P(q_0)\, dq \quad . \qquad (28)$$

331 q* can for example be a threshold incidence below which the disease can still be controlled. Obviously with
332 the monitoring programme, one wants to maximise the probability *P*(*q**) that the incidence is below this *q**.
333 Equation (28), which is a function of *N* and *X*, allows us to plot contour lines of equal value of *P*(*q**) in the *N*-
334 *X* plane (see figure 5). By superimposing the time constraint (27) on that plot it is possible to identify the
335 conditions under which it is time effective to verify reports of volunteer surveyors.

336

337 **RESULTS.**

338 **Approximations.**

339 Table 2 summarises the approximations to the upper limit of the *Z*% confidence interval of q_0, which we
340 termed the maximum plausible incidence. Given that the false-positive and false-negative rates of the
341 volunteer surveyor are known, these equations enable us to calculate the maximum plausible incidence, both
342 in the case of disease freedom and in the case of first detection. With this information, we can address the
343 question of the value for experts to verify reports of volunteer surveyors, instead of sampling themselves. If
344 in both cases the expert samples/verifies N_i trees, so N_i=x_i we see from table 2 that in both the situation
345 where the disease is absent and for the first detection case,

346

$$\widetilde{q_V} = \frac{\theta_{fp}}{1-\theta_{fn}}\ \widetilde{q_E}. \qquad (29)$$

347

348 Thus, the maximum plausible incidence becomes smaller or larger by a factor of $\theta_{fp}/(1-\theta_{fn})$ when the
349 experts verify reports of volunteer surveyors, than when the experts sample on their own. Figure 2a shows
350 lines of equal value of this factor as a function of the false-positive and the false-negative rate. We note that
351 θ_{fp} is also known as the false positive proportion, FPP, and, 1-θ_{fn} is also known as the true positive proportion,
352 TPP. The FPP/TPP ratio measures the value of volunteer surveillance.

353

354 **The accuracy of the approximations.**

355 The accuracy of the approximation of $\tilde{q}$ was quantified by ,

356 $$\frac{q_{\bullet\leftarrow\widetilde{exact}} - q_{\bullet\leftarrow\widetilde{approx}}}{q_{\bullet\leftarrow\widetilde{exact}}} \quad . \qquad (30)$$

357 Where $q_{\bullet\leftarrow\widetilde{exact}}$, is the upper limit of the *Z*% confidence interval for q_0 in the full model and $q_{\bullet\leftarrow\widetilde{approx}}$ is the

358 upper limit calculated for the approximation. The accuracy of the approximations $\tilde{q}$ for experts sampling on

359 their own has been studied (Parnell et al., 2015; Mastin et al., 2017). Therefore, we study the accuracy of the

360 scenario where experts verify the reports of volunteer surveyors only. Figure 3 shows the results of the

361 analysis. Clearly, both the approximation for the disease freedom case and for the first detection case are

362 more accurate for smaller epidemic growth rates, for shorter time intervals between samples, and for larger

363 sample sizes. The approximations are however surprisingly accurate. Even for survey intervals of 3 months,

364 for samples larger than, the difference between the approximation and the full model is less than 5%. For

365 samples larger than around 15 the difference is less than 10%.

366 <figure 3 around here>

367

368 **Accuracy of the $\frac{\theta_{fp}}{1-\theta_{fn}}$ ratio.**

369 The $\frac{\theta_{fp}}{1-\theta_{fn}}$ ratio quantifying the gain of involving volunteer surveillance into a programme to detect exotic

370 invaders is derived from the approximations. To see whether this ratio is also a good description of the gain

371 of involving volunteer surveyors when the full models are used, we calculated the upper limit of the

372 confidence intervals of q_0, for the full model of the expert sampling on their own, $\widetilde{q_E}$, and the full model for

373 the expert verifying reports of the volunteer surveyor, $\widetilde{q_V}$. The ratio of these two, $\widetilde{q_V}/\widetilde{q_E}$, was compared

374 with the $\frac{\theta_{fp}}{1-\theta_{fn}}$ ratio. Figure 4 shows the results of this analysis. As with the accuracy of the approximation of

375 the upper limit of the confidence interval for q_0, the $\frac{\theta_{fp}}{1-\theta_{fn}}$ ratio is less than 5% different from $\widetilde{q_V}/\widetilde{q_E}$ when

more than 35 samples are taken. The ratio is less than 10% different from $\frac{\widetilde{q_V}}{\widetilde{q_E}}$ when more than 20 samples are taken for less than 100 days between samples.

<Figure 4 around here>

Difference between disease freedom and first detection.

Figure 2b shows the maximum plausible incidence in the case of disease freedom sampling and in the case of first detection for the disease with a small epidemic growth rate (Ash dieback) and for the disease with a large epidemic growth rate (Citrus canker). The figure shows that the estimated incidence in the case of first detection is between 1.5 (for low epidemic growth rate) and 2.5 (for high epidemic growth rate) times the incidence in the case of disease freedom.

Time budgets of the expert and volunteer surveillance.

Disease freedom.

From (28) we find,

$$P(q^*) = \int_0^{q^*} Ae^{-Aq}\, dq = 1 - e^{-Aq^*}, \qquad (31)$$

and, with A given in Table 2, solving for *x* we get,

.

$$x = \frac{1}{\frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i}} \left(\frac{-1}{q^*} \ln(1 - P(q^*)) - N \sum_{i=0}^{K} \lambda^{-i} \right) \qquad (32)$$

Equation (32) is a straight lines in the x-N plane, with intercept $\frac{\frac{-1}{q^*}\ln(1-P(q^*))}{\frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i}}$ and slope $-\frac{\sum_{i=0}^{K} \lambda^{-i}}{\frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i}} = -\frac{\theta_{fp}}{(1-\theta_{fn})}$.

Figure 5 shows both (27) in orange and (32) in blue for different values of *P*. Maximising P maximises the probability that the epidemic has an incidence at or below q*, which can be taken as a threshold value below

which the epidemic can still be controlled. From the graphs we conclude that the optimal surveillance programme is to

(i) Survey by the expert only if $\frac{T_N}{T_X} < \frac{\theta_{fp}}{(1-\theta_{fn})}$

(ii) Verify volunteer surveyor reports only if $\frac{T_N}{T_X} > \frac{\theta_{fp}}{(1-\theta_{fn})}$ (33)

Readers should refer to the contour line values in Fig 2's left panel to get a sense of the values of the right hand side of these inequalities. Large values of T_N/T_X indicate longer verification time and large values of the right hand side indicate large error rates. This implies that error rates need to be quite low for very time intensive verification to be worthwhile.

<Figure 5 around here>

First detection.

From (28) we get,

$$P(q^*) = 1 + \frac{1}{\frac{1}{A} - \frac{1}{B}} \left(\frac{1}{B} e^{-Bq^*} - \frac{1}{A} e^{-Aq^*} \right), \quad (34)$$

where $A = \sum_{i=1}^{K} \lambda^{-i} N_i + \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=1}^{K} \lambda^{-i} x_i$ and

$$B = \sum_{i=0}^{K} \lambda^{-i} N_i + \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i} x_i .$$

In this case it is not possible to express *X* as function of *N* and model parameters. Contour lines were drawn numerically. An example is given in figure 5. The contour lines for equal P from (34) are virtually indistinguishable from straight lines (supplementary materials *I* gives a large set of graphs showing the generality of this statement). Moreover, the slope of the lines is virtually indistinguishable from $-\frac{\theta_{fp}}{(1-\theta_{fn})}$. This implies that in practice the same conclusion is reached for the case of first detection as that derived for the disease freedom case.

DISCUSSION.

In this paper we developed a method to include volunteer surveillance in efforts for the early detection of exotic invaders. Explicit equations relating the maximum plausible incidence to model parameters and number of volunteer reports were derived. These equations can be used by non-experts in the development of surveillance programmes to determine if volunteer data would be cost effective for a given species. We also quantified the value of volunteer surveillance and derived an expression showing how the ratio of time for an expert to sample a host and the time for an expert to validate a volunteer report, compared with the false positive and false negative rate of the volunteer, determines whether volunteer reports should be validated or left outside the regulatory survey. Volunteer surveillance accumulates potentially valuable datasets for research and outbreak response (Encarnacao et al., 20121). False-positive and false-negative observations are, however, a concern about the usefulness of the data. Using statistical techniques, it is possible to estimate false-positive and false-negative rates and correct for them as shown by Palmer et al. (2017), Brown et al. (2017) and Cruickshank et al (2019). All three of these examples use volunteer surveyor data, calculate a measure of the false-positive and false negative rate and in using the data include the measured error rates in the calculation of density and trend. The present case of early detection of invading exotic species is different in that a reported observation of an exotic invader cannot lead automatically to the assumption of the presence of the invader. The reporting will always need to be verified by an expert. The question thus is what the value is of volunteer surveillance reports. Should they be used as a preselection of sites/trees to be visited or is that not an effective use of the expert's time?

We have assumed that the expert has a zero or negligibly small false positive and false negative rate. For plant diseases this is often a valid assumption. The development of molecular diagnostics, culturing techniques, etc. is routine and the expert can take samples to the laboratory for diagnosis as needed. Nevertheless there will be cases where the assumption of zero false positive and negative rate is too crude an assumption. In that case the calculations become more involved as each possible series of correct positive, correct negative, false positive and false negative rate the volunteer reports needs to be taken into account since for each report there is a non-zero probability that an error-prone expert classifies it as positive. This

448 case does not lend itself for the derivation of an explicit expression for the maximum plausible incidence and
449 thus is not examined here.
450
451 We made a range of other assumptions that could be explored using the methods developed in the paper.
452 Inter-volunteer variation in false positive and false negative rates is an important aspects that we did not
453 include, again for the reason we did not manage to derive a simple explicit relationship between incidence
454 and survey results. We assumed the host population to have a constant exponential growth rate. Population
455 growth rates may vary substantially, particularly in early phases of invasion, due to stochastic environmental
456 or demographic effects, for example. We also assumed that q scales constantly with the population density,
457 which is a simplification because species may have variable false positive and negative rates the change as a
458 function of population density. What the effects of these factors is on the estimation of the maximum
459 plausible incidence is not clear at first sight and will be the subject of future work.
460
461 In previous work which was concerned with surveys by experts we also developed the work in two phases.
462 Parnell et al (2012) developed the simple rule of thumb that the expected incidence at first detection is the
463 population's intrinsic growth rate divided by the sampling rate. This, surprisingly simple equation was
464 subsequently, Parnell et al (2015), tested against a spatially explicit stochastic epidemiological model and it
465 was tested against a data set. It was shown that the simple equation performed well against model and data.
466 We envisage the same further work including variability in volunteer skills, variations in epidemic growth
467 rates and false positive and false negative rates of the experts involved to examine the robustness of the
468 simple approximations derived in the current work. Only after it has been shown that these simple
469 expressions give accurate estimates as compared with more elaborate, realistic models can the simple
470 expressions be widely used in practice.
471
472 We assumed that the expert will only verify reports of volunteers where they found an infected host.
473 Plausibly, the more common error among volunteers early in an invasion will be for false negatives, and it
474 may be worthwhile to verify some putative negative samples. However, negatives are seldom reported by

475 completely amateur volunteers. In some well trained volunteer groups, for example those associated with
476 horticultural societies, taxonomy groups, or state-run programs it may be possible to have volunteers report
477 when they did not find anything. Thus, the number of these negative reports is usually so small that we
478 ignored them in the present case.
479
480 We have used the homogeneous density as our prior. As motivated in Dixon (2005) a Beta -distribution may
481 be a more appropriate prior than a homogeneous density. Dixon was able to specify such a prior because,
482 although the species under consideration was very rare, it was observed on a number of occasions. In the
483 current work we used an uninformative prior in the derivation of the results since it made the analytical
484 solutions tractable. The question then arises as to whether the results are robust to the substitution of
485 informative priors. We do not have any prior information as the species has not been detected yet and thus
486 need a prior that is uninformative. Dixon (2005) suggested the use of a Beta distribution with both
487 parameters a little larger than 1 as a suitable uninformative prior. This places the prior close to the uniform
488 distribution (in the interval [0,1]), which corresponds to a Beta distribution with both parameters equal to 1.
489 Analytically, using the Beta distribution as a prior in the case where experts verify volunteer reports
490 (equations 8 and 9) we find the posterior distribution becomes ,

491
$$P(q_0|y_{xi}=0)=\frac{\frac{q_0^{\alpha-1}(1-q_0)^{\beta-1}}{B(\alpha,\beta)}\prod_{i=0}^{K}{k_1}^{x_i}}{\int_0^1\frac{q_0^{\alpha-1}(1-q_0)^{\beta-1}}{B(\alpha,\beta)}\prod_{i=0}^{K}{k_1}^{x_i}\,dq_0}\quad,\qquad(35).$$

492 Where B(α,β) is the beta-function and α and β are the parameters of the Beta distribution. It is noted that by
493 using a Beta prior distribution, certain special cases can be derived explicitly. For example, when there is a
494 single monitoring round in the disease freedom case of regulatory surveyor only, the posterior distribution
495 for q_0 is itself a Beta distribution with parameters α and β + N_0. However, upon extensions to incorporate
496 volunteer surveillance, multiple monitoring rounds or first case detection, such conjugacy is lost as in
497 equation (35) above. Thus, the use of the Beta prior lends itself to numerical calculations only. In general,
498 early in an invasive epidemic when disease incidence is low, the appropriate prior will be characterised with
499 α << β, resulting in a heavily right-skewed distribution with the majority of the probability density covering

500 an interval close to 0. Choice of the specific parameter values in particular cases may rely on expert opinion
501 and numerical analysis in such cases would allow for the sensitivity in predicted outcomes in the surveillance
502 effort to differences in the choice of parameter values to be examined. Exploration of that topic lies outside
503 the scope of this paper.
504
505 We have shown that the maximum plausible incidence of the disease when volunteer surveillance reports
506 are verified is a factor $\frac{\theta_{fp}}{1-\theta_{fn}}$ smaller (or larger) than the maximum plausible incidence when the expert
507 samples on their own. Given that both the false-positive and the false negative rates are small, including
508 volunteer surveyor into surveillance programmes can potentially be of great benefit. There is, however, a
509 possibility that including volunteer surveyors has a negative effect. When the false-negative rate is large, the
510 factor $\frac{\theta_{fp}}{1-\theta_{fn}}$ can be bigger than 1 (Figure 2a). It is not entirely clear whether that will happen in practice. If,
511 for example, the false positive rate is 0.2, as in the amphibian example (Cruickshank et al., 2019), the false
512 negative rate needs to be close to 0.8 before the $\frac{\theta_{fp}}{1-\theta_{fn}}$ ratio becomes larger than 1, which seems prima facie
513 unlikely. It is much more to be expected that false positive and false negative rates are smaller than 0.5; the
514 equivalent of flipping a coin. In that case the gain from including volunteer surveyors into surveillance
515 programmes for the early detection of exotic invaders will always be positive. This is a useful result since
516 doing better than coin flipping in assigning infected/infested status is the mildest minimum capability
517 criterion one could imagine for this type of activity and performance far in excess of this is likely to be a
518 requirement in any practical situation.
519
520 We have developed a range of approximations on the basis of which the maximum plausible incidence can
521 be calculated when the false-positive and false-negative rates are known.
522 For both types of calculations we need an estimate of the epidemic growth rate, *r*, and of the false-negative
523 and false-positive rates for the volunteer surveyor. For invading pathogens, the epidemic growth rate is not
524 normally known. In such cases information on past invasions and/or invasions at other places can be used

525 together with mechanistic insight into the effect of the difference in the environments is likely to have on
526 differences in epidemic growth rate (Parnell et al., 2015; Gottwald 2010) to produce estimates of *r*
527 heuristically.

528 Estimating false-positive and false-negative rates has been done in some recent cases (Falk et al.,
529 2019; Cruickshank et al., 2019; Brown et al., 2017). Spatial resampling techniques have been used in
530 ecology to approximate false positive and false negative rates of surveys of endemic species (Banks-Leite et
531 al 2014; Welsh et al 2013; Sólymos et al 2013). These models adjust for imperfect detection. They are
532 reviewed by Banks-Leite et al (2014). Since the false positive and false negative rates of volunteers cannot
533 be estimated in areas where the invader has not arrived yet, the case we are considering here, the
534 volunteers need to go to an area or country where the invader has established to use these approaches.
535 Other approaches could be that volunteer surveyors assess hosts, samples of hosts or photo material that
536 has also been assessed by experts. The expert assessment then can be used as the gold standard and the
537 false-positive and false negative rates of volunteer surveyors estimated. The need for expert assessment is
538 often the most costly part of the exercise. It would be worthwhile to investigate whether a technique to
539 estimate false positive and false negative rates for diagnostic tests, the latent class analysis (Turechek et al.,
540 2013), can be used in this case as well. For that analysis no gold standard is needed. A group of volunteers
541 is asked to assess the disease status of a group of hosts, the technique then both separates the hosts into
542 an uninfected and an infected group as well as estimating the false positive and false negative rates of each
543 of the volunteers.
544
545 Several authors (Parnell et al., 2015; Bourhis et al., 2019) have assessed the accuracy of the approximations
546 for the plausible mean incidence and the maximum plausible incidence in the cases where the expert samples
547 on their own. Here we have quantified the accuracy for the cases where experts verify reports of volunteer
548 surveyors (figure 3). In both cases it was shown that for the range of epidemic growth rates observed in
549 reality, (i.e. values of *r* between 0.002 and 0.02 per day) the approximations deviated less than 5% from the
550 full model when the number of samples assessed was larger than 50. The approximations deviated from the
551 full model by less than 10% when the number of samples exceeded 25. We conclude that the approximations

552 are accurate enough to be useful in a practical situation where other stochastic factors are likely to add
553 uncertainty to the detection process. The approximations resulted in equations relating the maximum
554 plausible incidence with the model parameters and survey results. This enables non-modelling-specialists to
555 use them in the development of surveillance programmes and in the evaluation of survey data sets. The
556 explicit relationship between the ratio of time needed for an expert to sample a host and the time needed
557 to verify a volunteer report compared with the false positive and false negative rate can help decide whether
558 including volunteer reporting in regulatory surveys is worth the effort. Parnell and Bourhis arrived at very
559 similar conclusions for the approximations to methods where the experts sample on their own.
560
561 Finally, we investigated whether verifying volunteer surveyor reports is time effective or whether the expert
562 going into the field on their own to sample hosts is the more time effective method. We have shown a very
563 simple rule for when reports of volunteer surveyors should be verified. This rule say that if the ratio of the
564 time an expert needs to sample a host themselves and the time needed to verify a report of a volunteer
565 surveyor and is larger than the factor $\frac{\theta_{fp}}{1-\theta_{fn}}$, the most time effective method is to dedicate experts' time only
566 to verification of the work of volunteer surveyors. This gives a clear criterion for when verifying reports by
567 volunteer surveyors should be included in the development of regulatory surveillance programmes.
568
569
570
571
572 **AUTHOR CONTRIBUTIONS:**

573 Frank van den Bosch, Kirsty Hassall and Neil McRoberts conceived the ideas and designed methodology;
574 Yoann Bourhis and Stephen Parnell developed programmes and made the graphs; Frank van den Bosch and
575 Neil McRoberts led the writing of the manuscript. All authors contributed critically to the drafts and gave
576 final approval for publication.
577

ACKNOWLEDGEMENTS:

FvdB and NMcR were funded by the California Department of Food and Agriculture, project title "Research Pest Detection Efficiencies", agreement number "20-0850-000-SA." YB and KLH were funded by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme grant Smart Crop Protection (SCP) grant number BBS/OS/CP/000001. SP received support from the UK government Department for Environment, Food and Rural Affairs funded "future proofing plant health" project (TH42222FR09).

DECLARATION OF INTEREST:

The authors have no conflicts of interest.

DATA AVAILABILITY STATEMENT:

We have created an online repository, https://gitlab.com/Yo-B/volunteer-surveillance-jtb-code-and-data, in which we have uploaded the data necessary to reproduce our figures, as well as the numerical integration code allowing one to reproduce our computations and regenerate the data. However, those computations are long to run, that is why we have uploaded their output results so that plots can be reproduced without wait. We have added reference to that repository at the end of the manuscript."

SUPPLEMENTARY MATERIALS

The R code needed to reproduce our results and figures is available at https://gitlab.com/Yo-B/volunteer-surveillance-jtb-code-and-data under GNUGPL3 licence.

REFERENCES.

1. Alonso Chavez, V., Parnell, S. & 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. (doi:10.1016/j.jtbi.2016.07.041).
2. Banks-Leite, C., Pardini, R., Boscolo, D., Cassano, C.R., Püttker, T., Santos Barros, C. & Barlow. J. (2014) Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. J. Appl. Ecol. 51, 849-859.
3. Bourhis, Y., Gottwald, T.R. & van den Bosch, F. (2018) Translating surveillance data into incidence estimates. Philosophical *Transactions of the Royal Society B Biological Sciences 374*(1776):20180262. doi: 10.1098/rstb.2018.0262.
4. Bourhis, Y., Gottwald, T.R., Lopez-Ruiz, F.J., Patarapuwadol, S. & van den Bosch, F. (2019) Sampling for disease absence-deriving informed monitoring from epidemic traits. *Journal of Theoretical Biology 461*, 8-16. doi: 10.1016/j.jtbi.2018.10.038.
5. Brown, N., van den Bosch, F., Parnell, S. & Denman, S. (2017) Integrating regulatory surveys and citizen science to map outbreaks of forest diseases: acute oak decline in England and Wales. *Proc. R. Soc. B 284*: 20170547. http://dx.doi.org/10.1098/rspb.2017.0547.
6. Cameron, A.R. & Baldock, F.C. (1998) Two-stage sampling in surveys to substantiate freedom from disease. *Prev. Vet. Med. 34* (1), 19–30. doi: 10.1016/ S0167-5877(97)0 0 073-1.
7. Cannon, R.M. (2002) Demonstrating disease freedom-combining confidence levels. *Prev. Vet. Med. 52* (3), 227–249. doi: 10.1016/S0167-5877(01)00262-8.
8. Coulston, J.W., Koch, F.H., Smith, W.D. & Sapio, F.J. (2008) Invasive forest pest surveillance: survey development and reliability. *Can. J. For. Res. 38* (9), 2422–2433. doi: 10.1139/X08-076.
9. Cruickshank, S.S., Bühler, C. & Schmidt, B.R. (2019) Quantifying data quality in a citizen science monitoring program: False negatives, false positives and occupancy trends. *Conservation Science and Practice.* 2019;e54. https://doi.org/10.1111/csp2.54.

10. Cuthbert, R. N., Diagne, C., Hudgins, E. J., Turbelin, A., Ahmed, D. A., Albert, C., ... & Courchamp, F. (2022). Biological invasion costs reveal insufficient proactive management worldwide. Science of the Total Environment, 819, 153404.
11. Dickinson, J.L., Zuckerberg, B. & Bonter, D.N. (2010) Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution and Systematics 41*, 149-172. http://doi.org/10.1146/annurev-ecolsys-102209-144636.
12. Dixon, P.J., Ellison, A.M. , Gotelli, N.J. (2005) Improving the precision of estimates of the frequency of rare events. Ecology, 86(5) 1114–1123.
13. Encarnacao, J., Teodosio, M.A. & Morais, P. (2021) Citizen science and biological invasions: a review. *Frontiers in Environmental Science* 8:602980. doi: 10.3389/fenvs.2020.602980.
14. Epanchin-Niell, R., Thompson, A. L. & Treakle, T. (2021). Public contributions to early detection of new invasive pests. Conservation Science and Practice, 3(6), e422.
15.
16. Falk, S., Foster, G., Comont, R., Conroy, J., Bostock, H., Salisbury, A., Kilbey, D., Bennett, J. & Smith B. (2019) Evaluating the ability of citizen scientists to identify bumblebee (*Bombus*) species. *PLoS ONE 14*(6), e0218614. https://doi.org/10.1371/journal.pone.0218614.
17. Ferguson, N., Donnelly, C. & Anderson, R. (2001) Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. *Nature 413*, 542–548. http://dx.doi.org/10.1038/35097116.
18. Gottwald, T.R. (2010) Current epidemiological understanding of citrus huanglongbing. *Annual Review of Phytopathology 48*, 19–139. http://doi.org/10.1146/annurev-phyto-073009-114418.
19. Larson, E. R., Graham, B. M., Achury, R., Coon, J. J., Daniels, M. K., Gambrell, D. K. & Suarez, A. V. (2020). From eDNA to citizen science: emerging tools for the early detection of invasive species. Frontiers in Ecology and the Environment, 18(4), 194-202.
20.

21. Mastin, A., van den Bosch, F., Bourhis, Y. & Parnell, S. (2022) Out of sight: Surveillance strategies for emerging vectored plant pathogens. *Scientific Reports, in press*. BioRxiv repository (doi: https://doi.org/10.1101/2022.01.21.477248).
22. Mastin, A., van den Bosch, F., Gottwald, T.R., Alonso Chavez, V. & Parnell, S.R. (2017) A method of determining where to target surveillance efforts in heterogeneous epidemiological systems. *PLoS Comput Biol. 28*, 13(8):e1005712. https://doi.org/10.1371/journal.pcbi.1005712.
23. Metz, J.A.J., Wedel, W. & Angulo, A.F. (1983) Discovering an Epidemic before it has Reached a Certain Level of Prevalence. *Biometrics 39*, 765-770.
24.
25. Palmer, J.R.B., Oltra, A., Collantes, F., Delgado, J.A., Lucientes, J., Delacour, S., Bengoa, M., Eritja, R. & Bartumeus, F. (2017) Citizen science provides a reliable and scalable tool to track disease-carrying mosquitoes. *Nature Communications 8,* 916 doi:10.1038/s41467-017-00914-9.
26. Parnell, S., Gottwald, T.R., Cunniffe, N.J., Alonso Chavez, V. & van den Bosch, F. (2015) Early detection surveillance for an emerging plant pathogen: a rule of thumb to predict prevalence at first discovery. *Proc. R. Soc. B 282*: 20151478. http://dx.doi.org/10.1098/rspb.2015.1478
27.
28. Reynolds, G.J., Gordon, T.R. & McRoberts, N. (2019) Projecting monterery pine (Pinus radiate) populations over time in the presence of various representations of pitch canker disease, caused by Fusarium circinatum. *Forest 10,* 437. doi:10.3390/f10050437.
29. Soroye, P., Edwards, B. P., Buxton, R. T., Ethier, J. P., Frempong-Manso, A., Keefe, H. E., ... & Bennett, J. R. (2022). The risks and rewards of community science for threatened species monitoring. Conservation Science and Practice, e12788.
30. Turechek, W.W., Webster, C.G., Duan, J., Roberts, P.D., Kousik, C.S. & Adkins, S. (2013) The use of latent class analysis to estimate the sensitivities and specificities of diagnostic tests for *Squash vein*

yellowing virus in cucurbit species when there is no gold standard. *Phytopathology 103*, 1243-1251. doi: 10.1094/PHYTO-03-13-0071-R.

31. Wikler, K., Storer, A.J., Newman, W., Gordon, T.R. & Wood, D.L. (2003) The dynamics of an introduced pathogen in a native Monterey pine (Pinus radiate) forest. *Forest Ecology and Management 179***,** 209-221. doiI:10.1016/S0378-1127(02)00524-8.

32. Welsh, A,H., Lindenmayer, D.B. & Donnelly, C.F. (2013) Fitting and interpreting occupancy models. PLoS One 8, e52015 (2013).

Figure 1: 1a, probability density of the incidence q_0. $\tilde{q}$ is the upper limit of the Z% confidence interval of q_0. This upper limit is termed the maximum plausible incidence. 1b, development of the incidence through time. The incidence growth exponentially. We are interested in estimating the incidence q_0 when our most recent sample takes place. A period Δ earlier a sample was taken ad the incidence at that time was $q_0 Z_1$, a period Δ before that a sample was taken and the incidence was $q_0 Z_2$, etc.

Figure 2: Left-hand panel: Contour lines of $\frac{\theta_{fp}}{1-\theta_{fn}}$ for rate values of θ_{fn} and θ_{fp}. Right-hand panel: The maximum plausible incidence (i.e. the right hand boundary of the 95% confidence interval for q_0) as function of the number of cases reported by the volunteer surveyor. Drawn lines are for a pathogen with an epidemic growth rate of 0.018 (comparable to Citrus Canker), the dashed line for a pathogen with growth rate 0.0024 (comparable to Ash Dieback). The maximum plausible incidence is shown both for (i) the case where during all monitoring rounds the expert, verifying reports of volunteer surveyors, does not find any host to be infected (disease freedom), and (ii) when in the last monitoring round the expert, verifying reports of the volunteer surveyor, detects an infection for the first time.

Figure 3: A comparison between the maximum plausible incidence, q_0, as calculated from the approximation and as calculated from the full model. Both disease freedom and first detection is considered for a range of false-positive and false-negative rates for two tree diseases, Citrus Canker and Ash Dieback. On the left-hand side of the black line, the value of q_0 calculated from the approximation is more than 10% different from that of the full model. On the left-hand side of the grey line, the value calculated from the approximation is more than 5% different from the full model.

Figure 4: The accuracy of the factor $\frac{\theta_{fp}}{1-\theta_{fn}}$.

From the full model we calculate $\widetilde{q_E}$ and $\widetilde{q_V}$. The ratio of these is compared to the factor $\frac{\theta_{fp}}{1-\theta_{fn}}$. Both disease freedom and first detection is considered On the left-hand side of the black line the value of the factor is more than 10% different from that calculated from the full model. On the left-hand side of the grey line the value of the factor is more than 5% different from that calculated from the full model.

Figure 5: Lines of equal probability that the disease is found before incidence q. each drawn line is the contour line for a value of q_0. The hashed line is the contour line for equal total time of the monitoring programme. The left hand graph shows a case where the optimal monitoring programme consists of experts only verifying the reported cases of the volunteer surveyor. The right-hand panel shows a case where the optimal surveillance programme consists of experts going into the field themselves to sample.

Table 1: The confusion matrix. Table of the disease status and the observation of infected, 1, and uninfected, 0, hosts. The incidence of disease in the host population is q. θ_{fp} is the false positive rate of the observations. θ_{fn} is the false negative rate of the observations.

		Disease status	
		1-q	q
		0	1
observation	0	(1-q)(1-θ_{fp})	$q\theta_{fn}$
	1	(1-q)θ_{fp}	q(1-θ_{fn})

Table 2: Approximations to the probability densities of the incidence of disease in the most recent survey round, $P(q_0)$. Densities are given for the case of disease freedom, where all survey rounds return no positive finds, and for first detection where in the most recent survey round one of more positives is found. For the disease freedom cases the right-hand column gives the upper limit of the Z% confidence interval for the incidence. For the case of first detection this upper limit is approximated using the z-score.

DISEASE FREEDOM	Probability density	Max. incidence
Expert only	$P(q_0) = \left(\sum_{i=0}^{K} \lambda^{-i} N_i \right) e^{-\left(\sum_{i=0}^{K} \lambda^{-i} N_i\right) q_0}$	$\widetilde{q_E} = \frac{-\ln(1-\frac{Z}{100})}{\sum_{i=0}^{K} \lambda^{-i} N_i}$
Volunteer surveillance only	$P(q_0) = \left(\frac{(1-\theta_{fn})}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i} x_i \right) e^{-\left(\frac{(1-\theta_{fn})}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i} x_i\right) q_0}$	$\widetilde{q_V} = \frac{\theta_{fp}}{1-\theta_{fn}} \frac{-\ln(1-\frac{Z}{100})}{\sum_{i=0}^{K} \lambda^{-i} x_i}$
Combined expert sampling and volunteer surveillance	$P(q_0) = C\, e^{-C\, q_0}$ $C = \sum_{i=0}^{K} \lambda^{-i} N_i + \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i} x_i$	$\widetilde{q_{EV}} = \frac{-\ln(1-\frac{Z}{100})}{C}$

FIRST DETECTION	Probability density	Max. incidence
Expert only	$P(q_0) = \frac{1}{\frac{1}{A} - \frac{1}{B}} \left(e^{-A\, q_0} - e^{-B\, q_0}\right)$ $A = \sum_{i=1}^{K} \lambda^{-i} N_i$ $A = \sum_{i=1}^{K} \lambda^{-i} N_i$ and $B = \sum_{i=0}^{K} \lambda^{-i} N_i$ $B = \sum_{i=0}^{K} \lambda^{-i} N_i$	$\widetilde{q_E} = \frac{1}{A} + \frac{1}{B} + \tilde{Z}\sqrt{\frac{1}{A^2} + \frac{1}{B^2}}$ $\tilde{Z}$ is the z-score for the standard normal distribution. For the 95% tail $\tilde{Z}$=1.64, for the 99% tail $\tilde{Z}$=2.33
Volunteer surveillance only	$P(q_0) = \frac{1}{\frac{1}{A} - \frac{1}{B}} \left(e^{-A\, q_0} - e^{-B\, q_0}\right)$ $A = \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=1}^{K} \lambda^{-i} x_i$ and $B = \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i} x_i$	$\widetilde{q_V} = \frac{1}{A} + \frac{1}{B} + \tilde{Z}\sqrt{\frac{1}{A^2} + \frac{1}{B^2}}$ $\tilde{Z}$ is the z-score for the standard normal distribution.
Combined expert sampling and volunteer surveillance	$P(q_0) = \frac{1}{\frac{1}{A} - \frac{1}{B}} \left(e^{-A\, q_0} - e^{-B\, q_0}\right)$ $A = \sum_{i=1}^{K} \lambda^{-i} N_i + \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=1}^{K} \lambda^{-i} x_i$ and $B = \sum_{i=0}^{K} \lambda^{-i} N_i \frac{1-\theta_{fn}}{\theta_{fp}} \sum_{i=0}^{K} \lambda^{-i} x_i$	$\widetilde{q_{EV}} = \frac{1}{A} + \frac{1}{B} + \tilde{Z}\sqrt{\frac{1}{A^2} + \frac{1}{B^2}}$ $\tilde{Z}$ is the z-score for the standard normal distribution.

742

743 **Table 3:** Epidemic growth rate of 6 tree diseases of natural forests and agricultural orchard.

Disease	organism	Mean epidemic growth rate day^{-1}	references
Ash dieback	*Hymenoscyphus fraxineus*	0.0024	Alonso Chavez et al. 2016
Sudden oak death	*Phytophthora ramorum*	0.0033	Alonso Chavez et al. 2016
Citrus canker	*Xanthomonas citri*	0.0184	Alonso Chavez et al. 2016
Huanglongbing	*Candidatus Liberibacter spp.*	0.0072	Alonso Chavez et al. 2016
olive quick decline syndrome	*Xylella fastidiosa*	0.0122	Mastin et al in press
Pine pitch canker	*Fusarium circinatum*	0.0019	Wikler et al 2003; Reynolds et al 2019

744

745

746

747

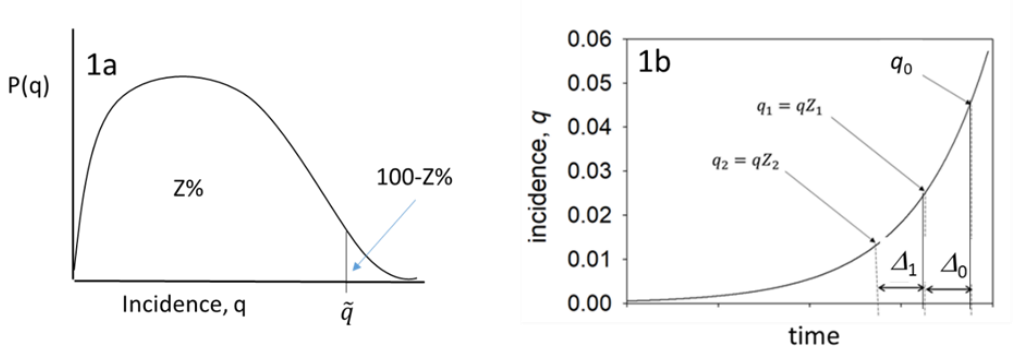


Figure 1.

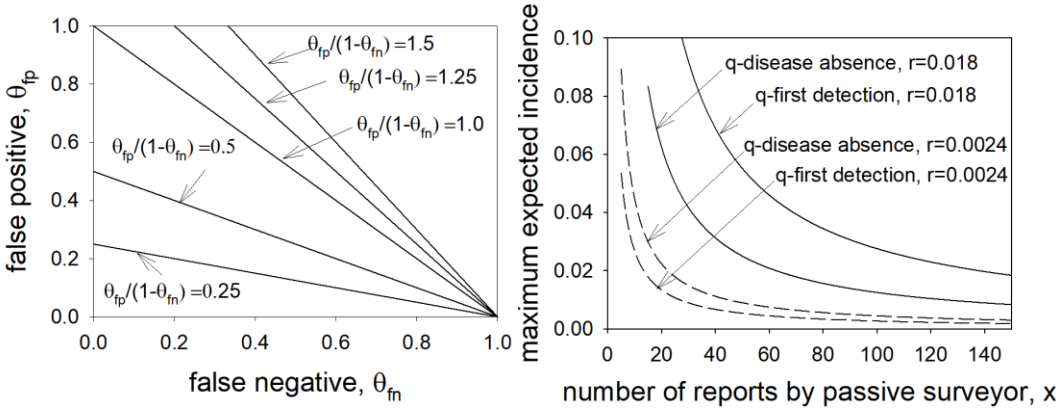
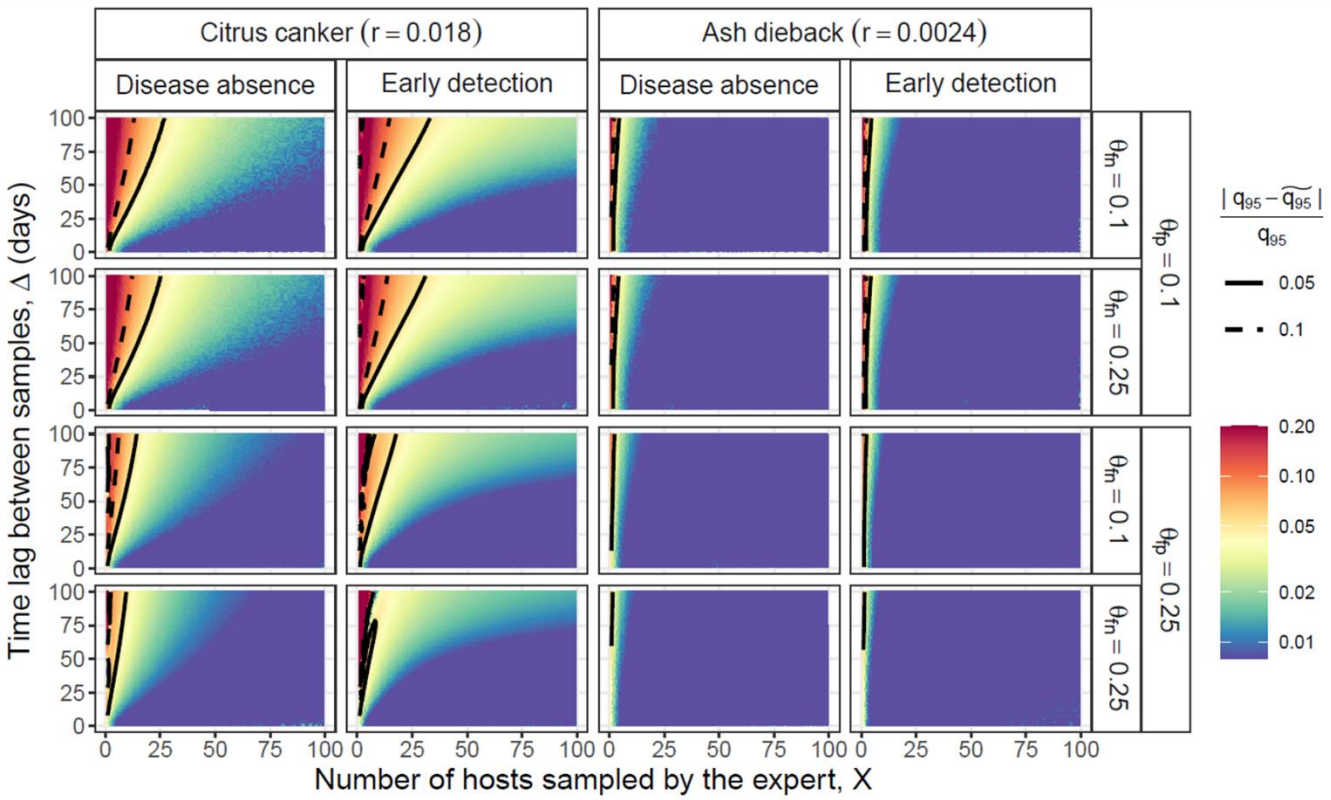


Figure 2



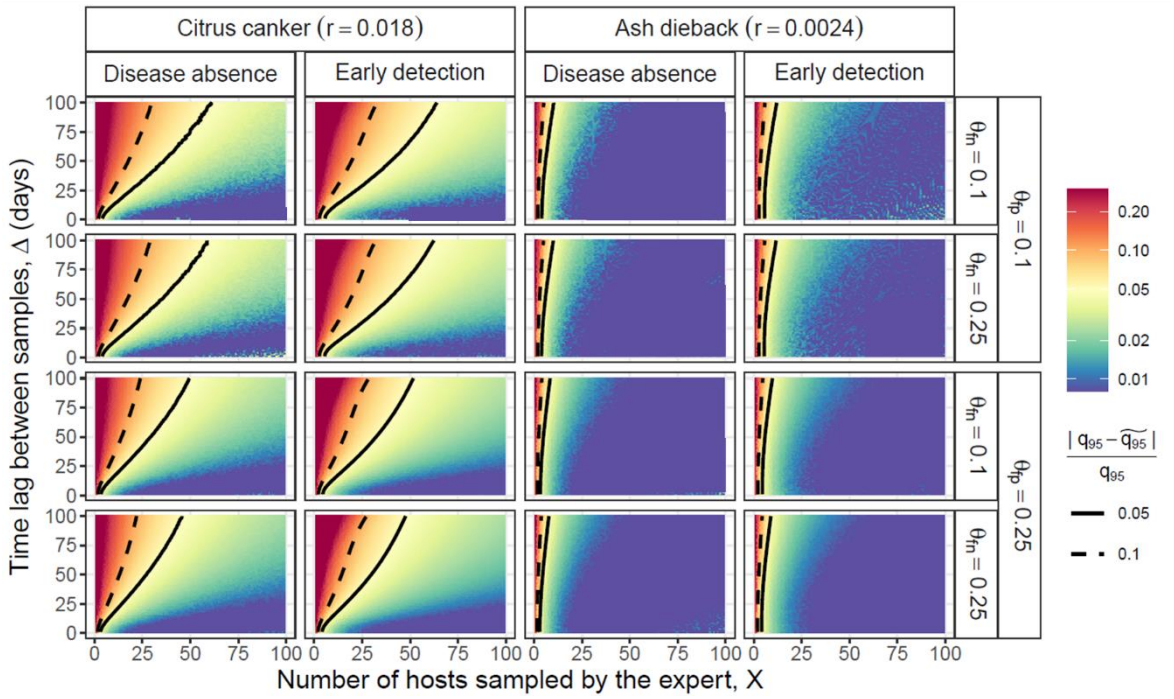
759

760

761 **Figure 3**

762

763

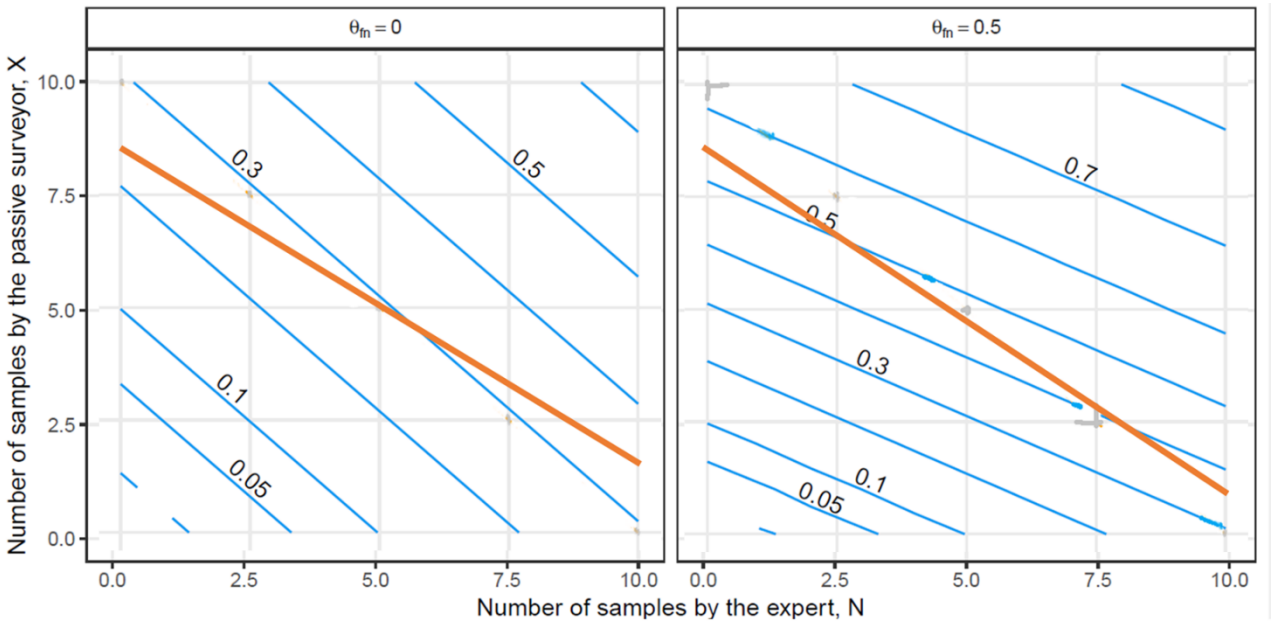


764

765

766 **Figure 4**

767



768

769 Figure 5