Technical Report

efsa SUPPORTING PUBLICATIONS

APPROVED: 19 November 2025
doi: 10.2903/sp.efsa.2025.EN-9788

Update of general guidelines for statistically sound and risk-based surveys of plant pests

European Food Safety Authority (EFSA),
Elena Lázaro, Stephen Parnell, Antonio Vicent Civera, Martijn Schenk, Jose Cortiñas Abrahantes, Juan Navas-Cortes, Hans-Hermann Thulke, Francesco Pecori, Joshua Koh, Jan Schans, Marc Aerts, Gabriele Zancanaro, Sybren Vos, Tomasz Kaluski

Abstract

At the request of the European Commission, EFSA prepared the general guidelines for surveys of plant pests, describing the legal, international and scientific context in which the surveys are designed, the basic principles implemented for surveillance of quarantine pests and introducing the concepts needed for the design of statistically sound and risk-based surveys. Three types of specific surveys are addressed: detection surveys for substantiation of pest freedom, delimiting surveys to determine the boundaries of a potential infested zone, and monitoring surveys for prevalence estimation when measuring the effectiveness of eradication measures or for the confirmation of a low pest prevalence area. For each type of survey, the survey parameters are introduced and their interactions analysed showing the importance of the assumptions that are taken for each one of them: (i) the aims of the survey are defined as achieving a certain level of confidence of detecting a given pest prevalence (design prevalence), this reflects the trade-off between the acceptable level of the risk and availability of resources that determine the strength of the evidence to support the conclusion of the survey; (ii) the target population is described by its structure and size, including the risk factors; and (iii) the method sensitivity is defined as the combination of the sampling effectiveness and the diagnostic sensitivity for each inspection unit. EFSA's RiPEST and RiBESS+ tools[1] are introduced for calculating the sample size using the survey parameters as input values for a statistically sound and risk-based survey design. The mathematical principles behind the tools are in line with the International Standards for Phytosanitary Measures. The survey design is flexible and can be tailored to each pest and specific situation in the Member States. Once the survey is implemented following this approach, the

[1] https://r4eu.efsa.europa.eu/

conclusions allow surveys to be compared across time and space, contributing to the harmonisation of surveillance activities across the EU Member States.

Keywords: confidence level, delimiting survey, design prevalence, detection survey, method sensitivity, monitoring survey, target population

Requestor: Directorate-General for Health and Food Safety, European Commission

Question number: EFSA-Q-2025-00467

Correspondence: Ask a Question

Acknowledgements: EFSA wishes to thank in particular the members of the EFSA working group on pest survey methods that contributed to the preparation of these guidelines. EFSA also wishes to thank Giulia Mattion, Alicia Culot, Naliny Feliu and Etienne Babin (trainees in the Environment, Plants and Ecotoxicology Unit – Plant Health Monitoring Team) for their support to the finalisation and publication of this scientific output. EFSA wishes to acknowledge the Netherlands Food and Consumer Product Safety Authority (NVWA), Instituto Valenciano de Investigaciones Agrarias (IVIA) and the Institute for Sustainable Plant Protection – National Research Council of Italy (CNR) for the support provided to this scientific output in the context of grants GP/EFSA/ALPHA/2021/11 and GP/EFSA/ALPHA/2021/08 – Lot 1 and Lot 2. EFSA also wishes to acknowledge the Member State institutions that contributed to the workshops in 2018 and 2019 held for the development of the survey approaches for *Phyllosticta citricarpa*, *Xylella fastidiosa* and *Agrilus planipennis*, and used when preparing this document. EFSA wishes to thank the reviewers of this document: Antonio Vicent Civera, Stephen Parnell and David Makowski.

Suggested citation: EFSA (European Food Safety Authority), Lázaro E, Parnell S, Vicent Civera A, Schenk M, Cortiñas Abrahantes J, Navas-Cortes J, Thulke H-H, Pecori F, Koh J, Schans J, Aerts M, Zancanaro G, Vos S, and Kaluski T. 2025. Update of general guidelines for statistically sound and risk-based surveys of plant pests. EFSA supporting publication 2025:EN-9788. 125 pp. doi:10.2903/sp.efsa.2025.EN-9788

ISSN: 2397-8325

Disclaimer: The designations employed and the presentation of material on any maps included in this scientific output do not imply the expression of any opinion whatsoever on the part of the European Food Safety Authority concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Any designation of Kosovo is without prejudice to positions on status and is in line with United Nations Security Council Resolution 1244 and the International Court of Justice Opinion on the Kosovo Declaration of Independence.

Summary

At the request of the European Commission, EFSA has prepared these general guidelines for designing statistically sound and risk-based surveys on plant pests to assist the EU Member States to carry out the different types of surveys as required by Regulation (EU) 2016/2031. The guidelines also aim to facilitate the harmonization of surveillance methods across the EU.

These guidelines for plant pest surveys are part of the EFSA Plant Pest Survey Toolkit[2] that is being developed to support the Member States in the preparation and the design of the surveys as well as to facilitate their implementation.

To plan a surveillance action for a specific pest it is necessary to approach any survey activity as a five-step process including five interrelated phases: *initiation*, *preparation*, *design*, *implementation* and *conclusion*.

In the survey *initiation* phase, general information should be collected about the aim of the survey, the target pest, the survey area and the host plants.

The second step is the *survey preparation* phase. This is required to gather the relevant epidemiological information and the landscape characteristics required to define the survey design, with the corresponding assumptions, i.e. the pest (characteristics and/or difficulties relevant to the survey), the target population (extension, structure: epidemiological units, risk factor areas, inspection units), and the sampling matrix, detection and identification methods. Within this scope, EFSA is preparing pest survey cards[3] for the regulated pests for which surveys are obligatory.

The third phase is the *survey design* (Section 4). This phase consists of quantifying each survey parameter as these are the input values which are needed to estimate the sample size and distribution. This document also describes the context in which the surveys are designed (legal, international standards, scientific knowledge) and the basic principles and approaches that are implemented for surveillance of Union quarantine pests. It introduces the surveyor to the requirements for the design of statistically sound and risk-based surveys. The concepts of general and specific surveillance are also introduced. Three specific types of survey are described: detection surveys for substantiation of pest freedom in an area (including buffer zone surveys); delimiting surveys to determine the boundaries of a potential infested zone; and monitoring surveys, for prevalence estimation, that can be applied in infested zones where the progress of eradication measures needs to be observed or where the confirmation of a low pest prevalence is required.

The *survey design* should start with setting the aims of the survey, deciding on the overall confidence level and design prevalence of the survey, based on the trade-off between the acceptable level of the risk and availability of resources. These two parameters need to be set by the risk managers as together they will determine the strength of the evidence to support the conclusion of the survey. It will then be necessary to estimate the other survey parameters

2 EFSA (European Food Safety Authority), online. Toolkit for plant pest surveillance in the EU. Available online: https://efsa.europa.eu/PLANTS/planthealth/monitoring/surveillance/index

3 Pest survey cards are published as part of the Toolkit for plant pest surveillance in the EU. Available online: https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/index-surveycards and https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/gallery

and consider the importance of the assumptions that are taken for each one of them. Detailed information on the host plants and their distribution in the survey area are needed to determine the size of the target population and its division into epidemiological units based on the homogeneity assumptions. By including risk factors, surveys will target those areas where the chances of finding the pest are higher. Determining the structure and size of the target population involves scientific knowledge on the epidemiology and detailed information of the local, regional and national landscapes. It will be required to determine the method sensitivity defined by the sampling effectiveness in the field and the sensitivity of the applied diagnostic method in the laboratory. To estimate the method sensitivity, the inspection protocol itself (i.e. that is applied to the inspection unit) and the experience and training of the inspectors should be taken into account as well as the laboratory methods and expertise available. The more precise and accurate the information used for selecting or estimating the survey parameters, the more reliable the conclusions of the survey will be.

After the survey parameters are determined, the survey design continues with the calculation of the sample size (i.e. number of 'inspection units' to be examined and/or tested) using the survey parameters as inputs of the statistical tool, which uses a statistically sound and risk-based approach (RiPEST, RiBESS+). The mathematical principles behind the tool are in line with the recommendations and guidelines provided by the International Plant Protection Convention (IPPC) in the various ISPMs and guidelines for pest surveys. The estimated number of inspections and/or samples should then be allocated to the epidemiological units and/or risk categories and the inspection units should be selected within the survey area.

An effective *survey design* relies on the technical aspects of the *survey preparation* and also on the involvement of the risk managers. The flexible approaches proposed in this document allow the survey design to be tailored to each specific situation in the Member States, taking into account the host plant distribution and available resources.

The fourth phase in the survey activity is the *survey implementation.* This includes the field inspections, sample collection and laboratory testing. Within this phase, specific field instructions for the inspectors need to be carefully formulated to indicate how to collect the required data.

Once the survey is conducted, the last step is the *survey conclusion* phase in which the results are reported, while considering the strength of the evidence to support this conclusion through the design prevalence and confidence level achieved. The underpinning assumptions made on homogeneity of the survey area, the method sensitivity, and the surveyed host plants should be included in the conclusion. The reliability of the conclusions of surveys designed using the proposed approaches depends strongly on the *survey preparation*. The clear formulation of the survey conclusion allows surveys to be compared across time and space, thus contributing to the harmonisation of surveillance activities across the EU Member States.

Considering that in the EU the surveys are implemented at Member State level, and that the data required for preparing the surveys are available at Member State or even regional level, the developed approach should be tailored to each specific situation in terms of host plants and resources.

Table of contents

Abstract 1
Summary 4
Table of contents 6
1 Introduction 9
2 General framework for surveillance 11
2.1 Legal basis in the EU 11
2.2 Responsibilities of the National Plant Protection Organisations 12
2.3 Existing guidelines and international standards 12
3 Surveillance 13
3.1 General surveillance 14
3.2 Specific surveillance 14
4 Survey design 16
4.1 Target population 16
4.1.1 Target population structure 17
4.1.2 Target population size 17
4.2 Epidemiological units 18
4.3 Risk factors 18
4.4 Inspection units 23
4.5 Survey components 24
4.6 Sampling matrix 24
4.7 Statistical sample size versus laboratory samples 25
5 Statistical background for sample size estimation 25
5.1 Conclusion on the pest status 26
5.2 Confidence level 27
5.3 Design prevalence 28
5.4 Selection of the confidence level and design prevalence 29
5.5 Method sensitivity 32
5.6 Relationship between confidence level, design prevalence and method sensitivity 34
5.7 Multiple component surveys 35
5.8 Unknown population size 37
5.9 EFSA RiBESS+ and RiPEST tools 38

6 Detection survey 38
6.1 Selection of the confidence level and design prevalence 39
6.2 Survey area and target population 39
6.2.1 Target population, epidemiological units and homogeneity 39
6.2.2 Risk factors 40
6.3 Inspection method and sample strategy 41
6.3.1 On-site visual examination of plants 41
6.3.2 Sample collection for laboratory testing 43
6.3.3 Detection by trapping 43
6.4 Method sensitivity 44
6.5 Sample size 44
6.6 Sample allocation 46
6.7 Reporting 47
6.8 Survey conclusion of detection surveys 47
7 Delimiting survey 49
7.1 General considerations 50
7.2 Natural and human-assisted spread 51
7.3 Identification of the potential source of infestation 53
7.4 Construction of the potential infested zone 54
7.4.1 A single potential source 54
7.4.2 Multiple source locations 56
7.5 Stepwise surveys in the potential infested zone 57
7.5.1 Narrowing down the potential infested zone 58
7.5.2 Enlarging the potential infested zone 58
7.6 Survey parameters for a delimiting survey 59
7.6.1 Target population size 59
7.6.2 Method sensitivity 62
7.6.3 Confidence level and design prevalence 62
7.7 Sample size 62
7.8 Sample allocation 63
7.9 Survey implementation 63
7.10 Survey conclusion for delimiting surveys 64

7.11 Reporting 64
7.12 Follow-up 65
8 Buffer zone survey 65
9 Monitoring survey 66
9.1 Monitoring surveys in infested zones 67
9.2 Survey design 68
9.2.1 General considerations on monitoring of eradication process 68
9.2.2 Definitions 68
9.3 Methodological framework 69
10 Plant pest survey Toolkit 72
10.1 Pest survey cards: 72
10.2 The Risk-based Pest Survey Tool (RiPEST): 72
10.3 Relational database (RDB): 72
10.4 Multi-pest Optimization Tool (OptiPest): 73
11 Conclusions 73
References 75
Abbreviations 79
Appendix A Specific changes and amendments in relation to General guidelines (EFSA, 2020) 80
References 82
Appendix B Delimiting strategies 83
References 95
Appendix C Sample survey design 96
References 106
Appendix D Two-Step Approach to Sample Size Calculations 107
Appendix E Monitoring survey 111
References 125

1 Introduction

At the request of the European Commission, EFSA was asked to support the EU Member States in the preparation and planning of the surveys of the EU quarantine pests (EFSA mandates on plant pest surveillance M-2017-0137, M-2020-0114 and M-2022-00069). In this context, EFSA prepared General guidelines[4] to assist the EU Member States to carry out the risk-based surveys on plant pests as required by Regulation (EU) 2016/2031[5]. This document presents an updated version of the existing general guidelines. For a detailed overview of the specific changes and amendments, please refer to Appendix A, where all modifications are clearly indicated.

The purpose of the document is to assist the EU Member States to plan survey activities of quarantine pests using risk-based and statistically sound pest survey approaches, in line with current international standards, and to facilitate the harmonization of surveillance across the EU. This document summarizes the framework in which the surveys are performed (legal, international standards, scientific knowledge), describes the basic principles and approaches that are used for surveillance of Union quarantine pests and introduces the surveyor to the principles for the design of statistically sound and risk-based surveys. The approach to surveillance should be in line with the guidelines for surveillance from the International Plant Protection Convention (IPPC) (EFSA, 2018; FAO, 2024a).

These general guidelines are part of the EFSA Plant Pest Survey Toolkit that has been developed to support Member States to carry out plant pest surveillance in the EU. The Toolkit also includes a glossary of surveillance terms which are used throughout this guidance and where the reader is referred . This toolkit consists of: (i) this document, the general guidelines; (ii) the pest survey cards (published in a dedicated gallery[6]) that guide the surveyor through gathering the relevant information needed for the preparation of a survey for a specific organism; (iii) several pest-specific guidelines that guide the surveyor in the design of statistically sound and risk-based surveys for specific pests, integrating the key information gathered using the pest survey card, and processing the information for the estimation and allocation of the sampling effort; (iv) the statistical software tools RiPEST[7], RiBESS+[8] and SAMPELATOR[9] that are used for the calculation of sample sizes;

4 https://efsa.onlinelibrary.wiley.com/doi/10.2903/sp.efsa.2020.EN-1919

5 Regulation (EU) 2016/2031 of the European Parliament of the Council of 26 October 2016 on protective measures against pests of plants, amending Regulations (EU) No 228/2013, (EU) No 652/2014 and (EU) No 1143/2014 of the European Parliament and of the Council and repealing Council Directives 69/464/EEC, 74/647/EEC, 93/85/EEC, 98/57/EC, 2000/29/EC, 2006/91/EC and 2007/33/EC. OJ L 317, 23.11.2016, p. 4–104.

6 https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/gallery

7 https://shiny-efsa.openanalytics.eu/app/surveillance

8 https://shiny-efsa.openanalytics.eu/app/ribess

9 https://shiny-efsa.openanalytics.eu/app/sampelator

v) relational Pest Surveillance Database[10] collecting data extracted from pest survey cards; and vi) OptiPest; a multi-pest crop survey optimisation tool[11] facilitating the planning of pest surveys at crop level and vii) webinars explaining the process of design of detection surveys for EU priority pests[12]. The preparation of a general data collection framework and the provision of a common reporting strategy on the pest surveys are outside the scope of this mandate (EFSA, 2018).

Surveillance on plant pests is primarily conducted to establish the pest status in an area and can be used for early detection of plant pests or to establish pest-free areas. Surveillance can be used to substantiate pest freedom in a country or a particular geographic area when pest absence provides the necessary safeguards for the trade of plants or plant products. Surveillance can also be used to monitor pest prevalence in an infested zone (e.g., to test the effectiveness of eradication measures) or buffer zone and can be used to delimit and monitor the size of an infested zone. The outcome of the survey activities should trigger the appropriate risk management decisions linked to the threat that a pest poses.

For any type of pest freedom survey, the aim is to detect a pest if it is present above a specified prevalence in a given area. A challenge is that it is statistically and practically impossible to conclude with 100% certainty that a pest is absent, even when it is not detected by a survey[13]. Similarly, if the pest is found in the survey sample, it is not possible to obtain the true pest prevalence in the total population but only to estimate the prevalence with a certain level of accuracy. To achieve absolute certainty on absence, every host plant in an area would need to be examined with an inspection procedure or sampling and diagnostic procedure that has perfect detection ability. Moreover, this would need to be repeated with a high frequency to ensure that the pest has not been introduced since the last survey. Clearly this is not feasible. In practice, it will only be possible to observe a relatively small proportion of the host plants at limited intervals, and with imperfect inspection procedures or sampling and diagnostic tests. Thus, the true absence or prevalence of a pest is uncertain even when a survey does not detect that pest.

Risk-based surveys target those locations or areas that are more likely to harbour a pest. This is more resource-efficient compared with simple random sampling, and risk-based surveys can either be used to increase the detection probability at the same level of resources or to reduce the survey effort while achieving the same level of confidence. However, appropriate design of risk-based surveys requires that the specific objective of each survey is clearly established, while knowledge should be available on the risk-factors that increase the probability of entry and establishment of the pest and on how the risk-factors are linked to the survey area (in terms of areas of climate suitability, host plants, habitats and epidemiology), thus allowing for the identification of high-risk areas. The

10 https://r4eu.efsa.europa.eu/app/pestDB

11 https://r4eu.efsa.europa.eu/app/crop-optimization

12 https://academy.europa.eu/courses/efsa-plant-pest-surveillance-toolkit;

13 This excludes situations in which a pest has no host plants in an area or when the climatic conditions do not allow for establishment of a pest. In these cases, it is neither recommended nor necessary to perform a survey. Article 22 of (EU) 2016/2031 provides the legal framework to refrain from surveillance of Union quarantine pests 'for pests for which it is unequivocally concluded that they cannot become established or spread in the Member State concerned due to its ecoclimatic conditions or to the absence of the host species.'

required survey methodology (inspection, sampling and laboratory procedures) and capacity to carry out the surveillance activities should be available as well.

These guidelines will first describe the general framework for surveillance, including the legal basis, and then cover various types of surveillance in a stepwise approach, while illustrating these types with several more specific scenarios.

2 General framework for surveillance

2.1 Legal basis in the EU

The general requirements for plant pest surveys (EFSA, 2023) in the EU territory are provided in Regulation (EU) 2016/2031. In particular, this regulation requires that surveys should be based on scientific and technical principles to detect the pest (EFSA, 2023) concerned.

Regulation (EU) 2016/2031 states that Member States should carry out risk-based surveys (EFSA, 2023) for the timely detection of Union quarantine pests in all areas where the pest concerned was not previously known to be present (Articles 22, 23 and 24). The Member States (MSs) need to define, within their multiannual survey programmes: (i) the objective and scope of each survey, (ii) the area and frequency of the surveys, (iii) the target population (EFSA, 2023) of the surveys, (iv) the survey methodology and the detection methods (EFSA, 2023) used, as well as (v) the data collection and reporting methods applied (Article 23). In Article 24a emphasis is given to surveys of priority pests, indicating that surveys on these pests should be performed annually and should include "a sufficiently high number of visual examinations (EFSA, 2023), sampling and testing, as appropriate for each priority pest, to ensure, as far as it is possible given the respective biology of each priority pest and the ecoclimatic conditions, with a high degree of confidence, the timely detection of those pests."

In addition, the regulation requires that MSs should perform surveys aiming at recognition and maintenance of protected zones as regards the presence of the protected zone quarantine pest concerned (Articles 32 and 34).

Finally, for demarcated infested zones, the regulation mentions that annual surveys should be carried out to monitor the development of the presence of the pest concerned. Modification of the pest status within demarcated areas should be based on the results of surveys (Articles 18 and 19).

The Union quarantine pests which should be included in the multiannual survey programmes are listed in Annex II of Commission Implementing Regulation (EU)

2019/2072[14], while the priority pests – for which a survey should be performed annually – are listed in Commission Delegated Regulation (EU) 2019/1702[15].

For some pests, certain implementing acts of Regulation (EU) 2016/2031 require to carry out statistically based surveys.

The current general guidelines for statistically sound and risk-based surveys of plant pests address all the above-mentioned requirements in Regulation (EU) 2016/2031.

2.2 Responsibilities of the National Plant Protection Organisations

According to Article IV of the International Plant Protection Convention (IPPC) (FAO, 2024a), the responsibilities of a National Plant Protection Organisation (NPPO) include surveillance of growing plants, both cultivated and uncultivated, including wild flora. The main objectives of these activities are to report the occurrence, outbreak and spread of pests, as well as their control (eradication, containment). NPPOs are also responsible for the designation, maintenance and surveillance of pest-free areas and areas of low pest prevalence (EFSA, 2023) in their country. Contracting parties to the IPPC should also, to the best of their ability, conduct surveillance of pests and develop and maintain adequate information on pest status (EFSA, 2023) in order to support categorisation of pests and risk assessments, and for the development of appropriate phytosanitary measures (Article VII).

Pest freedom substantiation can be used to enable the trade of specific plants or plant products while minimising phytosanitary risks. The phytosanitary guarantees are the responsibility of the exporting NPPOs. Therefore, it is also the responsibility of these NPPOs to provide evidence of the absence of a pest in a country or area, which can subsequently be used to certify that a consignment that originates in that country/area fulfils the regulatory requirements of the importing country.

2.3 Existing guidelines and international standards

EFSA's approach to survey design complies with the international standards and guidelines from the IPPC and the European and Mediterranean Plant Protection Organization (EPPO). In particular, these general guidelines have benefited from the information that is already available in several International Standards for Phytosanitary Measures (ISPMs) from the IPPC that refer to surveys (Table 1). Two ISPMs provide instructions on survey design, whereas eight ISPMs include procedures that either implement surveillance or refer to surveys. Two ISPMs provide instructions on collecting the data that are needed when designing the surveys, specifically on diagnostic sensitivity (ISPM 27) (FAO, 2016a) and risk factors (ISPM 32) (FAO, 2016b).

[14] Commission Implementing Regulation (EU) 2019/2072 of 28 November 2019 establishing uniform conditions for the implementation of Regulation (EU) 2016/2031 of the European Parliament and the Council, as regards protective measures against pests of plants, and repealing Commission Regulation (EC) No 690/2008 and amending Commission Implementing Regulation (EU) 2018/2019. OJ L 319, 10.12.2019, p. 1–279.

[15] Commission Delegated Regulation (EU) 2019/1702 of 1 August 2019 supplementing Regulation (EU) 2016/2031 of the European Parliament and of the Council by establishing the list of priority pests. OJ L 260, 11.10.2019, p. 8–10.

Table 1: International Standards for Phytosanitary Measures (ISPMs)* with reference to surveys on plant pests (as of September 2025)

Instructions on survey design	
ISPM 6	Surveillance
ISPM 31	Methodologies for sampling of consignments
Procedures employing surveys	
ISPM 1	Phytosanitary principles for the protection of plants and the application of phytosanitary measures in international trade
ISPM 4	Requirements for the establishment of pest-free areas
ISPM 8	Determination of pest status in an area
ISPM 9	Guidelines for pest eradication programmes
ISPM 10	Requirements for the establishment of pest-free places of production and pest-free production sites
ISPM 17	Pest reporting
ISPM 22	Requirements for the establishment of areas of low pest prevalence
ISPM 26	Establishment of pest-free areas for fruit flies (Tephritidae)
Procedures generating data for survey design	
ISPM 27	Diagnostic protocols for regulated pests
ISPM 32	Categorization of commodities according to their pest risk

* https://www.ippc.int/en/core-activities/standards-setting/ispms/

Several EPPO Standards are relevant for the design and implementation of surveys[16], in particular in the following series:

PM 3 – Phytosanitary procedures: methods to be followed for performing inspections, tests (EFSA, 2023) or treatments of commodities moving in trade, or surveys of quarantine pests.

PM 7 – Diagnostics: internationally agreed diagnostic protocols for regulated pests and horizontal standards on diagnostic issues.

PM 9 – National Regulatory Control Systems: procedures to be followed for official control with the aim of containing and eradicating pests.

3 Surveillance

16 https://www.eppo.int/RESOURCES/eppo_standards

Two kinds of surveillance are distinguished in ISPM 6 (FAO, 2018): general surveillance and specific surveillance.

General surveillance is an ongoing process 'whereby information on pests of concern in an area is gathered from various sources.'

Specific surveillance is defined as 'a process whereby information on pests of concern in an area is obtained by the NPPO over a defined period.'

3.1 General surveillance

General surveillance can provide data on the potential pest presence and data on the current threat of pest presence.

The potential pest presence includes the characteristics of the pest, e.g., identity, host range (EFSA, 2023), life cycle and current distribution; and the characteristics of the survey area (EFSA, 2023), e.g. the environmental suitability of the survey area for the pest, the distribution of host plants (EFSA, 2023) in the survey area and the presence of risk factors (EFSA, 2023) in the area such as trade hubs, storage facilities for plant products, and nurseries. The EFSA pest survey cards[17] present up-to-date information on the potential presence of EU quarantine pests. Evaluation of the potential pest presence may narrow down the target area for surveillance or can be useful for the identification of risk factors that need to be taken into account in a risk-based pest detection survey (EFSA, 2023).

Current information on the risk of introduction and spread of a particular pest can be obtained by NPPOs from several sources, such as the recurring assessment of international phytosanitary reporting (e.g. pest reports from trading partners, alert lists of regional plant protection organisations), analysis of trends in national trade records and inspection records (rate of influx of host plant material into the survey area, results of phytosanitary inspections for import, plant passports and export) and of 'general' observations (e.g. by phytosanitary inspectors (EFSA, 2023), farmers, professionals in tree and landscape maintenance and citizen science) or using EFSA Plant Health Horizon Scanning Newsletter. In particular, an analysis of the number of shipments of commodities that potentially carry regulated pests that enter the survey area provides useful information for the prioritisation of surveys.

For the Union quarantine pests, general surveillance provides the background information that allows the NPPO to undertake (risk-based) specific surveys to accurately determine the pest status. As mentioned in Regulation (EU) 2016/2031 Article 22, specific surveys are not required in MSs for pests for which the environment is not suitable for their establishment and spread (unsuitable climate and/or absence of host plants).

3.2 Specific surveillance

Specific surveillance is a targeted approach that is limited in time and relies on a survey collecting a sample or/and observations, being a subset of the total population (Eurostat, 2008). In line with ISPM No. 6 (FAO, 2018), EFSA distinguishes three types of surveys

17 https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/gallery

(Figure 2): the detection survey (see Sections 6 and 8), the delimiting survey (see Section 7) and the monitoring survey (see Section 9).

A detection survey is performed to support a conclusion on the presence or absence of a pest species in a specified area. Detection surveys may be designed for different phytosanitary objectives, such as substantiation of pest freedom in a country (ISPM 8) (FAO, 2017a), establishment of a pest-free area (ISPM 4) (FAO, 2017b) or early detection of a pest in a buffer zone surrounding an infested zone (EFSA, 2023; FAO, 1998). In the case of early detection, the survey is designed to detect the pest at an early stage of the epidemic.

A delimiting survey (EFSA, 2023) is an iterative procedure used to establish the boundaries of an area considered to be infested by or free from a pest. This procedure can be utilised to achieve phytosanitary objectives such as the demarcation of an infested zone for eradication (ISPM 9) (FAO, 1998) or containment measures, or for the establishment of a pest-free area.

Monitoring surveys are performed to determine the characteristics of a pest population that is present in an area, such as population density, aggregation or intraspecific diversity. Monitoring surveys may, for example, be applied as part of a procedure to maintain an official area of low prevalence of a pest (ISPM 22, FAO, 2005). EFSA has also developed a methodology of monitoring surveys to assess the effectiveness of an eradication programme (Section 9).

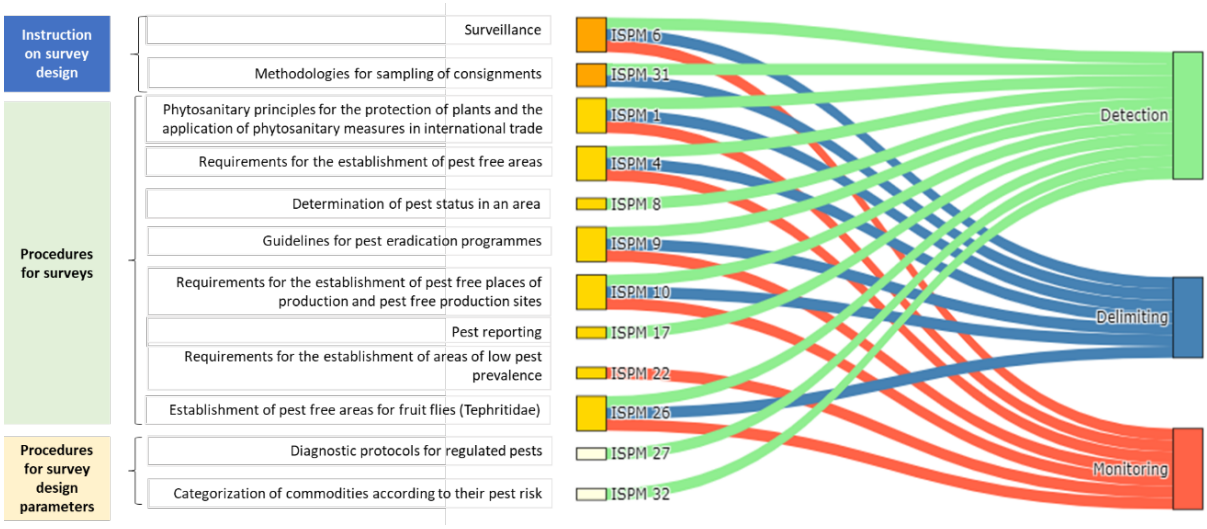


Figure 1: International Standards for Phytosanitary Measures (ISPMs) and the different types of phytosanitary surveys that are covered by those standards

Figure 2 illustrates the process that links the aim of the survey, the corresponding choice of survey type, the tools that could be used to calculate the sample size (EFSA, 2023) based on statistics (RiPEST, RiBESS+ and SAMPELATOR). Once it has been established which type of specific survey is needed, the actual survey design starts.

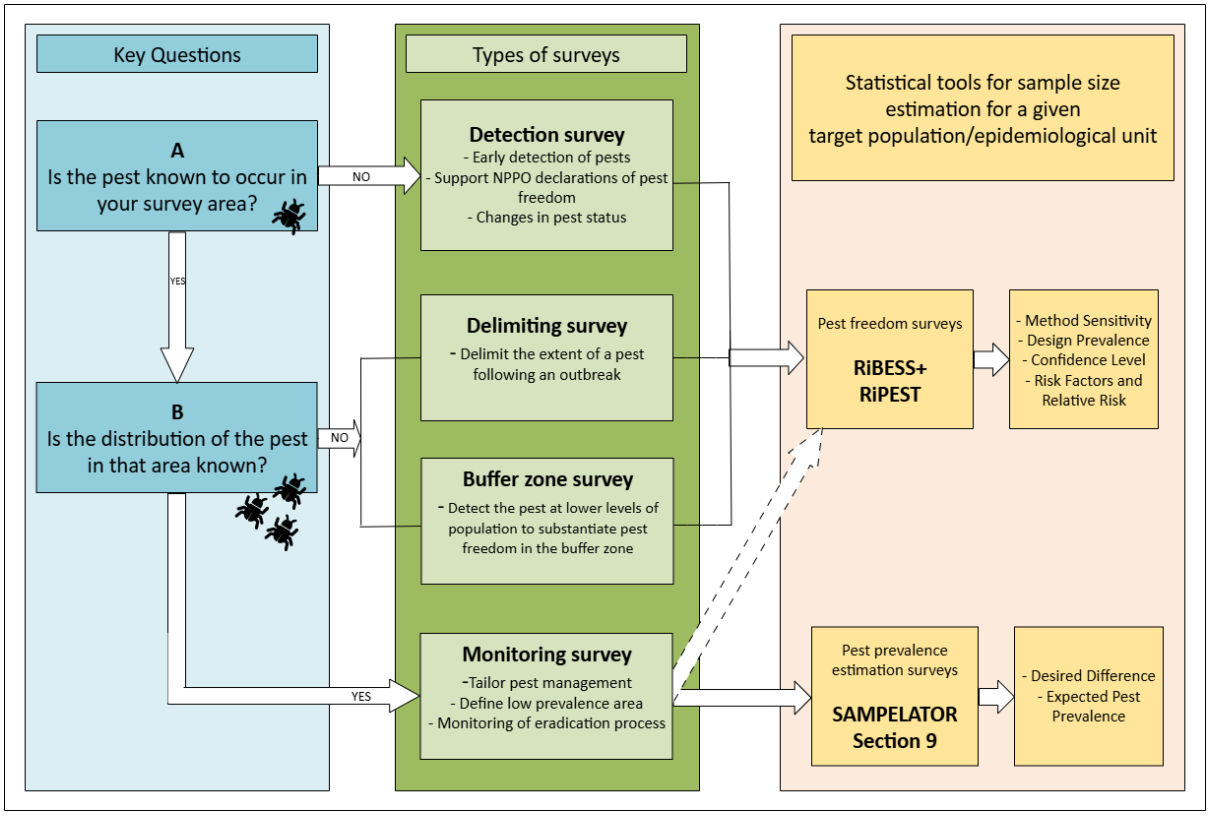


Figure 2: Choice of survey and corresponding tools for survey design. The input parameters for RiPEST, RiBESS+ and SAMPELATOR are shown in the boxes on the right of the image

4 Survey design

The survey design should be tailored to the situation in each EU Member State. This is needed because the conditions may vary across the EU Members States – or even within a Member State – in terms of the resources available for surveillance, the presence and importance of host plants, vector abundance, environmental conditions, the risk of entry of a pest, the occurrence of risk activities (EFSA, 2023) and presence of risk locations (EFSA, 2023), the availability of detection methods, etc.

In this document, the survey design is addressed in general terms, while detailed case studies of survey designs are presented in the guidelines for statistically sound and risk-based surveys as for *Xylella fastidiosa* (EFSA, 2020a) *Phyllosticta citricarpa* (EFSA, 2020b) and *Agrilus planipennis* (EFSA, 2020c). EFSA developed also a series of webinars to guide users through the Survey toolkit and its implementation for current priority pests[18].

4.1 Target population

When surveys are prepared and designed, it is necessary to first determine the distribution of hosts in the survey area where the pest can be present. The target population is defined as the set of inspection units: individual plants or commodities or vectors or a subdivision

[18] https://academy.europa.eu/courses/efsa-plant-pest-surveillance-toolkit

of the area covered by a trap in which the pest can be detected directly (e.g., by looking for the pest) or indirectly (e.g., by looking for symptoms indicating the presence of the pest) in a given habitat or survey area (EFSA, 2018).

When the aim of the survey is to substantiate pest freedom, it is recommended that the target population encompasses all or at least the most relevant hosts of the pest in order to minimise the risk of overlooking it. When pests are highly polyphagous, it will not always be practical to include all host species, and the target population needs to be limited to the plant species that are most likely to be infested.

To design a survey, the target population needs to be defined in terms of its structure (survey area, epidemiological units, risk areas and inspection units) and size.

4.1.1 Target population structure

To accurately target a survey, it is necessary to understand how the target population is structured and clearly define its subdivisions.

Usually, it is possible to define the target population structure subdividing it hierarchically into different levels. Figure 3 illustrates how to structure hierarchically the target population of *Elsinoë australis*, *E. citricola* and *E. fawcettii* (EFSA, 2022). Considering a unique survey component, within the survey area (level 1), epidemiological units are defined subdividing each NUT2 region according to two land-use categories (levels 2 and 3) which can be further subdivided according to different risk levels based on risk areas (level 4). Finally, for each one of the epidemiological units defined, the elementary units on which the detection methods are applied, that is, the inspection units, are distinguished and quantified (i.e., target population size) (level 5).

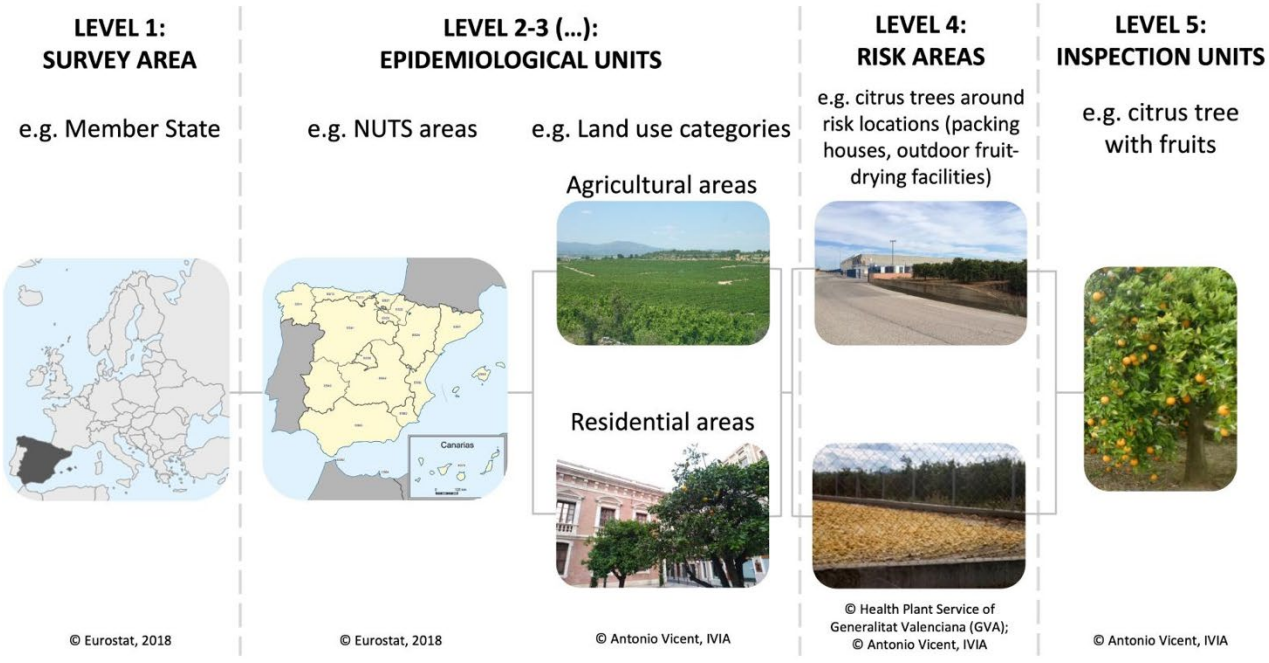


Figure 3: Example of hierarchical structure of the target population for *Elsinoë australis*, *E. citricola* and *E. fawcettii* in the EU (Sources: Eurostat, 2018 (levels 1–2), Plant Health Service of Generalitat Valenciana (GVA) (level 4, up), Antonio Vicent, IVIA (levels 3, 4 bottom,5)). Extracted from EFSA (2022).

4.1.2 Target population size

Once the target population has been identified in the survey area, the total number of inspection units for each epidemiological unit defined in each subdivision needs to be determined, thus defining the size of the target population. That is, the size of the target population is expressed as the total number of inspection units it contains. The exact size will often be unknown, and it is then required to reliably estimate it based on the available information in the EU Member State or alternatively provide a potential lower bound.

4.2 Epidemiological units

An epidemiological unit is defined as a homogeneous area where the interactions between the pest, the host plants and the abiotic and biotic factors would result in a similar epidemiology should the pest be present. The epidemiological units are homogeneous subdivisions of the target population in a given geographical area. They are the units of interest for which the sample size is estimated. Thus, for a statistically based survey it is therefore essential to clearly define these epidemiological units, indicating the related assumptions.

Epidemiological unit means a group of inspection units with a defined epidemiological relationship that share approximately the same likelihood of exposure to a pest. When multiple inspection units (plants, plant products or specimens) occur in a structured way in the landscape, sharing a similar likelihood of exposure to a particular plant pest, they can be included in epidemiologically homogeneous groups (i.e. epidemiological units), considering for example the geographical distribution (e.g. all apple trees in an orchard are usually of the same cultivar and are exposed to the same management practices and climatic conditions). Hence, should the pest be present, the individual plants would experience a similar exposure. It is also possible that a consignment of a particular plant product can be considered as an epidemiological unit. This would be the case when surveying lots of seed potatoes, for example, where each lot is harvested from a particular field and thus has experienced a similar exposure to a particular plant pest, should the pest have been present in that field.

Alternatively, the entire survey area may be one single homogeneous area in terms of interactions between the pest, host plants and (a)biotic factors. The individual plants within the survey area would then be subsequently targeted in order to draw conclusions for the entire area. This homogeneity assumption would rarely be fulfilled in practice as the epidemiology usually varies across larger areas in terms of ecology (habitats, environmental suitability, phenology, presence of host plants, etc.), exposure (pathways and entry points), and geographical characteristics.

When there is little information on the epidemiological homogeneity available for the target population, one case is to consider the landscape as a grid and that each cell within the grid is an epidemiological unit. Other natural spatial divisions may occur within a landscape depending on the host and pest e.g., individual orchards, fields or woodlands could be considered as epidemiological units. In these cases, the assumption of epidemiological homogeneity is likely to be fulfilled. This approach can be considered extreme as it will lead to a high number of epidemiological units and sample size. However, it provides a practical and systematic allocation in the absence of information to otherwise identify epidemiological units. This is termed the “two-step” approach, and a full description of this approach is available in Appendix D.

4.3 Risk factors

Consideration of risk factors in the survey design allows the survey efforts to be concentrated on those areas with a higher probability to be infested by the target pest.

Identification of risk factors and their relative risk (EFSA, 2023) estimation is essential for performing risk-based surveys. A risk factor is a biotic or abiotic factor (e.g., related to environment, ecosystem or a human activity) that increases the probability of infestation (EFSA, 2023) of an epidemiological unit by the pest. Each risk factor should have more than one level of relative risk for the target population. For each level, the relative risk needs to be estimated as the relative probability of infestation compared to a baseline set at 1 (EFSA, 2018).

Their inclusion in the survey design requires knowledge about the factors that influence the presence of a pest in specific parts of the area. This information, if available, allows the target population to be further subdivided into risk groups where the probability of infestation is similar. Their identification and their relative risk estimation need to be tailored to the situation in each EU Member State; the corresponding pest survey cards (EFSA, online) [19] provide pest-specific examples of risk factors that can be further developed by the MSs to address their particular situations.

Risk factors can be identified by general surveillance and may include the biological characteristics of the pest, the host range of the pest, the number and distribution of inspection units in the area to be surveyed, the environmental characteristics of the area and pathways for entry and spread of the pest in the area.

Examples of risk factors are the distance to an area where the pest is present, distance to import locations (Figure 4), distance to nurseries and garden centres where plant material from infested countries is introduced and parts of the survey area where the environmental conditions are more suitable for the pest than in other parts. When the risk factor is the distance to a particular risk location, the size of the area that is at higher risk (risk area) depends on the spread capacity of the pest and when the pest may have arrived, so risk factor identification may require knowledge on pest risk analysis (what is the risk activity?), knowledge on the territory (where do the risk activities take place?) and knowledge of the pest (how far can it spread?).

When the probability of infection (EFSA, 2023) by a pest is higher on a particular pathway, for the identification of risk areas, it is first necessary to identify the activities that could contribute to the introduction or spread of the pest (e.g. trade of plants for planting from areas where the pest is present). These activities should then be connected to specific locations (e.g. nurseries, garden centres). Around these risk locations, risk areas can be defined, knowing that their size depends on the spread capacity of the target pest and the availability of inspection units (e.g. host plants) around these locations.

Table 2 below shows some examples of risk activities and corresponding risk locations relevant for surveillance of *Aleurocanthus spiniferus*, *A. woglumi* and *A. citriperdus.*

Table 2: Example of risk activities and corresponding risk locations relevant for surveillance of *Aleurocanthus spiniferus*, *A. woglumi* and *A. citriperdus* (extracted from the pest survey card*)

Risk activity	**Risk locations**	**Risk areas**

19 The currently published pest survey cards are available at: https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/index-surveycards

Import, trade and storage of host plants for planting (e.g. mainly *Citrus* spp.).	Locations where imported *Citrus* spp. plants are stored or traded (e.g. nurseries, garden centres).	Areas surrounding risk locations where *Citrus* spp. plants are present (e.g. orchards, backyards and private gardens, parks).
Import, trade and storage of host plants for planting other than *Citrus* spp. originating from infested areas.	Locations where imported host plants other than *Citrus* spp. are stored or traded (e.g. nurseries, garden centres).	Areas surrounding risk locations where host plants other than *Citrus* spp. are grown (e.g. fields, forest edges, orchards, backyards and private gardens, parks).

* Pest survey card for survey of *Aleurocanthus spiniferus*, *A. woglumi* and *A. citriperdus* available at https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/aleurocanthus-spiniferus-woglumi-citriperdus

Another example is presented in Table 3 for risk activities and corresponding risk locations relevant for surveillance of Huanglongbing and its vectors.

Table 3: Example of risk activities and corresponding risk locations relevant for surveillance of *Ips typographus* (extracted from the corresponding pest survey card*)

Risk activity	**Risk locations**	**Risk areas**
Import and trade of the wood, bard and wood packing material of host trees.	Entry points (e.g. airports, harbours) and wood storage and trade facilities.	Areas surrounding risk locations where host plants are present.

* Pest survey card for survey of *Ips typographus* available at: https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/ips-typographus

With regards to the spread data, if available, related information is presented in the pest survey cards (EFSA, online)[19]. This information may specify the annual spread rate of the target pest or may provide additional information to guide this value estimation. For the priority pests, spread is expressed as the maximum of the rate of expansion (hereinafter referred to as annual spread rate) in pest prioritisation methodology (EFSA, 2025).

For detection surveys design using RiPEST, maximum, median or mean annual spread rate value can be used to define the risk areas around risk locations where the probability of infestation of the pest is higher, whereas for a delimiting survey, it is recommended to use the maximum spread rate for defining the potential infested zone around the source of an outbreak (Appendix B).

Figure 4: Examples of risk factors are the imports of wood packing material and imports of wood, which are known entry pathways of wood-boring insects and bark beetles. The presence of an airport itself may be a risk factor as illustrated by this trapping location for *Popillia japonica* given that adult beetles may hitchhike on vehicles and commodities (Source: NVWA, NL)

A risk factor may also be non-geographical. As an example, described in the survey card for surveys of *Ralstonia solanacearum*[20] the use of surface water for irrigation of potato fields can be considered a risk activity. Surface water may become contaminated when the potato industry uses infected potatoes and wastewater is discharged into surface water

[20] Pest survey card for survey of potato brown rot, *Ralstonia solanacearum* available at https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/ralstonia-solanacearum

without treatment (Janse et al., 2009). The presence of wild hosts increases the probability that the bacterium will become established in the surface water. Within this risk factor there may be different levels of risk, depending, for example, on the presence of industrial sites where ware and starch potatoes from contaminated areas are processed (with or without wastewater treatment) or the presence of wild host plants along the waterside.

The growing of susceptible varieties as a risk activity is illustrated in the Survey card for *Synchytrium endobioticum*[21] where the cultivation of non-resistant varieties is considered a risk activity. However, this needs to be tailored to the situation in each Member State, depending on which pathotypes are present locally and can be expected during surveillance. Initially, only pathotype 1 (D1) of potato wart disease occurred in Europe, and the use of resistant varieties provided a good level of control. Since new pathotypes have emerged, the effectiveness of varietal resistance depends on the pathotypes of *S. endobioticum* present in the soil. Only a few potato varieties are resistant to all the pathotypes that are widespread in Europe.

Similarly, for *Phyllosticta citricarpa*, the causal agent of citrus black spot, susceptibility varies between host species and the relative risks are provided in the pest survey card (EFSA, 2020d) (Table 4) for different citrus species. They were estimated based on a literature review, empirical observations, and expert consultation. The susceptibility of the citrus species is related to the likelihood of infection by the fungus and the resulting disease intensity. In particular, for lemon trees that have several flowerings during the year, the likelihood that young lemon fruit coincides with presence of inoculum and favourable weather conditions for the fungal infection is higher than for other citrus species that have only one flowering per year. For sweet orange, a very large range of varieties are grown, and each one has a very specific harvesting calendar. Late-maturing cultivars of sweet orange are considered to be more susceptible than the early-maturing ones.

Table 4: Example of relative risks for the various hosts of *Phyllosticta citricarpa* (extracted from EFSA, 2020d)

Botanical name	Common name	Relative risk (most susceptible)
Citrus limon (L.) Burm. f.	Lemon	1.5
Citrus sinensis Osbeck	Sweet orange (late-maturing cultivars)	1.4
Citrus sinensis Osbeck	Sweet orange (other cultivars)	1
Citrus reticulata Blanco	Mandarin	1
Citrus unshiu (Swingle) Marcow	Satsuma mandarin	1
Citrus paradisi Macfad	Grapefruit	1

To be able to apply a risk factor in the survey design, it is necessary to characterise both the relative risk and the proportion of the overall target population in the survey area to which it applies. This information can be entered in the RiPEST, RiBESS+ tools, which can

[21] Pest survey card for survey of *Synchytrium endobioticum* available at https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/synchytrium-endobioticum

subsequently calculate the required sample size of the survey and the distribution of the inspection units over the different risk groups.

4.4 Inspection units

The inspection units are the elementary units on which the detection method (EFSA, 2023) is applied (e.g. plants, commodities, pest vectors that are examined for detection of a pest) as part of a survey. They define the subset of the target population in each epidemiological unit for which the sample size is estimated. The identity of the inspection unit may differ from pest to pest and within each survey component, thus, it needs to be defined on a case-by-case basis (Figure 5).

For example, the inspection unit in an orchard could be a single citrus tree, or a soil sample from that orchard. In some cases, the inspection unit is defined directly as an individual tree that is visually inspected, as in the case of *Phyllosticta citricarpa* (EFSA, 2020b). However, other cases are more complicated, e.g. in the case of using a pheromone trap as a field detection method, the inspection unit would be defined as the area covered by the trap (or multiple traps) as in the case of surveying fruit flies such as *Bactrocera zonata*, and the method sensitivity (EFSA, 2023) associated with this inspection unit would be the probability that the pest will be caught by the trap if it is present in that area.

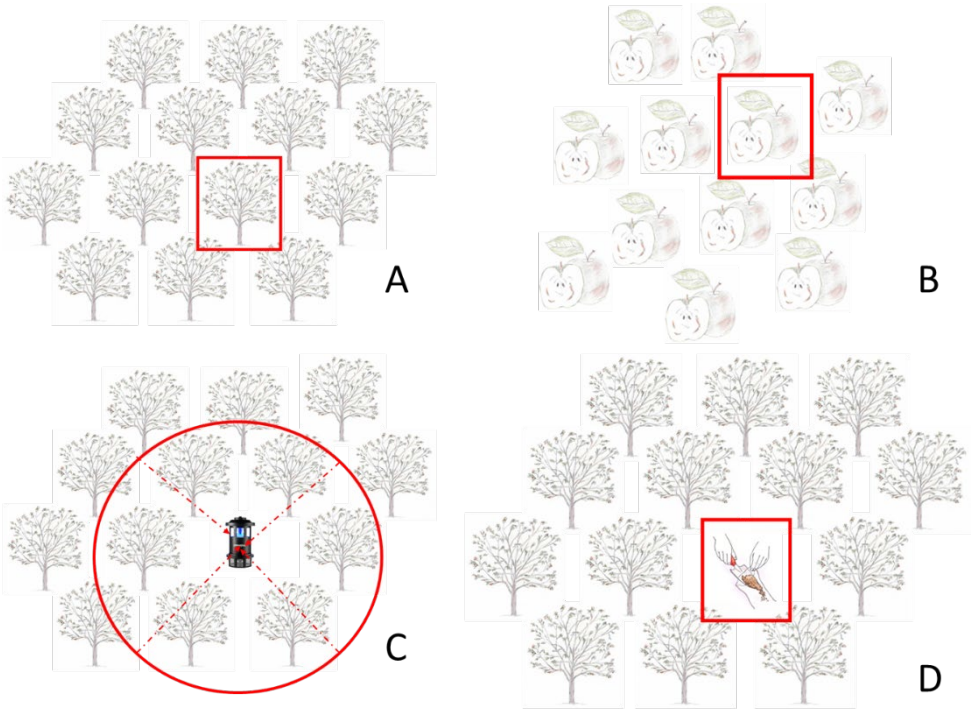


Figure 5: Different inspection units: A) plant, B) fruit (commodity), C) area, D) soil sample taken from the area

4.5 Survey components

A pest may be detected by different inspection targets. For example, when *Ralstonia solanacearum* is present in host plants it can be detected by visual examination followed by the collection and testing of symptomatic samples. *Ralstonia solanacearum* can also occur in water courses and can then be detected by testing water samples. In this example, the inspection unit can thus either be a plant or a water course. Because they require different inspection units, the respective survey components need to be designed separately in RiPEST. Other examples in which multiple types of inspection units can be considered are soil borne pathogens (that can be detected in host plants and soil) and vector borne diseases (that can be detected in host plants and vectors). Multiple inspection units can also be a consequence of using different detection methods for the same pest, for example when using both visual examination on individual plants and remote sensing or trapping (which are both area-based).

Hence surveys exist that are constructed of multiple components by addressing different inspection units (e.g., host plants, vectors, lots, soil, water; (EFSA, 2023)) due to different methods of identifying if a pest is present in the target population e.g. surveying plants for visual signs of the pest and trapping areas to catch the insect vector. This is because the sample size required for the survey are determined by the inspection units.

In some cases, the target population is comprised of only one component. This is the case that applies in a survey for *Agrilus planipennis*, in which all ash trees in a Member State that could potentially host the pest would constitute the single component of the survey.

When a pest is transmitted by a vector, the inspection of vector presence can be considered as a component of the survey. For example, in a survey for Huanglongbing, caused by '*Candidatus* Liberibacter spp.', which is spread by psyllid vectors, the survey can be split into two components, namely the host plants and the vectors. In these cases, these different components may imply a different structure in terms of definition of the epidemiological units (EFSA, 2023) and risk areas (EFSA, 2023).

The target population may also encompass plant products. Fruit is a potential pathway for spread of *Bactrocera dorsalis* and can also be considered as a survey component.

Even soil or water may represent a survey component. This would, for example, be the case when undertaking a survey for detection of potato cyst nematodes (*Globodera* spp.) and *Phytophthora* spp., respectively. For potato cyst nematodes can linger in the soil even in the absence of a host plant. Likewise, for *Phytophthora* spp. propagules can also occur in watercourses and be detected through water sampling.

In the situation where two or more survey components may be considered, although these surveys should be planned and conducted separately, under certain assumption the evidence obtained in each survey can be combined to provide an overall statement about the evidence for pest freedom (see Section 5.7 for further details).

4.6 Sampling matrix

The sampling matrix is the material that is examined and/or collected from the inspection unit based on the protocol for the specific pest on which the detection method is applied. Examples of sampling matrices include fruits or leaves visually examined and/or collected from a tree, soil samples collected around a plant, a pest specimen captured by a trap or collected from a plant.

4.7 Statistical sample size versus laboratory samples

For the purpose of this guidance document, the term 'sample size' refers to the number of inspection units to be examined in the survey to retrieve sufficient information on the pest presence or prevalence in the total population. This should not be confused with the number of 'laboratory (biological) samples' that are collected from the inspection units for laboratory testing (EFSA, 2023). That is, multiple laboratory samples may be collected from a single inspection unit.

5 Statistical background for sample size estimation

Survey designs should provide a level of confidence that the pest is absent or confidence to make inferences about the pest population characteristics.

The concepts and methodologies for demonstrating freedom from disease have been developed in the context of animal diseases and maximum residue level compliance, respectively (EFSA, 2012). Similarly, the same underlying principles have been applied to develop the concept of pest freedom in plant health. Pest freedom in plant health refers to a condition in which there is a high level of confidence - based on structured, representative sampling and well characterised detection methods - that a specific pest is absent from a defined plant population or area. Although absolute certainty is not achievable, the accumulated evidence from statistically sound and risk-based surveillance, when the pest is not detected, provides strong confidence in its absence. This evidence supports official pest status declarations under Article 92 and Annex IV of Regulation (EU) 2016/2031, which are essential for guiding phytosanitary measures, facilitating safe trade, and ensuring effective pest management within the Union and in international contexts.

The sample size, expressed as the number of inspection units, is calculated based on statistical principles that consider information about the population size (EFSA, 2023), the confidence required, the design prevalence and the sensitivity of the methods to detect the pest. In this context, EFSA developed online software tools (RiBESS+, RiPEST) applying the methodology for sample size calculations in risk-based pest freedom surveys i.e. for detection, delimiting and buffer zone surveys. Pest freedom surveys allow for conclusions in terms of probabilities of pest absence.

The application of an efficient sampling strategy that results in a representative sample (EFSA, 2023) of an adequate size should thus be addressed in the survey design. Potential approaches to conduct representative sampling for surveys are described in detail in Appendix C.

The key parameters needed for pest freedom survey design are:

- target population size. Indicates the number of inspection units (i.e. size of the host plant population) targeted by the survey to which the survey results will apply (EFSA, 2023).
- method sensitivity. This deals with how good the method is at detecting the pest when it is present. Method sensitivity combines sampling effectiveness and diagnostic sensitivity (EFSA, 2023) values.
- confidence level and design prevalence (EFSA, 2023). Both parameters define the strength of the evidence to support the conclusion of the survey.

Confidence level, design prevalence, method sensitivity and target population size are all interrelated. In this chapter, their relationship to each other and their relationship to the sample size and conclusion of the survey are presented with specific reference to plant pests.

When designing a survey, the selected confidence level and design prevalence need to be accepted by the risk managers. Their selection will have a large impact on the usefulness and the reliability of the survey conclusions.

As described in Section 4.1.1, it is important that the structure of the target population is well defined, otherwise the sample may not allow for reliable inferences about the entire population. When designing a survey, the underpinning assumptions related to the homogeneity of the survey area need to be clearly formulated and accepted by the risk managers given that sample size should be addressed independently for each epidemiological unit defined. These assumptions about the structure of the target population within the survey area will also have a major impact on the values of the survey parameters that will determine the sample size for the survey and will therefore have a strong impact on the reliability of the survey conclusions.

The concepts and methodologies for characterising an animal disease and affected population in terms of disease prevalence through surveillance activities were developed (Milanzi et al., 2015; Bourhis et al, 2019; Hester et al., 2015) and integrated in EFSA online software for the calculation of sample size in monitoring surveys (SAMPELATOR) for estimating mean and variance of a population prevalence. Similar concepts and principles are also applicable to the field of Plant Health (Parnell et al., 2017) where the same tool (SAMPELATOR) can be used for characterising the prevalence of plant pests. In addition, EFSA developed a methodology for monitoring the progress of eradication programmes by means of specific surveys. Section 9 addresses the methodologies for designing surveys for estimating pest prevalence in infested areas as well as for designing surveys for monitoring the progress and effectiveness of eradication programmes.

5.1 Conclusion on the pest status

For any type of pest freedom survey, the aim is to detect a pest if it is present above a specified prevalence in a given area. A challenge is that it is statistically and practically impossible to conclude with 100% certainty that a pest is absent, even when it is not detected by a survey[22]. Similarly, if the pest is found in the survey sample, it is not possible to obtain the true pest prevalence in the total population but only to estimate the prevalence with a certain level of accuracy. To achieve absolute certainty on absence, every host plant in an area would need to be examined with an inspection procedure or sampling and diagnostic procedure that has perfect detection ability. Moreover, this would need to be repeated with a high frequency to ensure that the pest has not been introduced since the last survey. Clearly this is not feasible. In practice, it will only be possible to observe a relatively small proportion of the host plants at limited intervals, and with

[22] This excludes situations in which a pest has no host plants in an area or when the climatic conditions do not allow for establishment of a pest. In these cases, it is neither recommended nor necessary to perform a survey. Article 22 of (EU) 2016/2031 provides the legal framework to refrain from surveillance of Union quarantine pests 'for pests for which it is unequivocally concluded that they cannot become established or spread in the Member State concerned due to its ecoclimatic conditions or to the absence of the host species.'

imperfect inspection procedures or sampling and diagnostic tests. Thus, the true absence or prevalence of a pest is uncertain even when a survey does not detect that pest.

To illustrate what can be concluded on the pest status when only a portion of the host plants can be surveyed, consider a population of trees in a given area. The trees are grown in orchards of 1,000 trees and there are 1,000 orchards in that area. Thus, in total there are one million trees in our example. When none of the trees in a survey sample of 1,000 trees shows signs of the pest, this provides a degree of evidence that the pest is not present. However, a sample of 1,000 trees in this area constitutes 0.1% of the entire population and there are still 999,000 trees that have not been inspected. Even if we increase our sample size to 10,000 trees, there are still 990,000 trees that have not been inspected, and which could be infested by the pest. The question thus becomes: how much evidence for pest absence does a sample of 1,000 trees provide? How does the weight of evidence increase if we sample 10,000 trees instead of 1,000?

The probability that at least one infested unit in the survey sample is detected is formally related to the sample size, the pest prevalence in the total population and the method sensitivity (see Section 5.5) of the detection method employed in the survey. Given a certain sample size and no detection of the pest in that sample, it can be concluded with a given amount of confidence that the pest is either absent or its true prevalence lies somewhere between zero and a maximum prevalence. The 'given amount of confidence' is the *confidence level* of our survey and the maximum prevalence that could have been reached is the *design prevalence*. The latter term is used to differentiate it from the true prevalence. The larger the confidence level and the smaller the design prevalence, the stronger the evidence for pest freedom and the more inspection effort is implied.

There is no one-size-fits-all approach to designing and conducting a survey and to determining the required number of samples. Consequently, the selection of confidence level and design prevalence values is a compromise between available resources in a Member State and the level of the risk of false conclusion that risk managers are willing to accept. Sometimes these values are also prescribed in legislation as is the case, for example, for delimiting surveys of *Xylella fastidiosa* in demarcated areas under the current measures laid down in Commission Implementing Regulation (EC) 2020/1201[23].

5.2 Confidence level

The *confidence level* reflects the level of reliability (confidence) of the conclusion of the survey. When it is stated for a given area that a pest is absent or present at a level below the design prevalence with 95% confidence, this means that given the methods and the assumptions taken, the statement is (on average) expected to be correct at least 95% of the time. In general, confidence levels are set at 95%. When setting the confidence level, the risk managers should consider the resources available and the epidemiological situation that might vary in the territory of the Member State.

An increase in sample size leads to higher confidence that the pest is indeed present at least below the maximum prevalence including possible absence) (Figure 6).

[23] Commission Implementing Regulation (EU) 2020/1201 of 14 August 2020 as regards measures to prevent the introduction into and the spread within the Union of *Xylella fastidiosa* (Wells et al.). C/2020/5520, OJ L 269, 17.8.2020, p. 2–39.

Although the above pest freedom approaches, and derivatives thereof, have been widely applied in animal health (Cannon, 2002), their application in plant health is relatively new (Bourhis et al., 2019; Hester et al., 2015; Parnell et al., 2017).

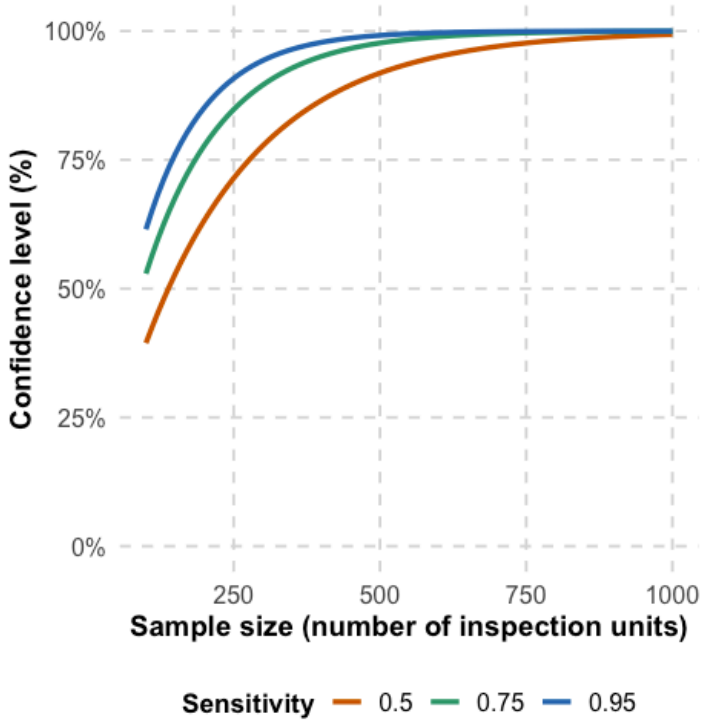


Figure 6: Confidence level increases with survey sample size. The confidence level depends both on the sensitivity of the detection method (red, green, and blue lines) as well as the design prevalence. The graphs are drawn with the design prevalence fixed at 1%

5.3 Design prevalence

The design prevalence is another variable that needs to be set by the risk manager, based on the trade-off between acceptability of the risk and availability of resources. The design prevalence refers to the prevalence of the pest in the survey area that the survey is designed to detect with the confidence requested by the confidence level.

In these guidelines, pest prevalence is defined as the 'fraction of infested inspections units in the total population'. In plant pathology the term 'incidence' is often used to represent this concept. Indeed, in ISPM 5 (FAO, 2024b), incidence is defined as the 'proportion or number of units in which a pest is present in a sample, consignment, field or other defined population.' However, in animal and human epidemiology, incidence is widely used to describe the rate of change of prevalence of infested units over time. Though the use of the term 'prevalence' has diverged in plant pathology (Nutter et al., 1991; Nutter et al., 2006), this use of the term is further confounded given that 'prevalence' is more commonly used in the study of arthropod pest populations. Moreover, the terms incidence and prevalence are often used interchangeably in plant pathology, leading to further confusion in their application. For example, ISPM 6 (FAO, 2018) uses both incidence and prevalence to describe pest populations, and this practice is repeated in other international plant protection standards.

For the purpose of substantiating or maintaining pest freedom, it is desirable to set the design prevalence to a value that accounts for the risk of overlooking the pest when the prevalence is low. This threshold value is usually based on scientific evidence, policy decisions and risk assessment.

Prevalence refers to a proportion of the population, and risk managers should note what this means for both the absolute number of infested hosts and their spatial distribution. That is, 1% of infestation of a small, clustered population has different implications for eradication and control than 1% of a larger widely distributed population.

In general, the higher the selected design prevalence is, the higher the risk that an outbreak remains undetected for a prolonged period of time and the more difficult it will be to eradicate the outbreak once detected. The lower the selected design prevalence is, the larger the survey sample size becomes to reach a given confidence level (Figure 7). The design prevalence of a survey is therefore a compromise between available resources and an acceptable level of the pest remaining manageable.

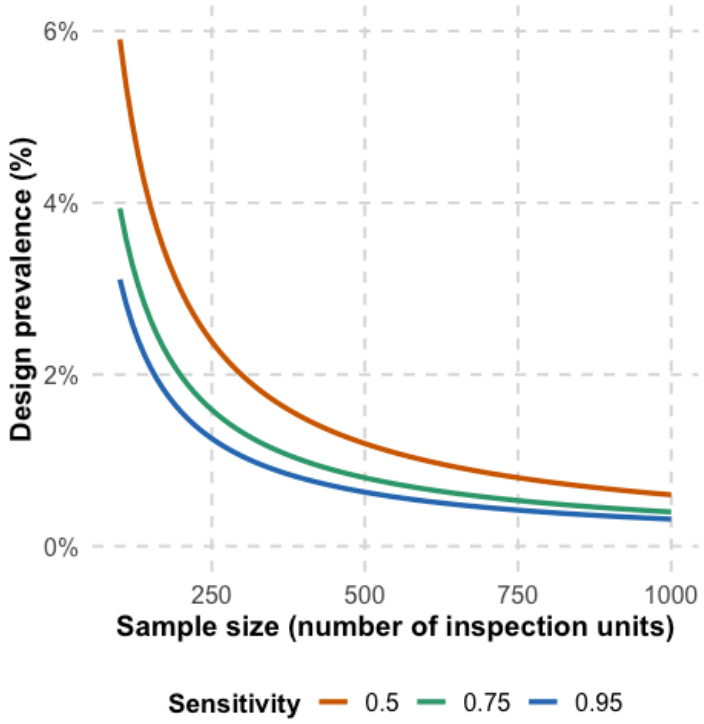


Figure 7: Relationship between design prevalence, survey sample size and sensitivity. The displayed graphs are calculated with the confidence level fixed at 95%

5.4 Selection of the confidence level and design prevalence

As indicated in previous sections, the confidence level and design prevalence must be set *a priori* by the risk managers. For the confidence level, a choice of 95% is commonly applied, implying that when the same survey is run 20 times, supposing that the target pest is present at the design prevalence, it is expected to yield a false negative result (e.g. overlook the pest) once. Lower values of the confidence level thus may lead to more false negative survey results.

With regards to the design prevalence, choosing lower values of design prevalence reduces the risk of false negative survey outcome i.e. overlooking the target pest although present.

However, a lower design prevalence comes at the cost of a higher sample size. The choice of the design prevalence is thus a trade-off between the risk of overlooking the pest being present and the cost of achieving the requested level of confidence. Several variables may influence the choice of design prevalence, as illustrated below.

(i) *Early detection and control*: Quarantine pests are subject to eradication and management measures following their discovery. Such measures are usually more cost effective and have a higher success rate if instigated when a pest population is still small. When there is a reliable estimate of the prevalence beyond which a pest can no longer be eradicated, a design prevalence that is below this threshold is desirable. The challenge may be that the required prevalence to achieve eradication is very low and thus requires an intensive survey effort. The risk manager must then balance the survey costs against the costs of failing to eradicate the pest.

(ii) *Arbitrary selection*: A common choice for the design prevalence is 1%. It is not entirely clear what motivates this choice, but 1% may be an appropriate value in the absence of other information. In this case, it is useful that risk managers consider what this percentage means in terms of the absolute number of infested hosts to ensure that they find this level of risk acceptable. Take, for example, an area that contains 1,000,000 host plants and a survey aiming to have 95% confidence of detecting a 1% design prevalence. Even if the survey has no positive findings, the pest still could be present with up to 10,000 host plants being infested. Thus, 1% seems to be a low design prevalence value, but when placed in an absolute context, the risk may still appear to be unacceptable to risk managers.

(iii) *Available survey resources*: When the amount of available resource for a survey is fixed, then the design prevalence that can be achieved with that available resource can be determined. Suppose there are resources to survey 100 epidemiological units in an area that contains one million host plants then, given information on the method sensitivity (see Section 5.5) (e.g. 80%), the design prevalence that will be achieved with a certain confidence level (e.g. 95%) can be calculated and amounts to 3.7% in this example. This strategy may result in a high design prevalence when resources are scarce and, thus to large absolute number of potentially infested hosts, but at least it is transparent to risk managers and can be used to reassess resource allocation for future surveys.

(iv) *Epidemiology*: Knowledge of epidemiological characteristics of the pest should be considered when selecting the design prevalence and confidence level for a detection survey. The presence and distribution of a pest population in an area is mainly determined by the multiplication rate and the spread rate of the pest species. The multiplication rate is the number of offspring per pest unit per time interval, as determined by several species-specific factors, e.g. the latency period, the longevity of multiplying life stages of the pest (e.g. sporulation stage for fungi, egg production stage for insects) and the survival rate of life stages of the pest. The spread rate is the maximum distance per time interval between the infested zone of a pest and a newly infested host plant outside that area. The spread rate depends on the mechanism (passive, e.g. wind or splash dispersal, or active, e.g. olfactory or visual search for host plants by flying insects).

When both the spread rate and multiplication rate are high (e.g. random wind-driven spread of abundantly sporulating fungi), a large part of the survey area may be rapidly infested, with a distribution over the area that matches the distribution of host plants. In this case, the design prevalence may be set higher than the default value because once the pest has been introduced, the prevalence in the survey area will increase rapidly and so it is expected to be relatively high. At this point, the pest may no longer be eradicated. When the spread rate is low (e.g. when host plants are abundant and pest behaviour is

guided primarily by olfactory and/or visual stimuli) and the multiplication rate is low (e.g. *Anoplophora glabripennis* in cool climates), the pest is likely to remain at a low prevalence, clustered in small parts of the survey area. In this case the design prevalence may be set lower than the default value.

An example for assessment of the eradicable level of pest prevalence is the modelling study by EFSA PLH Panel (2019), where a range of epidemiological scenarios for eradication of *Xylella fastidiosa* outbreaks in olive was investigated. The target population that was modelled constituted an area of 10 km × 10 km (10,000 ha), where 1,000,000 olive trees were planted on a regular grid at 10 m spacing. The modelling demonstrated that in the worst-case scenario (e.g. low vector control) an outbreak with 4,000 infected plants was eradicable. This occurred when the first detection of the pest was three years after the initial infection event. A survey, designed for timely (eradicable) detection of *Xylella fastidiosa* in this target population and epidemiological setting, should therefore be based on a design prevalence of 0.004 or less.

(v) *Substantiation of pest freedom*: A detection survey may be conducted to substantiate pest freedom for an area; for example, as a basis for recognition of EU protected zones (Article 32 of Regulation (EU) 2016/2031) or as a risk reduction option when exporting to non-EU countries. In this case, the survey conclusion which indicates pest presence below the design prevalence is interpreted as indicative of pest absence for the survey area (e.g. the area proposed as EU protected zone or the territory of the exporting country).

Models that combine import risk and potential establishment are used routinely in pest risk assessment by EFSA, Member States and other countries. These models describe the factors that influence the successful introduction of a pest into the EU such as the trade volumes, the risk of infestation at origin and the environmental suitability for establishment at destination in the EU (e.g. EFSA PLH Panel, 2017a, b) which could be considered when setting the confidence level and design prevalence.

(vi) *Delimiting surveys and buffer zone surveys*: Overlooking the pest in a delimiting survey will result in an inaccurate delineation of the infested zone and the failure of the eradication/containment programme. Therefore, the design prevalence for a delimiting survey should be lower than the design prevalence set for a detection survey. In the examples developed by (EFSA, 2020b), when the detection survey aims to detect 1% hosts infected by *Xylella fastidiosa* or *Phyllosticta citricarpa* with 95% confidence, the delimiting survey and buffer zone survey (EFSA, 2023) aim to detect 0.1% of infected hosts with 95% confidence. In the logical sequence of surveys, in areas where the pest is not known to occur, detection surveys are first conducted to confirm the pest-free status. Only once the first infestation is found, are delimiting and buffer zone surveys conducted. As an example, in Table 5 a gradient of design prevalence values is presented depending on the aims of the surveys of *Phyllosticta citricarpa* and the structure of the target population (extracted from (EFSA, 2020b)).

Table 5: Examples of design prevalence for the different types of survey for *Phyllosticta citricarpa* (EFSA, 2020b)

Phyllosticta citricarpa		Agricultural areas	Residential areas
Design prevalence for annual detection surveys	Pest absence confirmation	1%	1%
	Pest freedom in an area neighbouring an outbreak in an agricultural area	0.5%	1%

Design prevalence for delimiting surveys	0.1%	0.1%
Design prevalence for buffer zone surveys	0.1%	0.1%

The above scenarios illustrate that selection of the design prevalence depends on the situation in the survey area and the aim of the survey. As stated before, this value should be selected on a case-by-case basis and there is no universal value that can be applied in all EU Member States or areas within a Member State. For instance, when a pest is considered absent in an area and the risk of it being present or introduced is low, the design prevalence acts as a proxy for zero. In the absence of other information, the standard approach in animal health is to use a design prevalence of 1% at the herd level (FAO, 2014). This value will not suffice when the threat of a pest is more imminent, e.g. in areas adjacent to known outbreak areas where one wants to detect a pest when the population is still small enough to warrant rapid eradication. In outbreak areas – when the infested zone needs to be accurately delimited – an even lower design prevalence is needed to establish where to implement the eradication or containment measures to avoid further spread of the pest. To determine the prevalence of infested host plants at a level where eradication is still achievable depends on a range of interacting factors. These include the size of the infested zone, the environmental conditions, the host availability and distribution as well as the intensity of the eradication measures that will be implemented.

5.5 Method sensitivity

The method sensitivity is defined as the probability that a truly positive inspection unit will be confirmed as positive. The method sensitivity has two components, the *sampling effectiveness*, which evaluates the performance of detection activities conducted in the field (i.e. probability of selecting an infested sample from an infested inspection unit, probability that a truly positive inspection unit will be identified as positive by visual examination or other field detection method) and the *diagnostic sensitivity* (i.e. probability that a truly positive sample will test positive, which is a characteristic of the laboratory test used in the identification process).

The overall method sensitivity can be calculated by:

Method sensitivity = sampling effectiveness × diagnostic sensitivity

The sampling effectiveness depends on the specific characteristics of the pest and the host. For instance, when inspecting a tree for the presence of an insect pest, sampling effectiveness relates to the probability that the pest is indeed found and collected when present; when a trap is placed in an orchard, it relates to the probability that the pest is indeed caught when present; when inspecting a plant for the presence of a virus, it relates to the probability that the symptoms are indeed observed when the virus is present and that the plant material that is collected for laboratory analysis actually contains the virus.

Sampling effectiveness can be dependent on the ability of the inspector (training and expertise), specificity of symptoms, access to the part of the plant on which the pest can be observed (e.g. consider a pest that is present in the top of the tree; sampling effectiveness will be lower when observations are made from the ground than when using a ladder), weather conditions, or the growing stage of the plant. In general, the sampling effectiveness will be high when the pest is easy to observe or efficiently caught and symptoms are clear and specific. In some cases, where a pest has a long asymptomatic

or cryptic period, visual examination is not suitable for detection and asymptomatic material must be sampled for further analysis (Figure 8). Especially when dealing with asymptomatic infections that are unevenly distributed in a plant, the sampling effectiveness will be lower because only random collection of material that appears healthy can be relied on. When establishing sampling effectiveness, it is important to incorporate knowledge from laboratory testing into the inspection protocol. For example, a virus that is present asymptomatically might still have a high sampling effectiveness when the viral load is higher in the young leaves from the top of the plant.

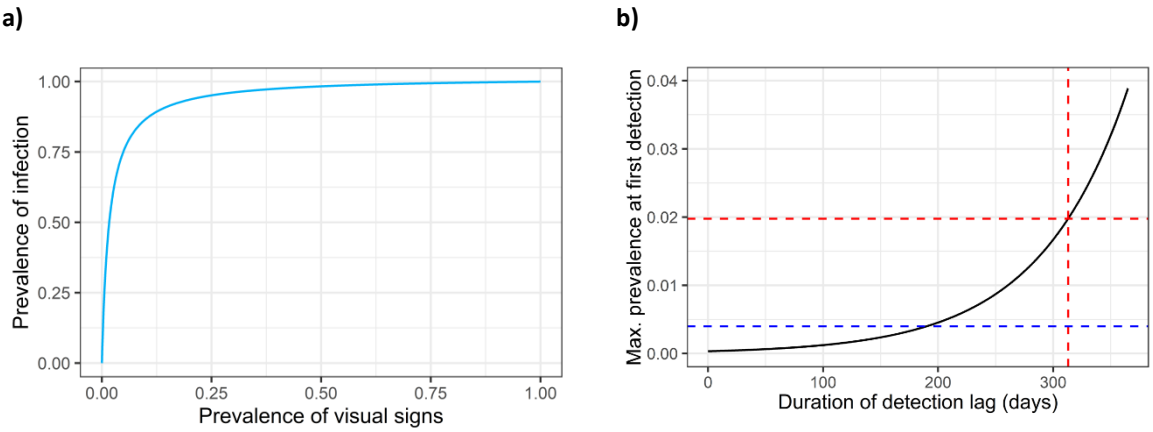


Figure 8: A long asymptomatic period poses a challenge for specific surveys. (a) The relationship between the true prevalence of an infection and the prevalence of visual symptoms of *Xylella fastidiosa* following the first discovery of a local outbreak. (b) The prevalence at first discovery assuming a visual survey inspecting 840 trees per day over a 50-day period. Despite this high inspection and sampling intensity, the prevalence at first detection is 0.02 (2%) of the population (red dashed line), whereas the estimated eradicable prevalence is much lower at 0.0004 (0.04%) (blue dashed line). Calculations are based on epidemic growth rate data from olive orchards in Apulia (Hornero et al., 2020), asymptomatic period data for olive trees (EFSA PLH Panel, 2019) and a mathematical framework linking epidemiological parameters and surveillance (Mastin et al., 2017; Mastin et al., 2019; Parnell et al., 2017). Source: Guidelines for statistically sound and risk-based surveys of *Xylella fastidiosa* (EFSA, 2020a).

Diagnostic methods used to classify the sample as positive or negative for a quarantine pest should be validated before being used. The diagnostic sensitivity should thus be provided by the laboratory performing the tests. For the purpose of sample size calculations, the diagnostic sensitivity is a key parameter. This parameter refers to the proportion of the truly infested samples that are diagnosed as positive relative to the overall number of true positives and false negatives (so, true positives/true positives + false negatives).

The method sensitivity has a direct effect on the ability to detect the presence of a pest and must therefore be considered when estimating survey sample size. For example, when the sampling effectiveness is 70% and the diagnostic sensitivity is 80%, the method sensitivity is 56%. This implies that in 44% of the cases, a truly infested host plants yields

a negative result. Such a number may seem problematic, but this is not necessarily the case, because this can be accounted for by a larger sample size. It is thus recommended that sampling effectiveness and diagnostic sensitivity are improved to high values because it will reduce the sample size and allow for a more efficient use of resources. If the above example of method sensitivity of 56% was applied in combination with a survey aim of 95% confidence to detect 1% design prevalence, the required sample size would be 533 for a population size of 100,000 (calculated using RiPEST or RiBESS+). Improvement of the sampling effectiveness to 85% and the diagnostic sensitivity to 95% is reflected in a method sensitivity of 81% which would result in a sample size of 368.

5.6 Relationship between confidence level, design prevalence and method sensitivity

The design prevalence, confidence level and method sensitivity are inextricably associated. Given a survey in which i) random sampling is applied and ii) each location or sampling unit has the same probability of being selected, the relationship between the method sensitivity (*MeSe)*, confidence level (*CL*) and design prevalence (*DP*) is as follows:

$$CL = 1 - (1 - MeSe.DP)^N$$

with *N* denoting the number of units inspected and/or sampled. On the other hand, assuming that both our sampling effectiveness and diagnostic sensitivity are perfect (thus *MeSe* = 1), the relationship between *DP* and *CL* is simplified as follows:

$$CL = 1 - (1 - DP)^N$$

Therefore, if the design prevalence that we select is high it will not take many inspections/samples to obtain a high confidence level. For example, if we choose a design prevalence of 5%, only 59 samples in an infinite population would suffice to obtain a confidence level of 95%. This implies that when we do not find any cases of the pest in such a survey, we can be 95% confident that the pest is either absent or has a true prevalence anywhere between 0% all the way up to 5% prevalence in the population. On the other hand, the lower the selected design prevalence, the harder it becomes to have confidence that the true prevalence is below that level, and thus a higher sample size would be needed to achieve it (Figure 9). Using the above example, if we choose a design prevalence of 1% (instead of 5%), the 59 samples in an infinite population would result in a confidence level of only 45%. This implies that when we do not find any cases of the pest in such as survey, we can be 45% confident that the pest is either absent or has a true prevalence anywhere between 0% and 1% prevalence in the population.

Furthermore, for a fixed design prevalence, the confidence reached will be further increased as the method sensitivity improves (Figure 9).

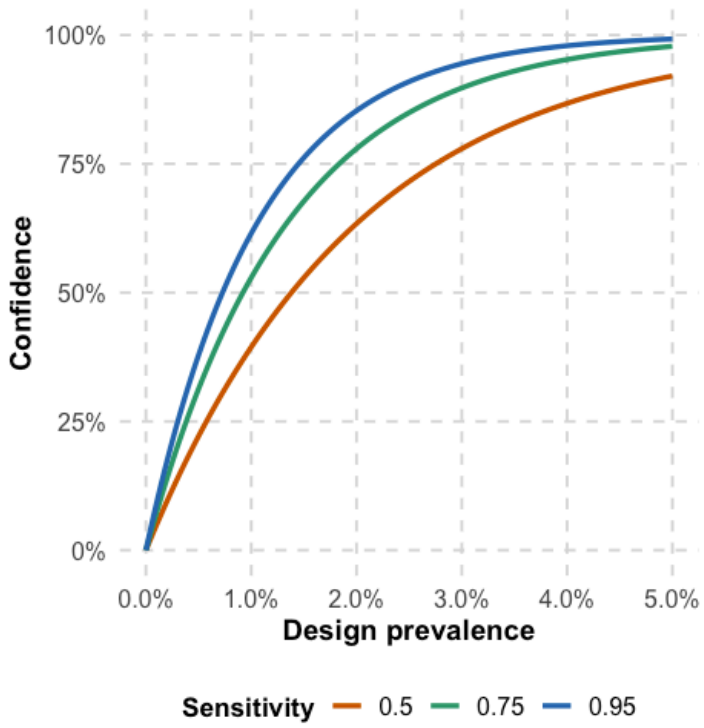


Figure 9: The relationship between confidence level and design prevalence. This relationship depends on the sensitivity of the detection method (fixed at either 0.5, 0.75 or 0.95) as well as survey sample size. This sample size is fixed at 100 in the figure, and the population is assumed to be of infinite size

5.7 Multiple component surveys

We have considered different situations in which the overall conclusion about the pest status of the survey area is constructed from different components (Figure 10). First, if the survey area comprises more than one epidemiological unit and each has to be surveyed (Figure 10 A, B).

Second, different inspection units considered in the survey result in multiple separate survey components. In both situations it is possible to combine the confidence achieved in individual components into a single overall confidence level.

Technically, the approach of combining multiple survey components into the confidence level achieved in survey can be calculated (Cannon, 2002) using the following formula:

$$CL = 1 - \prod_{i=1}^{n} (1 - CLi)$$

CL: overall confidence level of the survey

CLi: confidence level of the survey of component i

n: the number of different components.

This relation can be used to determine the necessary confidence to be achieved in either component to fulfil a confidence level required for the overall outcome of the survey. In the specific example of n = 2 components, CL = 0.95, and equal confidence levels for both

components, the confidence level to achieve within the survey components is obtained as follows:

$$0.95 = 1 - (1 - CLi)^2$$

$$CLi = 1 - \sqrt[2]{0.05} = 0.78$$

This means that when a 78% confidence level is achieved in both components then a 95% confidence is achieved for the complete survey. The general formula can be used to calculate overall confidence level for more than two components with different individual confidence levels achieved. However, the calculation requires that the related design prevalence has the same meaning for all components and therefore all components must consider the same inspection units. This can be satisfied for multiple epidemiological units (Figure 10A, B).

When survey components target different types of inspection units (Figure 10C), combining becomes more complex. In particular, it must be possible to convert the different inspection units into a common unit to which the combined survey conclusion and the design prevalence level can refer. For example, when the inspection unit of the components is either a host plant or a vector of the pest, it is possible to match these inspection units by using the field (or hectare) that harbors these plants/vectors as the shared inspection unit. Conversions need detailed elaboration of the overarching inspection unit and the referred design prevalence and go beyond the objectives of these general guidelines.

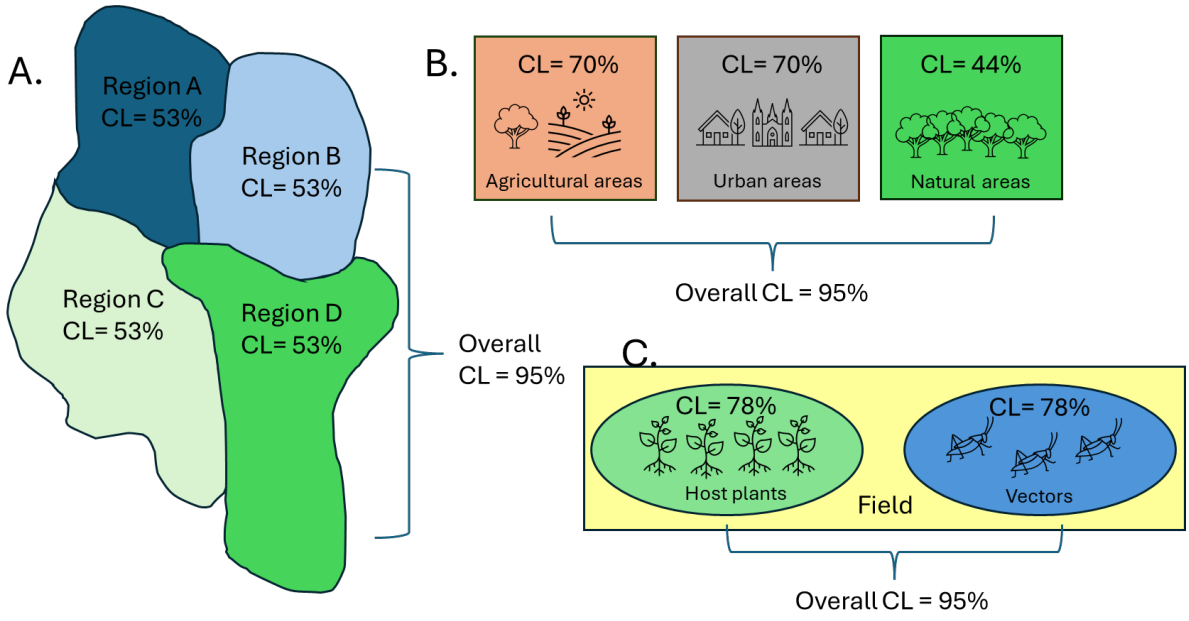


Figure 10: Examples of survey components. The single overall confidence level of a survey can be used to calculate the necessary confidence level achieved in either of the separate survey components e.g. 53% for four components in A. This can be applied when surveys make use of the same inspection unit but are conducted in separate (A) geographical regions or (B) land-use categories. This can also be applied when surveys use different inspection units, for example (C) when conducting separate surveys in host plants and vectors. However, combining these requires that

a unifying inspection unit can be identified, which in the example given would be a field, and the design prevalence can be meaningfully expressed therein.

5.8 Unknown population size

When limited information is available about the size of the target population within the survey area it is still possible to determine the sample size. Figure 11 shows how the estimated sample size for an unknown (statistically infinite – binomial distribution) target population size compares to a known target population size (finite population – hypergeometric distribution), given a fixed confidence level, design prevalence and method sensitivity. The sample size resulting from a hypergeometric and binomial distribution converges for large populations; above a certain target population size the estimates obtained are only marginally different. In other words, above a certain target population size, the population can be considered as infinite from a statistical point of view. In Figure 11, both curves converge around 15,000 host plants on approximately 370 samples. Above 15,000 plants, very few or no additional samples are needed to achieve the same confidence level and design prevalence (E.g. for the survey parameter values assumed in Figure 11: 15,000 – 370 samples; 20,000 – 371 samples; 60,000 – 373 samples). As a consequence, when information about the size of the target population is lacking, the sample size can be estimated using the binomial distribution. This will be at the cost of inspecting more units than needed, but the additional burden is negligible when the actual target population is large.

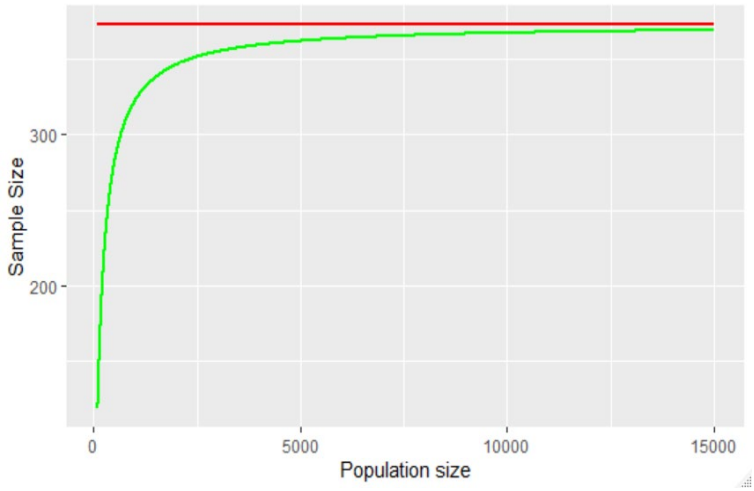


Figure 11: The sample size follows a hypergeometric distribution (green) for finite target population sizes. In the illustrated example, a method sensitivity of 80%, a confidence level of 95% and a design prevalence of 1% are applied. As the population size increases, the sample size converges on the red line, which represents a binomial distribution that assumes an infinite population size

5.9 EFSA RiBESS+ and RiPEST tools

In order to calculate the sample size, EFSA has made available two free online tools, RiPEST and RiBESS+, to support the surveillance programme managers[24]. A detailed manual for RiPEST (Bemelmans et al., 2023) and RiBESS+ are also published[25] as well as webinars to explain methodological framework and statistical tools[26]. Figure 12 shows a screenshot of RiBESS+ for calculating a sample size showing the five above-mentioned input parameters and the calculated output. Additional examples of the application of RiBESS+ are available in the pest-specific guidelines for survey design of *Xylella fastidiosa* (EFSA, 2020a), *Phyllosticta citricarpa* (EFSA, 2020b), and *Agrilus planipennis* (EFSA, 2020c).

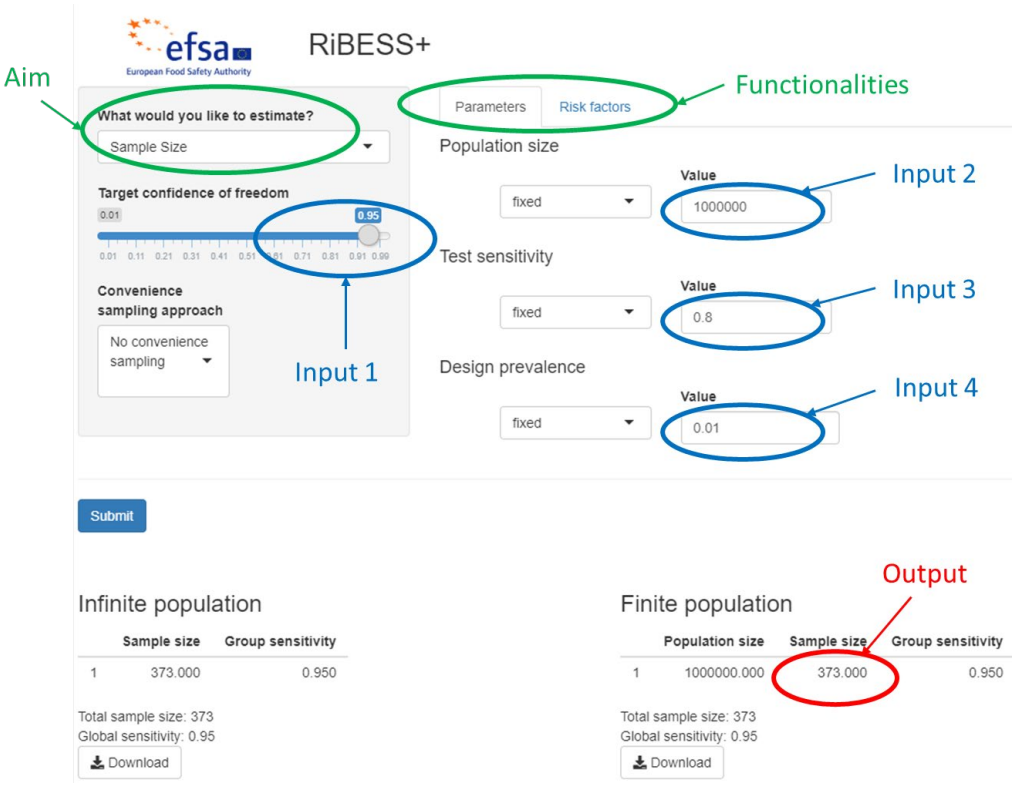


Figure 12: Screenshot of the sample size calculation (Output) using RiBESS+ with a 95% confidence level (Input 1), a 1% design prevalence (Input 4), assuming a population size of 1,000,000 host plants (Input 2), a method sensitivity of 80% (Input 3, combining sampling effectiveness and diagnostic sensitivity) and the risk factors tab (Functionalities). The green circles are the chosen aim and functionality, the blue circles are the input values of the survey parameters, and the red circle is the estimated output

6 Detection survey

[24] Available at https://shiny-efsa.openanalytics.eu/app/ribess

[25] https://zenodo.org/record/2541541/preview/ribess-manual.pdf

[26] https://academy.europa.eu/courses/efsa-plant-pest-surveillance-toolkit

A detection survey is performed to detect the presence of a pest in areas where it is currently not known to occur. Thus, the objective of a detection survey could be to substantiate pest freedom or to detect pests early. The design of the detection survey can be subdivided into several steps that are introduced below.

6.1 Selection of the confidence level and design prevalence

The first step is to determine the aim of the survey. The risk manager needs to decide on the overall confidence level and design prevalence of the survey based on the trade-off between the acceptable level of the risk and availability of resources (see Section 5.4).

6.2 Survey area and target population

In the next step, the size of the target population of the survey (see Section 4.1) should be determined and the target population divided into epidemiological units (see Section 4.2) based on the characteristics of the survey area.

6.2.1 Target population, epidemiological units and homogeneity

Within each epidemiological unit a number of inspection units can be distinguished which could potentially host the pests (see Section 4.4). When designing a survey, the underpinning assumptions related to the homogeneity of the survey area needs to be clearly formulated and accepted by the risk managers. To optimise the survey efforts in terms of the number of samples that represent the host population, it is essential to gather as much information as possible on the homogeneity of the territory and to choose an epidemiological unit size in which the homogeneity assumption is sufficiently realistic. The assumptions made about homogeneity will have a major impact on the values of the parameters that will determine the sample size for the survey and thus have a strong impact on the reliability of the survey conclusions.

Here, four examples are presented to illustrate how epidemiological units may be distinguished within the target population in a survey area. In all examples the target pest is not known to occur.

1 Homogeneous survey area

In this case, the entire survey area is a homogeneous area where the interactions between the pest, the host plants, abiotic and biotic factors would result in the same epidemiology should the pest be present. In other words, each inspection unit has the same probability to be infested. The entire target population of host plants is thus located within the same epidemiological unit. It is then sufficient to plan a single survey. Once a suitable design prevalence and confidence level have been set and the method sensitivity has been determined or estimated, the survey sample size can be calculated using RiPEST or RiBESS+. The resulting number of inspection units can be distributed over the area.

2 Survey area with multiple zones differing in pest epidemiological characteristics

In this case, the survey area is not homogeneous but instead harbours two or more different zones. Within each zone the interactions between the pest, the host plants, abiotic and biotic factors would result in the same epidemiology should the pest be present. However, these characteristics are different between zones. Each zone is then considered as a separate epidemiological unit. For each zone a separate survey is planned. Using RiPEST or RiBESS+ the survey sample size in each zone is calculated and the resulting number of inspection units are distributed within each zone.

A practical approach to implement this scenario is by using the NUTS regions (Eurostat, 2024) as the epidemiological units. Depending on the homogeneity of these areas regarding the pest–host combination under surveillance one can either select NUTS 1, NUTS 2 or NUTS 3 regions as epidemiological units. The different zones could also represent different types of environments, e.g. agricultural vs natural vs urban areas.

This approach is used and further detailed in case studies of the survey designs presented in the pest-specific guidelines for surveys of *Xylella fastidiosa* (EFSA, 2020a), *Phyllosticta citricarpa* (EFSA, 2020b) and *Agrilus planipennis* (EFSA, 2020c).

3 Survey area where the target population is subdivided into epidemiological units

In this case, the survey area is not homogeneous, but instead harbours many smaller units in which the interactions between the pest, the host plants, abiotic factors and biotic factors would result in the same probability of presence. An example would be a pest in an area where host plants are cultivated in an agricultural environment. Within each agricultural field, the epidemiological conditions are homogeneous and plants have an equal probability of being infested. Between fields these conditions may differ. Each field is then considered as a separate epidemiological unit. The survey is designed as a sample of fields from the total number of fields in the area. The inspection units for the survey are individual host plants in the sample of fields. To design this survey, the calculation of sample size is performed on two levels: the level of the epidemiological unit (the field) and the level of the area. At the field level, the number of host plants to be inspected in a field is calculated for a chosen confidence level and a design prevalence, using RiPEST or RiBESS+. This confidence level at field level is considered as the method sensitivity of the field inspection. At the second level of calculation, the number of fields to be included in the survey is calculated with the field confidence level as the method sensitivity, the target design prevalence, and the overall confidence level as the desired confidence level for the survey. As an overall result, the number of fields to be included in the survey and the number of host plants to be inspected per field has been calculated. The total number of host plants to be inspected for the survey area is the product of the number of fields and the number of host plants per field.

4 Survey area with multiple zones differing in pest epidemiological characteristics and where the target population is subdivided into epidemiological units

This is a combination of cases 2 and 3. In this case, the survey area is not homogeneous, but instead harbours two or more different zones. Subsequently, these zones harbour many smaller units in which the interactions between the pest, the host plants, the abiotic and biotic factors would result in the same epidemiology should the pest be present. For each zone a separate survey should be planned and within these zones the sample size calculations would be performed as for case 3.

6.2.2 Risk factors

Within each epidemiological unit a number of inspection units can be affected by a risk factor that, if properly characterised, allows them to be distinguished and grouped by their probability of infection. A risk factor affects the probability that a pest will be present in a specific portion of the target population. It may not always be possible to identify and include a risk factor in the survey design. Risk factors can only be included when both the relative risk and the proportion of the overall host plant population to which they apply are known or can be reliably estimated. The relative risk of the risk area(s) is estimated compared with a baseline (with a relative risk of 1). It is important not to overestimate the relative risk, because this will result in a smaller sample size than is actually needed to meet the strength of the evidence wanted for the survey. To estimate relative risks,

historical information on interceptions, trade volumes and the origins of the host plants and plant products can be considered. A procedure is suggested in the pest survey cards[27] to facilitate the characterisation of a risk factor. To identify risk areas, it is first necessary to identify the activities that could contribute to the introduction and/or spread of the pest. These activities should then be connected to specific locations. Risk areas can be defined around these locations, bearing in mind that their size depends on the spread capacity of the target pest and the availability of host plants around these locations. Table 6 shows an example for the surveillance of *Rhagoletis pomonella*, a fruit fly that affects apples but not known to occur in the EU (extracted from (EFSA, 2020e)).

Table 6: Example of a risk activity and corresponding risk locations relevant for the surveillance of *Rhagoletis pomonella* (EFSA, 2020e)

Risk activity	Risk locations	Risk areas
Imports of apples (and subsequent disposal of damaged fruit) from countries where the pest occurs.	Entry points, packing and sorting stations, and processing industries where such fruit is handled.	Areas surrounding the risk locations where *Malus* and *Crataegus* trees are present.
	Households, fresh markets and waste collection centres where apples are being consumed, sold and disposed of.	Residential areas with *Malus* and *Crataegus* trees receiving homemade compost.

6.3 Inspection method and sample strategy

Depending on the targeted pest and the targeted host plants, the appropriate inspection method and/or sampling strategy will need to be established. In view of having a univocal designation of the detection methods, sampling matrix and the diagnostic methods, EFSA prepared a report on the classification of plant pest detection and identification methods (EFSA, 2024) useful in the preparation and reporting of surveillance activities. The same classification is implemented in the tools (RiPEST and Surveillance database) developed by EFSA. The classification standardises the naming of these methods and provide the basis for developing a system for evaluating their performance. Field and laboratory methods, including trapping were classified at three levels, from general categories to detailed methods.

The main methods are explained below.

6.3.1 On-site visual examination of plants

This method applies to pests of which specimens can be detected directly on plants or pests that cause symptoms on plants which can be detected during on-site visual examination of the plant (Figure 13). Plants may need to be cut to detect specimens or symptoms inside plant tissues, e.g. the presence of larvae of wood-boring insects, discolouration of the xylem vessels or in potato tubers (Figure 14). A procedure for the optimal execution of the visual examination should be available, and if so, is presented in the survey cards. The detected specimens or symptomatic plant parts should be collected

27 https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/index

and sent to an official laboratory for diagnostic identification using identification keys, microscopy or molecular methods.

For visual examinations, it is particularly important to find the balance between the risk of overlooking a pest and the amount of resources spent on a single inspection unit. For example, by increasing the inspection time, the probability of finding the pest will also increase, but at some point looking for something that is probably absent will no longer be worthwhile. Many pests cause symptoms that might be confused with other biotic or abiotic stressors. If only samples are submitted for further identification to the laboratory when showing all typical symptoms, the chances of missing the target pest are increased. On the other hand, when the sample collection is not selective enough the laboratory resources might not be well used. The effort used in a single inspection is reflected in the sampling effectiveness that in turn affects the method sensitivity. It is thus necessary to establish optimal procedures for inspection during the design phase of the survey.

Figure 13: Some pests cause clear symptoms that can be detected during on-site visual examination of the plant. Tomato leaf curl New Delhi virus causes leaf curling, blistering and leaf distortion in young leaves of a field-grown zucchini plant. Chlorosis, vein banding and upwards rolling can be observed on older leaves (Source: Raffaele Giurato, EPPO Global Database, htpps://gd.eppo.int)

Figure 14: The bacterium *Ralstonia solanacearum* causes symptoms which can be detected during on-site visual examination of cut potato tubers. A transversally cut potato tuber shows a brown discoloured vascular ring with creamy white bacterial ooze emerging from the vascular ring (Source: NVWA, NL)

6.3.2 Sample collection for laboratory testing

Sample collection for laboratory testing is an inherent part of the inspection process in the event of suspicious findings and asymptomatic sampling. The material collected from the inspection unit that constitutes the laboratory sample should be sent to an official laboratory for diagnostic identification. When pooling, there needs to be good traceability of the samples and the dilution of the pest in the sample might call for the method to be appropriately adapted.

6.3.3 Detection by trapping

Traps may be employed in detection surveys for pests or their vectors (Figure 15). Depending on the organism, traps may be designed with an attracting factor such as a specific colour, shape, light, a chemical lure, or any combination thereof. The inspection unit in a trap-based survey is the area from which the target pest is attracted or passively collected, usually from a single trap. The sampling effectiveness is determined by the performance of the trap, which may depend on environmental and vegetation characteristics of the survey area, the placement of the traps and the sensitivity, quality and replacement frequency of chemical lures. The ensemble of pests (usually insects or fungal spores) caught by the trap is subsequently taken to the laboratory for identification.

Figure 15: A multi-funnel trap is used to catch *Monochamus* beetles which are known as vectors of the pine wood nematode *Bursaphelenchus xylophilus*, the causative agent of pine wilt disease (Source: Hugh Evans, Forest Research, UK)

6.4 Method sensitivity

Once the inspection and sampling procedures are established, following the procedures indicated by the competent authorities, the method sensitivity can be estimated (see Section 5.5). It might be necessary to determine the number of host plants that need to be inspected or sampled in a single field. Both RiPEST and RiBESS+ can be used to calculate this number when using predefined prevalence level (e.g. 1%) to obtain the confidence level at field level, which can be considered as the method sensitivity of the field inspection. Note that the more units that are inspected per field or the longer a single inspection unit is examined, the higher the confidence level will be. The increased time spent per inspection unit will improve the method sensitivity, which in turn will lead to a smaller sample size (see Section 5.5). However, as shown in Figure 6, there is a point where for a given sample size and design prevalence, the improvement of the method sensitivity has a limited effect on the achieved confidence of the survey. And to increase the confidence achieved at field level it might be more relevant to increase the number of field visits then to improve the method sensitivity. In this case, it will be necessary to balance the survey effort per field considering the overall estimated survey effort.

6.5 Sample size

RiPEST or RiBESS+ can be used to determine the number of inspection units to survey to achieve the objectives of the survey in terms of confidence level (e.g. 95%) and design prevalence (e.g. 1%), while also including the method sensitivity. Table 7 summarises the

input parameters that are needed to calculate the sample sizes using RiPEST or RiBESS+. Remember that when information about the size of the target population is lacking, the sample size can be estimated by assuming an infinite population size.

Table 7: Example of input values of the survey parameters needed for sample size calculations in RiPEST and RiBESS+ for a detection survey without risk factors. These survey parameters would result in a sample size of 373

Survey parameter	Input values
Confidence level	95%
Design prevalence	1%
Target population size	1,000,000
Method sensitivity	80%
Risk factors (optional)	none
Survey parameter	**Output values**
Sample size	373

When risk factors are included in the survey design, the RiPEST and RiBESS+ tool will require both the regular input parameters, but also the number of risk factors, number of levels of risk within each factor, the relative risk for each level, and the proportion of the overall plant population to which they apply.

When including a risk factor, the sample size is split across the different risk categories. Some samples will be taken in high risk-areas, while other samples will be taken in the baseline area. RiPEST and RiBESS+ will aim to ensure the same confidence across all risk categories. Fewer samples are needed to reach this confidence in the high-risk area.

Convenience sampling allows the sampling effort to be prescribed within each risk level according to a 'convenience' criterion. In this case, the survey designer may for instance, decide to sample four times more often in the high-risk areas and twice as often in the medium-risk areas compared to the baseline area. Note that, in this case, within each risk category the selection of inspection units will still be random, but that the proportion of inspection units that are within a high-risk area will be increased. Convenience sampling can only be applied when the size of the target population that is associated with each risk-level is known or can be estimated.

Placing more emphasis on surveillance of the high-risk areas will reduce the overall sample size (Table 8). This will either save resources or allows the surveyor to aim for a higher confidence to substantiate pest freedom. Hence, it will be more cost effective to apply risk-based convenience sampling and is therefore the recommend approach when risk factors have been identified. However, it should be kept in mind that if the risk factor identification or quantitative value of relative risk levels is violated then the outcome of the survey is as violated and the achieved confidence about the survey conclusion may not be met.

Table 8: Sample size and risk-factors. This example uses the same input variables for the survey parameters as in Table 7 and distinguishes two risk levels, being a high-risk area that is associated with a relative risk (RR) of 2 and contains 10% of the target population size (100,000 inspection units) and a baseline area. The table illustrates the effect of including a risk-factor on sample size depending on whether convenience sampling is applied or not.

Survey parameter	Samples in High risk area (RR=2)	Samples in base-line area (RR=1)
No risk factor	373	
One risk factor **No convenience sampling (delimiting & buffer survey)**	102	206
One risk factor **Convenience sampling (detection survey; high-risk:baseline = 2:1)**	164	82
One risk factor **Convenience sampling (detection survey; high-risk:baseline = 4:1)**	184	46

6.6 Sample allocation

Once the sample size is known for the overall survey area and, if applicable, the different epidemiological units or different risk areas, the inspection units should be distributed within the territory of a Member State and selected from the list of available locations.

When there are multiple epidemiological units, there are multiple options for allocating the samples (inspection units). A straightforward option would be to allocate the same confidence to each epidemiological unit ensuring that overall confidence level is satisfied, meaning that the same number of samples is allocated to each epidemiological unit in case the population size is large enough. For example, consider a survey aim of 95% confidence to detect 1% design prevalence, and an 80% method sensitivity; the overall required sample size would be 373 inspection units in a total population of 3 million plants. When these samples are allocated equally over five epidemiological units this would mean that 75 inspection units are to be visited in each epidemiological unit.

Another option when there are multiple epidemiological units would be that the samples are allocated proportionally according to the size of the target population in each epidemiological unit. This requires that the size of the host plant population within each epidemiological unit is known. For example, consider the above-mentioned example of 373 inspection units and five epidemiological units that respectively harbour 1,000,000, 750,000, 500,000, 500,000, and 250,000 host plants. When these samples are allocated proportionally to target population size this would mean that 125, 93, 62, 62 and 31 inspection units are to be visited in these epidemiological units, respectively.

The same procedure can be followed when there are high-risk areas and baseline areas within multiple epidemiological units. The total number of samples allocated to each risk

level should then be further allocated according to the size of the target population in each epidemiological unit.

Next, one needs to consider which data are needed and how these data will be collected, analysed and reported. Based on the needs, the specific instructions for the inspectors will need to be carefully formulated. Note that these activities are not addressed in these guidelines and fall within the remit of the competent NPPOs. However, these instructions should also include the list of equipment that is required to perform the inspections, for collecting, preparing and transporting the samples for further laboratory identification. As an example, Appendix D provides a list that was put together by the Plant Protection Service of the Federal State of Brandenburg.

6.7 Reporting

The requirements for reporting on the survey activities will depend on to whom the report is addressed. In general, the report of a detection survey should contain information on the type of survey, the target pest, the survey area, type and number of inspection units surveyed and whether the pest was detected or not. Because the sample size was determined using a statistical framework, the report should contain information on the level of confidence that the pest is absent. This can be done by including these parameters into the survey conclusion (see paragraph 6.8) or by reporting the parameters (confidence level, design prevalence) as such.

Because the reporting can also be used for documentation and evaluation purposes, it is recommended to include any information that was used when designing the survey. This includes information on the structure and size of the target population (epidemiological units, inspection units, number of inspection units within each epidemiological unit, and number of inspection units per site when the two-step procedure is applied), information on applied risk factors, and information on the applied survey method(s) (sampling effectiveness, methods in the field; diagnostic sensitivity, methods in the laboratory; method sensitivity).

Because the inspection units have a geographical component, tools for spatial mapping can also be used to present or report the survey results.

The reporting module was implemented in RiPEST, where two types of reports are available. One as a text document where all information used for survey design and maps are reported in a tabular form, and as an excel file. The latter, for detection surveys, is consistent with the EC templates for reporting to EUROPHYT.

6.8 Survey conclusion of detection surveys

When the target pest is detected, it can be concluded that the pest is present in the survey area and an infested zone must be demarcated based on this finding (Section 7). In most cases the pest will be absent, not only from the surveyed inspection units but from the full target population, but rather than simply reporting this as a conclusion it is important to consider the strength of the evidence to support it.

When the target pest is not detected it can be concluded that the target pest is absent from the survey area or if it is present its prevalence is below design prevalence. The formulation of the survey conclusion over the whole survey area requires the overall confidence level and the design prevalence to be reported. The basic format of the survey conclusion, when all inspection units are found free from the pest will be that:

'The survey area is free from the pest, based on a survey with a confidence level of XX% and a design prevalence of XX%'.

The survey conclusion allows the evaluation of surveillance activities to be compared across EU Member States, within a country, and from one year to another.

In principle, this formulation includes the original parameters that were set during the design of the survey. Conversely, when the actual inspected sample size differs from the sample size that was recommended during the survey design, the obtained confidence level of the survey should be re-estimated based on the actual number of inspections. This resulting confidence level should then be reported in the conclusion.

When drawing the survey conclusions, additional information can be considered that further substantiates the strength of the evidence for pest freedom, such as the underpinning assumptions made on homogeneity of the survey area, the method sensitivity, and the surveyed host plants.

For example, in Section 6.2, there is a difference between the scenario where it is assumed that the entire survey area is a single homogeneous area and the scenarios where it is assumed that the survey area harbours multiple epidemiological units. The latter scenarios will generally be more realistic, and thus provide stronger evidence for pest freedom (but at the expense of an increased sample size). The conclusion could thus be that:

The survey area was divided into four epidemiological units, and each of the unit of the survey area is free from the pest, based on a survey with a confidence level of 95% and a design prevalence of 1%.'

A second example is where the host plants targeted by the survey are specifically mentioned in the conclusion. This could be the case when less relevant hosts were not included in the target population. The conclusion could thus be that:

The survey area is free from the pest, based on a survey of Pyrus and Malus spp. with a confidence level of 95% and a design prevalence of 1%.

7 Delimiting survey

According to the International Standard for Phytosanitary Measures (ISPM 5), a delimiting survey is "…conducted to establish the boundaries of an area considered to be infested by or free from a pest" (FAO, 2024b). In the context of these guidelines, we consider that a delimiting survey is conducted to establish the boundaries of an infestation. This type of survey is usually carried out after a pest has been detected in an area where it was previously thought to be absent. Such a detection serves as the starting point to determine whether the pest is also present in the vicinity of the detection site. The methodology proposed by EFSA and implemented in RiPEST delimits the potential infested zone which can be further refined by MSs to set the infested zone. Delimitation is usually followed by demarcation, a process in which an area is established where phytosanitary measures are applied aimed at elimination or containment of a pest. In the EU, the establishment of a demarcated area after the official confirmation of a new outbreak of an EU quarantine pest is required by legislation according to Article 18 of Regulation (EU) 2016/2031 unless the pest concerned can be eliminated immediately. As such, the outcome of the delimitation determines how extensively the phytosanitary measures are to be applied. Delimitation may also inform the choice on whether eradication is still feasible, and thus the choice between eradication and containment measures or the abandonment of official measures.

Delimitation may also have consequences for the movement or trade of plants or plant products. A demarcated area usually consists of an infested zone and a buffer zone. The aim of a buffer zone is to prevent the spread of a quarantine pest out of the demarcated area. A demarcated area can also be established provisionally to avoid spread of a pest while the delimitation is ongoing.

Accurate delimitation is important because it affects both the allocation of resources on eradication measures (i.e. the size of the area where eradication measures must be applied) as well as the likelihood of success of the eradication programme (Mastin et al., 2020). If the delimited area is too large, this might result in imposing measures on an unacceptably high number of healthy plants. If the delimited area is too small, this will allow the pest to proliferate and spread further, necessitating more widespread measures later on or resulting in the permanent establishment of a pest. At the same time, the delimiting survey itself is also resource intensive, and a balance will need to be found to allow for a cost-effective procedure. This is eventually a risk management decision. The current guidelines aim to provide technical assistance to design a statistically sound and risk-based delimiting survey to support the MSs with the delimitation of a potential infested zone that corresponds to the smallest area within which the pest is circulating and within which subsequently the MSs can further delimit the infested zone for the demarcation of the area.

There is currently no standardized approach for a delimiting survey and no standardized approach to determine the area that is to be covered by this survey. In general, it could follow and outward, inward, grid-based or transect based-procedure. Combinations of these procedures are also possible. The best procedure can be case-specific and the choice of the procedure may be guided by the biology of the pest or the characteristics of the outbreak (size and extent of the pest population at the initial detection site, type of plant material affected, type of environment, etc.), and by the availability of resources. Note that in current practice, outward and grid-based procedures are often applied when dealing with plant pest outbreaks. Such outward procedures can then also be designed as an iterative procedure of multiple detection surveys. The advantage of an outward process is that surveillance starts in the area that is most likely to be infested, but the disadvantage is that the pest may have already moved well beyond that area, and spreads further while

the surveillance activities are lagging behind. "*Whether the inward or outward process is more effective depends on the certainty about the true position of the leading frontier of the outbreak"* (Sun et al., 2025). Whatever procedure is followed, a key aspect should be the substantiation of pest freedom in the area outside the infested zone.

These guidelines describe an inward approach for delimiting surveys, using an iterative procedure of multiple detection surveys in a potential infested zone. This procedure is also implemented in the RiPEST tool and can be used for one disease focus or multiple foci simultaneously. The inward iterative procedure is aimed at getting ahead of this front. The size of the potential infested zone is informed by the potential local spread of the pest, which has both a spatial and a temporal component, so the area will increase over time as the pest population grows and spreads. In order to define the potential infested zone, the timing of the introduction of the target pest needs to be known or estimated. The outcome of each detection survey will determine the next step in the procedure until pest freedom outside the potential infested area is substantiated. The process has been validated and tested under different epidemiological conditions (Koh et al., 2025, Sun et al., 2025).

7.1 General considerations

Once a new detection of a quarantine pest has been made, it is important to start the delimiting survey as soon as possible to minimise the risk of further spread. This means that typically there will be less time for the planning and design of the survey compared to a detection survey.

EU Member States should take the necessary phytosanitary measures for eradication when a Union quarantine pest is found in their territory. This typically means that the competent authorities need to decide which measures to take and where to apply them. To this end, it is necessary to establish in which area(s) the pest is or might be present. A delimiting survey is the appropriate tool to accurately establish this area. To this end, it is important that the authorities, or agencies that undertake the survey on their behalf, have the jurisdiction to carry out the survey activities. In particular, having access to all (potential) survey sites is an important prerequisite for which there should be a legal framework in place. For a detection survey, it might be acceptable from a practical point of view that accessibility is taken into account as a criterion for the selection of survey sites, but for a delimiting survey this would have an unacceptable impact on the reliability of the survey conclusions.

Moreover, fast and efficient planning of a delimiting survey is facilitated by contingency planning and by having an outbreak action plan in place. The action plan should describe the roles and responsibilities of the entities that are involved in the survey activities, including the roles of the competent authorities, agencies, other public authorities (e.g. states, municipalities), professional operators, and laboratories. The action plan should also describe the necessary field and laboratory equipment, personnel as well as a procedure on how to prioritize the survey activities relative to the regular activities. When the activities need to be scaled up, it should be known beforehand how additional equipment, personnel, and external expertise can be obtained. This also requires a procedure for the decision to scale up and information on how the supplementary budget is obtained. Preparation is enhanced by having good knowledge of the survey area, in terms of the environmental suitability for establishment of the pest, the distribution of host plants, and the presence of risk locations. Preparation is also enhanced by having protocols in place that describe the procedures for visual examinations, sampling, laboratory testing, and record keeping. Very often, this type of information and protocols are also needed for a detection survey and, thus, will be available for most of the potential

target pests, but protocols may need to be tailored to the delimiting survey (e.g., because additional hosts that are to be examined).

Each pest will have a specific life cycle, and knowledge of the epidemiology is needed for accurate delimitation. In particular, knowing which hosts can become infested by the pest and knowing how, and how far, the pest or its vector can spread is important for a delimiting survey. For several pests, this information is presented in the EFSA pest survey cards[28, 29]. The pest survey cards also contain useful information on the procedures for detection in the field and recommendations for sampling and laboratory procedures. The requirements for sampling are often linked to the subsequent tests that are to be carried out in the laboratory. Because a detection will typically be followed by the establishment of phytosanitary measures, official confirmation in a laboratory is usually required. As most of the target pests for a delimiting survey are currently absent in the survey area, it may be necessary to create beforehand a list of laboratories that have the diagnostic capacity for the identification of a given pest.

7.2 Natural and human-assisted spread

Delimitation takes into account the spread of a pest. Spread can occur through natural spread and human-assisted spread. Natural spread depends on the biological characteristics of a pest. Some pests actively spread by flying, walking, crawling or through vectors, while others are passively spread by water, wind, soil particles or plant seeds (Figure 16). The life stage of a pest may also influence its spread capacity. These biological characteristics will determine the spread rate, which can be used to guide the delimiting process. In practise, natural spread can be hampered by geographical barriers (mountains and water bodies) or can be amplified or reduced by environmental conditions (e.g., wind directions, weather conditions, and habitat fragmentation). Human-assisted spread of a pest can for example occur through movement of plants for planting, harvested products, plant materials, agricultural waste products, packaging or storage materials, vehicles or agricultural machines, soil, workers clothing or equipment. These materials can be infested or contaminated by the pest. Human-assisted spread may occur locally or over long-distances.

[28] Pest survey cards are published as part of the Toolkit for plant pest surveillance in the EU. Available online: https://efsa.onlinelibrary.wiley.com

[29] https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/index

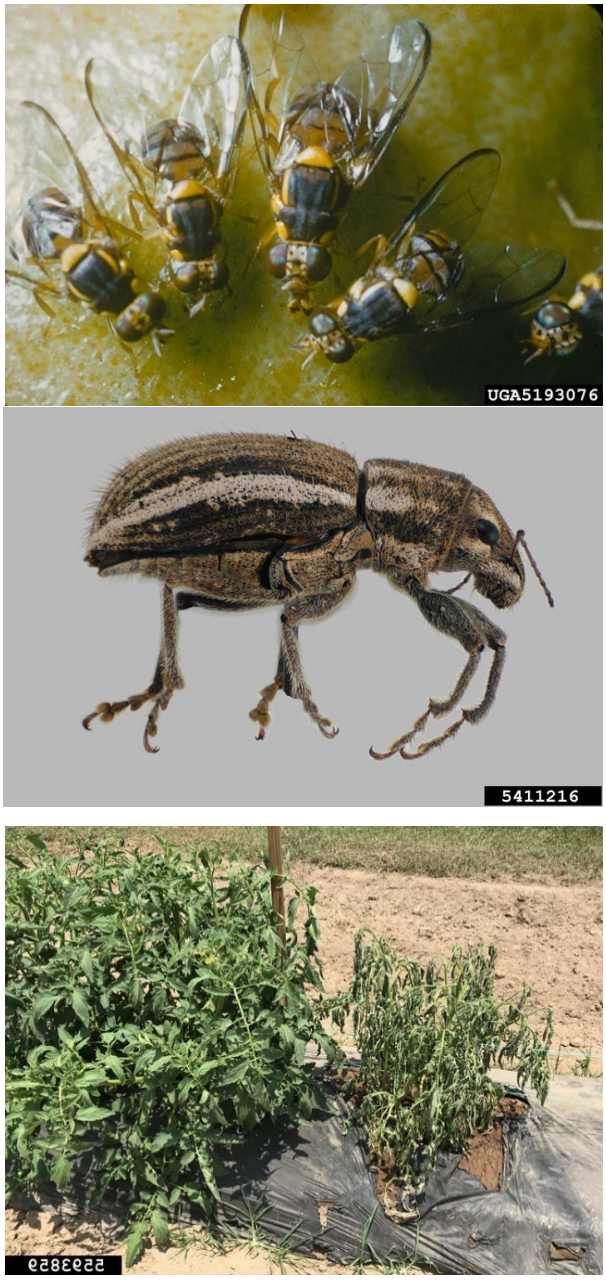


Figure 16: (a) *Bactrocera dorsalis* is a strong flyer that has the capacity to actively spread over multiple kilometres per year, while (b) the spread of *Naupactus leucoloma* is limited by its inability to fly. (c) *Ralstonia solanacearum* cannot actively spread, but can spread to new host plants by water. (Source: (a) Florida Division of Plant Industry, Florida Department of Agriculture and Consumer Services, Bugwood.org; (b) Anyi Mazo-Vargas, University of Puerto Rico, Bugwood.org; (c) wilting tomato plant, Rebecca A. Melanson, Mississippi State University Extension, Bugwood.org)

The delimiting procedure in these guidelines uses the spread rate of a target pest to establish the potential infested zone. This rate considers local spread (by a combination of natural spread and local human-assisted spread related to production practices) and often needs to be estimated. For example, during the EU Priority pests impact assessments (EFSA, 2025), rate of expansion estimates have been established by expert knowledge elicitation. Estimated values for the spread rate are subject to uncertainty, because of the

limited information available and because local conditions at the outbreak site may differ from conditions in locations for which information is available.

Besides the delimiting survey near an outbreak site, it may be necessary to consider the occurrence or long distance spread and perform trace-back and trace-forward activities. Such activities can be run in parallel to the delimiting survey. This is particularly relevant when an infestation is or has become associated with host plants for planting (e.g., seeds, seedlings, cuttings, nursery stock, or potted plants), because this greatly increases the likelihood of establishment. When infested plants for planting are transported, the associated pests move with them, potentially resulting in spread to multiple sites and over long distances. Trace-back activities can be used to determine the origin of an infested plant, while trace-forward activities can be used to find the destination of other potentially infested plants. If tracing activities lead to additional findings in other geographic locations this may trigger additional delimiting surveys.

7.3 Identification of the potential source of infestation

Following the detection of a target pest at a given location (the initial detection site), the first step of the delimiting survey procedure would be to identify the potential source of the infestation. Without human-assisted local spread (see Section 7.2), the source will be either the detection site itself or a site in its immediate vicinity. To this end, one should take into account the background information on the finding (e.g., its geographic location, type of environment, age and viability of the infested host(s), origin of the infested plant(s), pest population size, life stages of the pest), as well as the inspection history of the area surrounding the detection site. In particular, risk locations in the vicinity of the initial detection site should be inspected for the presence of the target pest because of the increased probability of introduction (e.g., harbours, airports, trade hubs, and processing plants). Several scenarios can be distinguished for the identification of the potential source, as illustrated in Figure 17.

    - In the absence of a risk location, the initial detection site itself should be considered as the potential source, and thus as the centroid of the potential infested zone.
    - If the target pest is also detected at a risk location or in its immediate vicinity, this location can be considered as the potential source. When the centroid of the potential infested zone is located at a risk location, the initial detection site may be located near its boundary. A more conservative approach would be to retain the initial detection site as a second potential source and retain the possibility that the target pest has spread from there.

- When the target pest is not detected at a risk location or in its immediate vicinity, the initial detection site itself should be retained as the potential source. However, as a conservative approach, both the initial detection site and the risk location may be considered as potential sources when constructing the potential infested zone. E.g., in the case of garden centres, nurseries, and processing plants, the plant material that caused the initial infestation may have been moved elsewhere between the time of introduction and the time of the initial detection.

When multiple risk locations are in the vicinity of the initial detection site and the target pest is detected at two or more risk locations or in their immediate vicinity, multiple potential sources have been identified. This results in the establishment of multiple potential infested zones.

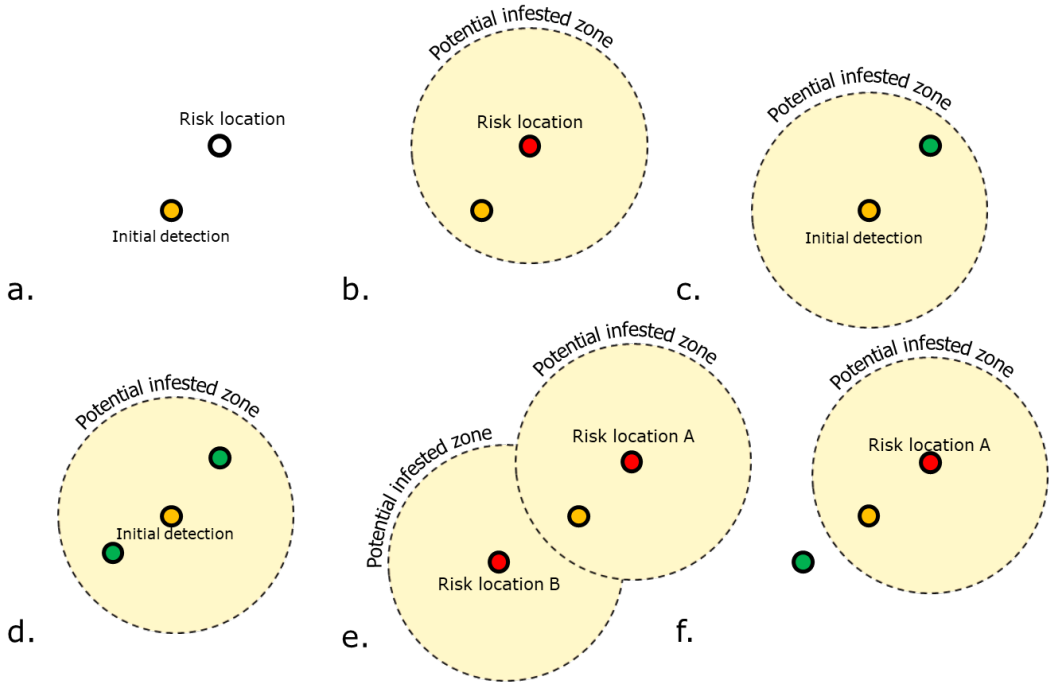


Figure 17: Different scenarios for the identification of the potential source of an infestation in a delimiting survey near a risk location. The large circles represent the potential infested zones. (a) A risk location (white dot) is in the vicinity of the initial detection site (orange dot). After inspection of the risk location, (b) the risk location can be used as the centroid of the potential infested zone when the target pest is detected there (red dot), whereas (c) the initial detection site is retained as the centroid when the target pest is not detected at the risk location (green dot). In case of multiple risk locations in the vicinity of the initial detection site, (d) the initial detection site is retained as the centroid when the target pest is not detected at both risk locations (green dots), (e) both risk locations can be used as centroids of potential infested zones when the target pest is detected at both of them (red dots), or (f) only one risk location becomes the centroid of the potential infested zone when the target pest is detected there (red dot).

7.4 Construction of the potential infested zone

7.4.1 A single potential source

In the next step of the delimiting survey procedure, a potential infested zone is constructed as a circular area around the most likely source of the infestation. The distance from the centroid to the circumference of the circle should reflect the maximum potential natural spread of the pest since the estimated timing of its introduction into the area. This can be based on current scientific knowledge about the spread rate of the target pest and previous surveillance activities in the area.

As indicated in Section 7.2, the natural spread capacity depends on the biological characteristics of the target pest. Detailed information on the spread rate will not be readily available for all pests. Moreover, estimates from literature might vary or need to be

tailored to the specific environmental conditions in the outbreak area. For several EU quarantine pests, the available information on annual spread capacity is presented in the pest survey cards. For a delimiting survey, it is recommended to use the median estimate of the maximum annual spread rate to reduce the risk that the pest has already spread beyond the survey area. When there is no information to estimate the spread rate of the target pest, the spread rate of a related species with similar dispersal capabilities can be used.

Providing that there have been previous detection surveys on the target pest in which the pest was not detected in an epidemiological unit, this means that pest freedom in the outbreak area was substantiated prior to the detection, and the upper bound of the time passed since its introduction can be estimated. Note that this is an estimate because it assumes that the pest would have been detected during the previous surveillance activities if present, which is not necessarily true. When there have been no surveillance activities on the target pest before its detection, the potential infested zone can, for example, be constructed by assuming that the target pest was introduced five years ago. A more conservative approach would result in a very large potential infested zone and although this increases the probability that the pest is restricted to the potential infested zone, the associated surveillance activities would be time-demanding and costly.

The maximum distance that the target pest may have spread is subsequently determined by modelling the maximum annual spread over time based on the number of years that have passed since its (potential) introduction. To this end, a dispersal kernel can be used. In RiPEST, the implemented approach considers an exponential dispersal kernel with a mean spread rate that is taken from a kernel whose 95% percentile equates to the maximum annual spread distance that was used as an input parameter. Using the exponential kernel with the corresponding mean (annual) rate, the multiannual distance limits for the spread of the pest can be calculated and provide the sequence of bands around the centroid of the potential infested zone (Figure 18). The methodological aspects underlying the RiPEST approach are described in detail in Appendix B. This provides the basis for understanding the rationale behind the use of the exponential kernel in the model.

An exponential dispersal kernel is often applied when modelling spread in a disease focus (Hyatt-Twynam et al., 2017; Meentemeyer et al., 2008; Zadoks & Van den Bosch, 1994). It should be noted, however, that the use of the exponential kernel is an approximation. In reality, some pests will require a dispersal kernel with a heavy-tailed distribution, while others will require a truncated distribution. The tail of a heavy-tail distribution extents for a much greater distance, with very low predicted probabilities, than an exponential distribution, and is thus characterised by rare long-distance dispersal events that involve much greater distances than those of the exponential kernels. The circular area around the centroid is considered to be the potential infested zone.

In principle, the potential infested zone could also be constructed using squares, but the use of squares with a side length equal to the diameter of the circle will result in an area that is 27% larger, while this additional area would have a lower likelihood of being infested compared to the area within the circle. For pragmatic reasons, it can nevertheless be useful to apply squares because these are easier to fit onto a gridded map. Throughout these guidelines and in the RiPEST tool, a circle-based strategy is applied.

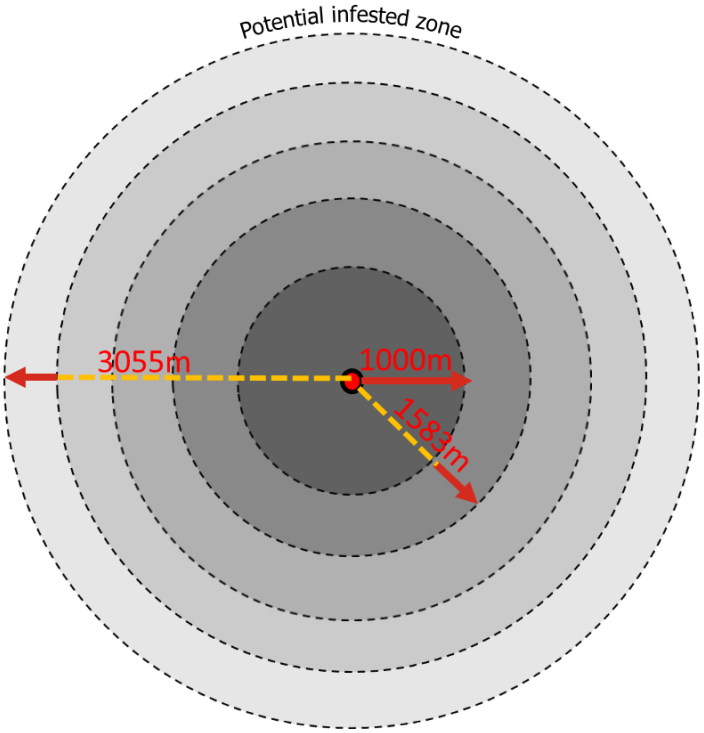


Figure 18: Construction of a potential infested zone for a delimiting survey of a hypothetical pest that has a maximum annual spread rate of 1000 meters. The size of the area is calculated using an exponential dispersal kernel and the number of years that have passed since the last detection survey in which pest freedom was demonstrated. The illustrated example shows the size of the potential infested zone after 1 to 5 years since the last detection survey.

7.4.2 Multiple source locations

When multiple source locations are in the vicinity of each other, the potential infested zones might overlap. The probability of this event increases when the time elapsed since the estimated introduction is longer or when the spread rate is high. When potential infested zones overlap, they can be merged into a single potential infested zone (Figure 19). There are two approaches for merging, namely by simply combining the areas covered by the initially defined potentially infested zones or by building a new circular potential infested zone around the centroid of the merged area. In the latter case, the merged area would need to encompass all initially defined potential infested zones. This latter approach results in a larger potential infested zone compared to the first approach and would thus generally result in more time-demanding and costly surveillance activities, but increases the probability that the pest is restricted to the potential infested zone. Once merged, the combined potential infested zone can be treated as a single entity in which the stepwise surveys are carried out.

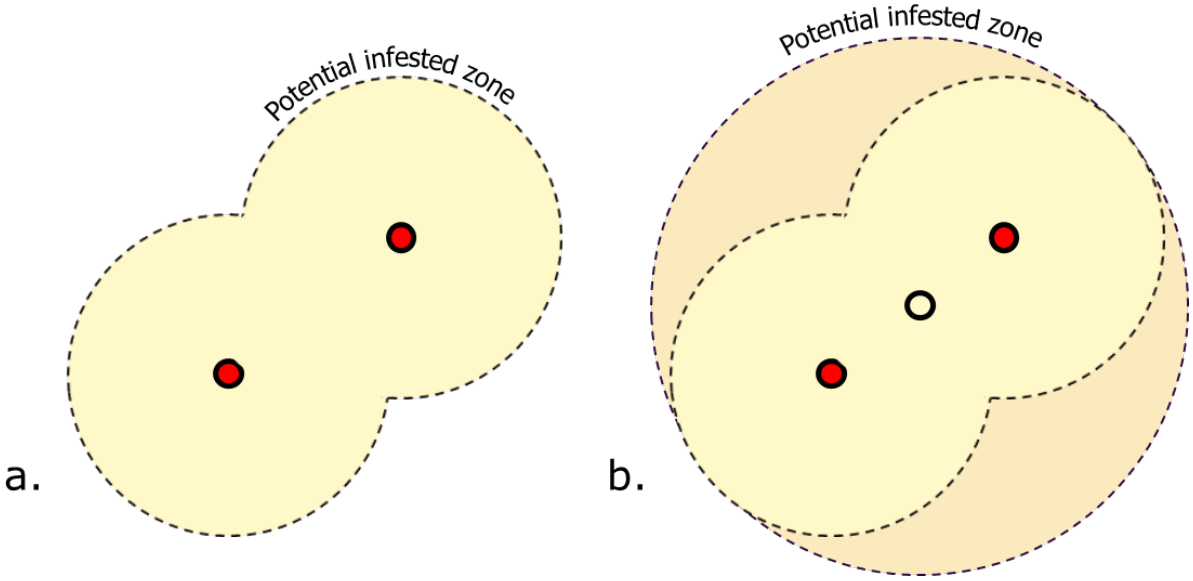


Figure 19: Graphical representation of the potential infested zone when dealing with multiple source locations that are in close vicinity. In case of overlap, the potential infested zones around individual findings can be merged (a) by combining the original areas, or (b) building a new potential infested zone around the centroid of the merged area that encompasses the original areas.

7.5 Stepwise surveys in the potential infested zone

When the potential infested zone has been constructed, this area can be subdivided into multiple bands. Each band represents the maximum distance that the pest may have spread in one year, starting from the source (Figure 20). Again, the exponential dispersal kernel is used to calculate the size of the area over time, and corresponding width of the bands. As a consequence, the width of the bands decreases when moving further away from the source, but this decrease becomes less pronounced in the outer bands. It should also be noted that the area covered by the bands increases when moving further away from the source.

Around the potential infested zone, one additional band is added in which pest freedom should be substantiated if the assumptions regarding the spread rate and time of introductions were met, and to account for ongoing dispersal of the target pest. The size of this band is calculated based on the exponential dispersal kernel by assuming one additional year of spread. The outer band of the potential infested zone is surveyed first. Within this band and the subsequent bands, the methodology for a detection survey can be applied, so a confidence level and design prevalence need to be set to determine the sample size within the target population of that band (see Section 7.6). Once the surveillance of the outermost band is completed, two situations can arise, depending on whether the target pest was detected or not (Figure 21).

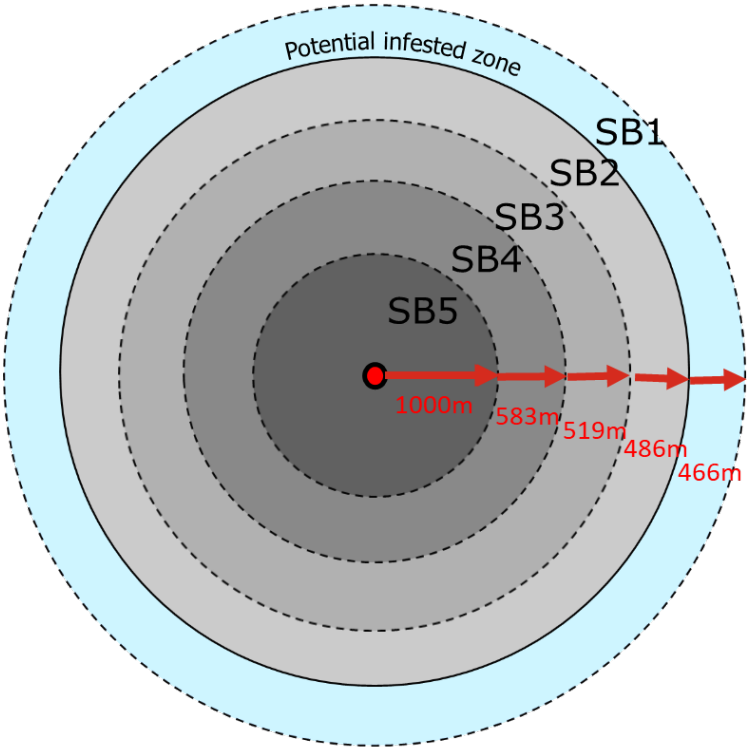


Figure 20: Subdivision of a potential infested zone for a delimiting survey of a hypothetical pest that has a maximum annual spread rate of 1000 meters, and for which the upper bound of the time passed since its introduction was four years. The potential infested zone is divided into survey bands based on an exponential dispersal kernel. Each of the five survey bands (SB1 to SB5) represents one year of spread starting from the potential source of the infestation. Note that the outermost survey band is located outside the potential infested zone.

7.5.1 Narrowing down the potential infested zone

When the target pest is not detected in any of the visual examinations or samples taken during the inspections in a survey band, it can be concluded that the target pest is absent from that band with the defined confidence level and design prevalence. The delimiting procedure then continues inwards to the adjacent band for a subsequent detection survey (Figure 21). This iterative procedure continues until the target pest is detected in one of the bands. Following a detection, it can be concluded that the area outside that band is still free of the target pest, whereas that pest is circulating in the area within. When the delimiting survey does not yield any new findings besides the initial one, it can be concluded that spread has not (yet) occurred.

7.5.2 Enlarging the potential infested zone

If the target pest is detected in the outermost survey band, it must be concluded that the established potential infested zone was too small. There could be several explanations for this. For instance, this could simply reflect variation around estimated means, an inaccurate estimate of the timing of introduction, an inaccurate estimate of the maximum spread rate of the pest or an incorrect localisation of the source of the introduction. The

delimiting procedure should then continue outwards by enlarging the potential infested zone by adding a new survey band around the band that was surveyed first (Figure 21). The width of this band is again calculated based on the exponential dispersal kernel by assuming one additional year of spread. The iterative procedure continues until a survey band is reached in which the pest is not detected. It can then be concluded that this band and the area outside is still free of the target pest, whereas that pest is circulating in the area within.

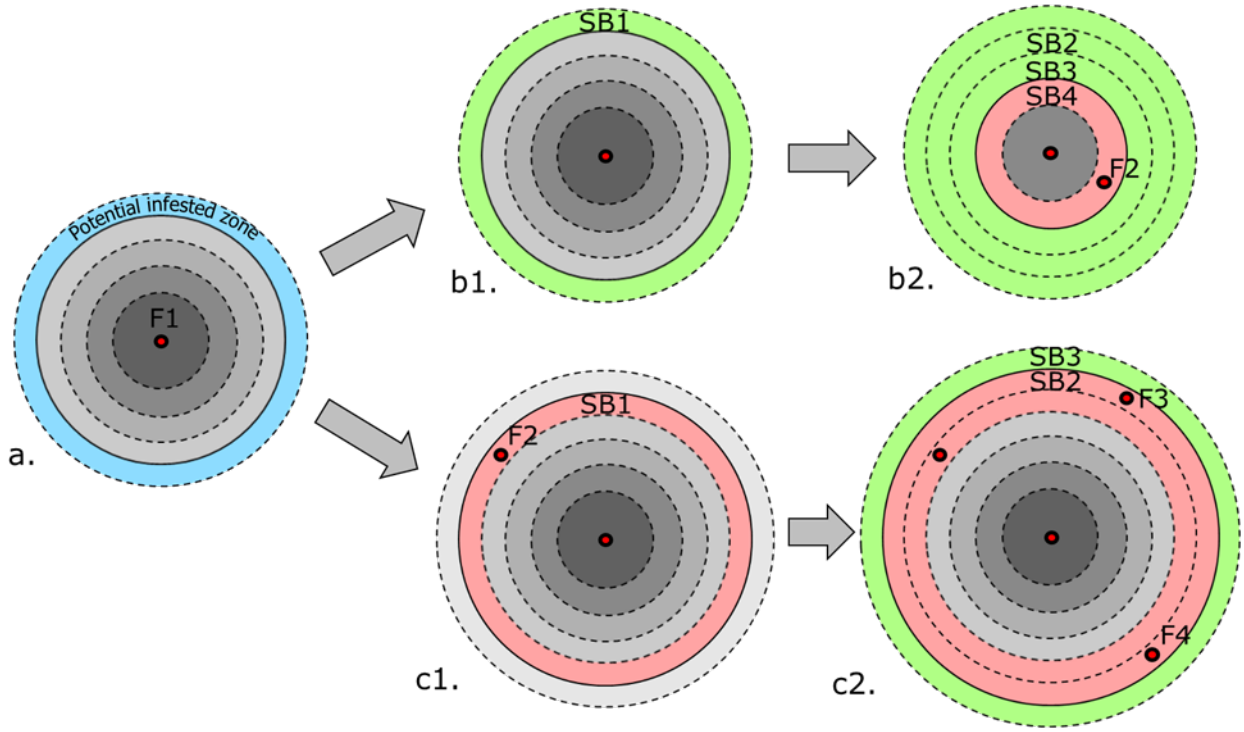


Figure 21: Delimiting survey strategy: a) A potential infested zone is established around the initial detection site (F1) and subdivided into multiple survey bands. An additional band is added on the outside which is then surveyed first. (b1) If the target pest is not detected in this band (SB1), the process continues inwards to the next survey band until a positive finding is made. (b2) For example, after clearing survey bands SB2 and SB3, the target pest was detected (F2) in survey band SB4. The area outside this band has thus been substantiated to be free from the target pest. (c1) If the target pest is detected in the outermost band (SB1), the process continues outwards by adding a new survey band until a survey band is found to be free from the pest. (c2) For example, after subsequent findings (F3 and F4) in survey band SB2, the target pest was not detected in survey band SB3. The area outside SB2 has thus been substantiated to be free from the target pest.

7.6 Survey parameters for a delimiting survey

7.6.1 Target population size

To calculate the sample size of the delimiting survey(s), one needs to determine the size of the target population in each survey band. All potential host species need to be identified

to establish the number of inspection units, and one needs to know at which sites these host plants can be found. This information can be obtained from a variety of sources, including official records on agricultural crops and forestry, public databases, aerial images (in the case of trees or field crops), information from local growers or property owners, or field inventories.

In general, the size of the potential infested zone will be relatively small compared to a detection survey that is used to substantiate pest freedom, and there will be no need for subdivision of the target population. However, the potential infested zone may already be quite large when the target pest has a high spread rate, has been present for several years, or has been found at multiple detection sites prior to the delimiting procedure. Then, there may be a need to structure and subdivide the target population (see Section 4.1) by defining multiple epidemiological units or considering risk factors.

During a delimiting survey, it will also be necessary to consider the minor and incidental host plants in the potential infested zone to ensure that the pest is not circulating in these hosts. By contrast, during a detection survey for pest freedom substantiation, the target population can typically be restricted to the main host plants, as the target pest is more likely to be detected on these hosts (Figure 22). However, for polyphagous pests there might still be a need to focus on those host plants that have a higher likelihood to be infested in a delimiting survey as well.

Figure 22: In a delimiting survey, additional hosts may need to be surveyed relative to a detection survey. E.g., the primary host for detection surveys in the EU on *Rhagoletis pomonella* would be (a) cultivated apples, whereas (b) *Crataegus* species should be included in the delimiting surveys as well. In the case of the cerambycid *Oemona hirta*, detection surveys could target host plants on which damage has been reported relatively frequently, such as lemon, mandarin, (c) sweet orange, persimmon, apple, poplar and gorse, while various (d) *Prunus* species should be included in delimiting surveys as damage by this pest has occasionally been reported on these species (Source: (a,b) Martijn Schenk, (c) *Citrus sinensis* orchard, Elena Lázaro (UV), (d) *Prunus persica* orchard, EPPO Global Database, courtesy of Ilya Mityushev).

In general, it will be possible to obtain accurate records for the target population size for agricultural hosts and trees, but this will be much more challenging for weeds and other small flowering plants. Accurate records are more likely to be obtained for open and cultivated spaces than for natural areas and gardens. The task of determining where host plants can be found becomes increasingly challenging as the size of the potential infested zone increases. For these reasons, it may become necessary to estimate the size of the target population. Estimation has some disadvantages. When the target population size is estimated, the exact locations of all host plants are not known, and one should take care not to introduce bias when allocating the sample to specific survey sites. When the size of the target population is underestimated the sample size will be too small and the required confidence level will not be met. This effect is more pronounced in small target populations.

7.6.2 Method sensitivity

The method sensitivity depends on the procedures and protocols that are applied in the field and the laboratory. The explanations provided in Sections 5.5 and 6.3.4 on detection surveys also apply to a delimiting survey.

7.6.3 Confidence level and design prevalence

As explained in previous sections, the confidence level and design prevalence need to be set by the risk manager based on the trade-off between acceptability of the risk and availability of resources. Because this will be the same for each survey band, the same values for confidence level and design prevalence should be used throughout the iterative process.

As stated in Section 5.2, confidence levels are generally set at 95%. By contrast, when choosing the design prevalence for a delimiting survey, it should be taken into account that the pest should be detected at such levels that eradication and containment strategies can still be implemented. An accurate delimitation of the infested zone is critical for the implementation of the eradication or containment programme. The higher the selected design prevalence, the higher the probability that the pest is overlooked in a survey band. Following the delimitation of the potential infested zone, the NPPO will demarcate the area. Once overlooked, the pest is eventually likely to be detected during surveys in the buffer zone, but in the meantime, precious time to eradicate or contain the outbreak will be lost.

In general, the design prevalence for a delimiting survey should be much lower than the design prevalence for a detection survey. In the examples developed by EFSA on *Xylella fastidiosa* (EFSA, 2020a) or *Phyllosticta citricarpa* (EFSA, 2020b), the delimiting surveys aim to detect 0.1% of infested hosts with 95% confidence, whereas the detection surveys aim to detect 1% of hosts infected with 95% confidence.

7.7 Sample size

The number of inspection units in the survey sample can be calculated using RiPEST or RiBESS+. Both tools calculate the sample size based on the confidence level and design prevalence, taking into account the method sensitivity. Table 9 gives an example of the input parameters for calculating the sample size of three survey bands using RiPEST or RiBESS+.

When the potential infested zone is large, the assumption of homogeneity may not be met, and there might be a need for a more complex design in which the survey sample is split across different epidemiological units and/or risk categories. RiPEST and RiBESS+ can be used to calculate this by treating the survey of an individual band as a separate detection survey.

Table 9: Example of the input values for the survey parameters that are used to calculate the sample size in RiPEST or RiBESS+ for three survey bands (SB) with a decreasing target population size.

Survey parameter	Input values		
	SB1	SB2	SB3

Confidence level	95%		
Design prevalence	0.1%		
Target population size of the survey band	200,000 host plants	20,000 host plants	2,000 host plants
Method sensitivity	80%		
Calculated sample size	3,716 inspection units	3,477 inspection units	1,941 inspection units

7.8 Sample allocation

Once the sample size of a survey band is known, the survey sites should be selected within that band from the list of available sites of inspection units. To avoid selection bias, all available sites need to be equally likely to be surveyed, and thus a random selection of the actual sites is recommended. The calculations in RiPEST or RiBESS+ are based on random sampling. It should be noted that random selection may sometimes lead to the selection of sites that are not easily accessible and may lead to clustered pests being overlooked.

In cases where the target population in a survey band has been divided into multiple epidemiological units and/or risk categories, the number of inspection units should be allocated to the subdivisions that have been defined before randomly selecting the survey sites within each subdivision. As described in Section 6.4, there are multiple options for allocating the samples.

7.9 Survey implementation

Before the survey implementation, one needs to consider what information is needed by the inspectors to carry out the inspections on the target pest. One also needs to consider which data are needed and how these data will be collected, analysed and reported. Based on the needs, the specific instructions for the inspectors will need to be carefully formulated. Typically, this type of information for inspectors will be provided in the form of a protocol or information sheet. This information should be concise and practical for the user. For example, such a protocol or information sheet would contain information on the host range, and basic information on life stages, symptoms and morphology. In particular, photographs or drawings showing typical symptoms on the main hosts and the morphology of the pest itself (if visible to the naked eye) should be included. When necessary, information on similar pests or pests causing similar symptoms can also be included. At the same time, the protocol would e.g. need to include information on how to select the inspection units within in a field or other survey location, how to inspect the host plant, how to collect samples, or how to place traps. Note that the development of such sheets is not addressed by these guidelines and falls within the remit of the competent authorities in the different Member States. Nevertheless, a lot of the key information can be retrieved from the pest survey cards.

Some flexibility on the design is needed during the implementation of the survey in the field. Particular survey sites that were selected during the design phase may no longer be available due to unforeseen circumstances. For example, because the host plants or crops have been removed or harvested in the meantime. If this is the case, it is recommended to replace these sites with new randomly selected survey sites in order to be able to

achieve the overall confidence level that was set by the risk manager. In addition, new information may become available that could trigger a re-evaluation of the survey design, e.g. reports of additional findings, results from tracing activities, or the identification of a more likely source of the infestation.

7.10 Survey conclusion for delimiting surveys

The delimiting survey is completed once the boundaries of the infested zone have been established. Part of the conclusion will consist of a map of the infested area, while the other part should indicate the strength of the evidence to substantiate the claim that the area outside the infested area is free from the pest.

Similar to a detection survey, the formulation of the survey conclusion on a cleared survey band, requires reporting of the overall confidence level and design prevalence. This allows for a conclusion that:

> ***'The survey band is free from the target pest, based on a survey with a confidence level of [X]% and a design prevalence of [X]%.'***

The survey conclusion allows for the evaluation of delimiting activities and comparison across EU Member States, within a country, and between outbreaks. The reported design prevalence is a proxy for pest absence that is decided upon by the risk managers when setting the aims of the survey. As soon as one positive finding is made in a survey band, this proxy of absence no longer holds.

In principle, this formulation includes the original parameters that were set during the design of the survey. Conversely, when the actual inspected sample size differs from the initially calculated sample size, the obtained confidence level of the survey should be recalculated based on the actual number of inspections.

This methodology delimits the potentially infested zone defining the boundaries of the area where the pest is circulating within which the MSs can further scrutinise the area to set the official boundaries of the infested zone that needs to be demarcated.

7.11 Reporting

The requirements for reporting on the survey activities will depend on to whom the report is addressed. The information generated during a delimiting survey will often be needed for pest reporting purposes.

The report of a delimiting survey should include information on the type of survey, the target pest, the initial detection site, the size of the survey area, the surveyed host plants, the number of surveyed inspection units and whether the pest was detected at additional sites or not. Because the sample size was determined using a statistical framework, the report should contain information on the level of confidence that the pest is truly absent in each of the bands that were cleared of the target pest by reporting the confidence level and design prevalence.

Because the reporting can also be used for documentation and evaluation purposes, it is recommended to include all information that was used when designing the survey. This includes information on the structure and size of the target population (epidemiological

units, inspection units, and number of inspection units within each epidemiological unit) and information on the applied detection methods (methods in the field, methods in the laboratory, method sensitivity).

The findings, survey bands, inspection units and the delimited area that results from the survey activities have a geographical component. Therefore, tools for spatial mapping can also be used to present or report the survey results.

Similarly to detection surveys also for delimiting survey the reporting module was implemented in RiPEST, where two types of reports are available (i) a text document where all information used for survey design and maps are reported in tabular form, and (ii) an excel file summarising all survey parameters and results that is aligned with the EUROPHYT template developed for the EU MSs reporting obligation of surveillance activities.

7.12 Follow-up

Based on the outcome of the delimiting survey, a demarcated area will generally be established, consisting of an infested zone and a buffer zone.

The infested zone will then be subject to phytosanitary control measures to minimize the probability of spread of the target pest out of the demarcated area and to eradicate the pest, whereas the buffer zone will be subject to surveillance activities. This buffer zone should be surveyed annually to ensure that the pest has not spread from the infested zone (see Section 8). If, based on the outcomes of the delimiting surveys, eradication is no longer considered feasible, the phytosanitary control measures can be aimed at containment.

The exact size of these zones is either defined by legislation or by experts based on the spread capacity of the pest (or its vector, if applicable). In cases where spread has not occurred and is not expected to occur, the infested zone can be limited to the initial detection site.

8 Buffer zone survey

In ISPM 5 (FAO, 2024b), a buffer zone is defined as 'An area surrounding or adjacent to an area officially delimited for phytosanitary purposes in order to minimize the probability of spread of the target pest into or out of the delimited area, and subject to phytosanitary or other control measures, if appropriate.'

According to Regulation (EU) 2016/2031, the extent of the buffer zone should be appropriate to the risk of the pest concerned spreading out of the infested zone naturally or by human activities in the infested zone.

A buffer zone survey is a particular type of detection survey. Although the aim of the survey is also to substantiate pest freedom, as the buffer zone surrounds and is adjacent to an infested zone, the hypothesis is that the pest might spread from the infested zone. Therefore, although the approach and the concepts developed for detection surveys are also applicable here, the aim is to detect the pest at much lower levels of population (e.g. 0.1% design prevalence and 95% confidence).

Once the boundaries of the infested zone are established, a buffer zone is defined around it. Ideally, its width is set according to the upper range of the yearly spread capacity of

the pest. Intensive surveillance is needed in the buffer zone to ensure the pest remains contained within the infested zone where an eradication programme can be implemented. If such a survey finds infested host plants in the buffer zone, delimiting surveys should then be conducted to establish the new boundaries of the infested zone, and the buffer zone should be extended accordingly.

The buffer zone could also be subdivided into different areas considering risk factors such as the proximity of the boundaries of the infested zone. The relative risk could be estimated using the spread capacity information of the pest, if available. For example, in (EFSA, 2020b), the outer band, 800 m wide, that surrounds the infested zone, which has been surveyed and found free from *Phyllosticta citricarpa*, was assumed to have double the risk of infection as the rest of the buffer zone.

Detailed examples of buffer zone surveys are set out in the guidelines for statistically sound and risk-based surveys of *Xylella fastidiosa* (EFSA, 2020a) and for *Phyllosticta citricarpa* (EFSA, 2020b). RiPEST and RiBESS can be used to support the design of buffer zone surveys.

9 Monitoring survey

According to the International Standard for Phytosanitary Measures (ISPM 5), a monitoring survey is defined as an "ongoing survey to verify the characteristics of a pest population" (FAO, 2024b). Monitoring surveys, also referred to as pest prevalence estimation surveys, are distinct from other survey types. In particular, differ from detection surveys, which aim to determine whether a pest is present in an area where the pest was not previously known to occur, and from delimiting surveys, which are carried out to define the boundaries of an infested zone following the detection of a new pest and to demarcate a buffer zone around it.

Thus, monitoring surveys are conducted to verify the characteristics of a known pest population in an area, based on a sample of inspection units. Under Regulation (EU) 2016/2031, such surveys may be used, for example, to verify low mean pest population densities in areas under containment measures, to track the reduction of the mean pest population density during long-term eradication programmes, or to maintain areas of low pest prevalence (ISPM 22) (FAO, 2005). In these guidelines, pest prevalence refers to the 'fraction of infested units in the total population' (see Appendix E.3 and Glossary (EFSA, 2023)).

Methods for monitoring pest populations and estimating prevalence are well established in plant health (Binns et al., 2000; Madden et al., 2007). Historically, these approaches have focused on endemic pests, where to estimate prevalence is important to inform decisions on control strategies. For example, determining whether pest density has reached a threshold that justifies pesticide application or assessing the likely impact of a pest in relation to prevailing environmental conditions in a given year. In contrast, methods for detection surveys of invading (i.e., non-endemic pests) pest populations have received considerably less attention (Parnell et al., 2017). A key methodological difference between a monitoring survey and a detection survey is that the former relies on representative surveys, and the latter on a targeted survey.

Methods for conducting representative sampling in monitoring surveys include simple random sampling, stratified random sampling, cluster sampling and multi-stage sampling.

These approaches are described in Appendix E. The EFSA SAMPELATOR tool[30] assists the design of monitoring surveys. In addition to estimating pest prevalence and its change over time, monitoring surveys can also provide data to estimate the spatial distribution of a pest. Although beyond the scope of the current guidelines, such estimation can be achieved using geostatistical approaches, which rely on statistical models to interpolate pest population data from a sample (Bouwmeester et al., 2012; Charest et al., 2002; Franke et al., 2009; Stonard et al., 2010; Tubajika et al., 2004), or through species distribution or niche models, which predict pest occurrence based on the similarity of environmental conditions to areas where a pest is known to occur (Bosso et al., 2017; Narouei-Khandan et al., 2016). Furthermore, monitoring surveys play an important role in providing spatial and temporal data for parameterising epidemiological models (Parnell et al., 2017). Well parameterized epidemiological models can be used to predict future spread and assess the effectiveness of pest eradication or other control strategies (Hyatt-Twynam et al., 2017).

9.1 Monitoring surveys in infested zones

Following the detection of a plant pest in a given area, there may be a need to eliminate that pest from that area. Except in situations where the conditions for establishment of a viable pest population are not met, eradication will generally only be possible through the active enforcement of mandatory phytosanitary measures. By measuring pest prevalence over time, it is possible to assess whether these measures are effective and whether the aim of eradication is being achieved. To this end, the competent authorities must establish clear criteria to determine whether eradication measures are effective or when eradication has been achieved. They must also determine in advance the required level of confidence in the survey results is sufficient.

The results of the pest prevalence surveys can be used to substantiate tightening of phytosanitary measures (when the prevalence does not decrease or decreases slower than expected), to declare successful eradication (when the evidence provides sufficient confidence on the absence of the pest), or to switch to a containment strategy (when eradication is no longer considered feasible).

The EU legal framework for protective measures against plant pests, Regulation (EU) 2016/2031, requires that, when a Union quarantine pest is detected in a Member State where the pest was previously not known to occur, the competent authority shall immediately implement all necessary phytosanitary measures to eradicate that pest from the area concerned (Article 17). When the competent authorities decide to establish a demarcated area, in which the eradication measures are applied (Article 18), there is an obligation to carry out annual surveys (Article 19). These surveys should cover both the infested zone and the buffer zone. Buffer zone surveys are needed to confirm that the pest has not expanded beyond the known infested area. The demarcated area can be lifted only once pest absence in the infested zone has been confirmed for a sufficiently long period.

As noted previously, while methods to monitor pest populations and estimate prevalence are well established in plant health (Binns et al., 2000; Madden et al., 2007), there is no standardized approach for applying these methods to plant pests that are subject to eradication or containment programmes, nor for comparing prevalence estimates over time. This section aims to address this gap by presenting possible methods for pest

[30] The SAMPELATOR tool is freely available with prior registration at: https://shiny-efsa.openanalytics.eu/

prevalence surveys. Further details on the proposed methodology can be found in Appendix E, which also highlight some of the advantages and disadvantages of these methods. Appendix E illustrates how these methods can be applied in practice and evaluates their performance through several simulation studies. Additional details are provided in the appendices, including the central sample size formula (Appendix E.1.1.), and the simulation results (Appendix E.2,1.).

9.2 Survey design

9.2.1 General considerations on monitoring of eradication process

An eradication process generally includes three main activities that focus either on surveillance, containment and treatment (ISPM 9) (FAO, 1998). Surveillance activities aim to determine the distribution and prevalence of the target pest at the onset of the eradication programme (e.g., through delimiting surveys or tracing activities), while the programme is ongoing (e.g., thorough buffer zone surveys), or to evaluate its effectiveness (e.g., through monitoring surveys). Containment activities are implemented to prevent the spread of the pest from the infested zone (e.g. by establishing a buffer zone or by imposing restrictions on the movement of infested/infected plants, infested/infected plant products, or contaminated materials). Treatment activities are carried out to eliminate any findings of the pest.

For monitoring the effectiveness of eradication measures of the pest, a well-defined infested area should be established at the start of the monitoring survey and remain fixed throughout the eradication period. Eradication measures for plant pests typically involve the removal of host plants. Such removal may also occur when pest infestation leads to plant death. However, host removal is not necessarily limited to infested plants; for example, when a clear-cutting zone is imposed around infested plants, healthy plants may also be removed. The removal of host plants introduces a challenge for monitoring pest prevalence, as the target population is not static. Any new pest findings outside the originally defined infested area should trigger a redefinition of the infested area. When a finite population correction factor can be applied in the survey design, the sample size might need to be adjusted accordingly (see Appendix E).

To measure a characteristic of a population over time, either longitudinal surveys or repeated cross-sectional surveys can be used. Longitudinal surveys in which the same units are followed throughout the survey period, require less extensive sample sizes. However, when eradication measures involve the removal of host plants, sample units are withdrawn from the targeted population. As a result, longitudinal surveys are not possible for monitoring eradication programmes. In such cases, repeated cross-sectional surveys must be applied, where the sample elements at time $t+1$ are different from and independent of those selected in previous sampling occasions $0, 1, \ldots, t$.

9.2.2 Definitions

The **target population** is defined as the set of individual plants, commodities, or vectors within the survey area in which the target pest can be detected (EFSA, 2023). The size of the target population corresponds to the number of inspection units (e.g., hosts) within the survey area. Typically, the target population is determined at the design stage of the preceding delimiting survey. The area where the pest is circulating should be demarcated to define where an eradication programme is to be implemented. The demarcated area should consist of an infested zone and a buffer zone (Regulation (EU) 2016/2031, Article 18). A delimiting survey should have been conducted to establish the boundaries of the infested zone, i.e. the area within which the pest is confined.

The objective of the monitoring survey is to assess the expected decline in the **prevalence,** $\pi(t)$, over time. Following the the EFSA Glossary (EFSA, 2023), pest prevalence corresponds to the fraction of infested units in the total population, that is, the probability that a randomly selected inspection unit from the target population is infested at time t. Prevalence is estimated from the start of the eradication programme at time $t = 0$ until a predetermined end point $t = T_E$, and measured at regular intervals (e.g., annually).

Even though the survey area of a monitoring survey is typically of limited size (assuming the target pest is still considered eradicable), it is generally not feasible to examine the entire target population with a detection method that has a method sensitivity of 100%, and at a frequency high enough to warrant that no new infestations have occurred since the last survey. Consequently, in practice, the true absence of the pest (i.e., a prevalence of zero) cannot be proven. Instead, the survey is designed to reach a specified design prevalence $\pi(T_E) = \pi_{DP}$.

9.3 Methodological framework

To monitor how pest prevalence changes over time, the proposed methodology considers a **logistic regression model.** This model describes how pest prevalence ($\pi(t)$) evolves throughout the eradication programme. If the programme is effective, pest prevalence is expected to decrease over time.

The model is expressed as:

$$\text{logit}(\pi(t)) = \log\left(\frac{\pi(t)}{1-\pi(t)}\right) = \beta_0 + \beta \times t,$$

where $\boldsymbol{\pi(t)}$ is the pest prevalence at time t; $\boldsymbol{\beta_0}$ corresponds to the logit of the **initial level of infestation** at t =0; $\boldsymbol{\beta}$ is the **rate of change in prevalence** over time in logit scale (expected to be negative during an effective eradication programme) and $\boldsymbol{t}$ = time (years: 0, 1, 2, …, T_E).

The key parameters used in this methodology are listed in the Table 10.

Table 10: Parameters used in logistic regression model.

Parameter	Definition	Typical_Value
π_0 (Initial Prevalence)	Starting pest prevalence in infested zone	Estimated from delimiting survey or from an additional survey if the sample size is not sufficiently large
π_{DP} (Design Prevalence)	Target final prevalence (risk manager defined)	0.001 - 0.01 (very low)
T_E (Time Horizon)	Eradication programme duration	3-5 years
n (Sample Size)	Number of inspection units per survey	Calculated statistically to detect the trend
β^* (Trend Parameter)	Minimal expected rate of prevalence declines to reach the design prevalence within the duration of the eradication programme	Negative value
α (Significance)	Type I error rate (falsely deciding the minimal decline is reached)	0.05 (5%)

Parameter	Definition	Typical_Value
Power	1 - Type II error control (falsely deciding the minimal decline is not reached)	0.90-0.95 (90-95%)

To evaluate the effectiveness of an eradication programme, it is expected to detect a **negative rate of change** ($\beta < 0, \beta^* < 0$) in pest prevalence. This can be formulated as a statistical test:

$$H_0: \beta = 0\ (no\ decline) \quad \text{versus} \quad H_1: \beta = \beta^*\ (decline\ expected),$$

where H_0 represents no change and H_1 represents the expected decrease according to the programme objectives.

This test is the **basis for determining the required sample size** and for evaluating whether the observed decline matches the expected rate of decline.

The **expected rate of decline** (β^*) is determined by two key quantities: the **initial prevalence** at the start of the eradication programme ($\pi(0) = \pi_0$) and the **design prevalence** at the end of the eradication period ($\pi(T_E) = \pi_{DP}$) :

$$\beta^* = \log\left(\frac{\pi_{DP}(1-\pi_0)}{(1-\pi_{DP})\pi_0}\right)/T_E$$

This expression defines the annual **rate of decrease in prevalence** that must be achieved to reach the target within the eradication period.

Once $\beta^*$is known, it is used to calculate the **sample size** (n) required to reliably detect the expected decline with a predefined **level of significance** α and **statistical power** (Demidenko, 2007):

$$n \approx \left(\frac{z_{1-\alpha}+z_{\text{power}}}{\beta^*}\right)^2 V,$$

where ***V*** is the variance of β^* and z_{prob} denotes the critical point (quantile) of the standard normal distribution corresponding to the probability $prob$, which in the formula are $1-\alpha$ or power.

In practical terms, this ensures that the sampling design is capable of reliably detecting the expected decline if the eradication programme is progressing as intended, allowing risk managers to assess whether control efforts are effective or need adjustment.

The methodology (more details can be found in Appendix E) briefly presented here can be further enhanced by:

- using fractional polynomials instead of linear trends,
- incorporating interim evaluations to adjust sample sizes,
- adjustment for finite population,
- design effect to account for heterogeneity,
- adjustments to account for method sensitivity.

The proposed methodological framework is summarized in the flow chart below, which outlines the main steps at the start, during intermediate sampling, and at the end of the eradication programme.

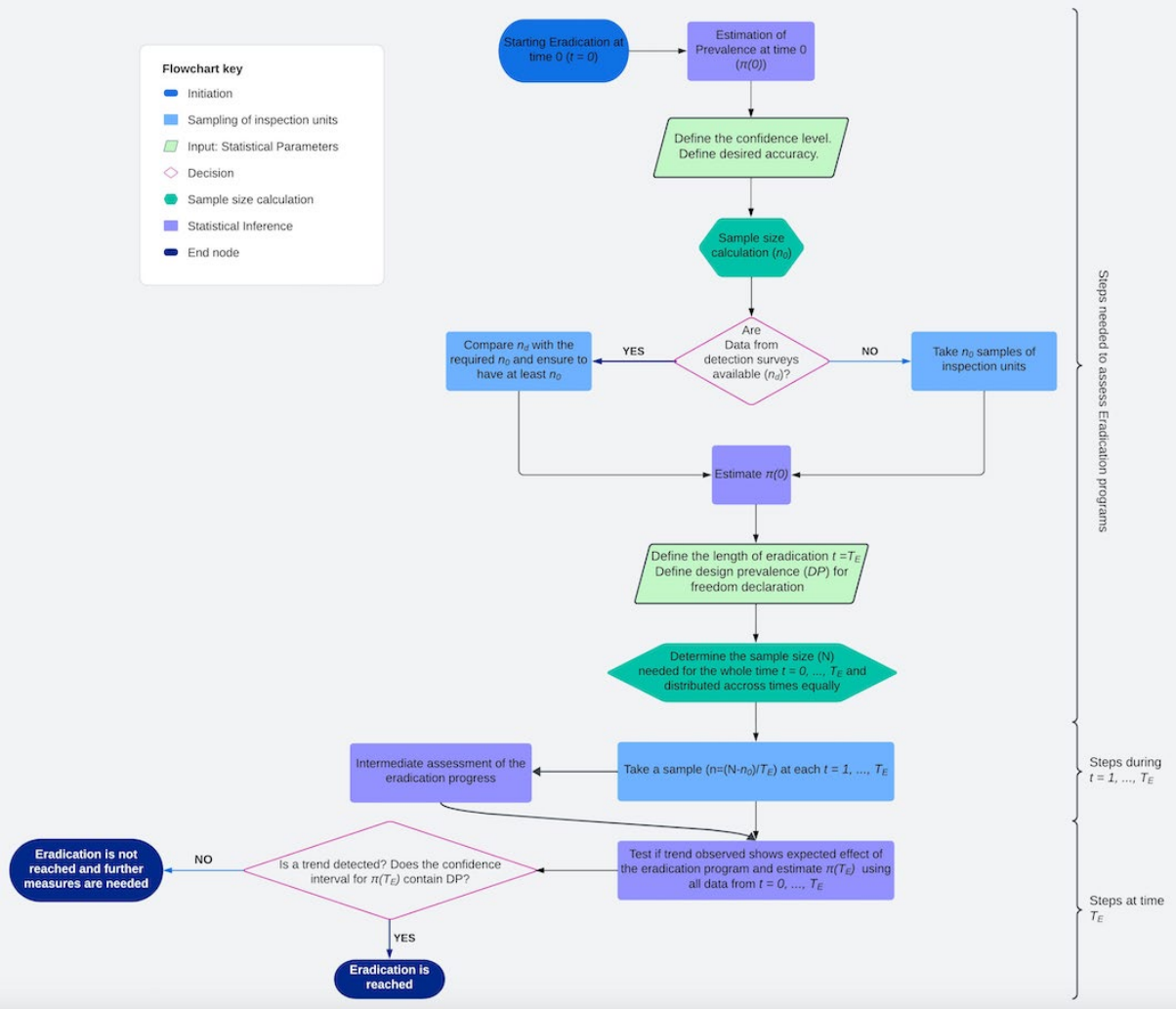


Figure 23: Flow chart of methodological framework for monitoring survey

This methodological framework provides EU-compliant, scientifically reliable methods for monitoring pest eradication programmes.

Key Benefits:

- Regulatory compliance with EU Regulation 2016/2031
- Statistical rigor through established hypothesis testing
- Objective decision-making based on quantitative evidence,
- Resource optimization via efficient survey design
- Transparent assessment with clear performance metrics.

Expected Outcomes:

- Improved effectiveness of eradication programmes
- Reduced subjectivity in performance assessment
- Enhanced credibility with stakeholders and regulators
- Better resource allocation and programme management
- Stronger scientific foundation for policy decisions.

The framework enables authorities to move beyond subjective assessments toward quantitative, evidence-based evaluation of eradication measure effectiveness. Ultimately

it strengthens pest management outcomes while ensuring full compliance with EU phytosanitary regulations.

10 Plant pest survey Toolkit

The EFSA Plant Pest Survey Toolkit has the purpose to assist EU Member States in the preparation and design of surveys for quarantine pests. It includes (i) documents relevant for all pests (e.g. General guidelines for statistically sound and risk-based surveillance - the current document); (ii) pest-specific documents, such as pest survey cards for the preparation of surveys[31] and several pest-specific and crop-based survey guidelines; (iii) the Risk-based Pest Survey Tool for designing the surveys (RiPEST), (iv) the multi-pest survey optimization tool OptiPest (v) the statistical software tools RiBESS+ and SAMPELATOR that are used for the calculation of sample sizes and (vi) the relational database as a repository of information needed for planning plant pest surveys. All these resources can be accessed through the index of the EFSA Plant Pest Survey Toolkit[31].

10.1 Pest survey cards:

The EFSA pest survey cards guide the EU Member States to gather the relevant information for the preparation of surveys of quarantine pests in the EU that conform to current international standards and EU regulation. Pest survey cards contain up to date information on the pest taxonomy, regulatory status, distribution, biology, plant hosts, potential establishment in the EU, factors associated with increased risk for entry and spread, and detection and identification methods. Pest survey cards also include all necessary information for preparing risk-based surveys (e.g., target population, epidemiological units, and inspection units). Pest survey cards are available as "story maps", an easy-to-use interactive format available online[31].

10.2 The Risk-based Pest Survey Tool (RiPEST):

RiPEST is an interactive expert system to guide users to plan and execute a statistically sound and risk-based survey on plant pests. Detection surveys (substantiating pest freedom of an area/country), delimiting surveys (defining the boundaries of a potential infested zone) and buffer zone surveys (monitoring the effectiveness of measures in or around the infested zone) can each be planned in three steps: preparation, design and implementation. The design can be adapted to the user's needs and is supported by information that is partially prefilled from the relational database. In RiPEST, the application RiBESS+ (Risk-based estimate of system sensitivity tool) applies statistical methods for estimating the sample size, design prevalence (achieved design prevalence), global (and group) sensitivity (achieved confidence level), and probability of pest freedom.

10.3 Relational database (RDB):

The EFSA RDB for pest surveys is a query tool and repository of information needed for planning a plant pest survey. In the RDB, information is divided into separate subject-based tables using table relationships to bring the information together as needed.

[31] https://efsa.europa.eu/PLANTS/planthealth/monitoring/surveillance/index

Separate tables include information on: host, pest, inspection unit, detection method, trapping method, laboratory testing, sampling matrix, asymptomatic period, timing, pest-vector, spread capacity, risk factors, risk locations and risk areas as well as relative risk levels. In the RDB specific queries can be defined to filter data from the tables upon user-defined needs. The tool is used to support (i) the crop-based multi-pest survey approach OptiPest, (ii) the single pest survey design using the RiPEST tool, and (iii) queries as a standalone search tool to find specific information for host (crop) x pest combinations on the sequence of operations to be performed from the field to the laboratory.

10.4 Multi-pest Optimization Tool (OptiPest):

OptiPest is a tool to support the optimisation process of surveys designed for multiple pests of the same host or crop. The implemented algorithm is designed to optimise the allocation of resources (number of inspection units to sample) for pest surveys in crop inspection protocols. Given the constraints of limited sampling (or testing) capacity per month, different sampling matrices (e.g. fruits, shoots) and the need to minimise redundant sampling, the current version of OptiPest aims to minimise the simultaneous inspection effort for multiple pests of a crop in terms of reducing the total number of month during which field visits should be planned and the total number of inspection units to be examined, while satisfying the survey requirements for each pest in the host or crop.

The RiBESS+, RiPEST and OptiPest tools and the RDB are all available on the r4eu platform (https://r4eu.efsa.europa.eu/).

11 Conclusions

At the request of the European Commission, EFSA has prepared these general guidelines for designing statistically sound, risk-based surveys on plant pests to assist the EU Member States to carry out the different types of survey that are required under Regulation (EU) 2016/2031.

These guidelines for plant pest surveys are part of the EFSA toolkit for pest surveys[2] that is being developed to support the Member States in the preparation and the design of the surveys as well as to facilitate their implementation.

This document describes the context in which the surveys are designed (legal, international standards, scientific knowledge), the basic principles and approaches that are implemented for the surveillance of EU quarantine pests and introduces the surveyor to the requirements for the design of statistically sound and risk-based surveys. The concepts of general and specific surveillance are introduced. Three specific types of survey are described: detection surveys to substantiate of pest freedom in an area; delimiting surveys to determine the boundaries of an infested zone; and monitoring surveys, to estimate prevalence, that can be applied in infested zones where the progress of eradication measures needs to be observed or where the confirmation of a low pest prevalence is required.

The survey parameters are defined and their interrelations described. During the survey design, the survey parameters need to be set. The survey design should start with setting the aims of the survey, deciding on the overall confidence level and design prevalence of the survey, based on the trade-off between the acceptable level of the risk and availability of resources. These two parameters together determine the strength of the evidence to support the conclusion of the survey. Detailed information on host plant distribution in the

survey area is needed to determine the size of the target population and its division into epidemiological units based on the homogeneity of the area. By including risk factors, surveys will target those areas where the probability of finding the pest is higher. Determining the structure and size of the target population involves scientific knowledge on the epidemiology and detailed information of the local, regional and national landscapes. The method sensitivity is estimated by integrating the sampling effectiveness in the field and the sensitivity of the diagnostic method applied in the laboratory. The more precise and accurate the information used for selecting or estimating the survey parameters, the more reliable the conclusions of the survey will be. Considering that in the EU the survey is implemented at Member State level, and that the data required for preparing the surveys are available at Member State or even regional level, the developed approach should be tailored to each specific situation in terms of host plants and resources. Therefore, it is essential that the assumptions made in estimating each one of the survey parameters are well formulated and accepted/recognised by the competent authorities.

After the survey parameters have been determined, the sample size (i.e. number of inspections and/or samples thereof to be examined and/or tested) can be calculated using EFSA's dedicated statistical tool RiPEST/RiBESS+. The mathematical principles behind the tool are in line with the recommendations and guidelines provided by the IPPC in the various ISPMs and guidelines for pest surveys. The number of inspections and/or samples should then be allocated to the epidemiological units and/or risk categories and the inspection units should be selected within the survey area. Further, specific instructions for the inspectors need to be carefully formulated to indicate how to collect which data. The flexible approaches proposed in this document allow the survey design to be tailored to each specific situation in the Member States taking into account the host plant distribution and available resources. The success of a good survey design relies on the technical aspects of the survey preparation and on the involvement of the risk managers.

Once the survey has been implemented and the inspections conducted and/or the samples collected and analysed, the survey conclusions need to be formulated, while considering the strength of the evidence to support this conclusion. The underpinning assumptions made on the homogeneity of the survey area, the method sensitivity, and the surveyed host plants should be included in the conclusion. The reliability of the conclusions of surveys designed using the proposed approaches depends strongly on the survey preparation. The proposed formulation of the conclusion allows surveys to be compared across time and space, thus contributing to the harmonisation of surveillance activities across the EU Member States.

References

Aerts M, Geys H, Molenberghs G and Ryan L, 2002. Topics in Modelling of Clustered Data. Chapman & Hall, London.

Bemelmans D, Cortinas Abrahantes J, Kaluski T and Verbeke T, 2023. RiPEST – Risk-based PEst Survey Tool (1.0.0). Zenodo. https://doi.org/10.5281/zenodo.8335473

Binns M, Nyrop J and van der Werf W, 2000. Sampling and Monitoring in Crop Protection: The Theoretical Basis for Designing Practical Decision Guides. CABI publishing, 304 pp.

Bosso L, Luchi N, Maresi G, Cristinzio G, Smeraldo S and Russo D, 2017. Predicting current and future disease outbreaks of *Diplodia sapinea* shoot blight in Italy: species distribution models as a tool for forest management planning. Forest Ecology and Management, 400, 655–664.

Bourhis Y, Gottwald TR, Lopez-Ruiz FJ, Patarapuwadol S and van den Bosch F, 2019. Sampling for disease absence—deriving informed monitoring from epidemic traits. Journal of Theoretical Biology, 461, 8–16.

Bouwmeester H, Heuvelink GBM, Legg JP and Stoorvogel JJ, 2012. Comparison of disease patterns assessed by three independent surveys of cassava mosaic virus disease in Rwanda and Burundi. Plant Pathology, 61, 399–412.

Cannon R, 2002. Demonstrating disease freedom – combining confidence levels. Preventive veterinary medicine, 52, 227–249. doi: 10.1016/s0167-5877(01)00262-8

Charest J, Dewdney M, Paulitz T, Philion V and Carisse O, 2002. Spatial distribution of *Venturia inaequalis* airborne ascospores in orchards. Phytopathology, 92, 769–79.

Demidenko E, 2007. Sample size determination for logistic regression revisited, Statistics in Medicine, 26, 3385–3397.

EFSA (European Food Safety Authority), 2012. A framework to substantiate absence of disease: the risk-based estimate of system sensitivity tool (RiBESS) using data collated according to the EFSA Standard Sample Description. Supporting Publications 2012: EN-366. 44. https://doi.org/10.2903/sp.efsa.2012.EN-366

EFSA PLH Panel (EFSA Panel on Plant Health), 2017a. Scientific Opinion on the pest risk assessment of *Eotetranychus lewisifor* the EU territory. EFSA Journal 2017;15(10):4878, 122 pp. doi:10.2903/j.efsa.2017.4878

EFSA PLH Panel (EFSA Panel on Plant Health), 2017b. Scientific Opinion on the pest risk assessment of *Diaporthe vaccinii* for the EU territory. EFSA Journal 2017;15(9):4924, 185 pp. doi:org/10.2903/j.efsa.2017.4924

EFSA (European Food Safety Authority), Ciubotaru RM, Cortiñas Abrahantes J, Oyedele J, Parnell S, Schrader G, Zancanaro G and Vos S, 2018. Technical report of the methodology and work-plan for developing plant pest survey guidelines. EFSA supporting publication 2018: EN-1399. 36 pp. https://doi.org/10.2903/sp.efsa.2018.EN-1399

EFSA PLH Panel (EFSA Panel on Plant Health), 2019. Update of the Scientific Opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory. EFSA Journal 2019;17(5):5665, 200 pp. doi:10.2903/j.efsa.2019.5665

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Vicent Civera A, Schans J, Schenk M, Schrader G, Cortiñas Abrahantes J, Zancanaro G and Vos S, 2020a. Guidelines for statistically sound and risk-based surveys of *Xylella fastidiosa*. EFSA Supporting publications, 17 (6), 1873. doi.org/10.2903/sp.efsa.2020.EN-1873

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Vicent Civera A, Schans J, Schenk M, Schrader G, Cortiñas Abrahantes J, Zancanaro G and Vos S, 2020b. Guidelines for statistically sound and risk-based surveys of *Phyllosticta citricarpa*. EFSA supporting publications, 17 (7), 1893. https://doi.org/doi:10.2903/sp.efsa.2020.EN-1893

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Vicent Civera A, Schans J, Schenk M, Cortiñas Abrahantes J, Zancanaro G and Vos S, 2020c. Guidelines for statistically sound and risk-based surveys of *Agrilus planipennis*. EFSA supporting publication 2020:EN-1983. 65 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1983

EFSA (European Food Safety Authority), Parnell S, Schenk M, Schrader G, Vicent A, Delbianco A, and Vos S, 2020d. Pest survey card on *Phyllosticta citricarpa*. EFSA supporting publication 2020:EN-1863. 35 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1863

EFSA (European Food Safety Authority), Schenk M, Dijkstra E, Delbianco A and Vos S, 2020e. Pest survey card on *Rhagoletis pomonella*. EFSA supporting publication 2020:EN-1908. 26 pp. https://doi.org/10.2903/sp.efsa.2020.EN-1908

EFSA (European Food Safety Authority), Lázaro E, Vicent Civera A and Delbianco A, 2022. Pest survey card on *Elsinoë australis*, *E. citricola* and *E. fawcettii*. EFSA supporting publication 2022:EN-7561. https://doi.org/10.2903/sp.efsa.2022.EN-7561

EFSA (European Food Safety Authority), 2023. General glossary for surveys of quarantine organisms. Available online: https://efsa.europa.eu/plants/planthealth/monitoring/surveillance/glossary. Last updated: 20 July 2023.

EFSA (European Food Safety Authority), Martino ME, Delbianco A, KaluskiT and Vos S, 2024. Technical report on the classification of plant pest detection and identification methods. EFSA supporting publication 2024:EN-8381. 27 pp. doi:10.2903/sp.efsa.2024.EN-8381

EFSA (European Food Safety Authority), de la Peña E, Dehnen-Schmutz K, Gilioli G, González-Moreno P, Makowski D, Mastin A, Nougadère A, Paracchini ML, Parnell S, Portaluri A, Purse B, Ribaya M, Sánchez B, Scala M, Tramontini S and Vos S, 2025. EFSA methodology for assessing candidate priority pests under EU Regulation 2016/2031. EFSA Supporting publication 2025:EN-9230. 42 pp. doi: 10.2903/sp.efsa.2025.EN-9230

Eurostat, 2024. NUTS – Nomenclature of Territorial Units for Statistics 2024 classification. Available online: https://ec.europa.eu/eurostat/web/nuts

FAO (Food and Agriculture Organization of the United Nations), 1998. ISPM (International Standards for Phytosanitary Measures) 9. Guidelines for pest eradication programmes. IPPC, FAO, Rome, 14 pp. Available online: https://www.ippc.int/en/publications/611/

FAO (Food and Agriculture Organization of the United Nations), 2005. ISPM (International Standards for Phytosanitary Measures) 22. Requirements for the establishment of areas of low pest prevalence. IPPC, FAO, Rome, 12 pp. Available online: https://www.ippc.int/en/publications/599/

FAO (Food and Agriculture Organization of the United Nations), 2014. Risk-based disease surveillance - A manual for veterinarians on the design and analysis for demonstration of freedom from disease. FAO Animal Production and Health Manual No. 17. FAO, Rome.

FAO (Food and Agriculture Organization of the United Nations), 2016a. ISPM (International Standard for Phytosanitary Measures) 27. Diagnostic protocols for regulated pests. In (pp. 12): FAO (Food and Agriculture Organization of the United Nations).

FAO (Food and Agriculture Organization of the United Nations), 2016b. ISPM (International Standard for Phytosanitary Measures) 32. Categorization of commodities according to their pest risk. In (pp. 16): FAO (Food and Agriculture Organization of the United Nations).

FAO (Food and Agriculture Organization of the United Nations), 2017a. ISPM (International Standards for Phytosanitary Measures) 8. Determination of pest status in an area. FAO, Rome, 16 pp. Available online: https://www.ippc.int/en/publications/612/

FAO (Food and Agriculture Organization of the United Nations), 2017b. ISPM (International Standards for Phytosanitary Measures) 4. Requirements for the establishment of pest free areas. IPPC, FAO, Rome, 12 pp. Available online: https://www.ippc.int/en/publications/614/

FAO (Food and Agriculture Organization of the United Nations), 2018. ISPM (International Standard for Phytosanitary Measures) 6. Surveillance. FAO, Rome. Available online: https://www.ippc.int/static/media/files/publication/en/2019/02/ISPM_06_2018_En_Surveillance_2018-05-20_PostCPM13.pdf

FAO (Food and Agriculture Organization of the United Nations), 2024a IPPC Secretariat, International Plant Protection Convention (1997). Second edition. Rome, FAO. https://openknowledge.fao.org/handle/20.500.14283/cd0175en

FAO (Food and Agriculture Organization of the United Nations), 2024b. ISPM (International Standards for Phytosanitary Measures) 5. Glossary of phytosanitary terms. In (pp. 38). FAO, Rome.

Franke J, Gebhardt S, Menz G and Helfrich HP, 2009. Geostatistical Analysis of the Spatiotemporal Dynamics of Powdery Mildew and Leaf Rust in Wheat. Phytopathology, 99, 974–84.

Hester S, Sergeant E, Robinson AP and Schultz G, 2015. Animal, vegetable, or...? A case study in using animal-health monitoring design tools to solve a plant-health surveillance problem. Biosecurity Surveillance: Quantitative Approaches. F. Jarrad, S. Low-Choy and K. Mengersen. 6: 313.

Hornero A, Hernandez-Clemente R, North PRJ, Beck PSA, Boscia D, Navas-Cortes JA and Zarco-Tejada PJ, 2020. Monitoring the incidence of *Xylella fastidiosa* infection in olive orchards using ground-based evaluations, airborne imaging spectroscopy and Sentinel-2 time series through 3-D radiative transfer modelling. Remote Sensing of Environment, 236, 111480.

Hyatt-Twynam SR, Parnell S, Stutt ROJH, Gottwald TR, Gilligan CA and Cunniffe NJ, 2017. Risk-based management of invading plant disease. New Phytologist, 214 (3), 1317–1329.https://doi.org/10.1111/nph.14488

Janse JD, Bergsma-Vlami M, Beuningen Av, Derks H, Hendriks H, Horn N, Janssen F, Kavelaars J, Roenhorst A, Schoeman M, Steeghs M, Tjou-Tam-Sin NNS, Verdel M and Wenneker M, 2009. Brown rot in potato. Gewasbescherming, 40, 176-187.

Koh JMJ, Cunniffe NJ and Parnell S, 2025. Assessing delimiting strategies to identify the infested zones of quarantine plant pests and diseases. Sci Rep **15**, 5610. doi.org/10.1038/s41598-025-90343-2

Madden LV, Hughes G and van den Bosch F, 2007. The Study of Plant Disease Epidemics. APS Publications, 421 pp. doi:10.1094/9780890545058

Mastin AJ, van den Bosch F, Gottwald TR, Alonso-Chavez V and Parnell SR, 2017. A method of determining where to target surveillance efforts in heterogeneous epidemiological systems. PLoS Computational Biology, 13, 1–23. https://doi.org/10.1371/journal.pcbi.1005712

Mastin AJ, van den Bosch F, van den Berg F and Parnell S, 2019. Quantifying the hidden costs of imperfect detection for early detection surveillance. Philosophical Transactions of the Royal Society B, 374, 2018261. https://doi.org/10.1098/rstb.2018.0261

Mastin AJ, Gottwald TR, van den Bosch F, Cunniffe NJ, Parnell S, 2020. Optimising risk-based surveillance for early detection of invasive plant pathogens. PLoS Biol 18(10): e3000863.

Meentemeyer RK, Anacker BL, Mark W and Rizzo DM, 2008. Early detection of emerging forest disease using dispersal estimation and ecological niche modeling. Ecological Applications, 18 (2), 377-390.

Milanzi E, Njeru Njagi E, Bruckers L and Molenberghs G, 2015. Data Representativeness: Issues and Solutions. EFSA supporting publication 2015:EN-759, 159 pp.

Narouei-Khandan HA, Halbert SE, Worner SP and van Bruggen AHC, 2016. Global climate suitability of citrus huanglongbing and its vector, the Asian citrus psyllid, using two correlative species distribution modeling approaches, with emphasis on the USA. European Journal of Plant Pathology, 144, 655–670. doi:10.1007/s10658-015-0804-7

Nutter, J. F., Teng, P., & Shokes, F. M., 1991. Disease assessment terms and concepts. *Plant disease*, *75*, 1187-1188.

Nutter Jr F, Esker PD and Coelho Netto RA, 2006. Disease Assessment Concepts and the Advancements Made in Improving the Accuracy and Precision of Plant Disease Data. European Journal of Plant Pathology, 115, 95–103.

Parnell S, van den Bosch F, Gottwald T and Gilligan CA, 2017. Surveillance to inform control of emerging plant diseases. Annual Review Phytopathology, 55, 591–610. doi:10.1146/annurev-phyto-080516-035334

Stonard JF, Marchant BP, Latunde-Dada AO, Liu Z, Evans N, Gladders P, Eckert MR and Fitt BDL, 2010. Geostatistical analysis of the distribution of *Leptosphaeria* species causing phoma stem canker on winter oilseed rape (*Brassica napus*) in England. Plant Pathology, 59, 200–10.

Sun, H., Douma, J. C., Schenk, M. F. and van der Werf, W., 2025. Comparing inward and outward strategies for delimiting non-native plant pest outbreaks. Journal of Pest Science, 1-16.

Tubajika KM, Civerolo EL, Ciomperlik MA, Luvisi DA, Hashim JM. 2004. Analysis of the spatial patterns of Pierce's disease incidence in the lower San Joaquin Valley in California. Phytopathology, 94, 1136–44.

Zadoks JC and Van den Bosch F, 1994. On the spread of plant disease: a theory on foci. Annual review of phytopathology, 32(1), 503-521.

Abbreviations

CL	Confidence level
DP	Design prevalence
EFSA	European Food Safety Authority
EPPO	European and Mediterranean Plant Protection Organisation
EU	European Union
FAO	Food and Agriculture Organization of the United Nations
IPPC	International Plant Protection Convention
ISPM	International Standards for Phytosanitary Measures
MeSe	Method sensitivity
NPPO	National Plant Protection Organisation
NUTS	Nomenclature of Territorial Units for Statistics
PIZ	Potential infested zone
SRS	Simple random sampling

Appendix A Specific changes and amendments in relation to General guidelines (EFSA, 2020)

These updated guidelines for statistically sound and risk-based surveys of plant pests present several key differences compared to the EFSA (2020) publication. These differences reflect advancements in methodologies, tools, experiences from pest surveys in Member States and detailed approaches to enhance the precision of pest survey designs.

One significant update is the introduction of the RiPEST tool alongside the OptiPest tool. This addition provides surveyors with more advanced statistical tools to accurately estimate sample sizes and facilitate the process of planning surveys at crop level.

In the detection survey section, the updated guidelines offer a more detailed approach to sample allocation and reporting. These additions ensure that survey results are accurately documented and communicated, enhancing the overall quality and transparency of the survey process.

The delimiting survey section has been significantly expanded to provide more comprehensive steps for delimiting surveys. This includes considerations for natural and human-assisted spread and the construction of the potential infested zone. These additions provide a more thorough approach to identifying and managing potentially infested zones, thereby improving the effectiveness of delimiting surveys. Furthermore, the updated guidelines include a new section on follow-up activities after delimiting surveys, emphasizing the importance of continuous monitoring and management of pest situations.

The monitoring survey section introduces information on pest prevalence surveys and simulations to improve the estimation of prevalence over time. These enhancements allow for better tracking of pest eradication efforts and more accurate prevalence estimation.

These guidelines introduce new section on Plant pest surveillance toolkit explaining all tools and documents developed by EFSA to support risk managers in the process of planning, designing and execution of statistically sound and risk-based surveys of regulated pests.

The updated guidelines also include additional appendices, such as Appendix B on delimiting strategies and Appendix E on prevalence survey simulations. These appendices provide further detailed methodologies and examples to support survey design and implementation.

Overall, the updated guidelines present a more comprehensive and detailed framework for conducting statistically sound and risk-based surveys of plant pests. These enhancements reflect the latest scientific advancements and provide EU Member States with the necessary tools and methodologies to conduct effective pest surveys, ensuring the protection of plant health and the harmonization of surveillance activities across the EU.

Key differences:

1. **Tools:**
   - The updated version introduces the **RiPEST tool** alongside the **OptiPest tool**.
2. **Detection survey:**

- The updated version has a more detailed approach to **sample allocation** and **reporting**.

3 **Delimiting survey**:

- The updated version provides more comprehensive steps for delimiting surveys, including considerations for **natural and human-assisted spread** and **the construction of the potential infested zone**.
- The updated version includes a new section on **follow-up** after delimiting surveys.

4 Monitoring survey:

- The updated version adds information **on pest prevalence surveys** and **simulations** to improve the estimation of prevalence over time.

5 Appendices:

- The updated version includes additional appendices, such as **Appendix B** on delimiting strategies and **Appendix E** on prevalence survey methodology.

References

EFSA (European Food Safety Authority), Lázaro E, Parnell S, Vicent Civera A,Schans J, Schenk M, Cortiñas Abrahantes J, Zancanaro G and Vos S, 2020. General guidelines forstatistically sound and risk-based surveys of plant pests. EFSA supporting publication 2020:EN-1919.65 pp. doi:10.2903/sp.efsa.2020.EN-1919

Appendix B Delimiting strategies

B.1. Potential infested zone

The proposed method for constructing the potential infested zone is based on modelling the maximum annual spread over time. This approach defines the potential infested zone as a circular area (i.e. circle-based strategy) around the most likely source of infestation. The radius of this circle represents the maximum potential natural spread of the pest since its (potential) introduction. It is calculated based on the number of years since the last detection survey that confirmed pest freedom, and on available knowledge about the pest's annual spread rate.

To estimate this maximum potential natural spread, it is essential to obtain information on the annual spread rate. In the proposed approach, assuming an exponential kernel, the annual spread distance is considered to follow an exponential distribution, Exp(λ), where λ (lambda) is a parameter that indicates how quickly the probability of dispersal decreases with distance. A higher λ means the pest tends to spread shorter distances; a lower λ means it may reach farther distances. The mean annual spread distance is given by 1/λ.

The value of λ is calculated in different ways, depending on the type of information available:

- If the pest has been assessed in the EFSA Scientific Report on the Impact Assessment of EU Priority Pests (EFSA, 2025), λ is calculated as:

$$\lambda = -\frac{Ln(1-0.95)}{distance}, \quad \text{(equation 1a)}$$

  where ***distance*** corresponds to the 95th percentile (0.95) of the annual spread rate and is set using the **median** of the elicited **maximum annual spread distance** (i.e., the 50th percentile).

- If the pest has not been assessed in EFSA Scientific Report on the Impact Assessment of EU Priority Pests (EFSA, 2025), but the **mean annual spread distance** is available, λ is calculated as:

$$\lambda = \frac{1}{distance}, \quad \text{(equation 1b)}$$

  where ***distance*** is the **mean** annual spread.

- If the pest has not been assessed in the EFSA Scientific Report on the Impact Assessment of EU Priority Pests (EFSA, 2025), but the **median annual spread distance** is available, λ is calculated as:

$$\lambda = -\frac{Ln(1-0.5)}{distance}, \quad \text{(equation 1c)}$$

where ***distance*** is the median annual spread. To characterise the spatial component of the spread process, the method considers the number of years (n) since the last detection survey confirming pest freedom. In this context, n is interpreted as the number of years the pest may have been present and potentially spreading. Assuming the spread each year is independent and follows an identical distribution, i.e., $X_1, X_2, \ldots, X_n \sim \text{Exp}(\lambda)$, the density function is given by:

$$f_{X_i}(x) = \lambda \cdot e^{-\lambda \cdot x} \text{ for all } 1 \leq i \leq n. \quad \text{(equation 2)}$$

Thus, the total spread over n years is represented by the sum $Z = \sum_{i=1}^{n} X_i$, which follows a gamma distribution with shape parameter n and rate λ:

$$f_Z(x) = \frac{\lambda^n}{\Gamma(n)} e^{-\lambda \cdot x} \cdot x^{n-1}. \quad \text{(equation 3)}$$

Using this gamma distribution and predefining the shape (n) and rate (λ), the method supports two main steps:

1. **Determine the radius of the potential infested zone** as the 95th percentile of the Gamma(n, λ) distribution.
2. **Subdivide the zone into concentric survey bands**, each representing one year of potential spread. This supports **progressive surveys** that allow refinement of the infested area based on survey results (see Section 7.5). Each survey band (SBn) is defined by the 95th percentile of a Gamma distribution with shared λ and year-specific shape parameter n.

B.1.1. Illustrative example

To illustrate the method, a scenario is presented in which the target pest was assessed in the EFSA Scientific Report on EU Priority Pests (EFSA, 2025), and the median maximum annual spread rate was estimated at 1,000 meters based on expert judgement. The last survey confirming pest freedom was conducted 4 years ago, meaning the pest could have been present and spreading naturally for up to 4 years.

Construction of the potential infested zone

In this case, λ is calculated as:

$$\lambda = -\frac{Ln(1-0.95)}{distance} = -\frac{Ln(1-0.95)}{1000} \approx 0.003$$

Thus, to determine the area of the potential infested zone, it is considered the 95th percentile of a Gamma distribution with shape parameter n=4 (years) and rate parameter λ =0.003:

$$P_{95}\left(Ga(4,0.003)\right) \approx 2{,}588\, m.$$

This value defines the radius of the circular area within which the pest may have naturally spread since its possible introduction ("potential infested zone") (Figure 1).

Subdivision of the potential infested zone

After defining the potential infested zone, it is subdivided into multiple concentric survey bands. Each band represents the maximum distance that the pest may have naturally spread in one year from the source of infestation. This band-based approach facilitates iterative and progressive surveys aiming to either confirm pest absence or adjust the size of the infested area, depending on survey outcomes.

Assuming spread follows Gamma(n, λ) with λ = 0.003 and n from 1 to 5 (to test one year beyond the assumed introduction time), the survey bands are defined as follows:

Table 1: Survey bands defined by the 95th percentile of a Gamma (n, 0.003). The corresponding radius (in meters) represents the outer boundary of each band. *Position* indicates whether the band falls within or outside the potential infested zone, which was determined based on the initial assumption of n=4.

Survey Band	Year	95th Percentile Radius (m)	Position
SB5	Year 1	$P_{95}\left(Ga(1,0.003)\right) \approx 1{,}000\, m$	*Within the potential infested zone*
SB4	Year 2	$P_{95}\left(Ga(2,0.003)\right) \approx 1{,}583\ m$	*Within the potential infested zone*
SB3	Year 3	$P_{95}\left(Ga(3,0.003)\right) \approx 2{,}101\, m$	*Within the potential infested zone*
SB2	Year 4	$P_{95}\left(Ga(4,0.003)\right) \approx 2{,}588\, m$	*Within the potential infested zone*
SB1	Year 5	$P_{95}\left(Ga(5,0.003)\right) \approx 3{,}055\, m$	*Outside the potential infested zone*

The width of the bands narrows with distance, but area increases due to the circular shape (Figure 1 and Table 2). The outermost band (**SB1**) represents one additional year of potential spread beyond the initial assumption (i.e., 5 years instead of 4). Following the EFSA methodology implemented in RiPEST, this outermost band (SB1) serves to validate the assumptions on the pest's spread capacity and the estimated time of introduction, while accounting for any ongoing undetected dispersal. This band is surveyed first, and if no pest is detected, surveillance can proceed progressively inwards (see Section 7.5 for further details of this delimiting strategy).

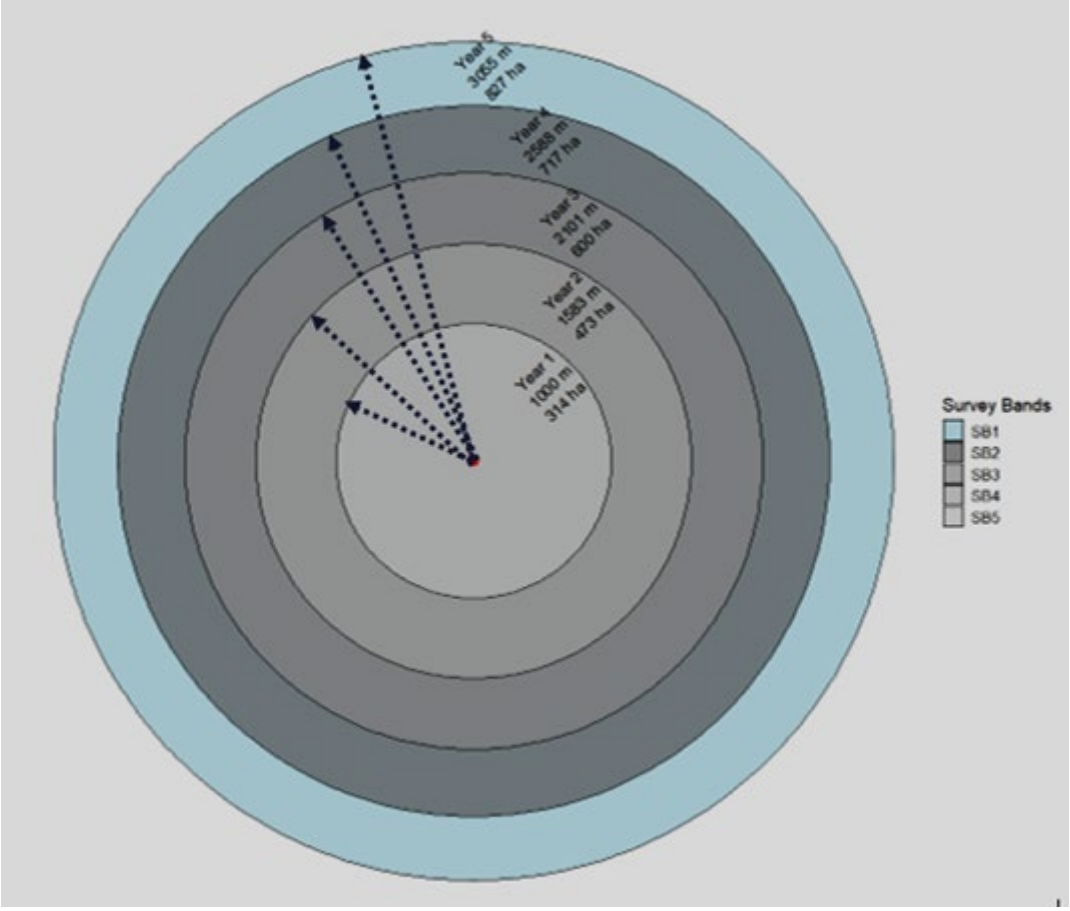


Figure 1: Concentric survey bands around the infestation source, with radii corresponding to the 95th percentiles of Gamma distributions (Gamma(n, λ = 0.003)). Each band (SB1 to SB5) is colour coded. The red dot marks the infestation source. Labels show year, radius (m), and band area (ha). Arrows connect the centre to external labels. The area of SB4 corresponds to the initially defined potential infested zone under the assumption of n = 4 with a radius ≈ 2,588 m.

Table 2: Radius, width, and area of successive survey bands defined from concentric circles centred at the source of infestation. The area of each band (in hectares) was calculated as the difference in area between two consecutive circles. Note: all calculations are based on geometric properties, assuming circular symmetry.

Survey Band	Year	Circle Radius (m)	Band Width (m)	Band Area (ha)
SB5	Year 1	1,000	1,000-0 = 1,000	≈ 314
SB4	Year 2	1,583	1,583-1,000= 583	≈ 473
SB3	Year 3	2,101	2,101-1,583= 518	≈ 600

SB2	Year 4	2,588	2,588-2,101= 487	≈ 717
SB1	Year 5	3,055	3,055-2,588= 467	≈ 828

B.2. Testing the performance of delimiting strategies

In this appendix we outline the methods and results of our study (Koh et al., 2025) which aimed to assess the performance of various delimiting strategies, including the recommended strategy. In addition to comparing the performance of the recommended strategy with two other delimiting strategies, we also aimed to assess the strategies' performance under different scenarios where important parameters, such as the duration of the pest spread and the generational/yearly spread distance, are not estimated perfectly. The full study can be found here: https://doi.org/10.1038/s41598-025-90343-2

B.3. Epidemiological model and case study pest

To assess the performance of the recommended delimiting strategy (hereafter referred to as the EFSA strategy), we first created an individual-based model (IBM) that simulated the spread of a pest in a finite, spatially explicit population of trees (i.e. the model keeps track of the coordinates and pest-status of each tree over time). For our case study pest, we selected Huanglongbing (HLB), or citrus greening, which is a disease caused by the EU priority pest '*Candidatus* Liberibacter spp.' and spread by psyllid vectors, spreading in an urban population. The IBM was built in R 4.3.1 (R Core Team, 2023) and tested using a landscape with a similar density of citrus trees to that of Seville, a large city in Spain. It is difficult to unambiguously quantify the number of citrus trees in Seville and estimates can range from 25,000 sour orange (*C. aurantium*) trees in 1996 (Oliva & Bonells, 1996) to 50,000 sour orange trees in 2020 (Cabanillas, 2020). In line with these estimates, we used the value quoted by Galvañ, et al. (2023) who estimated there are approximately 46,000 citrus trees in the city of Seville which has an area of 141.4 km2. To approximate the tree density in Seville, the IBM generates 15,941 trees and randomly distributes them in a 7 km x 7 km plot. As with other studies, the rate of disease transmission between a pair of trees separated by distance d_{ij} was modelled with an exponential dispersal kernel, $K(d;\, \alpha) = \exp\left(-\frac{d}{\alpha}\right)$ where α is the scale parameter (Cunniffe et al., 2015; Milne et al., 2020; Parry et al., 2014). We used the data presented in Arakawa & Miyamoto (2007) to estimate the mean distance travelled by psyllids (345,946m), under controlled conditions, and chose the scale parameter of the exponential dispersal kernel in our model, α, such that $2 \times \alpha = 345.946$, invoking the standard relationship between the mean dispersal distance in two dimensions and the exponential dispersal kernel scale parameter (Fabre et al., 2021). With a value of 172,973 for α, we then parameterized the baseline infection rate (β) by conducting a line search and obtained a value of 0.0001424 which would achieve a median disease prevalence of 50% after 5 years. With both values of α and β, we ran 1500 iterations of the IBM to estimate the true mean maximum yearly spread distance (1056.91m), the true mean generational spread distance (737.69m) and the true mean number of generations per year (5). The mean number of generations per year of our simulated HLB pest corresponds with previous research that estimated the psyllid vectors had 9-10 generations per year (Djeddour et al., 2021) and the latent period of HLB is approximately equal to 1 generation of the psyllid vectors (Lee et al., 2015).

B.4. Simulating different delimiting strategies

Along with the EFSA strategy, we also analyzed two other strategies: namely the In-to-Out strategy and the multi-foci strategy (Figure 2). Like the EFSA strategy, the In-to-Out strategy is essentially circles of varying radii drawn around the first detected infected tree, except that it always starts with the smallest circle surveying the immediate area around the first infected tree (Figure 2A), while the Multi-foci strategy never changes the length of the radius, but considers the multiple detections within each survey round and draws new circles around each new detection (Figure 2C). Additionally, we considered three different versions of the In-to-Out and EFSA strategies. The first version (Linear) assumes that the pest spreads to the estimated spread distance every year. However, this is likely an overestimate and the effects of an exponential spread that is compounded yearly can be approximated by a gamma distribution. The simplest approach is to parameterize the shape parameter with the estimated number of years the pest has been spreading for and use the corresponding mean maximum yearly spread distance to parameterize the rate parameter (Gamma Year). However, the underlying assumption of the Gamma Year version is that the pest has only one generation per year. Therefore, in the case of polycyclic pests, it would be more accurate to parameterize the shape parameter with the estimated number of pest generations since the first infection and use the mean maximum generational spread distance to parameterize the rate parameter (Gamma Gen).

All seven strategies adopt the equation from EFSA's risk-based estimate of system sensitivity (RiBESS+) tool to calculate the number of trees that need to be sampled in each survey band and use the standard values of 0.95 and 0.01 for the confidence limit and design prevalence respectively.

To test the results, we assessed the performance of all seven strategies in multiple scenarios of increasing complexity, varying the sensitivity of the detection method, the inclusion or exclusion of a 1-year asymptomatic period, the assumed mean maximum yearly/generational spread distance, the assumed duration of pathogen spread and whether the strategies started at the origin of the epidemic or a random symptomatic tree. The performance of the delimiting strategies was measured with four metrics. The first is Capability, which is the number of infected trees delimited by the strategies divided by the total number of infected trees present at the end of the delimiting survey. The second is Efficiency, which is the area of the delimited potential infested zone divided by the area of the convex hull (minimum area needed to delimit all the infected trees present at the end of the delimiting survey). The third is Effort, which is the total number of trees surveyed, and the fourth is simply the number of survey rounds taken to delimit a potential infested zone.

B.5. The performance of the different strategies

The RiBESS+ equation performed well and negated any changes in method sensitivity on the Capability and Efficiency of all seven strategies. Therefore, a decrease in method sensitivity only resulted in increased Effort levels and more survey rounds to delimit a potential infested zone. Throughout the various scenarios, the EFSA strategies constantly outperformed the others and achieved the highest levels of Capability while requiring the lowest amount of Effort and survey rounds. The multi-foci strategy only matched the performance of the EFSA strategies in the "perfect scenario" when the origin of the epidemic could be traced, asymptomatic trees could be detected, and the spread distance and duration of pest spread was estimated perfectly. The performance of the multi-foci strategy was most negatively affected when the strategy started on a random symptomatic tree. This is because when the strategy starts at the edge of the epidemic, subsequent detections are made towards the origin of the epidemic, causing the multi-foci strategy to neglect new infections that occur beyond its starting point. The In-to-Out strategies had very similar Capability levels to the EFSA strategies but required much

higher levels of Effort and more survey rounds. This is because the In-to-Out strategies start by surveying the immediate area surrounding the first detection, whereas the EFSA strategies start by surveying the estimated edge of the epidemic.

Of the three different versions of the EFSA strategy, the Gamma Gen version performed the best. Compared to the other versions, it consistently had the highest Capability levels and even had an average Capability of > 90% when the generational spread distance was underestimated (Figure 3 and Table 3). Because the Gamma Gen version had wider survey bands than the other two versions (Table 4), matching or overestimating the generational spread distance resulted in poorer Efficiency and tended to overestimate the area of the convex hull (Figure 3 and Table 3). However, even when the generational spread distance was overestimated, the Gamma Gen version, on average, only delimited an area approximately 2 times the convex hull (Figure 3 and Table 3). The high performance of the Gamma Gen version is likely due to it being more suitable than the other two versions for delimiting a polycyclic pest like the one we simulated. This was further supported when looking at how often each version correctly estimated the spread distance of the epidemic and moved inwards (Figure 4). Of the three versions, the Gamma Gen version was the only one that moved inwards > 90% of the time when the estimated spread distance was matched, regardless of whether the duration of pest spread was overestimated or matched or underestimated (Figure 4). Even when the assumed number of generations per year was underestimated to 3 generations per year, as long as the duration of the pest spread was matched or underestimated, the EFSA Gamma Gen version could achieve Capability levels of > 90% even when the generational spread distance was greatly underestimated (350m). Therefore, where possible and when the target pest is polycyclic, the EFSA strategy should be parameterized with the mean number of generations per year and the mean maximum generational spread distance.

While the inclusion of a 1-year asymptomatic period to the "perfect scenario" only resulted in a drop in Capability levels from 1.0 to 0.97 (EFSA Gamma Year and Linear) and 0.99 (EFSA Gamma Gen), it is likely that a longer asymptomatic period would have a greater effect. Therefore, when delimiting pests with long asymptomatic periods (e.g. *Xylella fastidiosa*), the potential infested zone should be supplemented with additional bands to account for asymptomatic individuals.

Figures and Tables

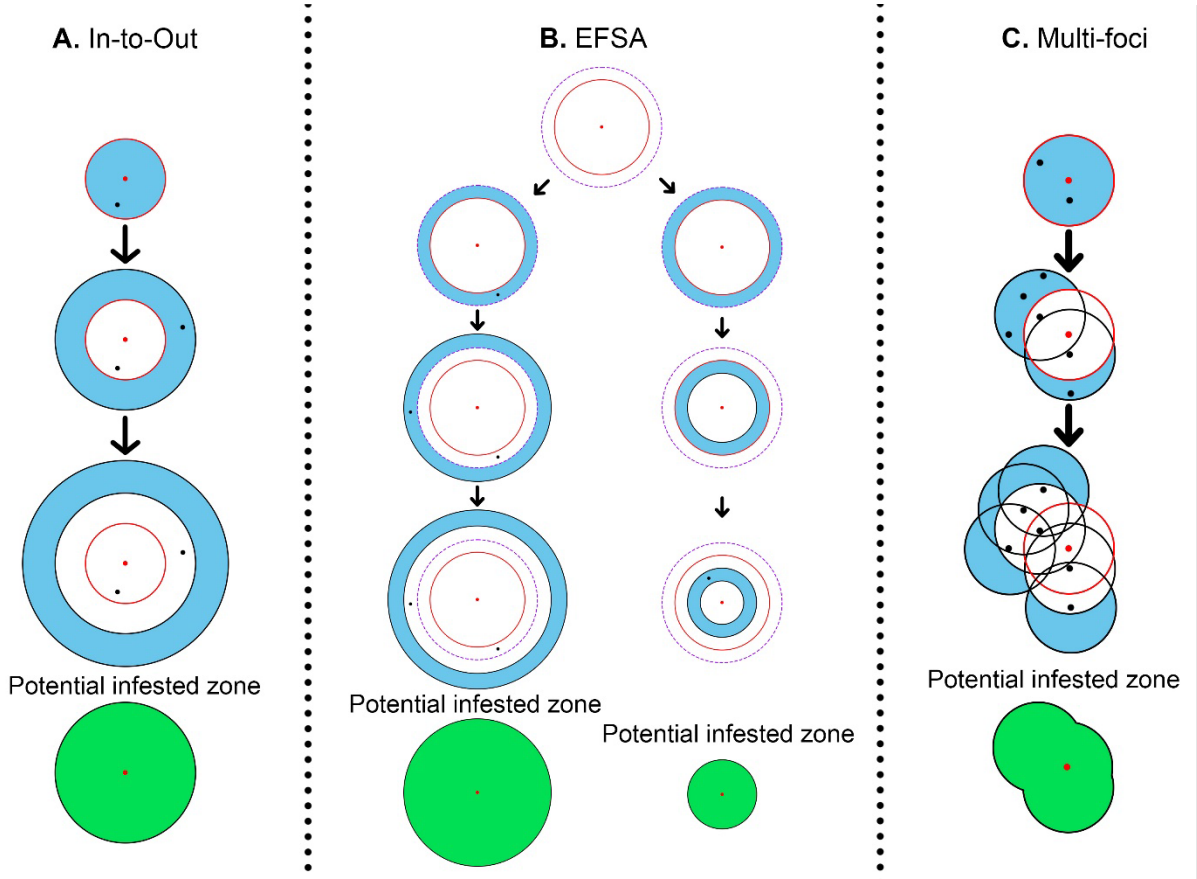
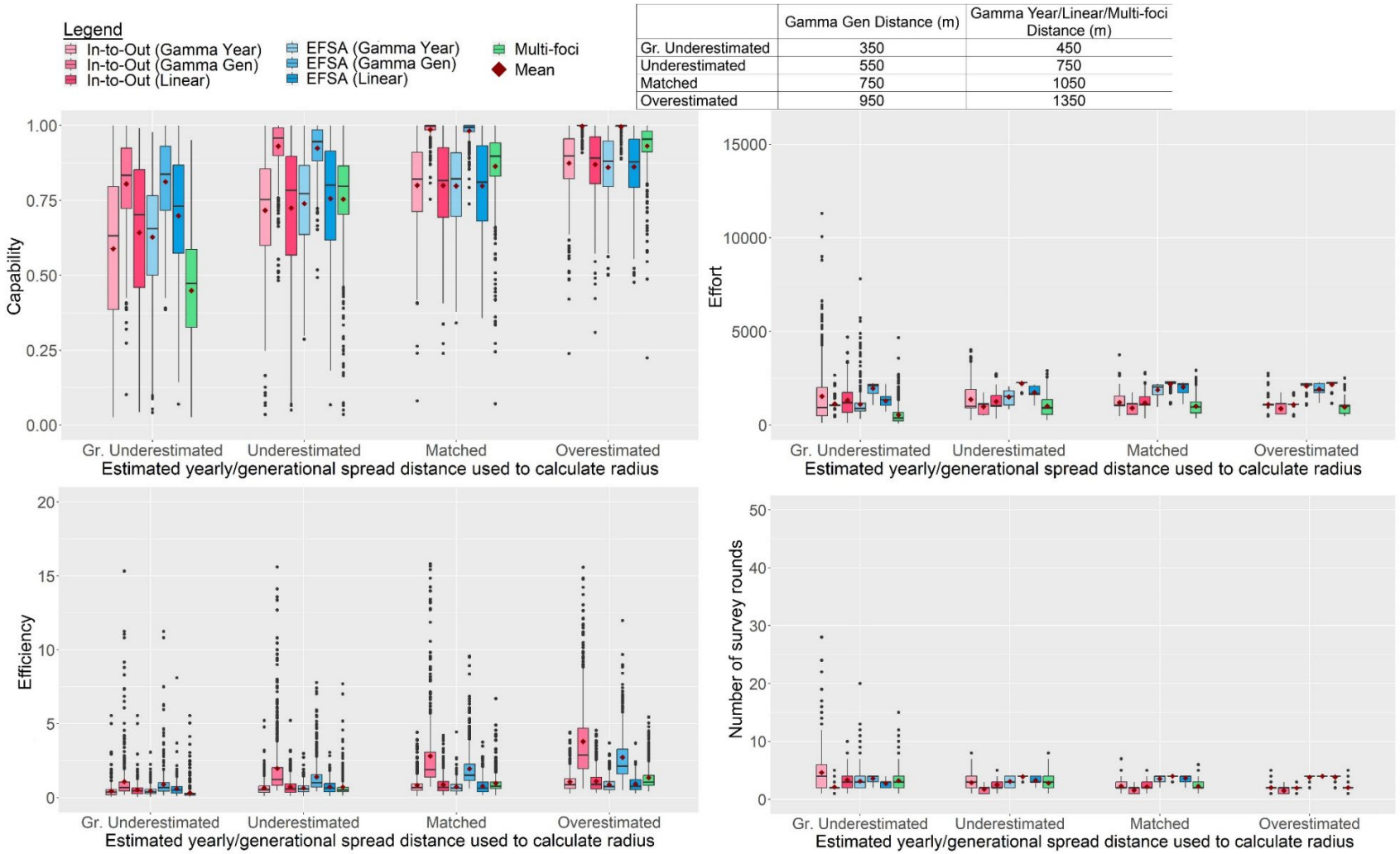


Figure 2: Diagram of how the three delimiting strategies work. Red dots represent the first infected tree detected, black dots represent subsequent detections of infected trees, and red circles represent the first circle drawn for each strategy.



	Gamma Gen Distance (m)	Gamma Year/Linear/Multi-foci Distance (m)
Gr. Underestimated	350	450
Underestimated	550	750
Matched	750	1050
Overestimated	950	1350

Figure 3: Boxplots showing the performance of the seven different strategies when Method Sensitivity was 0.5 and when the duration of the pest spread was 2 years while the assumed duration of pest spread for all seven strategies was 3 years (overestimated duration of pest spread).

Table 3: Median and interquartile values (in brackets) for the boxplots shown in Figure 2

CAPABILITY

Estimated yearly/generational spread distance used to calculate radius

	Gr. Underestimated	Underestimated	Matched	Overestimated
In-to-Out (Gamma Year)	0.63 (0.39 - 0.80)	0.75 (0.60 - 0.86)	0.82 (0.71 - 0.91)	0.90 (0.82 - 0.96)
In-to-Out (Gamma Gen)	0.83 (0.72 - 0.92)	0.96 (0.90 - 0.99)	0.98 (0.93 - 1.00)	1.00 (1.00 - 1.00)
In-to-Out (Linear)	0.70 (0.46 - 0.85)	0.78 (0.57 - 0.90)	0.82 (0.69 - 0.93)	0.89 (0.81 - 0.96)
EFSA (Gamma Year)	0.66 (0.50 - 0.77)	0.77 (0.64 - 0.87)	0.82 (0.70 - 0.91)	0.88 (0.80 - 0.95)
EFSA (Gamma Gen)	0.84 (0.72 - 0.93)	0.95 (0.88 - 0.99)	0.99 (0.98 - 1.00)	1.00 (1.00 - 1.00)
EFSA (Linear)	0.73 (0.57 - 0.87)	0.80 (0.62 - 0.91)	0.81 (0.68 - 0.93)	0.88 (0.79 - 0.95)
Multi-foci	0.47 (0.33 - 0.59)	0.80 (0.70 - 0.87)	0.90 (0.83 - 0.94)	0.95 (0.91 - 0.98)

EFFORT

Estimated yearly/generational spread distance used to calculate radius

	Gr. Underestimated	Underestimated	Matched	Overestimated
In-to-Out (Gamma Year)	931 (506 - 2006)	1004 (910 - 1894)	1069 (1010.5 - 1560)	1099 (1056 - 1118)
In-to-Out (Gamma Gen)	1076 (1021 - 1089)	1120 (567 - 1139)	1139 (578 - 1156)	1139.5 (582 - 1164)
In-to-Out (Linear)	1197 (683 - 1743)	1017 (975 - 1574)	1090 (1054 - 1520)	1118 (1086 - 1131)
EFSA (Gamma Year)	890 (750 - 1180)	1525.5 (1074.75 - 1835.75)	2033 (1628.5 - 2141)	2186 (2100 - 2231)
EFSA (Gamma Gen)	2111 (1681 - 2191)	2290 (2271 - 2285)	2286 (2209 - 2294)	1865 (1718 - 2214)
EFSA (Linear)	1357 (1053 - 1527)	1681.5 (1620 - 2088)	2189 (1723 - 2232)	2261 (2225.75 - 2278)
Multi-foci	377 (224 - 677)	917 (573 - 1364.5)	962 (646.5 - 1241.5)	995 (613.5 - 1076)

EFFICIENCY

Estimated yearly/generational spread distance used to calculate radius

	Gr. Underestimated	Underestimated	Matched	Overestimated
In-to-Out (Gamma Year)	0.36 (0.20 - 0.53)	0.53 (0.36 - 0.78)	0.66 (0.47 - 0.93)	0.87 (0.63 - 1.29)
In-to-Out (Gamma Gen)	0.67 (0.45 - 1.05)	1.23 (0.82 - 2.03)	1.89 (1.39 - 3.08)	2.88 (1.96 - 4.69)
In-to-Out (Linear)	0.45 (0.25 - 0.66)	0.58 (0.34 - 0.92)	0.69 (0.46 - 1.08)	0.89 (0.60 - 1.37)
EFSA (Gamma Year)	0.38 (0.26 - 0.57)	0.57 (0.38 - 0.80)	0.62 (0.43 - 0.90)	0.75 (0.54 - 1.09)
EFSA (Gamma Gen)	0.67 (0.42 - 1.01)	1.00 (0.73 - 1.55)	1.51 (1.15 - 2.26)	2.13 (1.62 - 3.27)
EFSA (Linear)	0.50 (0.30 - 0.75)	0.64 (0.40 - 0.95)	0.63 (0.40 - 1.04)	0.76 (0.54 - 1.21)
Multi-foci	0.21 (0.15 - 0.31)	0.51 (0.41 - 0.70)	0.76 (0.61 - 1.05)	1.04 (0.82 - 1.49)

NUMBER OF SURVEY ROUNDS

Estimated yearly/generational spread distance used to calculate radius

	Gr. Underestimated	Underestimated	Matched	Overestimated
In-to-Out (Gamma Year)	4 (2 - 6)	3 (2 - 4)	2 (2 - 3)	2 (2 - 2)
In-to-Out (Gamma Gen)	2 (2 - 2)	2 (1 - 2)	2 (1 - 2)	2 (1 - 2)
In-to-Out (Linear)	3 (2 - 4)	2 (2 - 3)	2 (2 - 3)	2 (2 - 2)
EFSA (Gamma Year)	3 (2 - 4)	3 (2 - 4)	4 (3 - 4)	4 (4 - 4)
EFSA (Gamma Gen)	4 (3 - 4)	4 (4 - 4)	4 (4 - 4)	4 (4 - 4)
EFSA (Linear)	3 (2 - 3)	3 (3 - 4)	4 (3 - 4)	4 (4 - 4)
Multi-foci	3 (2 - 4)	3 (2 - 4)	2 (2 - 3)	2 (2 - 2)

Table 4: Comparison of the different radii length and band width between the three different EFSA versions.

	RADIUS LENGTH (M)			BAND WIDTH (M)		
YEARS	Gamma Year	Gamma Gen	Linear	Gamma Year	Gamma Gen	Linear
1	1056.91	2254.026	1056.91	1056.91	2254.026	1056.91
2	1673.66	3867.362	2113.82	616.75	1613.336	1056.91
3	2221.189	5389.481	3170.73	547.529	1522.119	1056.91
4	2735.531	6865.178	4227.64	514.342	1475.697	1056.91
5	3229.409	8311.427	5284.55	493.878	1446.249	1056.91

	Gamma Gen Distance (m)	Gamma Year/Linear Distance (m)
Gr. Underestimated	350	450
Underestimated	550	750
Matched	750	1050
Overestimated	950	1350

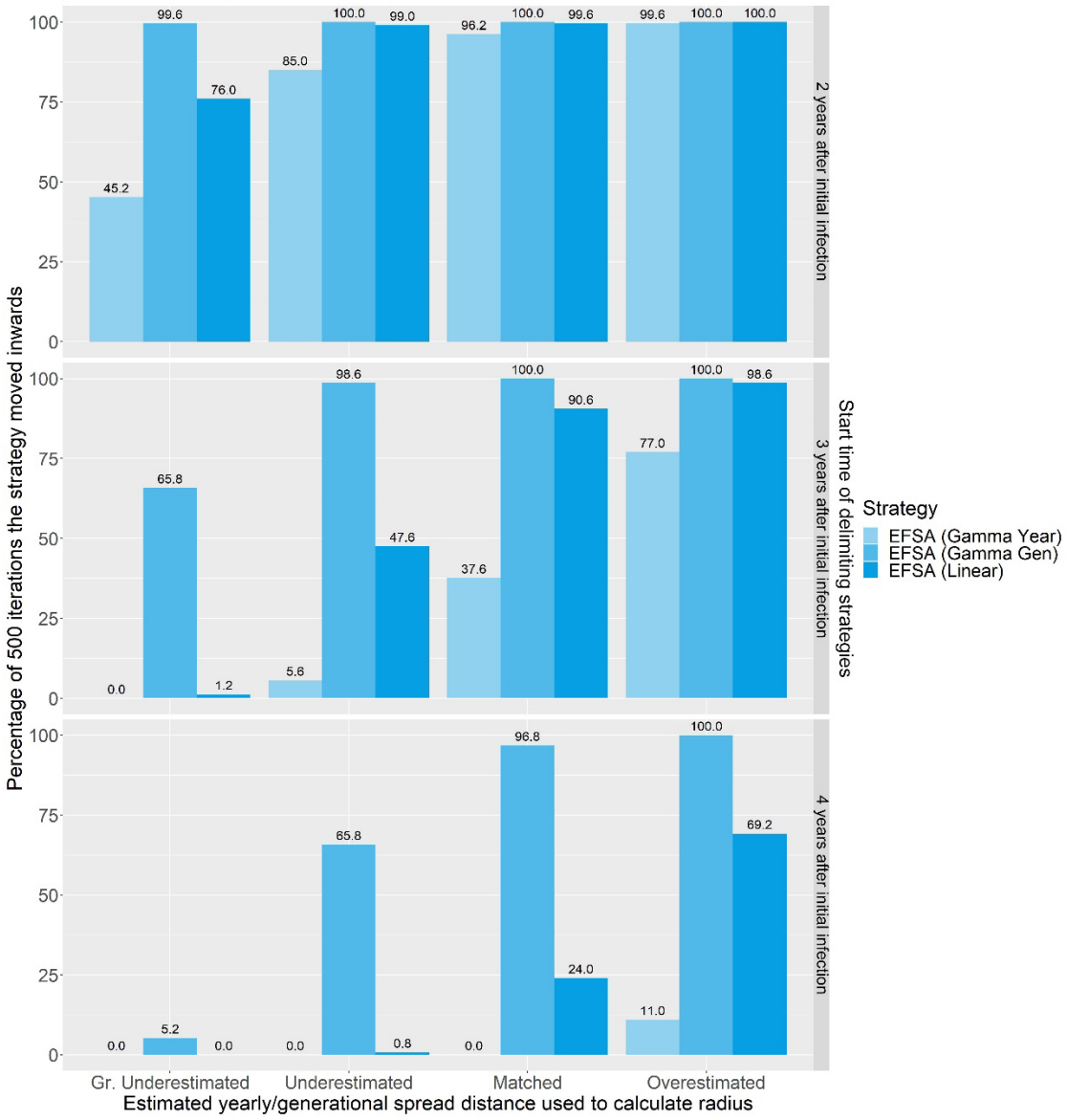


Figure 4: Bar graphs showing how often each of the different EFSA versions moved inwards to delimit the potential infested zone.

References

Arakawa K and Miyamoto K, 2007. Flight ability of Asiatic citrus psyllid, *Diaphorina citri* Kuwayama (Homoptera; Psyllidae), measured by a flight mill. *Research Bulletin of the Plant Protection Service (Japan)*, (43).

Cabanillas, F. (2020). La nueva (y energética) vida de las naranjas. La Vanguardia. Available at: https://www.lavanguardia.com/vida/20200219/473659497199/la-nueva-y-energetica-vida-de-las-naranjas.html (Accessed: 25/11/2025)

Cunniffe NJ, Stutt RO, DeSimone RE, Gottwald TR and Gilligan CA, 2015. Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. PLoS computational biology, 11(4), e1004211.

Djeddour D, Pratt C, Constantine K, Rwomushana I and Day R, 2021. The Asian Citrus Greening Disease (Huanglongbing).

EFSA (European Food Safety Authority), de la Peña E, Dehnen-Schmutz K, Gilioli G, González-Moreno P, Makowski D, Mastin A, Nougadère A, Paracchini ML, Parnell S, Portaluri A, Purse B, Ribaya M, Sánchez B, Scala M, Tramontini S and Vos S, 2025. EFSA methodology for assessing candidate priority pests under EU Regulation 2016/2031. EFSA Supporting publication 2025:EN-9230. 42 pp. doi: 10.2903/sp.efsa.2025.EN-9230

Fabre F, Coville J and Cunniffe NJ, 2021. Optimising reactive disease management using spatially explicit models at the landscape scale. In Plant diseases and food security in the 21st century (pp. 47-72). Cham: Springer International Publishing.

Galvañ A, Bassanezi RB, Luo W, Vanaclocha P, Vicent A and Lázaro E, 2023. Risk-based regionalization approach for area-wide management of HLB vectors in the Mediterranean Basin. Frontiers in Plant Science, 14. URL: https://doi.org/10.3389/fpls.2023.1256935

Koh JMJ, Cunniffe NJ and Parnell S, 2025. Assessing delimiting strategies to identify the infested zones of quarantine plant pests and diseases. Sci Rep **15**, 5610. doi.org/10.1038/s41598-025-90343-2

Lee JA, Halbert SE, Dawson WO, Robertson CJ, Keesling JE and Singer BH, 2015. Asymptomatic spread of huanglongbing and implications for disease control. Proceedings of the National Academy of Sciences, 112(24), 7605-7610.

Milne AE, Gottwald T, Parnell SR, Alonso Chavez V and Van den Bosch F, 2020. What makes or breaks a campaign to stop an invading plant pathogen?. PLOS Computational Biology, 16(2), e1007570.

Oliva SR and Bonells JE, 1996. El naranjo amargo de Sevilla. URL: https://www.sevilla.org/servicios/medio-ambiente-parques-jardines/e-articulos-tecnicos/naranjo_amargo.pdf

Parry M, Gibson GJ, Parnell S, Gottwald TR, Irey MS, Gast TC and Gilligan CA, 2014. Bayesian inference for an emerging arboreal epidemic in the presence of control. Proceedings of the National Academy of Sciences, 111(17), 6258-6262.

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Appendix C Sample survey design

C.1. Introduction

Sample survey design comprises all steps carried out in order to obtain a descriptive or inferential statistic of a specific population of interest by just studying a portion of that population (Kalton, 1983; Kish, 1965). When the whole population is examined then inferences are unnecessary, as estimation methods are not needed, given that the whole population was scrutinized. Although, census provides the ideal answer to any specific question, a survey has several advantages such as cost-effectiveness. Moreover, it is not always feasible to study the whole population. These advantages only apply if the survey is designed in accordance to statistical guidelines that ensure the control of specific errors that may arise due to studying only part of the population (Stopher and Meyburg, 1979). The guidelines are a collection of decisions such as the way in which data is collected, the method used for processing the data and sample design (Kalton, 1983).

As it was already mentioned in the work plan in designing a sample survey (EFSA, 2018) it is of the utmost importance to have a clear definition of the target population, and its elements, i.e. the units that make up the population from which information is sought. For instance, EFSA is mandated to collect data from EU Member States on a wide range of topics, like pesticide monitoring in food items, monitoring zoonoses and food-borne outbreaks in humans, food and animals, residues of chemical elements in foods of animal origin, and antimicrobial resistance in zoonotic and indicator bacteria from food-producing animals and food. Although the data for each topic are obtained from EU Member States, the elements of the targeted population are different. However, the principles of survey design are universal and the guidelines could also be applied to other surveys in which EFSA is involved. For pesticide monitoring, the elements are obtained from food items only, while for zoonoses and food-borne diseases, humans are also sample elements. In addition to recognising the elements, a clear definition of the population has to be stated.

The definition of the population should be linked with the objectives of the sample survey. Objectives can broadly be divided into two groups: estimation and inference. In order to evaluate control measures that are in place, the objectives of the survey should be centred on estimation, mainly used to produce quantitative and numerical descriptions (estimation) of relevant aspects of a target population, like the proportion of the population with a trait of interest (population infested with a specific pest). For a survey conducted to estimate a parameter of interest in a population, some level of certainty (usually expressed as a confidence interval) is associated with the estimate obtained. Confidence intervals provide a range of values in which we believe the true value of the parameter lies, and we commit a type I error if the true value does not lie within the estimated range. The probability of committing this type of error is specified in advance and incorporated into the sample size calculation during the survey design in order to keep it under control.

After clearly defining the target population and the aims of the survey, issues on how to decide on the portion of the population that needs to be included in the survey can be addressed. Such issues are collectively referred to as sample design. A choice has to be made between using probabilistic or non-probabilistic sampling methods.

The main characteristic of non-probabilistic sampling methods is that elements are chosen arbitrarily and it is not possible to associate each element with a probability of being selected. Examples of non-probabilistic methods include: (i) convenience sampling, where elements are selected if they can be easily and conveniently accessed; (ii) volunteer sampling, where elements

are included upon volunteering; (iii) judgement sampling, where the researcher decides on the elements that are likely to be representative of the population and hence included in the survey; and (iv) quota sampling, where sampling is carried out until a specific number of units (quotas) for various sub-populations have been selected. Non-probabilistic methods are prone to subjectivity and may affect the representativeness of the resulting sample. Due to arbitrariness in the selection of elements, it is difficult to quantify the impact that a non-representative sample would have on the survey results. Nevertheless, in some instances non-probabilistic methods may be the only option.

In probabilistic methods, every element in the population has a non-zero probability of being selected, thereby minimising subjectivity, and several choices exist that ensure representativeness of the sample. All probabilistic methods assume the existence of a sampling frame, from which elements can be selected. This can be in form of a list of all elements in the population or some equivalent procedure identifying the elements in the population. Within the sampling frame, sampling units also have to be defined, these are the units that will actually be selected, and these might be the individual elements or groups that contain the population elements. The definition and organisation of the sampling frame/units is one of the factors that influence the choice of the sample design.

The importance of selecting a sample that will achieve the pre-specified aims cannot be overemphasised. In general, choosing a sample design that will require input from several interested parties and trade-offs is inevitable. These trade-offs should be well documented and be integrated (if possible) into production of the statistics of interest. Note that estimates of the population characteristics and sampling variability approximation depend on the sample design; thus a survey is basically identified by its sampling design. A more detailed description of the sampling designs is given in the following sections and further details can be found Milanzi et al., (2015).

C.2. Sampling designs used for pest prevalence surveys

C.2.1. Simple random sampling

Simple random sampling (SRS) is the simplest form of drawing elements from a targeted population. It involves drawing elements successively such that each population member has an equal and non-zero probability of being selected (Barnett, 2002). Assume we have a population with N elements and we would like to draw a sample of n elements. For selection with replacement, i.e. a selected element is returned in the population and thus can be selected more than once, each population element has a selection probability of $\frac{1}{N}$ at each sampling turn. Otherwise, if selection is without replacement, selection probability changes at each sampling turn, i.e. at first sampling turn each element has $\frac{1}{N}$ selection probability, $\frac{1}{N-1}$ at the second turn, etc. When sampling is without replacement n cannot exceed N, while if sampling is with replacement n can be any value. Many statistical theories assume sampling with replacement (Kish, 1965). Sampling units are the individual population elements (inspection units).

Though the theory and mathematical properties of SRS are well developed, it is rarely used in practice, mostly because it is not feasible. For example, in the monitoring of a specific pest, a numbered list of inspection units would be required to perform a randomised selection process. When the population is too large and sparse, selected elements may be very far apart, thereby decreasing efficiency in executing the survey and increasing the costs. These and other practical considerations make SRS the least popular design in practice. Nevertheless, it is the basis of all the other designs such that in some situations computations from SRS can be used to approximate those from other more complex designs by adjusting with some known factors;

hence its properties are useful. The precision of other sampling designs is usually compared with the precision in SRS.

When estimation of some characteristics of the targeted population is of interest, it is important that the estimate be obtained with the highest precision practically possible. Sample size is thus calculated with the aim of obtaining a desired level of precision. Let $Y_1, Y_2, \ldots, Y_i \ldots, Y_N$ be elements from the targeted population with variance, $\mathrm{var}(Y_i) = S^2$, $let\ \bar{Y}$ be the population mean and $\mathrm{var}(\bar{Y}) = S^2/N$. Likewise, let $y_1, y_2, \ldots, y_i, \ldots, y_n$, $\bar{y}$, s^2, and $\mathrm{var}(\bar{y}) = s^2/n$ be the corresponding quantities from the sample (s^2 is the element variance). It can be shown that the sample size n can be obtained as:

$$n = \frac{4\,z^2 s^2}{w^2} = \frac{z^2\,s^2}{d^2}, \text{where } \mathrm{s}^2 = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n-1}$$

Where w^2 is the desired width of the confidence interval for the estimated mean, d is the margin of error, defined as the error which the researcher is willing to accept in estimating the statistic of interest and z is the normal quantile of $\alpha\ (type\ I\ error)$, the risk that a researcher is willing to accept that the true margin of error exceeds the acceptable margin of error (Bartlett et al., 2001). In general, it is clear that a smaller sample size corresponds to a larger margin of error. While d and z are usually fixed, $\mathrm{var}(\bar{y})$ depends on the sampling design.

C.2.2. Multi-stage sampling

In SRS the sampling units are individual population elements, i.e. each sampling unit has only one element. As noted in the section for SRS, this method of selecting elements is not always viable. In such cases it may be useful to select groups of elements rather than individual elements; such groups are known as clusters. Clusters are a composition of several population elements. For example, in selecting a sample for monitoring a pest in crops within a Member State, it is more practical to select fields/orchards (these can be easily enumerated) and then include a certain number of hosts from a selected field/orchard in the survey. In this example, fields and orchards are clusters since they are both made up of a group of population elements (host plants). Note that each population element can only be in one cluster at a particular time. It is important that the defined clusters do not overlap.

The obvious advantage of multi-stage sampling over SRS is its cost-effectiveness in terms of listing and locating the elements (Kalton, 1983). The major drawback is the increase in element variance.

The nature and size of the selected clusters determine whether all the elements in the selected cluster are included in the survey or further sampling within the cluster is needed. When the clusters are very large, as would be the case with fields or orchards in the example given above, elements can be sampled in two phases: firstly, the Member State is divided into clusters and SRS is used to select the required clusters; secondly, within each selected cluster a sample of elements is drawn. This is referred to as sub-sampling and it can be extended to more than two phases when necessary. When the final cluster size is small, all the elements in the selected clusters can be included in the sample, otherwise another stage of sampling may be required.

C.2.3.Impact of clustering

It is well known that the information contained in such a sample is less than the information that would have been in the same sample, assuming independence. Elements within the same cluster are likely to be more similar than elements between clusters. The strength of this similarity is quantified using the coefficient of intraclass correlation (ρ). This amount of information in

clustered data depends not only on the cluster size, but also on the structure and strength of the correlations among observations from the same cluster (Faes et al., 2009).

The impact of clustering is assessed through the design effect (D), defined as the ratio of variance of the estimate under SRS to variance under the design of interest, multi-stage sampling in this case. For $\rho > 0$, $D > 1$ and this means variance from the cluster sampling sample is larger than variance from SRS sample with the same sample size. For the maximum value of $\rho = 1$, D equals the cluster size. Thus, we would require a sample size D times larger under cluster sampling than would be required under SRS, and $\rho = 0$ corresponds to $D = 1$. That is, the variance in the cluster sample is the same as that in SRS for the same sample size. In general, bigger cluster sizes and large intraclass correction give high values of D. Design effect can also be used to obtain the effective sample size, i.e. the sample size one would need in an independent sample to equal the amount of information in the actual correlated sample.

Assume we have a targeted population for which sampling in groups (clusters) would be cost effective enough to justify the loss in precision. Define $Y_{it,}$ as the i^{th} element in the t^{th} cluster and denote $\overline{Y_c}$ as the population mean. It can be shown (Kish, 1965) that $\mathrm{var}(\overline{Y_c}) = S_c^2/A$, where A is the total number of clusters in the population. Likewise, let y_{it}, $\overline{y_c}$ and $\mathrm{var}(\overline{y_c}) = s_c^2/a$ be the corresponding quantities for the sample. Selection of elements proceeds by selecting a out of A clusters using SRS and including all elements from the selected clusters in the survey. It follows that the total sample size $n = aB$, where B is the total number of elements sampled in each cluster. The number of clusters to be selected can be obtained as:

$$a = \frac{4z^2 s_c^2}{w^2} = \frac{z^2\, s_c^2}{d^2},$$

Where

$$s_c^2 = \sum_{t=1}^{a} \frac{(\bar{y}_t - \bar{y}_c)^2}{a-1}$$

Where $\bar{y}_t$ is the estimated mean in t^{th} cluster, the rest of the parameters are as defined in the SRS scenario. Note that s_c^2 computes the variability of cluster means from the overall mean or alternatively the variance between clusters.

Alternatively, if instead of including all elements from the selected clusters, only a selection b out of the B elements within a cluster is taken, the sample size can be obtained by adjusting the SRS variance with the design effect. Design effect was defined as the ratio of variances under the SRS design and the design of interest (i.e. cluster sampling), and for the mean estimate, this implies:

$$D = [1 + \rho(b-1)] = \frac{s_c^2/a}{s^2/n}$$

$$\Rightarrow \frac{s_c^2}{a} = [1 + \rho(b-1)] \times \frac{s^2}{n}.$$

The margin of error for the mean estimate under cluster sampling is given by:

$$d = z\sqrt{\frac{s_c^2}{a}} \Rightarrow d^2 = z^2 \frac{s_c^2}{a}$$

$$\Rightarrow d^2 = z^2[1 + \rho(b-1)] \times \frac{s^2}{n}$$

$$\Rightarrow n = [1 + \rho(b-1)] \times \frac{z^2 s^2}{d^2}.$$

C.2.4.Stratified sampling

When the population of interest falls naturally into groups, sampling may be organised within each of these groups. Such groups are known as strata. As an example, for a sampling exercise encompassing an EU Member State, which is subdivided into NUTS regions, each of these NUTS regions could be considered a stratum.

In this type of sampling, the characteristic of interest is surveyed and analysed within each stratum, after which the results are combined, to provide an overall sample result. Within each stratum, various sampling procedures may be used; for instance, SRS, or multi-stage sampling.

An important consideration in stratified designs is the allocation of the total sample size to the various strata. There are different approaches to this:

- Proportional allocation

In this approach, a uniform sampling fraction is used across the strata. The sample size allocated to each stratum is proportional to the stratum size.

- Neyman allocation

Assuming equal costs across strata, the allocation that focuses on minimising sampling variance is called the Neyman allocation. Strata which have more variability are allocated a larger sample size.

Given a specified allocation scheme, and the desired precision (here represented by the margin of error), the required overall sample size, as well as the allocation to strata, can be determined.

Suppose a population of size N is stratified into H strata, each of size N_h, $h = 1, \ldots, H$. The 'weights' $W_h = \frac{N_h}{N}$ denote the population proportion of the strata. Simple random samples are drawn separately within strata.

For the estimation of the population mean $\bar{Y}$, the stratified estimator is given as

$$\bar{y}_{st} = W_1\bar{y}_1 + W_2\bar{y}_2 + \cdots + W_H\bar{y}_H = \sum_{h=1}^{H} W_h\,\bar{y}_h,$$

with $\bar{y}_h$ the stratum sample means. The variance of this estimator, ignoring the finite population correction factor, can then be expressed as

$$Var(\bar{y}_{st}) = \sum_{h=1}^{H} W_h^2 \frac{S_h^2}{n_h},$$

with n_h the stratum-specific sample size; $\sum_h n_h = n$, where n is the 'total' sample size, and S_h^2 the population variance in stratum h. The finite population correction factor can be incorporated as in Barnett (1991, p. 110), Kalton (1983, p. 20) and Groves et al. (2004, p. 112).

Estimators for the population variances in the strata are given as

$$s_h^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}(y_{hi} - \bar{y}_h)^2$$

A proportion is just a special case of a mean, and, therefore, estimation of a population proportion P follows similar logic:

$$p_{st} = W_1 p_1 + W_2 p_2 + \cdots + W_H p_H = \sum_{h=1}^{H} W_h p_h,$$

with p_h the stratum sample proportions. An estimator for the variance of p_h, ignoring the finite population correction factor, can then be expressed as

$$var(p_{st}) = \sum_{h=1}^{H} W_h^2 \frac{p_h(1-p_h)}{n_h - 1}.$$

This is usually approximated to

$$var(p_{st}) = \sum_{h=1}^{H} W_h^2 \frac{p_h(1-p_h)}{n_h}.$$

To estimate the mean with a margin of error of size d, the sample size required is derived as follows. A sample size n is required, such that

$$z\sqrt{Var(\bar{y}_{st})} = d.$$

Substituting for $Var(\bar{y}_{st})$, we have that

$$z\sqrt{\sum_{h=1}^{H} W_h^2 \frac{S_h^2}{n_h}} = d,$$

$$z^2 \sum_{h=1}^{H} W_h^2 \frac{S_h^2}{n_h} = d^2.$$

Now, $w_h = \frac{n_h}{n}$, the sample proportion of the stratum. Note that this is different from W_h, the population proportion of the stratum. From the sample proportion of the stratum, $n_h = nw_h$. Substituting for n_h above, we get

$$z^2 \sum_{h=1}^{H} W_h^2 \frac{S_h^2}{nw_h} = d^2.$$

We solve for n in this equation, obtaining

$$n = \frac{z^2}{d^2} \sum_{h} \left(\frac{N_h}{N}\right)^2 \frac{S_h^2}{w_h}.$$

- Proportional allocation

Under proportional allocation, the proportions of the sample in the stratum, w_h, are set equal to the proportions of the population in the stratum, W_h; i.e. $w_h = W_h = \frac{N_h}{N}$. The formula to calculate the sample size is then

$$n = \frac{z^2}{d^2} \sum_{h} \frac{N_h}{N} S_h^2.$$

- Neyman allocation

Neyman allocation (Groves et al., 2004, p. 117; Som, 1996, p. 211; Kalton, 1983, p. 24; Barnett, 1991, p. 120) is the allocation that minimises sampling variance, assuming equal costs across strata. It is sometimes referred to as the optimum allocation (Som, 1996, p. 211).

Neyman allocation requires the following:

$$n_h = \frac{W_h S_h}{\sum W_h S_h} n$$

This implies that for a margin of error of size d, the following sample size is required:

$$n = \frac{z^2}{d^2}\left(\sum W_h S_h\right)^2.$$

C.2.5. Designs for measuring change over time

It is often tempting to compare the results of a particular survey with similar surveys from the past with the aim of assessing change over time. This should essentially be possible if the same variable was measured in the different surveys. If measuring change over time is the main objective of the survey, it is important to outline this clearly from the beginning, because measuring change based on surveys designed to measure a different quantity, e.g. population mean, may result in less precise estimates or low power to detect the change.

A distinction is usually made between gross and net change. For example, in a pest outbreak monitoring survey, the change in the number of infested hosts from 2007 to 2010 may be measured as follows: select host elements for the 2007 survey and estimate the number of infested host cases by the specific pest; follow the same elements at some pre-specified time intervals for the whole period 2007–2010. At each time interval, estimate the number of infested hosts. By the end of study period (2010), the change in the number of infested hosts between 2007 and 2010 can be estimated. The crucial characteristic of this method of measurement is that it allows changes of an individual element to be tracked. Alternatively, after estimating the number of infested hosts in 2007, we can collect another independent sample in 2010 and compute the required estimate. The change in the number of infested hots is obtained as the difference between the estimates from the two years. The choice of which measure to use totally depends on the objective(s) of the survey.

In general, repeated survey designs are recommended for measuring change. These can either be panel designs or repeated cross-sectional surveys. Panel designs allow measurement of both net and gross change while repeated cross-sectional surveys only allow for gross change.

A longitudinal survey is a well-known form of panel design where the initial selected sample is followed for the whole period of the survey and they produce precise net change estimates. To put things in perspective, consider a survey conducted at time t and $t+1$, where the interest is in estimating change in the mean of variable y ($\bar{y}$) between time t and $t+1$, $\Delta = \bar{y}_{t+1} - \bar{y}_t$. It can be shown that

$$\mathrm{Var}(\Delta) = \mathrm{Var}(\bar{y}_{t+1}) + \mathrm{Var}(\bar{y}_t) - 2\sqrt{\mathrm{Var}(\bar{y}_{t+1})\mathrm{Var}(\bar{y}_t)}\,\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t).$$

It follows that the Δ will be estimated more precisely if $\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t)$ is high and positive. The best way to attain high and positive correlations is to use the same elements both at time t and $t+1$. This is achieved with a longitudinal survey. Medium and positive correlations can be obtained if there is some degree of overlap between elements at time t and $t+1$, which can be realised through another form of panel design, referred to as rotating panel surveys. In this design, the sample at $t+1$ will partially be composed of elements from the sample at t, hence change will be estimated with medium precision. Given that negative correlations are very rare in surveys,

$\mathrm{Var}(\Delta)$ will be the highest when $\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t) = 0$, and this corresponds to repeated cross-sectional surveys where the elements for the sample at t are different from (independent of) the elements in the sample at $t+1$. Note that while repeated cross-sectional surveys will lead to less precise change estimates than longitudinal surveys, the former produces highly precise population mean estimates ($\bar{y}_t$ and $\bar{y}_{t+1}$) than the latter. Indeed, note that

$$\mathrm{Var}(\bar{y}_{t+1}) = \mathrm{Var}(\Delta) - \mathrm{Var}(\bar{y}_t) + 2\sqrt{\mathrm{Var}(\bar{y}_{t+1})\mathrm{Var}(\bar{y}_t)}\,\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t), \text{ and}$$

$$\mathrm{Var}(\bar{y}_t) = \mathrm{Var}(\Delta) - \mathrm{Var}(\bar{y}_{t+1}) + 2\sqrt{\mathrm{Var}(\bar{y}_{t+1})\mathrm{Var}(\bar{y}_t)}\,\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t),$$

will have low values when $\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t)$ is close to zero. Thus, if the main interest is on the individual population estimates at each time point, then repeated cross-sectional surveys are recommended; otherwise, panel designs should be used.

In longitudinal surveys, both $\bar{y}_t$ and $\bar{y}_{t+1}$ estimate the population mean for the population defined at time t since the same elements are followed for the whole survey period. If the population is dynamic, $\bar{y}_{t+1}$ does not estimate the population mean for the population at time $t+1$. Population elements included in the survey at time t are likely to be selected such that the resulting sample is representative for the population at that time, which might not necessarily be representative of the population at time $t+1$.

On the other hand, in repeated cross-sectional surveys, each estimate, i.e. $\bar{y}_t$ and $\bar{y}_{t+1}$, estimates the population mean for the population at that particular time. This is because at each time point a fresh sample is selected. This important difference between the two types of survey and practicality are important determining factors for deciding which type of survey to use. For example, longitudinal surveys are impractical for the pesticide monitoring study since it is not possible to follow samples of commodities over a period of time. For pest monitoring, we could consider following up the host plants.

In addition to precision and representativeness considerations, the power to detect expected change is also a crucial factor in surveys that are meant to measure change. Say we need a power of $1-\beta = 0.80$ to detect a change $\delta = 0.05$ (in proportions) for a two-sided alternative hypothesis at $\alpha = 0.05$ type I error level. Further, let $\mathrm{Var}(\bar{y}_t) = \mathrm{Var}(\bar{y}_{t+1}) = 0.25$.

Let the survey elements at the two-time intervals be the same such that $\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t) = 0.95$. It follows that

$$\mathrm{Var}(\Delta) = 0.25 + 0.25 - 2 \times 0.25 \times 0.95 = 0.025$$

The required sample size is

$$\mathrm{n} = \frac{0.025(\,1.96 + 0.84)^2}{0.05^2} \approx 78$$

Let the survey elements at the two-time intervals overlap such that $\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t) = 0.5$. It follows that

$$\mathrm{Var}(\Delta) = 0.25 + 0.25 - 2 \times 0.25 \times 0.95 = 0.25$$

The required sample size is

$$\mathrm{n} = \frac{0.25(\,1.96 + 0.84)^2}{0.05^2} \approx 784$$

Let the survey elements at the two-time intervals be independent such that $\mathrm{Corr}(\bar{y}_{t+1}, \bar{y}_t) = 0.01$. It follows that

$$\mathrm{Var}(\Delta) = 0.25 + 0.25 - 2 \times 0.25 \times 0.01 = 0.495$$

The required sample size is

$$\mathrm{n} = \frac{0.495(\,1.96 + 0.84)^2}{0.05^2} \approx 1552$$

Thus, a repeated cross-sectional survey will require a sample size about 20 times larger than a longitudinal survey to detect a change of 0.05 with 80% power.

For a pest from which the host recovers and eradication measures are not applied, longitudinal surveys would be the preferred option to measure changes over time as the sample size could be drastically reduced. However, for those pests where the host does not recover from infestation, or eradication measures that withdraw the host plant from the targeted population are in place, cross-sectional surveys should be conducted instead.

C.3. Principles used in a surveillance design to substantiate pest freedom

The various sampling designs could be used for this purpose as well. In general, SRS schemes are discussed when sample size calculations are presented. In this document the sample size needed to detect an infestation when its prevalence is at or above a so-called design prevalence is based on the principles developed by Cannon (2002), assuming the binomial or the hypergeometric probability distributions, depending on the size of the population under investigation.

If the population can be considered infinite[32] (i.e. the individual probability of being positive does not change throughout the sampling exercise; also referred to as 'sampling with replacement'), the binomial distribution can be used:

$$CL = 1 - (1 - DP \cdot MeSe)^n$$

where CL is the confidence achieved when all test results are negatives, DP is the design prevalence, $MeSe$ is the method sensitivity and n is the sample size.

From which *n* can be derived as follows:

$$n = \frac{log(1 - CL)}{log(1 - DP \cdot MeSe)}$$

While, if the population is finite, the hypergeometric adjustment is needed. In this case, the confidence achieved is given by:

$$\mathrm{CL} \cong 1 - \left(1 - \frac{n \cdot MeSe}{N - 0.5 \cdot (N \cdot DP \cdot MeSe - 1)}\right)^{N \cdot DP}$$

where *N* is the total population size.

[32] Though a universal definition of 'infinite' and 'finite' population does not exist, the rule of thumb is that a population can be considered 'infinite' when $n/N < 0.1$ (Evans M, Hastings N and Peacock B, Statistical Distributions, Third Edition. Wiley Interscience, 2000)

In this case the sample size is given by:

$$n \cong \frac{\left(1-(1-\mathrm{CL})^{1/(N\cdot DP)}\right)\cdot\left(N-\frac{1}{2}\cdot(N\cdot DP\cdot MeSe-1)\right)}{MeSe}$$

The sample size calculation assumes a diagnostic test with 100% specificity (Sp = 1), given that the design of any survey to demonstrate the absence of a pest should specify a sequence of further testing that would be performed to clarify the true status when a positive reaction is reported.

An important assumption when SRS is considered is that no risk factor plays a role in the distribution of the infestation, meaning that the target population is homogeneously distributed in the study area (e.g. a Member State) and that the infested units are homogeneously distributed across the target population. If these assumptions are not valid, the scenario-tree modelling techniques introduced by Martin et al., (2007) can be used, which explicitly account for non-representative sampling approaches. These methods capture the effect of differential sampling from population strata with different risks of infestation, allowing quantification of the benefits of risk-based sampling. The risk-based sampling refers to the consideration of infestation risk factors when determining the sampling pressure applied in different strata of the population under surveillance (Cameron, 2012).

The principle is that the design prevalence (DP), as a single value, implies that all units within the target population have the same average probability of being infested. Scenario-tree modelling effectively divides the population into different risk groups, using the relative risk of infestation in each group to adjust *DP* to reflect the group-level probability of infestation (Cameron, 2012).

Once risk factors are identified and the levels for each risk factor determined, the target population can be subdivided into subgroups of host plants that could be considered to have the same risk of infestation. The proportion of the population within each risk subgroup i (PP_i) could be used together with the relative risk (RR_i) associated with each of the subgroups relative to a baseline subgroup. The DP is then adjusted for each subgroup considering these two parameters, relative risk and population proportion, using the following formula:

$$WR_i = \frac{RR_i}{\sum_{j=1}^{r} PP_j \cdot RR_j}$$

where WR_i is the weighted risk for subgroup i of the population. To calculate the adjusted DP (EPI_i) the original DP should be multiplied by the weights WR_i as follow:

$$EPI_i = WR_i \cdot DP$$

The EPI_i is then used in combination with method sensitivity, confidence and population size to estimate the number of samples that need to be collected in each subgroup to achieve the desired confidence for each subgroup. Where an overall confidence is desired, then convenience sampling schemes could also be considered to ensure that samples are distributed by subgroups in order to ensure the expected convenience proportions.

References

Barnett V, 1991. Sample Survey: Principles and Methods. Arnold, London. 173 pp.

Barnett V, 2002. Sample Survey: Principles & Methods. 3rd edition. Hodder Arnold, London. 241 pp.

Bartlett JE, Kotrlik JW and Higgins CC, 2001. Organizational Research: Determining Appropriate Sample Size in Survey Research. Learning and Performance Journal, 19, 43–50.

Cameron A, 2012. The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. Preventive veterinary medicine, 105, 280–286.

Cannon R, 2002. Demonstrating disease freedom - combining confidence levels. Preventive veterinary medicine, 52, 227–249. doi: 10.1016/s0167-5877(01)00262-8

EFSA (European Food Safety Authority), Ciubotaru RM, Cortiñas Abrahantes J, Oyedele J, Parnell S, Schrader G, Zancanaro G and Vos S, 2018. Technical report of the methodology and work-plan for developing plant pest survey guidelines. EFSA supporting publication 2018: EN-1399. 36 pp. doi:10.2903/sp.efsa.2018.EN-1399

Milanzi, E., Njeru Njagi, E., Bruckers, L., & Molenberghs, G. (2015). Data representativeness: issues and solutions. EFSA Supporting Publications, 12(2), 759E. https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/sp.efsa.2015.EN-759

Faes C, Molenberghs G, Aerts M, Verbeke G and Kenward MG, 2009. The effective sample size and an alternative small-sample degrees of freedom method. The American Statistician, 63, 389–399.

Groves RM, Fowler FJ Jr, Couper MP, Lepkowski JM, Singer E and Tourangeau R, 2004. Survey Methodology. John Wiley & Sons, New York.

Kalton G, 1983. Introduction to Survey Sampling. Sage, California.

Kish L, 1965. Survey Sampling. Wiley, New York. 643 pp.

Martin MAJ, Cameron AR and Greiner M, 2007. Demonstrating freedom from disease using multiple complex data sources. 1: A new methodology based on scenario trees. Preventive veterinary medicine, 79, 71–97.

Som RK, 1996. Practical Sampling Techniques, 2nd Edition, Revised and Expanded. Dekker, New York.

Stopher P and Meyburg AH, 1979. Survey sampling and multivariate analysis for social scientists and engineers. Heath, Lexinton, MA. 385 pp.

Appendix D Two-Step Approach to Sample Size Calculations

As described in Section 4.2, an epidemiological unit is a homogenous area where the interactions between the pest, the host plants, the abiotic and biotic factors and conditions would result in similar epidemiology should the pest be present.

When there is little information available on the homogeneity of the target population from which to distinguish epidemiological units, the 'two-step approach' can be applied as an alternative. This involves splitting the population into spatial divisions from which we would expect pest populations to cluster. That is, pest populations are rarely distributed evenly across a landscape, but tend to cluster in natural spatial divisions such as fields, orchards or woodlands etc. For example, a pest of deciduous trees may cluster within woodlands, and there are multiple woodlands within a landscape. In this case the epidemiological unit is an individual woodland and the overall landscape within the target population contains multiple epidemiological units (i.e. the total number of woodlands). If there are no natural spatial divisions such as fields, orchards or woodlands, or these are too heterogenous in size, the landscape can simply be considered as a grid. In this case each grid cell (e.g. hectare) is considered the smallest spatial level, and the larger spatial level is the gridded landscape that constitutes the target population.

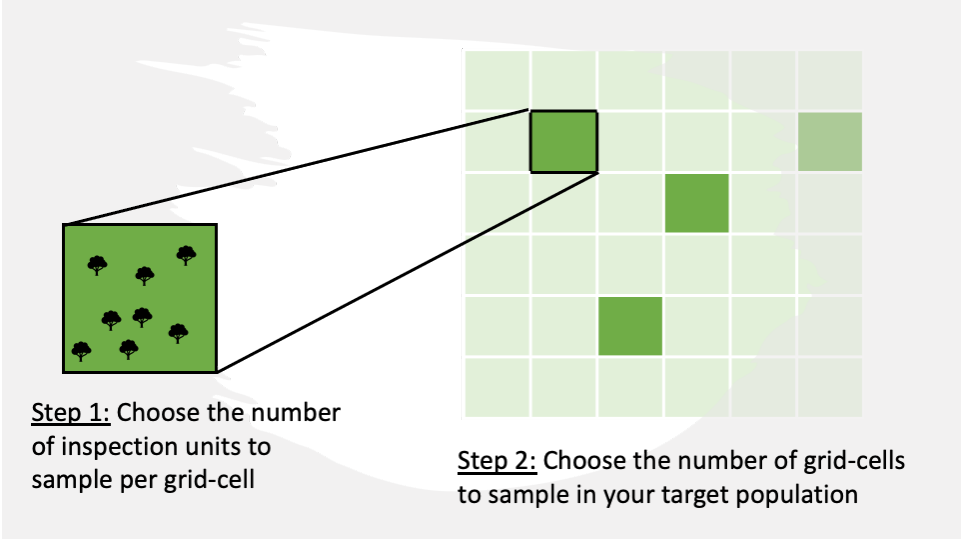
The two-step approach leads to a high number of epidemiological units since there are typically many fields, orchards, woodlands or grid cells in a survey area. So, a target population comprised of 1000 orchards would contain 1000 epidemiological units. The high number of epidemiological units will lead to a high number of samples. However, this provides a practical and systematic approach in the absence of other information to identify epidemiological units. The two-step approach involves firstly calculating the number of inspection units required to achieve a desired confidence level and design prevalence within each individual epidemiological unit (woodlands, fields, orchards, grid-cells, ...). Secondly, the total number of epidemiological units that need to be sampled in order to achieve a defined overall confidence level and design prevalence within the target population, is calculated.

D.1. Grid-based approach

The simplest case is to consider the landscape as a grid. In this case, the 'two-step' survey design is developed by dividing the target population into grid-cells and first estimating the number of inspection units that need to be sampled within each grid-cell (Step 1) and then by estimating the number of grid-cells that need to be inspected and sampled (Step 2) to achieve the desired overall confidence level for detecting the pest above the design prevalence.

The confidence level that is achieved when calculating the number of inspection units to sample within each grid-cell in the first step is used as the method sensitivity when calculating the number of grid-cells that need to be inspected in the second step.

As a consequence, the higher the confidence within a grid-cell, the fewer grid-cells need to be inspected, and inversely if the confidence at grid-cell level is decreased, the more grid-cells need to be inspected to achieve the overall pre-determined confidence level of the survey.



Step 1: Number of inspection units to inspect within each grid-cell. In this first step, the objective is to calculate the number of inspection units that need to be sampled within each grid-cell. The confidence level and design prevalence for a grid-cell need to be set. The number of inspection units to sample within each grid-cell will then be determined based on the method sensitivity and the total number of inspection units within the grid-cell. **Step 2: number of grid-cells to inspect.** In this second step, the objective is to calculate the number of grid-cells to select. Notably, the confidence-level that was used for each individual grid-cell becomes the method sensitivity to be used when calculating the total number of grid-cells that need to be sampled.

Note that the risk manager can choose whether to sample a low number of grid-cells with an intensive sampling effort, or a high-number of grid-cells at low sampling effort. Which approach is most cost-effective will e.g. depend on the costs per sample within a grid-cell and the travelling costs associated with sampling multiple grid-cells. If these costs are known, RiPEST or RiBESS+ can be used to determine the optimal strategy.

D.2. Example calculations using the two-step approach

Consider a target population for a pest which is comprised of 1,500,000 trees. Information to divide the population into broad epidemiological units is not available and so the area is split into

one-hectare grid cells. The inspection unit is the individual tree and we have estimated a method sensitivity of 0.7 (i.e. there is a probability of 0.7 to detect the pest in an individual tree, if it is present). Our survey will inspect 200 trees in each grid-cell, and we will survey 100 grid-cells. We assume an overall design prevalence of 1% and wish to calculate what overall confidence level our survey will achieve.

Table 1: Survey design parameters for a hypothetical application of the two-step approach.

Survey design		
Target population	Survey area	1,500,000 trees
	Epidemiological unit	1 ha
	Inspection unit	Individual tree
Survey parameters	Design prevalence (DP)	1%
	Confidence level (CL)	95%
	Method sensitivity (MSe)	0.7
	Within grid-cell confidence (β)	To be determined
	Within grid-cell design prevalence (α)	1%
	Within grid-cell sample size (N1)	200
	Number of grid-cells sampled (N2)	100

Step 1a. We first calculate the probability to detect the pest if it is present within an individual hectare. If we assume no risk factors are available, then this is calculated by RiPEST as:

$$\beta = 1 - (1 - \alpha \mathrm{MSe})^{N1}$$

Our choice of survey parameters in Table 1 outlines that we will sample 200 trees within each hectare grid-cell, each with a sensitivity of 0.7, and we wish to achieve a design prevalence of 1%. This leads to a within grid-cell confidence level, β, of approximately 0.75.

Step 2a. Next, we wish to calculate the overall confidence level of the survey. The sensitivity of the overall survey is given by the within grid-cell confidence level, β, which we estimated in step 1 to be 0.5, and the design prevalence (DP) is the same as in Step 1. To calculate the overall confidence level, in the simple case where there are no risk factors, we have:

$$\mathrm{CL} = 1 - (1 - \beta \mathrm{DP})^{N2}$$

.

Using our survey design parameters (Table 1) this leads to an overall confidence level of 90%.

This approach also allows the risk manager to optimise their survey design as the same overall confidence level can be achieved with different combinations of N1 and N2.

Appendix E Monitoring survey

E.1. Sample sizes

There are two paradigms in statistical methodology for determining the sample size for a representative sample (i.e. the number of inspection units to be sampled). Both focus on one or more statistical parameters of interest. In this case, the parameter of interest is either the prevalence $\pi(t)$ at time t or the parameter(s) reflecting the trend in the prevalence over time. The two paradigms are:

1. **CI-paradigm** (confidence interval paradigm). If no prior research hypotheses have been formulated about these parameter(s), the confidence interval paradigm is the obvious approach: the accurate estimation of the parameter(s) of interest by a confidence interval (region) of a predetermined width (area) and confidence (typically 95%).
2. **HT-paradigm** (hypotheses testing paradigm). If research hypotheses can be formulated about these parameters, hypotheses testing is the preferred approach: testing the hypotheses of interest with a predetermined level of significance (by controlling the probability of a type I error of wrongly rejecting the null hypothesis, which is typically set at 5%) and a predetermined power (by controlling the probability of a type II error of wrongly failing to reject the null hypothesis).

In the context of pest prevalence surveys, the hypothesis of interest states that the eradication programme will reach the design prevalence π_{DP}. It is then the objective to determine whether or not a downward time trend in the prevalence from its initial value $\pi(0)$ to the design prevalence π_{DP} is reached within the scheduled time frame $(0, T_E)$. Given a specific target pest and target population of host plants, the sample size (i.e., the number of 'inspection units' to be examined and/or tested) should allow the detection of a particular downward time trend with a predetermined level of significance and power.

When the initial prevalence $\pi(0)$ is unknown, it is necessary to design a survey to accurately estimate the initial prevalence (using the CI-paradigm for its design). Therefore, the proposed procedure combines both paradigms.

Further details on the proposed methodology can be found in Milanzi et al. (2015).

E.1.1. Logistic regression model

The logistic regression model gives a basic model for a time trend in the prevalence

$$\log\left(\frac{\pi(t)}{1-\pi(t)}\right) = \beta_0 + \beta t, \;\; t = 0,1,\ldots T_E, \qquad (1)$$

where $\pi(t)$ represents the prevalence at time t, β_0 is the intercept, and β the slope parameter that indicates the time trend.

For an eradication programme, it is expected that the trend is negative, meaning $\beta < 0$. To ensure that a **specific negative trend $\boldsymbol{\beta^* < 0}$** is detected and the null hypothesis (*) of no trend is rejected in favour of this negative trend

$$H_0: \beta = 0 \quad \text{versus} \quad H_1: \beta = \beta^*, \qquad (*)$$

with a sufficiently high **predetermined power** and at a **predetermined level of significance** α, the sample size needs to be sufficiently large (Demidenko, 2007):

$$n \approx \left(\frac{z_{1-\alpha} + z_{\text{power}}}{\beta^*}\right)^2 V.$$

z_{prob} denotes the critical point (quantile) of the standard normal distribution corresponding to the probability $prob$, which in the formula are $1-\alpha$ or power, and V is the variance of the slope estimate at the negative slope β^*. This variance V is the 2nd diagonal element of the inverse of the Fisher information matrix. This information matrix depends on the relative distribution of the time points $t = 0,1,\ldots T_E$ as scheduled in the survey programme and on the unknown intercept β_0 and slope β. We assume an even distribution of the planned survey activities over time (implying the same sample size at each time point). The logit($\pi(0)$) and the alternative value β^* need to be inserted for β_0 and β.

E.1.2. Determining the slope β^*

The specific negative time trend β^* is determined by the prevalence at the start of the eradication programme $\pi(0) = \pi_0$ and the design prevalence $\pi(T_E) = \pi_{DP}$ at the end of the eradication period:

$$\beta^* = \log\left(\frac{\pi_{DP}(1-\pi_0)}{(1-\pi_{DP})\pi_0}\right)/T_E$$

which can be interpreted as the log odds ratio per unit of time.

E.1.3. Estimation of the initial prevalence

The exact value of the design prevalence π_{DP} should be established by the risk manager, but the value of the prevalence π_0 is not known. In some cases, suitable data to estimate π_0 are already available, like the samples from the delimiting survey. However, it is important that this initial prevalence is estimated sufficiently accurately and is representative of the entire target population, as this prevalence is the starting point used for all sample size calculations. A separate sample size calculation for a cross-sectional survey at $t = 0$ might indicate that additional data have to be collected (complementing the data from the delimitation). The sample size for such an initial prevalence survey can be determined using the R-function ssize.propCI() of the MKpower R-package or EFSA's SAMPELATOR app.

E.1.4. Sample sizes for each time period

Once the slope β^* has been determined, the sample size for each time point of the survey programme ($t = 1,\ldots T_E$) can be calculated by distributing the total calculated sample size over the T_E time periods. The default approach here is to include the data at time 0 and distribute the total size evenly. Other choices are possible, but they complicate the calculation of time-specific sample sizes and the computational implementation.

E.1.5. Estimation of the slope and the final prevalence

The data collected during the consecutive time points of the survey programme can be used to evaluate whether the imposed eradication measures are effective. Once the entire survey programme is completed and data have been collected from the prevalence survey at the start of the eradication programme (at time 0) and from the repeated cross-sectional surveys, the logistic regression model (1) can also be fitted to all data to estimate the slope and the final

prevalence at time T_E. The confidence interval for $\pi(T_E)$ is particularly interesting, because it allows us to check whether the eradication objective has been achieved by verifying whether this confidence interval contains the design prevalence π_{DP}. Declaring eradication measures successful requires an additional detection survey to be conducted after time T_E, which yields no findings.

E.1.6. Improving the logistic regression model by fractional polynomials

The pattern of the decreasing trend from $\pi(0) = \pi_0$ to $\pi(T_E) = \pi_{DP}$ is unknown and can theoretically take many shapes. Formula (1) represents the most straightforward trend: a constant decrease over time (linear) on the logit (log odds) scale. However, the actual trend might be more complex. A fractional polynomial (FP) is very flexible and is defined as

$$\text{logit}(\pi(t)) = \begin{cases} \beta_0 + \beta(t+1)^p & p \neq 0 \\ \beta_0 + \beta\log(t+1) & p = 0 \end{cases} \qquad (2)$$

with p as an unknown (fractional) power that can take any real value. Taking $p = 1$, model (2) reduces to model (1) (although with a different intercept β_0). This simple but flexible model can take many different shapes, as depicted in Figure 1.

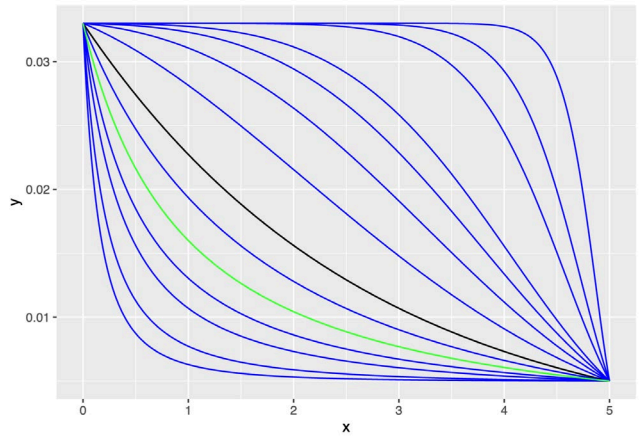


Figure 1: Different shapes of the fractional polynomial for $\pi(0) = \pi_0 = 0.033$ and $\pi(T_E) = \pi_{DP} = 0.005$ and p ranging over the values -3,-2,-1,-0.5,0,0.5,1,2,3,4,5,10,15,30; the black curve corresponds to model (1) with $p = 1$ and the green curve corresponds to $p = 0$ (log-model). Axis x is the time step (in years), and axis y the simulated prevalence.

The FP-model (2) can be used at different stages.

E.1.6.1. Using the FP model for the final evaluation

The FP model (2) can be used to analyse all collected data at the end of the eradication programme (Section E.1.5). The R function gamlss() (in the homonymous R-package), using the function fp() in the formula, can be used for fitting fractional polynomial logistic regression. This will improve the estimation of the actual trend over time and of the final prevalence. However, when the sample sizes were calculated on a possibly poor-fitting model (1) and were not

determined to guarantee an accurate estimation of final prevalence $\pi(T_E)$, the obtained inference results might lack sufficient power and accuracy.

E.1.6.2. Using the FP model at an intermediate design stage

Ideally, the FP model is already applied at the initial design stage, but the power p in model (2) is unknown and cannot be estimated based on the initial sample at time $t = 0$. A possible strategy is to estimate p from the data at an early intermediate time point $t = 1$ or 2, and then to recalculate the sample sizes for the subsequent time points. This recalculation can be done once or repeatedly at every next time point, based on a new estimate for the power p using all available data.

E.1.6.3. Using the FP model with an epidemiological transition model

If a transition model can be used to predict the expected sequence of prevalences over time, these can be used to determine the best fitting model (2) with the corresponding value for the power p. The sample sizes can then be calculated using the FP model (2) with this value for p.

E.1.7. Prediction of $\pi(T_E)$ at intermediate time points

When the prevalence data collected at intermediate time points is used for evaluation of the FP model (subsection E.1.6.2), the fitted FP model can also be used to predict the prevalence $\pi(T_E)$ at the end of the eradication programme. This is helpful for assessing whether the progress of the eradication programme and the objective to reach the design prevalence π_{DP} are on course. If not, this predication may trigger adoption of additional phytosanitary measures.

E.1.8. Estimation of the prevalence $\pi(T_E)$

As described in Section E.1.5, all collected data can be used to estimate the final prevalence based on the applied model and to test whether the design prevalence was reached by verifying whether the obtained confidence interval for $\pi(T_E)$ contains the design prevalence π_{DP}. However, as the sample sizes were not determined to accurately estimate $\pi(T_E)$, this confidence interval might be too wide. In that case, an additional sample could be taken at (or after) time $t = T_E$. The size of that sample could be calculated with the objective to use only the data at $t \geq T_E$ without the data obtained at $t < T_E$ and any model or using all available data over the 5-year period and using the assumed model.

E.1.9. Accounting for heterogeneity

When data are hierarchically structured (clustered) or other sources of overdispersion or heterogeneity are known to be in effect, the additional parameter accounting for this phenomenon can be estimated when data from the first pest prevalence surveys become available. The sample size for the next periods t can then be adapted by the approximate formula (see e.g. Faes et al., 2009)

$$n_t \leftarrow n_t \times \varphi$$

with $\varphi \geq 1$ the overdispersion (heterogeneity) factor, also called the *design effect*, which for clustered data takes the form

$$\varphi = (1 + (m - 1)\rho)$$

where m is the (average) cluster size, and ρ the intra-cluster correlation. In case there is some beforehand knowledge about φ from literature, this can already be used when planning the first survey.

The design effect can have a large impact. Table 1 shows values for the design effect for some values of the cluster size and the intra-cluster correlation. For instance, a very moderate correlation of 0.1 and a small cluster of size 10 imply almost a doubling of the required sample size. For a cluster of size 50, it increases to about sixfold the sample size with no design effect (intra-cluster correlation 0). At a local scale, pests can have a clustered distribution when patches of infested plants (disease foci) are surrounded by healthy plants. This pattern results from an initial introduction of the pest in a field (or a forest stand) followed by spread to neighbouring plants. Also at a larger geographical scale, plants that are closer to the initial source of an infestation have a higher probability to become infested because the natural spread capacity is a limiting factor in reaching new host plants. However, the degree of clustering, and hence the design effect, is often not known beforehand.

Table 1: Design effect φ for different intra-cluster correlations and different (average) cluster sizes.

	rho=0	rho=0.05	rho=0.1	rho=0.25	rho=0.5	rho=0.75	rho=1
m=2	1	1.05	1.1	1.25	1.5	1.75	2
m=10	1	1.45	1.9	3.25	5.5	7.75	10
m=25	1	2.20	3.4	7.00	13.0	19.00	25
m=50	1	3.45	5.9	13.25	25.5	37.75	50
m=100	1	5.95	10.9	25.75	50.5	75.25	100

For fitting the available data, the model's distributional component must also be adapted. The binomial likelihood function can be replaced by extensions such as the beta-binomial likelihood function (see e.g. Aerts et al., 2002).

E.1.10. Accounting for method sensitivity

The sample size calculations and the model need to account for the method sensitivity. Assuming the method sensitivity does not change over time, a pragmatic approach would be to base all data-based estimates on adjusted prevalences (by dividing the observed prevalences by *MeSe*). More advanced methods are available which are based on the EM algorithm (see e.g. Magder and Hughes, 1997), but these would complicate the sample size calculations.

E.1.11. Finite population adjustment

In case the target population size N is finite and known, and in case the sample comprises more than 5% of the total population, the sample size n can be adjusted downward by the population correction factor:

$$n \leftarrow n \times \frac{N}{N+n}.$$

Population size might change over time during eradication, making such corrections less straightforward. A pragmatic approach would be to use the largest population size (correction factor closest to 1) in all calculations, i.e. the population size at the start of the eradication programme.

E.2. Simulations

This section describes four scenarios in an infinite target population that are used for indicative simulation studies. We consider an eradication programme of $T_E = 5$ years with the intended design prevalence $\pi(T_E) = \pi_{DP}$ equal to 0.005. The true (but considered unknown) prevalence at the start of the eradication programme $\pi(0) = \pi_0$ is taken as 0.0333 (based on the data used in van Woensel et al., 2025).

The initial prevalence π_0 must be estimated sufficiently accurate (see Section 9.3.3). Assuming that the target pest is not widespread and $\pi_0 \leq 0.1$, and by taking a confidence level 0.95 and width 0.025, the required sample size for estimating π_0 would be 2292.

Simulating a random sample of size 2292 at $t = 0$ (with $\pi(0) = 0.0333$), we obtain the estimate 0.03098.

Using this estimate for π_0, $\pi_{DP} = 0.005$ and the logistic regression model (1), the slope of interest $\beta^* < 0$ is computed as -0.3701. As discussed in Section 9.3.1, the sample size for the survey at each time point $t = 1, \ldots, 5$ can be calculated, taking a level of significance 0.05 and power 0.99. These assumptions result in a total sample size of 1680 that is then evenly distributed over the 5 years. This results in a sample of size 336 in each year of the eradication programme. The calculated sample size would increase rapidly when the initial prevalence $\pi(0)$ gets closer to the design prevalence (see Figure 1 in Appendix E.3.1).

This latter sampling in $t = 1, \ldots, 5$ is repeated 1000 times. For each of these 1000 simulation runs, the test for detecting the trend is performed (rejecting the null hypothesis or not), and the confidence interval for the final prevalence $\pi(T_E)$ is computed to check whether this interval contains the design prevalence. The width of the interval is determined as well.

The scenarios are:

1. Scenario 1: a linear trend and design prevalence reached

This is the ideal setting in which the eradication programme reaches the design prevalence, and the model used is the true model (Model 1).

2. Scenario 2: a linear trend and design prevalence not reached

The correct model is used (Model 1), but the eradication programme does not reach the design prevalence, instead reaching a prevalence of 0.015

3. Scenario 3: a misspecified trend

The eradication programme is effective (reaches the 0.005 prevalence), but Model (1) is used, whereas the eradication pattern follows that of a fractional polynomial of degree 3 (Model 2 with power $p = 3$).

4. Scenario 4: using fractional polynomials

Identical to Scenario 3, but a fractional polynomial model is used for the analysis at the end of the eradication programme. As the true power is unknown, this value is estimated from the data.

E.2.1. Results

Table 2 shows the results. For all scenarios, except for scenario 2, the trend was detected in nearly all runs (79% for scenario 2). The average point estimate for the final prevalence $\pi(T_E)$ is close to the design prevalence 0.005 in scenario 1 and 4. In scenario 2 it is about 0.015, which was indeed the true prevalence in this scenario (the eradication programme was unsuccessful in reaching the predetermined design prevalence 0.005). So, the data and the analysis show that the eradication programme was unsuccessful and further efforts are needed to reach the design prevalence 0.005. The left coverage was almost 0, as almost all confidence intervals were lying beyond 0.005. In scenario 3, the average point estimate for the final prevalence $\pi(T_E)$ is about twice 0.005, and the left coverage is too low (about 36%). The reason is that model (1) was incorrect and not flexible enough to follow the eradication pattern. That is also the case for the sample size calculations in scenario 4, but when a fractional polynomial is used to analyse the data this seems to improve the results considerably.

Section E.3.2 shows some more details of the results of the simulations.

Table 2: Simulation results for the four scenarios. Out of 1000 simulation runs: % runs in which a trend was detected, the average of the point estimate for the final prevalence, % runs in which the lower limit of the confidence interval (CI) was below the design prevalence 0.005, % runs in which the upper limit of the CI was above the design prevalence 0.005, % of CI containing the design prevalence 0.005, the average width of the 1000 CI's.

Scenario	% Trend detections	Average point estimate $\pi(T_E)$	% Left coverage CI	% Right coverage CI	% Two-sided coverage	Average width CI
1	100	0.0055	92.6	99.1	91.7	0.010
2	79	0.0158	0.3	100	0.3	0.018
3	99.8	0.0101	35.6	100	35.6	0.014
4	100	0.0059	88.9	98.8	87.7	0.013

E.2.2. Further discussion

Even though the explored scenarios are limited in number, the results of the simulations indicate that the sample size methodology is acting appropriately. Further simulations are needed to investigate:

1. The effect of heterogeneity and method sensitivity (Section 9.3.9 and 9.3.10). Especially intra-cluster correlation and cluster size can have a huge multiplicative effect on the required sample sizes (see Table 1).

2. The simulations were conditional on a single realization of the sample size calculation to estimate the initial prevalence π_0. A double-nested simulation design is required to investigate the effect of variation in that initial estimation step.

3. The effects of intermediate adaptation of the sample size calculations (Section 9.3.6.2), the use of an epidemiological transition model (Section 9.3.6.3), and the intermediate prediction of the prevalence at the end of the eradication programme (Section 9.3.7).

4. The methodology is detecting the downward trend in pest prevalence, which is assumed to be attributed to the control measures applied. Nevertheless, there might be additional factors that could contribute to the decline of the pest population (e.g. environmental conditions).

E.3. Further details on the simulations

E.3.1. Sample size

The sample size formula shown in Section E.1.1 with the expression for β^* from Section E.1.2 heavily depends on the required power and on the difference between the initial value π_0 and final prevalence π_{DP}. Figure 2 below shows this relationship for the choices taken in the simulation section (solid black line; level 0.05 and power 0.99), but also for power 0.95 (dashed line) and 0.90 (dotted line). Sample sizes are shown as $\pi(0) = \pi_0$ takes values from 0.01 to 0.05. A smaller decrease in prevalence is more difficult to detect and a larger (annual) sample size is needed when the initial prevalence is closer to the design prevalence. The blue line marks the required annual sample size of 336 corresponding to an estimate of 0.03098 for the initial prevalence. Finally, the lower the required power, the smaller the required sample size.

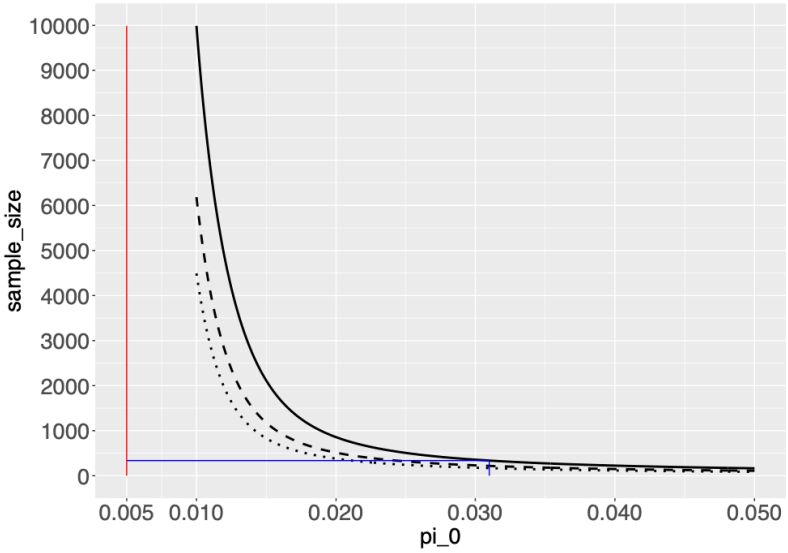


Figure 2: The required annual sample size as a function of the initial prevalence (taking values from 0.01 to 0.05) for level 0.05 and power 0.99 (solid black), 0.95 (dashed black) and 0.90 (dotted black). The blue line marks the required annual sample size of 336 corresponding to an estimate of 0.03098 for the initial prevalence.

E.3.2. Different scenarios

This supplementary material shows some individual results of five randomly selected simulation runs. The figures 3-6 shows the true unknown eradication pattern (curve in red, solid line) and the pattern used for the sample size calculations (curve in black, solid line). In Figure 3 (scenatio 1) these two curves only differ in their starting value $\pi(0) = \pi_0$ (as this value is estimated in the initial survey at $t = 0$). The five dashed broken-line curves connect the annual observed prevalences for the samples taken in five randomly selected simulation runs. These curves are not necessarily monotone, as random fluctuations can cause next year's observed prevalence to be coincidentally higher (although the true underlying pattern is decreasing). They show how observations might look in practice. Tables 3-6 show some results of the analyses at the end of the eradication programme: the selected run in the first column; the estimated slope of the fitted logistic regression model; the p-value of the trend test; the left limit, the point estimate, and the right limit of the confidence interval for the final prevalence $\pi(T_E)$. The objective is to have 99% of the runs detecting the trend and to have point estimates close to and confidence intervals containing the design prevalence 0.005.

E.3.2.1. Scenario 1

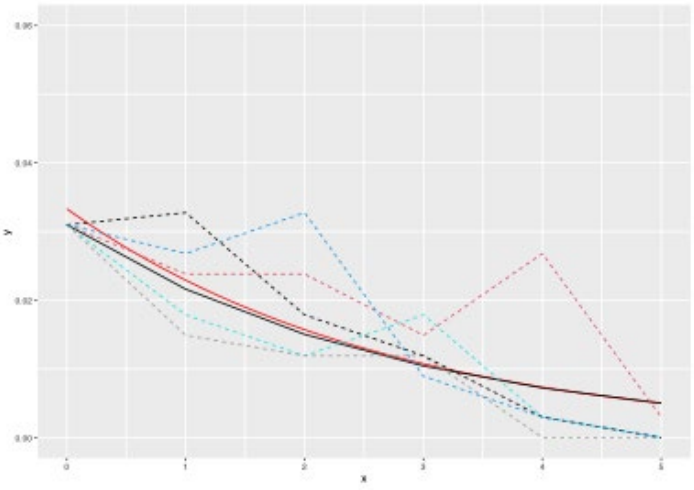


Figure 3: Scenario 1: the true unknown eradication pattern (curve in red, solid line) and the pattern used for the sample size calculations (curve in black, solid line). The five dashed differently coloured broken-line curves connect the annual observed prevalences for the samples taken in five randomly selected simulation runs.

Table 3: Scenario 1: results of the analyses at the end of the eradication programme for the samples taken in five randomly selected simulation runs: the selected run (1st column); the estimated slope of the fitted logistic regression model (2nd column); the p-value of the trend test (3rd column); the left limit, the point estimate, and the right limit of the confidence interval for the final prevalence (4th, 5th and 6th columns respectively).

simulation	slope	p-value	left	prev t=5	right
61	-0.4552	0.0000	0.001	0.0033	0.009
449	-0.4455	0.0000	0.001	0.0036	0.009
288	-0.5873	0.0000	0.001	0.0017	0.006
178	-0.1967	0.0026	0.006	0.0119	0.022
684	-0.4055	0.0000	0.002	0.0045	0.011

In this scenario, all runs i) detect the trend, ii) have point estimates for the final prevalence relatively close to 0.005 (except run 178), and iii) all confidence intervals (except run 178) contain the value 0.005. Of course, this scenario reflects the ideal situation, in which a correct model has been used, and the eradication programme was effective in reaching the design prevalence 0.005.

E.3.2.2. Scenario 2

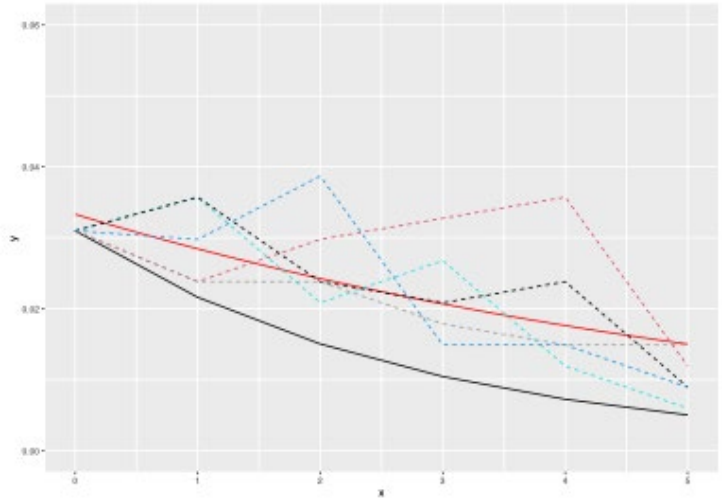


Figure 4: Scenario 2: the true unknown eradication pattern (curve in red, solid line) and the pattern used for the sample size calculations (curve in black, solid line). The five dashed differently coloured broken-line curves connect the annual observed prevalences for the samples taken in five randomly selected simulation.

Table 4: Scenario 2: results of the analyses at the end of the eradication programme for the samples taken in five randomly selected simulation runs: the selected run (1st column); the estimated slope of the fitted logistic regression model (2nd column); the p-value of the trend test (3rd column); the left limit, the point estimate, and the right limit of the confidence interval for the final prevalence (4th, 5th and 6th columns respectively).

simulation	slope	p-value	left	prev t=5	right
61	-0.2133	0.0013	0.006	0.0114	0.021
449	-0.1507	0.0099	0.009	0.0154	0.026
288	-0.1672	0.0068	0.008	0.0136	0.024
178	-0.0557	0.1632	0.015	0.0238	0.037
684	-0.1872	0.0030	0.007	0.0130	0.023

The broken line curves (and the red solid curve) show that the eradication programme did not succeed in reducing prevalence to 0.005. As expected because of the smaller difference between the design prevalence and the initial prevalence, the trend is not always detected (run 178 shows a p-value above 0.05), and the estimates for $\pi(T_E)$ are too high and show that eradication was not effective and further measures are needed.

E.3.2.3. Scenario 3

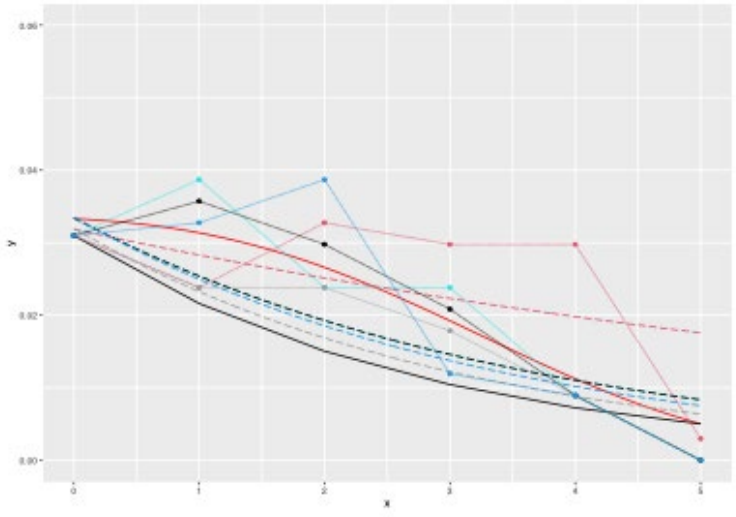


Figure 5: Scenario 3: the true unknown eradication pattern (curve in red, solid line) and the pattern used for the sample size calculations (curve in black, solid line). The five solid differently coloured broken-line curves connect the annual observed prevalences for the samples taken in five randomly selected simulation runs. The five dashed differently coloured smooth curves show the fitted models, estimated from the observed prevalences and in the same colour as these respective prevalences.

Table 5: Scenario 3: results of the analyses at the end of the eradication programme for the samples taken in five randomly selected simulation runs: the selected run (1st column); the estimated slope of the fitted logistic regression model (2nd column); the p-value of the trend test (3rd column); the left limit, the point estimate, and the right limit of the confidence interval for the final prevalence (4th, 5th and 6th columns respectively).

simulation	slope	p-value	left	prev t=5	right
61	-0.2828	0.0001	0.004	0.0083	0.017
449	-0.2828	0.0001	0.004	0.0083	0.017
288	-0.3280	0.0001	0.003	0.0064	0.014
178	-0.1220	0.0247	0.010	0.0176	0.029
684	-0.3041	0.0001	0.004	0.0075	0.015

The eradication now follows a different pattern (red solid line), whereas the design assumed the same pattern (black solid line) as in scenarios 1 and 2 (misspecified design model). The observed data (broken-line curves) do follow the red line, but are analysed with the wrong model. Additional dashed curves show the fitted models (in the same colour as the data). As a consequence of the misspecified model, the estimates for $\pi(T_E)$ in the Table 5 are biased upwards.

E.3.2.4. Scenario 4

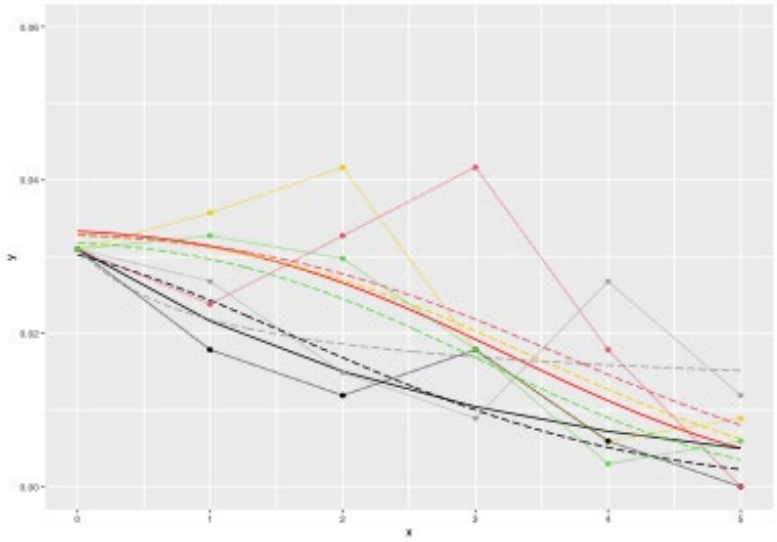


Figure 6: Scenario 4: the true unknown eradication pattern (curve in red, solid line) and the pattern used for the sample size calculations (curve in black, solid line). The five solid differently coloured broken-line curves connect the annual observed prevalences for the samples taken in five randomly selected simulation runs. The five dashed differently coloured smooth curves show the fitted models, estimated from the observed prevalences and in the same colour as these respective prevalences.

Table 6: Scenario 4: results of the analyses at the end of the eradication programme for the samples taken in five randomly selected simulation runs: the selected run (1st column); the estimated slope of the fitted logistic regression model (2nd column); the p-value of the trend test (3rd column); the left limit, the point estimate, and the right limit of the confidence interval for the final prevalence (4th, 5th and 6th columns respectively).

sim	power	slope	p-value	left	prev t=5	right
432	-0.5	1.2457	0.0036	0.010	0.0151	0.024
891	3.0	-0.0104	0.0002	0.001	0.0035	0.012
825	2.0	-0.0751	0.0000	0.001	0.0023	0.008
287	3.0	-0.0078	0.0007	0.002	0.0063	0.016
866	3.0	-0.0067	0.0015	0.003	0.0080	0.019

This scenario is the same as the previous scenario 3, but the model used for the analysis is a fractional polynomial with power estimated from the data. The Figure 6 now shows fitted curves that do reflect the true pattern (red solid line), and they do no longer overestimate the prevalence $\pi(T_E)$ so much. The additional (second) column in the table shows the estimated power of the fractional polynomial (true power is 3). All runs detect the trend, and all confidence intervals (except that of run 432) contain the design prevalence 0.005. So, although the wrong model was used at the design stage, the use of the estimated fractional polynomial at the analysis stage corrected the estimation in the right direction.

References

Aerts M, Geys H, Molenberghs G and Ryan L, 2002. Topics in Modelling of Clustered Data. Chapman & Hall, London.

Demidenko E, 2007. Sample size determination for logistic regression revisited, Statistics in Medicine, 26, 3385–3397.

Faes C, Molenberghs G, Aerts M, Verbeke G and Kenward MG, 2009. The effective sample size and an alternative small-sample degrees-of-freedom method. The American Statistician, 63, 389-399.

Magder LS, and Hughes JP, 1997. Logistic regression when the outcome is measured with uncertainty. American Journal of Epidemiology, 146, 195-203.

Milanzi, E., Njeru Njagi, E., Bruckers, L., & Molenberghs, G, 2015. Data representativeness: Issues and solutions. *EFSA Supporting Publication 2015; 12(2):EN-759, 159 pp. doi:10.2903/sp.efsa.2015.EN-759*

van Woensel E., van der Werf W., Parnell S., van Vliet A., Douma B.J.C.,2025. Exploring the contribution of citizen science to statistically sound and risk-based surveillance of insect pests in the EU. EFSA supporting publication 2025:EN-9186. 94 pp.doi:10.2903/sp.efsa.2025.EN-9186