**Food and Agriculture Organization of the United Nations**

17

manual

# RISK-BASED DISEASE SURVEILLANCE

A manual for veterinarians on
the design and analysis of surveillance for
demonstration of freedom from disease

FAO **ANIMAL PRODUCTION AND HEALTH**
**manual**

# RISK-BASED
# DISEASE SURVEILLANCE

A manual for veterinarians on
the design and analysis of surveillance for
demonstration of freedom from disease

**Authors**
A. Cameron, F. Njeumi, D. Chibeu and T. Martin

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS
Rome, 2014

# Contents

# Acronyms

| | |
|---|---|
| **ABCRC** | Australian Biosecurity Cooperative Research Centre for Emerging Infectious Diseases |
| **AGID** | agar gel immunodiffusion |
| **ALOP** | acceptable level of protection |
| **BSE** | bovine spongiform encephalopathy |
| **CBPP** | contagious bovine pleuropneumonia |
| **CFT** | complement fixation test |
| **CSF** | classical swine fever |
| **ELISA** | enzyme-linked immunosorbent assay |
| **EU** | European Union |
| **FAO** | Food and Agriculture Organization of the United Nations |
| **FMD** | foot-and-mouth disease |
| **HPAI** | highly pathogenic avian influenza |
| **ISVEE** | International Society for Veterinary Epidemiology and Economics |
| **NSP** | non-structural protein ELISA |
| **OIE** | World Organisation for Animal Health |
| **PDS** | participatory disease searching |
| **PCR** | polymerase chain reaction |
| **RBT** | Rose Bengal test |
| **SNT** | serum neutralization test |
| **SPS** | sanitary and phytosanitary |
| **WTO** | World Trade Organization |

# Preface

*"Science may be described as the art of systematic over-simplification."*
**Karl Popper (1902–94)**

*"We must plan for freedom, and not only for security, if for no other reason than that only freedom can make security secure."*
**Karl Popper (1902–94)**

## INTRODUCTION

From a disease point of view, the world is becoming a more dangerous place. Increasing global population and improvements in the standard of living mean that there is a rapidly increasing demand for animal protein. To meet this demand, animal production has intensified. The international movement of animals and animal products has been made cheaper and faster through improved transport infrastructure. Increasing human and livestock population has placed pressure on wildlife habitats, resulting in closer contact between wildlife, domestic animal populations and humans.

This complex mix of factors means that 'traditional' livestock diseases have the opportunity to spread and multiply much more quickly, and that 'new' diseases, arising from wildlife populations or genetic changes in existing pathogens, have a much greater chance to impact on animal and human populations.

Managing these disease threats poses enormous challenges and requires inputs from many disciplines. Good quality information is one essential requirement: what diseases exist; where they are found; what impact they are having; which populations are at risk; how we can prevent, control or eradicate these diseases. Animal disease surveillance plays a central role in providing this information.

Risk-based surveillance is not a particular technique; rather, it describes a general approach to undertaking disease surveillance. The principle is simple and self-evident: the most efficient way to find disease is to survey the animal populations that are most likely to be affected. This is in contrast to the more traditional statistically-based approach of taking representative samples from a population.

While the idea of risk-based surveillance is simple, the implications are complex. The approach can be much more cost-effective for some purposes, but if misused, it can lead to serious errors or it can be more expensive than traditional approaches. At the heart of risk-based surveillance is an understanding of risk factors and their impact on disease distribution. If these factors are poorly understood, risk-based surveillance may not be an appropriate tool. The analysis of data collected through risk-based surveillance has required the development of new analytical techniques.

## PURPOSE OF THIS MANUAL

This manual seeks to present a comprehensive overview of the issues relating to risk-based surveillance for the purpose of demonstrating freedom from disease. It is targeted at veterinarians

who are interested in surveillance and the analysis of surveillance data. While a number of the concepts are necessarily complex and technical (particularly in relation to statistical data analysis and modelling), the manual assumes no prior knowledge of these areas and tries to explain them in an easy-to-understand manner.

An attempt has been made to keep the language relatively simple, so that the manual is accessible to those whose first language is not English. It also aims to be relevant to the animal health situation in developing countries, as well as more developed countries. In order to make certain concepts more accessible, some explanations use non-technical terminology. Statisticians and other experts may occasionally take exception to this rather loose use of language, but we have deliberately used this style of language and terminology so as to avoid unnecessary complexity.

## HOW TO USE THIS MANUAL

The manual contains a number of different types of material as follows:

- General background on disease surveillance. In order to understand risk-based approaches, it is important that the statistical foundations of traditional approaches to surveillance are well understood.
- Specific background to risk-based surveillance;
- A description of techniques for the analysis of risk-based surveillance;
- Examples of the implementation of the analytical techniques using spreadsheet software;
- A guide to using web-based software for the analysis of risk-based surveillance data, as well as data featuring practical examples.

Concepts are developed in a progressive manner. The later chapters make frequent reference to concepts introduced and explained in the earlier chapters.

The manual may be used in a number of ways: as a self-learning tool, as a reference book, or as a training course resource.

### Self-learning tool

Interested and motivated individuals may use the manual to teach themselves about the design and analysis of risk-based surveillance. In this case, it is recommended that readers start at the beginning and work their way through each chapter. Any of the introductory material that the reader is already familiar with can be skipped.

### Reference book

Those who are already familiar with some aspects of risk-based surveillance and associated analytical methodologies may wish to use the manual as a reference book, to dip into for specific information, as required. The extensive examples may be helpful in this regard.

### Training course resources

The manual may also be used as a resource for structured training courses. The course presenter may use the manual when preparing the course syllabus, and course participants may be given copies to enable them to read more detailed descriptions of material touched on during the course.

## ACKNOWLEDGEMENTS

Finally, FAO deserves special thanks for the support it has provided in the development of this manual. It is hoped that, with FAO's assistance, the manual and the associated web-based software will provide veterinarians working in disease surveillance with the tools they need to implement and effectively analyse risk-based surveillance.

Chapter 1
# Introduction to animal disease surveillance

*"To know that you do not know is the best.*
*To pretend to know when you do not know is a disease."*
Lao-tzu (604 BC–531 BC)

## INTRODUCTION

The aim of this manual is to assist those working in animal health to design and analyse appropriate disease surveillance systems, particularly for early warning, detection of disease, or demonstration of freedom from disease. In order to design an effective surveillance system, two things are required:

- an understanding of available surveillance options, and
- an ability to compare and evaluate the different options, so that you can decide on the best combination.

This chapter discusses the characteristics of animal disease surveillance systems that allow us to compare and evaluate their use for a variety of purposes; it also introduces a range of different possible approaches to surveillance. Some of the material in this chapter and in Chapter 3 is based on information contained in another book by Angus Cameron.[1]

## TERMINOLOGY

A range of different terms with specific meanings are used in this manual. The appendix lists the main abbreviations used, but the key terms and meanings are set out in Table 1.

## CHARACTERISTICS OF A SURVEILLANCE SYSTEM

In order to design, evaluate and compare surveillance options, it is important to understand the different characteristics of a surveillance system. This section discusses a number of characteristics that can be used to describe surveillance.

### Origin of surveillance information

#### Active surveillance

Active surveillance describes an activity that is designed and initiated by the prime users of the data. The main purpose of the activity is disease surveillance, and examples include:

- a serological survey to assess the prevalence of antibodies to brucellosis;
- a farmer questionnaire to identify the level of mortality in their animals.

The **term 'active' is employed here** because the users of the surveillance data (e.g. the veterinary authorities) are actively involved in generating the data.

---

[1]  Cameron, A.R. (in press) *Surveillance* in the Animal Health Management Essentials series. Under publication by the OIE Regional Coordination Unit, Bangkok.

TABLE 1

| Term | Meaning |
|---|---|
| Surveillance | Surveillance is the systematic, ongoing collection, collation and analysis of information related to animal health, and the timely dissemination of information to those who need to know, so that action can be taken. (OIE, 2012) |
| Surveillance system | For a particular disease, this refers to the range of different activities that are able to produce data about the status of that disease in the population. |
| Component (SSC or surveillance system component) | A surveillance system may have one or many components. A component is a single activity that generates surveillance data. A component may be thought of as a single source of surveillance data. |
| Disease | **While this manual primarily deals with infectious diseases, certain parts of the text may refer to other types of diseases**. Although technically, 'disease' refers to the clinical manifestations of an infection or other physiological abnormality, the term is often used more widely. In the context of 'freedom from disease', it is often used synonymously with 'infection'. For disease control purposes, it is the presence of the pathogen (infection), rather than clinical signs (disease), that is normally most important. |
| Infection | Technically, this means that a pathogenic agent has entered and is multiplying in an animal. Less formally, this can be generalized to mean that an animal has the characteristics of interest. For instance, when considering antibody tests, it **is worth taking into account that an animal may have antibodies which indicate previous** exposure to an agent. |
| Country | Surveillance applies to a defined geographical region. For simplicity, this manual has used the example of surveillance at the country level, but the techniques apply equally to a range of different levels. 'Country' can therefore be used interchangeably with terms such as zone, region, province, state, enterprise or compartment. |

One of the significant advantages of active surveillance is that the activity is designed by the users of the information. Therefore, it is possible to ensure that both the nature and the quality of the data collected are adequate to meet the users' surveillance requirements.

***Passive surveillance***

Passive may be thought of in two ways. First, passive surveillance describes surveillance systems where information on disease events is brought to the attention of the veterinary authorities without them actively seeking it. Another way of thinking about passive surveillance is that it uses data that have already been collected for some other purpose; in such circumstances, veterinary services do not initiate the data collection.

Examples of passive surveillance include:

- **A farmer disease reporting system**. In the process of seeking advice, diagnosis or treatment for sick animals, farmers 'report' disease. The reason the farmers make the report is not to help the surveillance system, but to seek veterinary assistance for the problem with their animals. The use of the data for surveillance is secondary.

- **Abattoir meat inspection**. The reason for meat inspection is to ensure the safety and quality of meat sold to consumers. If the data were not used for surveillance, meat inspection would still be required.

The main advantage of passive surveillance systems is that they are **inexpensive**. As a result, they often can ensure much greater coverage of the animal population. However, the data may not fully meet the veterinary services' needs and there is little control over data quality. Data quality may be improved if farmers and veterinarians are provided with education or rewards to improve reporting for specific conditions.

## Disease focus

### Targeted surveillance

Targeted surveillance describes surveillance that is focused on a specific disease or pathogen.

For example, a serological survey for brucellosis may use the Rose Bengal test (RBT). Blood from each sampled animal is tested, and the result of the test is classified as RBT positive or RBT negative. An animal that has tuberculosis or foot-and-mouth disease (FMD), but not brucellosis, would be simply classified as RBT negative, as these other diseases are not of interest in the surveillance activity.

The term *targeted surveillance* can be used in two different senses. In this case, it is referring to surveillance targeted at a specific disease. Later in this manual we will use the term in a different sense – surveillance targeted at a high-risk portion of the population. To differentiate between the two, it is preferable to refer to the second situation as *risk-based surveillance* rather than targeted surveillance.

### General surveillance

General surveillance is not focused on a particular disease, but can be used to detect any disease or pathogen. For example, the farmer disease reporting system is a general surveillance system, as any disease may be reported. However, not all diseases will be reported with the same reliability. Farmers are more likely to report diseases that show clear signs and have a significant impact (for example, many animals are infected, or the disease results in death, such as haemorrhagic septicaemia) than they are to report diseases that display few signs or do not result in an immediate economic impact (such as that caused by intestinal parasitic infections).

Some laboratory tests, such as histopathology, allow detection of many different diseases, rather than just a single disease.

An important feature of general surveillance is that it is not only able to detect known diseases of interest, but it may also be able to detect new, emerging, exotic or unknown endemic diseases. In other words, it is not necessary to be looking for a specific disease in order to find it.

The distinction between general and targeted surveillance depends on the disease detection system used. Targeted surveillance is based on the use of tests that are able to provide a yes/no answer for a specific disease. Examples include:

- polymerase chain reaction (PCR)
- enzyme-linked immunosorbent assay (ELISA)
- agar gel immunodiffusion (AGID).

General surveillance is based on tests that are able to identify multiple diseases (in some cases, all diseases). These tests include:

- clinical examination
- disease investigation
- post-mortem investigation
- meat inspection
- histopathology
- various syndromic surveillance activities.

### Purpose of the surveillance, and nature of the disease

Although there may be some special cases, the purpose of most animal health surveillance can be divided into the following four categories:

- surveillance for diseases that are present
  - describing the level or distribution of disease (or a pathogen or risk factors for disease)
  - detecting cases of disease (at the animal or group/herd level), in order to take action at that same level.
- surveillance for diseases that are absent
  - detecting the incursion of new, emerging or exotic diseases (or pathogens or their risk factors)
  - demonstrating freedom from disease or pathogens.

### Surveillance stakeholders

The general purpose of animal disease surveillance is to collect information to support decision-making to improve or maintain animal health or welfare. Surveillance stakeholders are those who are either involved in the generation and collection of surveillance data, or who make decisions that could be assisted by access to surveillance data.

The range of people responsible for making decisions about animal health is very wide. Obviously, it encompasses those involved in setting national disease control policy, but it also encompasses regional and local veterinary staff, including field staff involved in individual animal care; livestock owners; livestock traders, and those supporting the livestock industries. It also includes international organizations such as the World Organisation for Animal Health (OIE) and the Food and Agriculture Organization of the United Nations (FAO), as well as neighbouring countries and trading partners. All of these groups need reliable surveillance information in order to make decisions.

#### Population coverage

Population coverage refers to the proportion of the population that is actually examined as part of the surveillance system. Two approaches can be used.

#### Sampling

When sampling, only some animals in the population are examined. For example, a sentinel herd system involves a relatively small number of herds; a small number of animals from these herds are tested or examined at regular intervals – animals that are not in the sentinel herds are not examined at all, and therefore these herds are used as a sample of the population.

FIGURE 1

Surveillance

Type of disease → Disease absent from the country/zone

Purpose of surveillance →

Disease present in the country/zone

Detect exotic, new or emerging diseases

Demonstrate freedom from disease

Describe the level and distribution of disease

Detect cases of the disease and take appropriate action
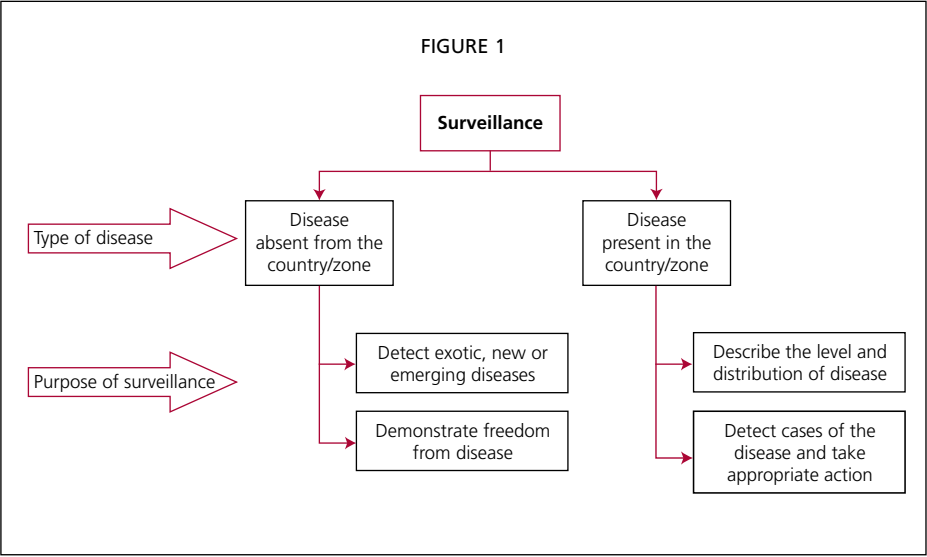
A structured survey may involve randomly selecting a number of villages or farms, and then randomly selecting some animals from these villages or farms to test.

### *Comprehensive coverage (census)*

In a census, all animals in the population are examined. For example, if the population of interest is 'all farmed pigs in the country', a passive disease reporting system covers the entire population, as every single pig in the country is examined (even if only superficially) at more or less regular intervals. If a particular animal becomes diseased, there is a chance that that disease event will be captured by the surveillance system – the probability depends on many factors (for example, severity of the disease, relationships between farmers and veterinarians, whether a report is made). But each pig, if it becomes sick, has a chance of being recorded in the system.

### Representativeness

The representativeness of a surveillance system describes how well the information that is gathered describes the population of interest.

If the level of a characteristic of the animals in our surveillance system (for example, the percentage of animals with protective antibody titres) is approximately the same as the level in the source population, the system is representative of the population with respect to that characteristic. If there is a difference between the animals in the surveillance system and the animals in the source population – for instance, 90 percent of animals with protective antibodies, compared with 60 percent in the source population – the surveillance system is not representative, but is biased.

Bias is the difference between the real value in the population and the value we measure through our surveillance. In many cases, bias due to a non-representative surveillance system can cause significant problems.

**Example**

Consider abattoir surveillance to assess the level of contagious bovine pleuropneumonia (CBPP) in a population: the system uses a sample of the population; the population of interest is 'all farmed cattle', but the surveillance examines only cattle that are sent to the abattoir. Animals infected with CBPP are likely to get sick or die on the farm, so an animal with the disease is much less likely to be sent to the abattoir than a healthy animal. As a result, the proportion of cattle with CBPP in the abattoir is likely to be much lower than the proportion on farms. Therefore, abattoir surveillance is biased.

As the surveillance system is likely to detect a lower proportion of animals than the proportion that is truly infected, this type of system is negatively biased.

The meat inspection system in some developing countries may be less developed than in other countries. This means that it is more common in developing countries for sick animals to be sent to an abattoir than it is in countries where **more stringent** controls are in place. As a result, abattoir surveillance is more useful for detecting clinical disease in some less developed countries than in more developed countries.

Making animal health management policy decisions on the basis of biased information can be very hazardous. If this information were being used to monitor the progress of a control programme, or to prioritize spending on future disease control programmes, the wrong decisions could be made, which, in turn, might have a negative effect on the health of the population. Such a situation could occur where, for example, the level of disease may seem low, and therefore no action is taken, but the true level of disease is high.

Surveillance systems that provide comprehensive coverage of a population are more likely to be representative. For example, in a passive reporting system, theoretically, virtually all animals are seen by their owners on a regular basis. Such a system is comprehensive, as it represents a census of the population. However, if some farmers are more likely to report disease than others, such a system may not be representative.

**Example**

A surveillance system for brucellosis may be based on farmer reporting of abortions or arthritis. If a control programme is in place – specifically one that involves modifying the management systems around calving, in order to limit the spread of the disease – then farms that adopt good management practices are less likely to have the disease. Farms that do not use good management practices may have higher levels of disease. However, farmers with poor management may also be less likely to report disease than farmers with good management.

The disease rates may be higher, but the reporting rates may be lower from farms with poor management than from farms with good management. The outcome is that, even with a system where every infected animal has a chance of being reported, differences in disease and reporting probabilities can result in a bias – in this case, making the total level of disease appear lower than it actually is.

Surveillance systems that aim to provide an accurate assessment of the level of disease typically produce results in terms of proportions, for example, the percentage of animals with CBPP, or the percentage of animals with protective antibody titres against FMD.

If you are making decisions (for example, evaluating the progress of a disease control programme) based on data expressed in the form of a proportion or percentage, it is important that the surveillance system is set up to avoid bias.

## Type of data collected

### Diagnoses

Diagnoses refer specifically to disease, usually clinical disease. At the level of an individual animal, a diagnosis tells us what disease an animal has. In surveillance, a diagnosis is used to classify some animals as having a particular disease and other animals as not having that disease.

In order to make a diagnosis, the animal should be examined by a veterinarian. If necessary, specimens should be submitted for laboratory testing.

### Classifications

Often, we are not solely interested in clinical disease, but in some characteristic of the animal that is related to disease, such as the characteristics described in the following examples:

- A serological survey to demonstrate freedom from FMD will classify animals as seropositive or seronegative. In this case, seropositive animals (due to vaccination or previous exposure at some time in the past) are unlikely to have the disease – we are simply using the serological status as an indicator of whether the animal has been exposed to the virus (or possibly to a vaccine) at some time in the past.
- Surveillance to evaluate the progress of a vaccination programme for FMD can be carried out by estimating the proportion of animals that have protective antibodies. This is based on the antibody status of the animals rather than a diagnosis of disease.

Any measurable characteristic may be used to classify animals for the purposes of surveillance.

### Analysis of specimens

Both the diagnosis of disease and the classification of animals according to some characteristic (for example, antibody status) are usually achieved using some type of test. Certain tests are laboratory based, such as:

- enzyme-linked immunosorbent assay (ELISA) to measure antibody levels;
- virus isolation;
- PCR to detect a pathogenic agent.

Other tests can be performed in the field; such tests may include clinical diagnosis by a veterinarian, or meat inspection in an abattoir.

In cases where laboratory testing is being used, what is collected for surveillance is normally not just information alone; rather it may include a specimen from the animal (blood, milk, a tissue sample, etc.). This specimen is subjected to one or more tests, in order to produce test results – the data required.

*Signs and syndromes*

In circumstances where disease is shown to be present, the most commonly collected information is the diagnosis. As a definitive diagnosis is not always possible, some surveillance systems are designed to collect uninterpreted data, rather than the diagnosis that would result from its interpretation.

To make a diagnosis, a veterinarian will observe the **signs** exhibited by a sick animal (for example, lameness, coughing, increased heart rate) and will interpret these signs in order to determine what disease is causing the problem.

Many signs are easily observed by people without veterinary training. Although non-veterinarians may be unable to make a definitive diagnosis, people who work with livestock are often very good at identifying clinical signs in their animals. Village animal health workers are not usually veterinarians, but they have been trained to recognize disease signs. However, there may be legal restrictions on who can **officially confirm** a diagnosis (for example, qualified veterinarians only).

Therefore, a surveillance system may collect data on the signs of disease observed. Changes in the patterns of signs observed in a population may indicate changes in the diseases that cause those signs. For instance, even if the diagnosis is not known, a sudden increase in the number of cases of coughing indicates the potential introduction and spread of a respiratory disease. This information can be used to initiate a detailed disease investigation to determine the cause of the coughing.

In order to make interpretation and reporting of this type of surveillance simpler, cases are often classified into **syndromes** according to the key sign or group of signs.

A syndrome is simply a defined collection of signs. In the examples on page 6, the syndrome may be 'respiratory disease' and may include any case of disease that manifests as coughing, difficulty in breathing and so on. Other syndromes include:
- acute febrile illness
- diarrhoea
- skin lesions
- sudden death
- lameness.

Both reporting of signs and reporting of syndromes are referred to as **syndromic surveillance**.

Syndromic surveillance is usually designed to help with the detection of changes in disease patterns or the early detection of new diseases. When a change is detected, it must be followed up by more detailed investigations to diagnose the disease causing the change.

Surveillance may collect data on the signs or the general syndrome associated with a disease. The use of syndromes in data collection and reporting is more common than the use of signs because, with syndromes, there is only one data item per case (for example, respiratory disease). With signs, a single animal may exhibit many different symptoms (for example, coughing, difficulty in breathing, standing with neck extended, increased heart rate) – and this makes reporting, collation and analysis of the data more complicated.

### *Negative reporting*

Negative reporting is a special form of disease reporting. The data item in this type of surveillance is the fact that an animal *does not* have a specified disease.

Negative reporting data may be used in two ways:

- To rule out a disease in a laboratory-based reporting system.
  - For instance, in a country seeking to demonstrate freedom from bovine spongiform encephalopathy (BSE), laboratory results may be collected from BSE tests on neurological cases. The results may all be negative. This does not provide any information on what neurological diseases are present, but it does provide evidence that BSE is not present.
- To rule out a disease in a clinical reporting system.
  - This is common for diseases, such as FMD, which show clearly evident clinical signs, and spread quickly in a naive, susceptible population.

For example, a system may be established in which veterinarians complete a report after every farm or village visit, indicating that FMD was not present at the time of the visit. No special examination is necessary because, if FMD were present, it would normally be very easy to identify just by looking at the animals. The fact that the veterinarian visited the farm and did not see any evidence of disease provides information that the disease was absent. (While there is a small chance that the veterinarian was mistaken, such errors can occur in any type of testing or surveillance.) A surveillance system that collates large numbers of negative reports from across a wide area can provide objective evidence that there are unlikely to be any animals with clinical signs of FMD.

Documentation of a clinical negative reporting system can provide valuable reassurance to trading partners about continued freedom from disease in a particular zone, compartment or country.

### *Indirect indicators*

Some surveillance systems do not collect data on the disease or health status of animals directly; rather, they adopt a more indirect approach.

For instance, information provided by drug companies, distributors and feed supply stores on sales of particular types of veterinary drugs and/or feeds can be used for indirect surveillance. Changes in the patterns of drug sales and commercial feed sales are likely to be good indicators that there is a change in the pattern of disease. However, this does not confirm what the disease is or that any observed changes must be followed up by a detailed investigation to assess if there is really a change in disease incidence and, if so, what the disease is.

Surveillance for indirect indicators of disease is often described as an aspect of syndromic surveillance, and is commonly used to assist with the early detection of disease. The ideal indicators are, therefore, those that change early in the disease process, as the following examples demonstrate.

The most common surveillance system used to detect disease is based on farmers reporting to veterinarians when they have a disease problem. However, before the farmer calls the veterinarian, they may try to treat the problem themselves. If a new, widespread problem affects a livestock population, it may be possible to detect the problem through

the use of drug sales and/or commercial feed sales, rather than having to wait for veterinary reports, which may come sometime later.

In human disease surveillance, thermometer sales and workplace sick leave records can serve as useful early indicators of disease patterns in the population.

Indirect indicator surveillance normally refers to active surveillance. The veterinary authorities establish a relationship with the holders of the data (for example, drug suppliers), and ask that updates on sales be provided for analysis at regular intervals (e.g. daily or weekly).

### *Risk factors*
Most surveillance, including indirect surveillance, entails collecting information about disease or a disease-related state. Another approach to surveillance is to measure the risk factors that may be involved in causing the disease. This type of surveillance seeks to provide alerts before an outbreak of disease, so that preventive measures can be put in place.

Examples of risk factor surveillance are:
- Vector surveillance for vector-borne diseases. The vector for bluetongue is the *Culicoides* biting midge. Insect trapping sites provide surveillance information on the presence or absence of the disease vector.
- Surveillance for risk factors for the development of algal blooms, which may produce toxins that kill farmed aquatic animals or contaminate aquatic products, thus making them unsafe for human consumption.
  - Surveillance systems can be established to monitor sunlight and water temperature, in order to assess the risk of the development of blooms. This is risk factor surveillance for the development of algal blooms.
  - Surveillance may directly measure the amount of algae present, and whether they are toxic or not. This is risk factor surveillance for aquaculture or food safety.

External risk factors, or factors not having a direct biological effect on the occurrence of disease in animals, may be considered for surveillance activity. For example, in some regions, movement of animals from one area to another during religious festivals has resulted in an increase or resurgence of FMD outbreaks and other transboundary animal diseases. Data on prices and livestock movements may be used to predict times of increased risk and the location of potential new disease outbreaks.

## Quality
The way the quality of surveillance is measured depends on the type of surveillance carried out.

### *Surveillance to demonstrate disease freedom or detect disease*
When surveillance is undertaken to demonstrate freedom from disease, or for early detection of disease, the conclusion is either that disease has been detected and is therefore known to be present, or that disease has not been detected, and is therefore believed not to be present.

With this 'yes/no' result, two types of errors may be made. First, it is possible to falsely conclude that disease is present when, in fact, it is not (a false alarm). False alarms may cause concern and expense, but do not ultimately endanger the disease status of the population (because no disease is present). A good surveillance system would be expected to generate a false alarm from time to time.

The second error is to falsely conclude that disease is not present when, in fact, it is (surveillance failure). Missing a genuine case of disease can be a serious error, as the disease may spread undetected.

A surveillance system can be thought of as a type of diagnostic test on the entire population: the population has or does not have a disease and the surveillance data are used to make a decision. The ability of a surveillance system to correctly identify a diseased population is analogous to the ability of a diagnostic test to identify a diseased animal. It is measured quantitatively by the sensitivity of the surveillance system.

Sensitivity is the key measure of the quality of a surveillance system that aims to detect disease or demonstrate freedom from disease. Sensitivity is discussed further in Chapter 3. The evaluation of the quality of the surveillance system therefore depends on an estimation of the sensitivity of the surveillance system.

While the sensitivity of the surveillance system is the key measure of system performance, the sensitivity of the system is based primarily on the number of animals (or herds) sampled and the sensitivity of the test(s) applied to them. The other important consideration when evaluating a surveillance system designed for demonstrating freedom or detecting disease is some assessment of the representativeness or otherwise of the sample, and whether or not any biases in the sample have been accounted for in calculating sensitivity.

---

**Example**

Let us assume that we are undertaking a survey to demonstrate freedom of the cattle population from bovine brucellosis. We know that only 5 percent of cattle herds in the region are dairy herds and the other 95 percent are beef herds. However, because we can use a bulk milk test for dairy cattle, it is easier to test dairy cattle than it is to test beef herds, should we find ourselves in a situation where we are obliged to do serological testing.

Let us also assume that we decide to test 320 herds, and we have calculated that this will give us the level of system sensitivity we need. For this calculation, we are assuming that our sample is representative of the population. However, if the likelihood of infection differs between beef herds and dairy herds and we test mostly dairy herds (because testing these herds is easier and cheaper), then our sample is unlikely to be representative of the population. In this case, our estimate of the sensitivity of the surveillance will be incorrect and we may either overestimate or underestimate the true value. To overcome this, we need to either ensure that the sample is representative of the population, or make adjustments to our calculations to take account of the difference in risk between herd types.

---

### Surveillance to measure the level or distribution of disease

In order to determine the level of disease, a surveillance system most often measures prevalence (the proportion of infected animals in a population). Various other measures, such as incidence, may be used, but for the purpose of this discussion, prevalence will serve as an example.

Assessing the quality of a measure of prevalence involves assessing the two types of error that can occur: systematic error and random error.

*Systematic error*

Systematic error is the error produced by some systematic problem in the surveillance system. If the same surveillance is conducted repeatedly on the same population, the error will always be present, and the result will be the same.

Systematic error is measured by bias, which is defined as the difference between the true result and the expected result of the surveillance system (the expected result is the average of all results you would get if you repeated the same surveillance many times).

---

**Example**

Abattoir surveillance might be used to assess the prevalence of clinical paratuberculosis (Johne's disease) in cattle. This disease causes chronic diarrhoea and weight loss. Therefore, in some countries, infected animals may be less likely to be sent to an abattoir than healthy animals.

As a result, the prevalence of clinical cases of Johne's disease in an abattoir will always be lower than the prevalence in the general population. Abattoir surveillance for Johne's disease is therefore biased.

---

*Random error*

Random error is due to the fact that the result of our surveillance can vary according to the chance of selecting one animal or the next animal. With small sample sizes, the random error can be large.

---

**Example**

Consider a population of 1 000 animals with a true prevalence of 10 percent. If only three animals are chosen at random, it is quite likely that all three would be healthy animals. This means that our estimate of the prevalence from our sample would be 0 percent. The random error is 10 percent.

Consider the same population, but a sample of 300 animals instead of three animals. It would be much less likely to select all healthy animals for the entire sample of 300. It is more likely that the sample would have 10 percent of 300 (30) infected animals, but due to random sampling, the actual number of infected animals in the sample might be a little higher or a little lower. It is quite likely that we would select one or two infected animals more or less than the expected number. It is much less likely that we would select a number of infected animals that is very different from the expected number (e.g. selecting as few as 4 or as many as 60 infected animals by chance).

---

The precision of an estimate describes how much random error there is. When calculating the results, the size of the random error is described by the confidence intervals around an estimate.

## Cost and practicality

An important characteristic of surveillance systems is their cost. The precision (when measuring disease) or sensitivity (when detecting disease) of a surveillance system increases as the number of animals examined increases, but so also does the cost. A good surveillance system should be cost-effective.

In addition to cost, the resources to undertake surveillance must be available. Practicality should always be considered.

## SURVEILLANCE OPTIONS

Surveillance can be carried out in a range of different ways. This section describes a number of different approaches.

## Passive disease reporting systems

The term passive disease reporting systems describes the surveillance that is achieved when farmers identify that they have some sick animals, and they contact a veterinarian seeking help.

Passive disease reporting systems are the most common and probably the most important form of surveillance in any country. They are a form of passive surveillance, as the reason farmers contact a veterinarian is not because they are seeking surveillance, but because they are seeking help with the treatment of their sick animals.

These systems are also classified as general surveillance, as they can be used to identify a wide range of diseases.

Passive disease reporting systems have a number of key advantages:

- The coverage of the animal population is usually very good because the person responsible for identifying the disease is the farmer. Most animals in the population are closely observed by their owners relatively frequently. This is in contrast to surveys, where only a very small proportion of the animal population is examined.
- The system is relatively inexpensive – farmers need to contact the veterinarian anyway, and so therefore the main additional cost incurred is related to collecting the information for surveillance purposes.

Passive disease reporting systems are often the means by which new diseases – either incursions of exotic diseases or emerging diseases – are first discovered; this is because these systems achieve high coverage of the population, and are capable of detecting any disease (unlike the situation that applies where targeted surveillance is used).

Therefore, passive disease reporting systems play a very important role in any national surveillance system. These systems are far from perfect, however, due to the possibility of:

- farmers not observing their animals;
- farmers not recognizing signs of disease;
- farmers being afraid to report disease, due to the fear of negative consequences;
- farmers being unable to report disease because they live in a remote area;
- failure of the reporting system within the veterinary services to correctly register or diagnose the disease.

Efforts to address these limitations can significantly improve early detection of disease.

While there are many variations in the detailed operation of farmers' disease reporting systems, a typical system may operate as described below:

- An animal gets sick, and this fact is noticed by the farmer. The chances of a farmer noticing a sick animal depends on the signs it manifests with (more spectacular signs, such as sudden death, unusual neurological signs, or large, visible lesions are easier for a farmer to notice) and the number of animals affected (if more than one animal is affected, a sick animal will be easier to notice). Sometimes a farmer may experience problems that are not associated with clinical signs; for example, subclinical disease, nutritional deficiencies or mastitis at a herd level may cause production losses that are noticed by the farmer, thus prompting a call to the veterinarian.

- The farmer contacts somebody about a sick animal or animals. The simplest case is when the farmer contacts the local government veterinary officer directly. Alternatively, the farmer may contact a private veterinarian, who then contacts a government veterinarian. The process may also involve a number of other steps, such as contacting neighbours, the village head or the local animal health worker for assistance. Ultimately, if the official veterinary service knows about the case, the information can be used for surveillance purposes.

- Information about the case is recorded. Normally, this is done by the local government veterinarian, but it can happen at other stages. Information may be recorded in a number of ways, but most often, a standard paper form is used.

- The written disease report is passed through a reporting hierarchy. If a report is compiled by the local village animal health worker, it will be passed to the district veterinary office. The information may then be passed from the district office to the provincial office, and from there perhaps to a regional office, before eventually arriving at the national office. At each stage, the information in the disease report may be analysed, summarized, or transformed into a different format. One common approach is for reports to be collated at the district level, with a summary report indicating the number of cases of different diseases sent to the provincial office each month. The provincial office combines all district reports into a single provincial summary of the number of cases, which is then sent to the national office. The national office then collates all the provincial reports.

Once surveillance data have been collected at the national level, they are available for use. Routine use of farmer reporting data often includes annual reports of the number of cases of different diseases reported each year, as well as information necessary to meet international reporting obligations.

Diagnostic laboratories are often seen as alternative sources of surveillance data. However, the process by which samples arrive at the laboratory is basically the same as for the passive disease reporting system. For example:

- The farmer notices that an animal is sick and seeks veterinary help.
- A diagnostic specimen may be collected and sent to the laboratory.
- Data from the laboratories are summarized and sent to the provincial or national offices for reporting, either linked to field reports or independent of them.

## Abattoir

Abattoir surveillance is commonly used as a form of passive surveillance. The main advantages of this type of surveillance are:

- It is inexpensive – animals are processed and inspected for other purposes, and therefore the abattoir surveillance costs are primarily related to data capture and any laboratory tests performed.
- It can cover a very large number of animals.
- It allows collection of diagnostic specimens, such as blood or tissue samples, for laboratory testing.
- It provides a relatively constant supply of surveillance data.
- It enables data to be collected from a relatively small number of abattoirs that slaughter animals from a large number of farms or villages (thereby decreasing the data collection costs).

Active, targeted surveillance can also be carried out at abattoirs, in order to take advantage of some of these benefits.

Abattoirs vary significantly from country to country and from area to area. Highly industrialized commercial abattoirs are sophisticated factories with large workforces and tightly controlled food hygiene and safety requirements. Village abattoirs may operate outdoors and slaughter only a very small number of animals under poor hygiene conditions.

The types of surveillance information that can be collected from an abattoir include:

- routine meat inspection findings
- targeted specimens for laboratory analysis
- enhanced inspection findings.

### Routine meat inspection findings

In all but the smallest abattoirs, there is some form of meat inspection. Normally, a limited number of parts of the carcase and viscera are examined.

The aims of meat inspection are to ensure that the meat is fit for human consumption, or to detect or exclude a limited number of specified conditions. For instance, specific lymph nodes may be examined to detect granulomas, in order to be sure that the animal is not infected with tuberculosis.

If the findings of routine meat inspection are recorded and captured by the surveillance system, they may provide a useful source of surveillance data about diseases that can be detected.

In many abattoirs, animals are also examined before slaughter, and this information may be used to supplement the meat inspection findings. These examinations, which are rarely detailed, aim to detect obvious injuries or lesions, as well as signs (such as depression or fever) which may indicate that an animal is clinically ill.

### Targeted specimens for laboratory analysis

Abattoirs offer a valuable opportunity to collect specimens that cannot be collected easily from live animals. The simplest method is the collection of blood, but tissue specimens may also be collected. Large numbers of samples can be collected very rapidly at a busy abattoir, thus making this task simpler and cheaper than collecting similar specimens in the field.

The ability to collect specimens depends on the nature of the abattoir and the type of specimen required.

*Blood*

Blood is best collected as soon as possible **after the animal has been slaughtered**, and while it is being bled. In a busy commercial abattoir, this is one of the most dangerous, and therefore one of the most strictly controlled areas of the plant, because it is the only place inside the abattoir where there are live animals, and this may place **workers at serious risk of injury**.

Even if there is plenty of blood available to be collected, it is necessary to consider carefully how it can be collected without endangering or disrupting normal abattoir operations. Collecting blood at smaller, less busy abattoirs may be easier.

*Tissue samples*

Tissues samples can often be collected during or after removal of the viscera from the carcase. The ability to take tissue samples depends on the way in which tissues are used by the abattoir. If whole organs (such as livers) are going to be sold into the market, the abattoir may be reluctant to allow samples to be taken, and may require them to be purchased.

*Enhanced inspection findings*

Routine inspection may detect only a limited number of **animal health problems**. It may be possible to carry out special inspections at the abattoir for a specific disease that can be detected during post-mortem examination. This may be done by:

- external research
- surveillance staff
- existing meat inspectors, who have been trained to carry out more detailed examinations in order to detect disease.

These more detailed examinations may be further improved by the collection of specimens by the meat inspectors for laboratory confirmation.

## Sentinel herds

A sentinel herd usually consists of a relatively small number of animals, which are kept together and are visited on a regular basis and tested. Testing usually involves blood tests to check for antibodies to specific diseases. It may also involve clinical examination or tests for a specific disease agent.

The typical operation of a sentinel surveillance system is as follows:

- A relatively small number of sentinel herds are established in areas considered to be at high risk of disease incursion.
- Where possible, animals are individually identified.
- When animals are first introduced into the sentinel group, they are tested to ensure that they are susceptible to the target disease (i.e. they do not already have antibodies).
- At each subsequent test, the antibody status of the animals is assessed.
- If an animal is antibody positive, this indicates that the animal has been exposed to the disease in the time between the current test and the previous (negative) test.

Sentinel herds or flocks are therefore distinguished from other systems by being a relatively small group of identified animals, placed in a fixed strategic location, and monitored over time.

## Surveys

Surveys are often seen as the best way to carry out surveillance, but they can be costly and logistically challenging. They are a form of active surveillance, and therefore the veterinary services have full control over the design of the survey and the data collected.

The key advantage of surveys is that the sampling strategy can be developed to exactly meet the needs of the veterinary services and decision-makers. Many other forms of surveillance involve a compromise between the data needed to support decision-making and the data that are available.

Surveys may be representative or risk-based (targeted at a subpopulation with a higher risk of having the disease).

Representative surveys are the most common form. With this approach, it is possible to confidently calculate measures of the level of disease, or probabilities of disease freedom, without the fear of error due to bias.

*Survey Toolbox* (Cameron 1999), Parts I and II (Chapters 2 to 9) deals with most aspects of livestock disease surveys; Chapter 3 concentrates on techniques to ensure a representative sample.

Risk-based sampling is used to detect disease or to demonstrate freedom from disease. Animals are chosen from high-risk groups, so that if the disease is present, there is a better chance of detecting it than would be the case if purely representative sampling were used.

## Syndromic and indirect surveillance

Various forms of syndromic surveillance have been used for many years. However, recent interest from the field of human surveillance has led to a great deal of research in the area.

A syndrome is defined as a collection of signs that indicate the presence of a disease. Syndromic surveillance is therefore concerned with the signs and groups of signs that are associated with disease. These may be clinical signs, such as fever, lameness and diarrhoea, or indirect signs, such as a decrease in the feed consumption at the pen level in a piggery, or an increase in antibiotic feed additive sales from a supplier. When the signs do not relate to clinical signs, this type of surveillance is known as indirect surveillance.

Syndromic surveillance involves the identification of specific signs or groups of signs, and analysis of the patterns of these signs in space and time.

The purpose is not to diagnose a specific disease, but to detect abnormal patterns of signs that may be due to one of a large number of diseases. When an abnormal pattern is detected, a disease investigation follows, in order to diagnose the actual cause of the disease.

Patterns of signs and syndromes are often much less clear than direct diagnoses of disease.

**Example**

If diarrhoea is used as an indicator of the presence of classical swine fever (CSF), a syndromic surveillance system might collect farmer reports of diarrhoea in their pigs, or sales of treatments for diarrhoea.

Diarrhoea can have many causes, so there would be a constant stream of reports coming into the surveillance system. A single case of CSF would just be one more report among many others. However, CSF usually occurs as major outbreaks, and can spread from farm to farm. When it enters the population as a new cause of diarrhoea, the normal pattern of reports of diarrhoea may change.

In order to detect these changes, large amounts of data are required to establish the normal patterns of the sign or syndrome being analysed. These patterns describe how much illness there is, seasonal variations, and normal random variations (in the absence of the target disease). An understanding of the normal patterns makes it possible to detect a change in these patterns when the new disease appears.

The source of data for syndromic surveillance systems should normally be fast, simple and inexpensive, and should allow the routine collection of large amounts of data.

**Example**

Commercial poultry farms expect a certain level of mortality each day, and they routinely record the levels of daily mortality in their sheds. Since death is a syndrome that can be used to detect disease, the data on mortality (if collected centrally for analysis) could easily be used to detect unusual patterns of mortality in the population and trigger a rapid investigation.

The above examples illustrate the three types of data that can be collected by a syndromic surveillance system:

- individual clinical signs (diarrhoea, fever, lameness, agitation, etc.) – farmers or veterinarians record the clinical signs they observe, without making a diagnosis on the basis of these signs; patterns and combinations of the signs are analysed to determine what is normal, and to detect what is abnormal.
- syndromes (respiratory, gastrointestinal, neurological, death, etc.) – cases are classified according to the dominant organ system involved; these classifications can be analysed, in order to ascertain unusual patterns.
- indirect indicators of disease (feed consumption, drug usage, etc.) – indicators that are not observed directly in sick animals, but are observed indirectly.

## Negative reporting (zero reporting)

A veterinary negative reporting system is a specialized surveillance system designed to provide evidence of freedom from disease. This system is a type of passive surveillance which aims to document information that is being generated for other purposes.

Veterinary staff routinely visit farms, villages and other places where animals are kept for a range of reasons, such as examining and providing treatment of clinical cases, vaccination as well as other control activities or inspections and certifications. During the course of these visits, there is normally an opportunity to meet with the livestock owners and to observe the other animals.

If the veterinary services are aiming to demonstrate that a country or zone is free from a disease that normally shows clear and obvious clinical signs, each visit by veterinary staff provides the requisite  evidence. Even if specific examination of animals is not undertaken, it is very unlikely that a disease such as FMD exhibiting its normal manifestations in cattle or pigs could be present without the farmer asking the veterinarian about it, or the veterinarian noticing the disease in the animals. The fact that disease is not noticed during a routine visit can therefore be seen as evidence that the disease is not present.

After each visit, the veterinarian completes a brief report, which includes the location, the date, and confirmation that the target disease was not seen or reported during the visit.

The 'test' in this case (talking to the owner, and inspecting the animals from a distance) is not very sensitive and has very low sensitivity in early cases of disease. However, it is very inexpensive.

Information from the veterinary negative reporting system can be used in response to Carl Sagan's often quoted phrase: "Absence of evidence is not evidence of absence." In other words, to provide evidence that the disease is absent, a simple absence of reports is not adequate. The veterinary negative reporting system generates documented evidence that the disease is not present. Over time, the number and coverage of these reports can provide significant evidence that the country or zone is free from the disease in question.

## Participatory disease surveillance

Participatory disease surveillance (or participatory disease searching, (PDS)) is a relatively new term used to describe an approach to surveillance involving the engagement of farmers.

The method arose out of earlier work on participatory epidemiology and participatory rural appraisal. The common features of all these approaches are the use of trained teams to conduct semi-structured or unstructured interviews with farmers, and the use of a variety of tools to get an overall assessment of the problems and needs of the farmers. Typical tools include:

- participatory disease or risk mapping
- brainstorming
- participatory piling
- development of calendars
- prioritization or ranking exercises
- open discussions.

The prime objective of participatory approaches remains surveillance, and a key output is quantitative data on the occurrence of disease. The participatory approaches from which PDS evolved are specifically designed to give investigators a general understanding of issues and problems from the point of view of the farmers; they are also designed to help address these problems without any preconceptions of what the most important issues might be.

PDS may be used in two ways:

- targeted surveillance, investigating the occurrence of a single disease (for example, highly pathogenic avian influenza (HPAI) in Indonesia, or rinderpest in Pakistan or Somalia). This application is at odds with the participatory philosophy, as the prime concern of investigators is finding out about the disease of interest – although they may be happy to learn about disease in general (or indeed other problems) from the farmers' point of view, they are not in a position to do anything about these more general problems.
- general surveillance, in which information about all diseases of importance to farmers can be collected and prioritized. The investigators are limited by their preconception that animal disease is a key problem.

Because PDS is a surveillance activity, rather than a component of a rural development activity, and because its main purpose is to collect data, it is better to separate it from the associated methods from which it evolved, and to assess its value in terms of surveillance.

PDS is active surveillance (general or targeted). Trained teams visit villages and talk to farmers, and the reason they do this is to generate surveillance data. Farmers are the source of the information, and the way data are collected is through discussion with farmers.

PDS may be thought of as an alternative approach to the passive disease reporting system, which overcomes some (but not all) of the problems of low farmer reporting rates.

The participatory tools used in PDS are not something special for this activity; rather, they are simply a documented approach to collecting good information from farmers. Aspects of this approach can and should be used (to the extent appropriate) whenever veterinary staff are discussing disease-related issues with farmers.

Chapter 2
# Probability theory

*"If thus all events through all eternity could be repeated, by which
we would go from probability to certainty, one would find that
everything in the world happens from definite causes and according
to definite rules, and that we would be forced to assume amongst
the most apparently fortuitous things a certain necessity."*

Jakob Bernoulli (1654–1705)

Chance governs many of the events related to surveillance. When an infection enters a herd, not all animals become infected, but chance determines those animals that are infected and those that are not. When animals are selected in a survey, chance determines those that are selected and those that are not.

Understanding and analysing surveillance requires an understanding of the chance processes that govern a whole range of events. Probability theory provides us with a number of rules that help us understand and predict the outcome of chance events.

The early study of probability was concerned with games of chance, legal decisions and life insurance. The Bernoulli brothers, Jakob and Johann, were brilliant mathematicians who made significant contributions in these areas.

One of the brothers, Johann, forced his son Daniel Bernoulli (1700–82) to study medicine, on the grounds that there was no money in mathematics. Sharing his father's and his uncle's mathematical talent, Daniel applied concepts of probability to the problem of smallpox vaccination. At that time, vaccination involved inoculation of a cut on the skin with live smallpox virus. A vaccinated person had a chance (about 1 in 200) of contracting the disease and dying because of the vaccination. On the other hand, not being vaccinated meant that a person had a chance of 1 in 7 of dying of smallpox in the longer term. Using principles developed in the analysis of lotteries, and applying them to estimates of life expectancy, Daniel Bernoulli concluded that smallpox vaccination was, despite the risks, the best course of action.

## RANDOM VARIABLES

In mathematics, a variable is something that varies, or that can assume a number of different values. For example, the presence of the sun is variable – sometimes it is present (during the day) and sometimes it is not (during the night). In this case, the variable follows a clearly predictable pattern (night follows day regularly).

> Random variable: an unpredictable event that follows a long-run patten.

A *random* variable describes an unpredictable event. For example, the toss of a die, the flip of a coin, or the gender of a baby all represent random, unpredictable events. We can never tell if a single toss of a coin will result in a head or a tail; neither can we tell if a particular natural conception will result in a male or a female.

However, as indicated in the quotation from Jakob Bernoulli on page 21, while the outcome of individual random events may not be predictable, if they are repeated many times, a pattern becomes apparent. Probability theory provides rules through which we can understand and predict the outcome of repeated random events, or determine the likelihood of different outcomes.

---

**Example**

When a six-sided die is thrown, it is not possible to predict what the result will be, whether it is a ⚀ or a ⚅ or any other outcome. However, there are six possible results (⚀, ⚁, ⚂, ⚃, ⚄, ⚅) and, with a fair die, each result is equally possible. The probability of getting a ⚂, for instance, can therefore be calculated as:

$$\frac{\text{The number of ways in which the required outcome can be achieved}}{\text{The total number of possible outcomes}} = \frac{1}{6}$$

Even if we know the probability of getting a ⚂ is 1/6, we still cannot tell whether or not we will get a ⚂ for a single roll of the die. However, if we roll the die 60 times, we can calculate the expected number of times that we would get a ⚂. This is:

$$\text{The number of rolls} \times \text{the probability of 3 for each roll} = 60 \times \frac{1}{6} = 10$$

This means that if we throw the die 60 times, we would expect to get a ⚂ ten times. But you do not always get what you expect. What this means is that you could also get a ⚂ more than ten times or less than ten times. Getting ⚂ 9 times or 12 times (for example) are both possible, but the most likely result is ten times.

**Conclusion**

The individual result is unpredictable, but probability allows us to predict the pattern of results for a random variable when an event is repeated many times.

---

## NOTATION AND SYMBOLS

Random variables or events are often assigned a short name or letter. For instance,

> The probability of an event X is written as: **P(X).**

following the random event of tossing a coin, where the possible result is either a *head* or a *tail*, getting a head could be called **H**. In probability formulae, the probability of a specified event is expressed as P( ). Thus, the probability of getting a head when tossing a coin would be written as P(H).

## THE RULES OF PROBABILITY

If the probability of an individual event is known, the rules of probability allow us to calculate the probability of various combinations of events. These rules can be summarized as:

- Range – the possible range of probability values
- AND – the probability of one event AND another event

- OR – the probability of one event OR another event
- SUM – the probability of all possible events
- NOT – the probability of an event NOT happening
- Conditional – the probability of an event, given that another event has already happened.

## Range

Probabilities are always between zero and one.

Probability values are proportions. As shown in the example on page 22, when all outcomes are equally likely, a probability is calculated as:

$$\frac{\text{The number of ways in which the required outcome can be achieved}}{\text{The total number of possible outcomes}}$$

The number on the top of this equation (the numerator) is always a part of the number on the bottom (the denominator). This proportion can be thought of as *the proportion of all possible outcomes that are the required outcomes.*

As the numerator is always less than or equal to the denominator, proportions (and therefore probabilities) are always in the range from zero to one. If, in probability calculations, the result is less than zero or larger than one, it is a clear indicator that you have made a mistake.

An event with a probability of one is certain to occur. An event with a probability of zero can never occur – it is an impossibility.

Probabilities are often expressed as percentages, ranging from 0 percent to 100 percent.

## AND

Consider the example of flipping a coin. There are two outcomes: heads (H) and tails (T).

AND rule:
P(A and B) = P(A) × P(B).

The probability of each is ½. Let us calculate the probability of first throwing a heads and then throwing a tails.

There are two ways to calculate this. In the first, for two throws of the coin, all possible outcomes can be listed:

- H, T
- H, H
- T, T
- T, H

There are now four possible outcomes, and only one of these matches our required result (H,T). Therefore, the probability of throwing a heads and then a tails is one quarter: P(H,T) = ¼.

The second way of calculating this is to consider the individual probabilities. Remember that all probabilities are between one and zero. If we are calculating the probability that first one event occurs, and then another event occurs, the probability of both occurring must be smaller than either event happening on its own (it is harder to throw a heads followed by a tails than it is to just throw a single heads, or just throw a single tails). With numbers between zero and one, they get smaller when you multiply them together. The AND rule therefore says that to calculate the probability of one event AND then another event, you **multiply** the probability of the first by the probability of the second.

For our coin flip, this means:

$$P(H,T) = P(H) \times P(T)$$
$$= \frac{1}{2} \times \frac{1}{2}$$
$$= \frac{1}{4}$$

---

**Example**

**Question:** The prevalence of disease in a herd is 15 percent. A pen-side test for the disease has a sensitivity (probability of giving a positive result in a diseased animal) of 85 percent and a specificity (probability of giving a negative result in a non-diseased animal) of 100 percent. If one animal is randomly selected from the herd and tested, what is the probability of getting a positive test result?

**Answer:** In order to get a positive test result, two events must occur. First, an infected animal must be selected, and then that animal must give a positive result in the test.

P(infected) = prevalence = 15%

P(test positive) = sensitivity = 85%

P(infected AND test positive) = 15% × 85% = 12.75%

---

It is important to know that this rule only holds true if the two events are independent.

The AND rule assumes events are **independent**.

Independence means that the probability of one event occurring is not influenced by whether or not the other event occurs. When tossing a coin, the probability of getting H on the second throw is unrelated to whether you get H or T on the first throw.

Many other probabilities are not independent: consider the probability of the weather being windy P(wind), and the probability of rain P(rain). A meteorologist may tell us that, for a particular day:

- P(wind) = 20%, and
- P(rain) = 40%

We may therefore calculate

$$P(\text{wind AND rain}) = 20\% \times 40\% = 8\% \qquad \textbf{WRONG}$$

However, experience tells us that wind and rain often go together and that they are therefore not independent. In reality, the probability of wind and rain together is likely to be higher than 8 percent, but the probability cannot be calculated using just the individual probabilities.

## OR

Ten-sided dice (or decimal dice) are often used to help with random sampling. These have ten faces numbered 0 to 9 and each side therefore has a probability of being selected of 1/10 or 10 percent.

For a single throw of the die, what is the probability of either throwing a 2 OR a 9? There are ten possible outcomes, but now there are two that meet our requirements. The

<div style="float:left">

OR rule:
P(A or B) = P(A) + P(B).

</div>

probability is therefore 2/10 or 20 percent.

When we say OR we mean that there are several different ways to achieve the outcome required, so the probability is greater than the probability of each individual outcome. The OR rule states that to calculate the probability of either one outcome or another outcome, we **add** the probabilities of each of the outcomes. This can be written as:

$$P(A \text{ OR } B) = P(A) + P(B)$$

---

**Example**

**Question:** A village contains the following numbers of animals:

Cattle: 40

Goats: 30

Sheep: 20

Pigs: 10

If one animal is chosen at random, what is the probability that it will be either a sheep or a goat?

**Answer:** There are 100 animals. The probability of choosing a goat is 30/100 (30 percent) and the probability of choosing a sheep is 20/100 (20 percent).

**P(sheep or goat) = P(sheep) + P(goat) = 20% + 30% = 50%**

---

The OR rule also has an important requirement. This rule only holds if the two events are

<div style="float:left">

The OR rule assumes that events are mutually exclusive.

</div>

*mutually exclusive*. In the case of our example, this means that it must not be possible to select an animal and for it to be both a sheep and a goat. In some situations, events are not mutually exclusive. If we again consider our example of the weather, it is possible to have both wind and rain together. We could calculate the probabilities of having either wind or rain:

$$P(\text{wind OR rain}) = P(\text{wind}) + P(\text{rain}) = 20\% + 40\% = 60\% \qquad \textbf{WRONG}$$

This overestimates the probability of wind or rain because it does not take into account the occurrence of wind and rain, and so the answer is not correct. This can be demonstrated using a Venn diagram, as shown in Figure 2 on page 26. The overlapping area represents the chance of having both wind and rain. To correctly calculate the probability of wind or rain, you should use:

$$P(\text{wind OR rain}) = P(\text{wind}) + P(\text{rain}) - P(\text{wind AND rain})$$

This removes the overlap and prevents it from being 'double counted' in the probability calculation.

FIGURE 2
**Venn diagram showing two non-mutually exclusive events**
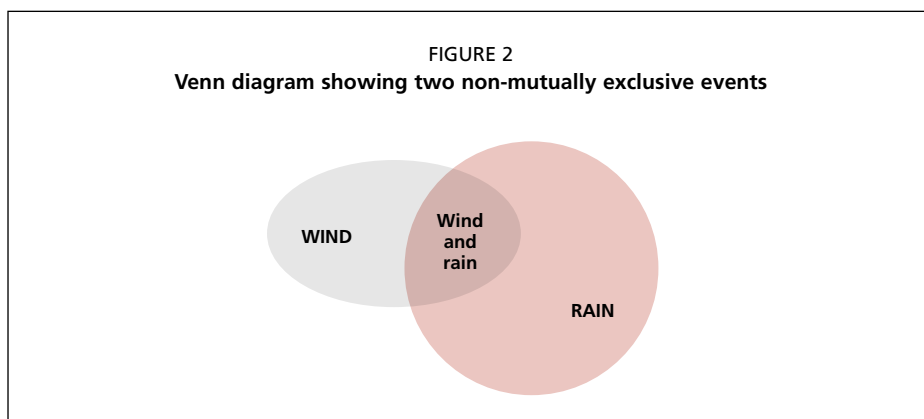


## Sum of all possible outcomes

This simple rule says that the sum of all possible outcomes for a random event must add up to 1. For example, there are six possible outcomes when rolling a six-sided die. Each side has a probability of 1/6. The sum of the probabilities of the six sides is therefore 1.

The SUM rule:
the sum of all possible outcomes equals 1.

This can be written as:

$$\sum_{n=1}^{N} P(A_n) = 1$$

This is read as "the sum of the probability of all *n* outcomes ($A_1$, $A_2$, $A_3$… up to $A_n$) where *n* is from 1 to N (the total number of possible outcomes) is equal to 1".

## NOT

With our ten-sided die, what is the probability of not getting a 3? We can calculate this using our OR rule, as it is equivalent to getting a 0 or a 1 or a 2 or a 4 or a 5 or a 6 or a 7 or an 8 or a 9. This can be expressed as:

The NOT rule:
P(NOT A) = 1 - P(A).

$$P(NOT\ 3) = P(0) + P(1) + P(2) + P(4) + P(5) + P(6) + P(7) + P(8) + P(9)$$
$$= \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10}$$
$$= \frac{9}{10}$$

A simpler way of calculating this is to use the previous SUM rule. If the probabilities of all possible outcomes add up to 1, then the probabilities of all possible outcomes except for a single outcome A, must add up to 1 - P(A). So:

$$P(NOT\ 3) = 1 - P(3)$$
$$= 1 - \tfrac{1}{10}$$
$$= \frac{9}{10}$$

**Example**

**Question:** In a surveillance system with the objective of detecting cases of disease, the probability that disease will be detected in a single animal is 10 percent. What is the probability of detecting disease in at least one animal, if a group of eight animals are tested?

    **Answer:** This problem needs to be addressed in several steps.

    What is the probability that disease will not be detected in a single animal?

    P(detected) = 10%

    Therefore P(not detected) = 1-10% = 90%

    What is the probability that disease will not be detected in eight animals?

    This is an application of the AND rule. Restating the question, what is the probability that disease will not be found in the first animal, AND disease will not be found in the second animal AND… etc. up to the eighth animal.

$$
\begin{aligned}
\text{P(not detected in 8 animals)} \quad = \quad & \text{P(not detected in animal 1)} \times \\
& \text{P(not detected in animal 2)} \times \\
& \text{P(not detected in animal 3)} \times \\
& \text{P(not detected in animal 4)} \times \\
& \text{P(not detected in animal 5)} \times \\
& \text{P(not detected in animal 6)} \times \\
& \text{P(not detected in animal 7)} \times \\
& \text{P(not detected in animal 8)} \times \\
= \quad & \text{P(not detected)}^8 \\
= \quad & 0.9^8 = 0.43
\end{aligned}
$$

    What is the probability that disease will be detected in at least one of the eight animals?

    This is an application of the NOT rule, as it is simply one minus the probability that it would not be detected in any of the eight animals. This makes the final calculation:

$$
\begin{aligned}
\text{P(detected in at least one of 8)} &= 1 - (1 - \text{P(detected in 1)})^8 \\
&= 1 - (1 - 0.1)^8 \\
&= 1 - 0.43 \\
&= 0.57
\end{aligned}
$$

## Conditional probabilities

Consider a bowl containing balls that are selected at random. The bowl contains five yellow balls and two green balls. When a ball is selected, it is not replaced in the bowl. If three balls are selected at random, what is the probability that all three will be yellow?

    For the first ball selected, there are five yellow balls out of the total of seven balls, so the probability of drawing a yellow is 5/7.

    When the second ball is drawn, if the first ball was yellow, then there are only four yellow balls left, out of a total of six, so the probability is 4/6. However, if a green ball

was drawn in the first draw, there would still be five yellow balls and the probability would be 5/6.

In this example, the probability of drawing a yellow ball in the second draw depends on (or is *conditional* on) what ball was selected in the first draw. Conditional probability is expressed using the following notation:

P(A|B)

which is read as "the probability of A given B" or "the probability of A conditional on B". In our example,

P(ball 1 is yellow) = 5/7
P(ball 2 is yellow | ball 1 is yellow) = 4/6
P(ball 3 is yellow | ball 1 and 2 are yellow) = 3/5

The probability that all three balls are yellow is therefore:

$$P(3\text{ balls are yellow}) = \frac{5}{7} \times \frac{4}{6} \times \frac{3}{5}$$
$$= \frac{60}{210}$$
$$= 0.286$$

The sensitivity of a diagnostic test is an example of a conditional probability. Sensitivity is

| Sensitivity is a conditional probability. |

the probability that a test will give a positive result if the animal is truly infected.

This can be expressed as:

P(T+|D+)

where T+ means that the test gives a positive result and D+ means that the animal is diseased.

Conditional probabilities refer to the situation where the probability of one event depends on whether another event has occurred. In the previous section on the AND (multiplication) rule on page page 23, it was noted that the rule is only valid if the two events are independent. Independence is the opposite of conditional probabilities – this is when the probability of one event does not depend on whether another event has occurred. Two events, A and B, are considered to be independent when:

P(A|B) = P(A)

Or in other words, the probability of A given that B has occurred is the same as the probability of A, regardless of whether B has occurred or not.

## GENERAL RULES

The AND (multiplication) and OR (addition) rules are only valid in certain cases. These rules can be extended so that they are valid in all cases.

### General AND rule

The probability of both A and B happening depends on whether A and B are independent. If B is conditional on A, then the AND rule can be rewritten as:

$$P(A \text{ and } B) = P(A) \times P(B|A)$$

For example, the probability that an animal is infected and that it gives a positive test result is equal to the probability that it is infected (prevalence of disease) multiplied by the probability that it will give a positive test result, given that it is infected (sensitivity of the test). This can be written as:

$$P(D+ \text{ and } T+) = P(D+) \times P(T+ \mid D+)$$
$$= \text{prevalence} \times \text{sensitivity}$$

### General OR rule

The OR rule depends on the two events being mutually exclusive. If they are not, the rule can be expressed as:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

This is illustrated in Figure 2 on page 26.

## PROBABILITY DISTRIBUTIONS

The values for a random variable are unpredictable individually, but form some sort of pattern in the long run. We can use such patterns to make predictions about how likely various events are.

---

**Example**

The prevalence of infection in a population is 20 percent. We take a random sample of 40 animals from the population. How many infected animals will there be in our sample?

---

The expected number of infected animals is the probability of getting an infected animal multiplied by the number of animals selected, or 20 percent × 40 = eight infected animals. However, as this is a random process, we will not always get exactly eight infected animals after selecting 40. Instead, we may get a few more or a few less. However, if we repeat this experiment many times, a pattern begins to emerge.

FIGURE 3
**Number of infected animals selected based on different numbers of surveys randomly selecting 40 animals from a population with a prevalence of 20 percent.**
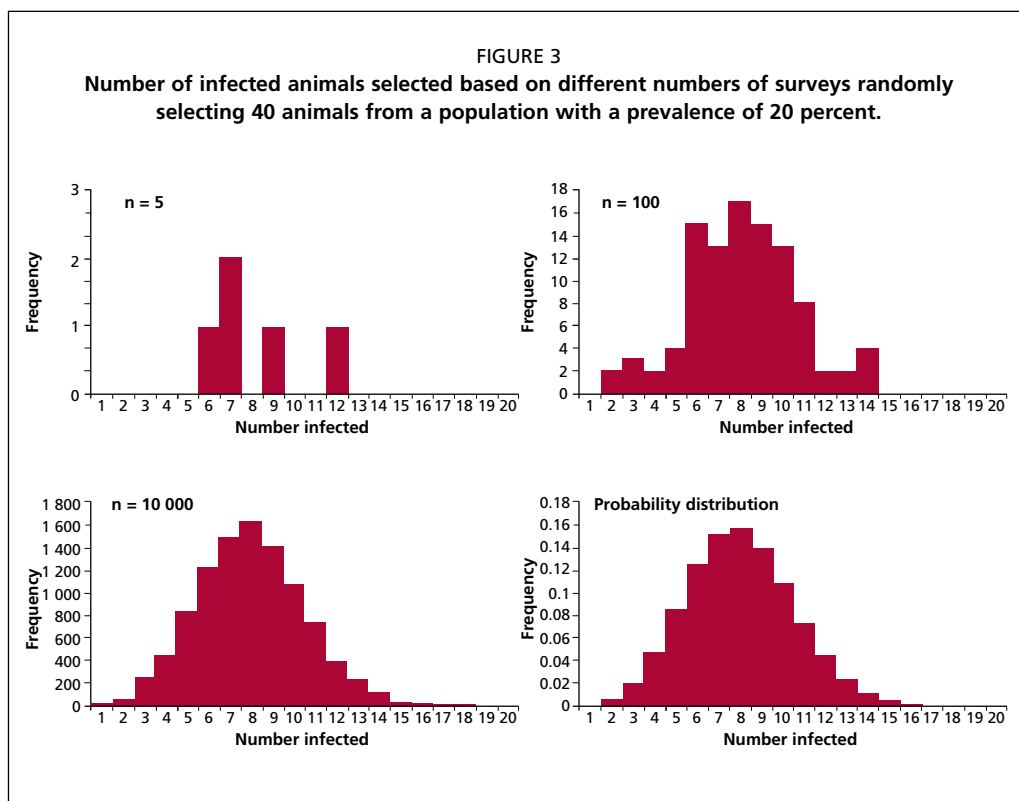
Figure 3 shows what might actually happen in this example. In the first image (n=5) the sampling has been repeated five times, and each time a different number of infected animals was selected (12, 9, 6 and twice 7). The expected number of eight was not selected at all. The actual number of infected animals selected is a random variable, and this shows how it is not possible to predict individual outcomes, or even a small number of outcomes.

In the second image, sampling has been repeated 100 times. This time, eight is in the middle of the values obtained, but the results are still rather irregular. The third shows the result of repeating the sampling 10 000 times, and reveals a distinct pattern.

The final image is based not on multiple repetitions, but on the theoretical probability of selecting each different possible outcome.

## BAYES' THEOREM

Conditional probabilities can be expressed in the form P(A|B). Sometimes it is useful to be able to calculate the inverse probability i.e. P(B|A). This calculation is achieved by another probability rule, known as Bayes' theorem.

FIGURE 4

---
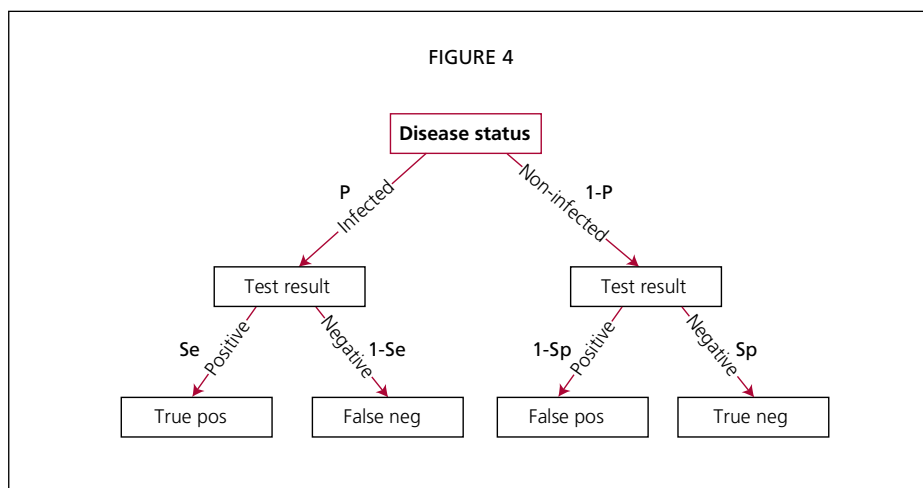
**Example**

An example of this situation arises when using tests for clinical diagnosis. As mentioned above, the sensitivity of a test is a conditional probability: P(T+|D+). This is the probability of getting a positive test result, given that the animal is truly infected.

Knowing that diagnostic tests can sometimes deliver erroneous results, it would be useful to be able to calculate P(D+|T+), or the probability that the animal is truly infected, given that we have tested the animal and have got a positive test result. This value is known as the positive predictive value.

---

There are two ways we can get a positive test result when testing an animal, as shown in Figure 4.

A positive test result could be a true positive (the animal is infected, and gives the right test result) or a false positive (the animal is not infected, and gives the wrong test result).

The probability that an animal is truly infected if we get a positive test result is the proportion of these positive outcomes that are true positives:

$$P(D+|T+) = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

The probabilities are indicated in Figure 4, and therefore this can be calculated as:

$$P(D+|T+) = \frac{P \times Se}{(P \times Se) + (1 - P) \times (1 - Sp)}$$

This simple result has surprisingly far-reaching implications. In this case, P stands for the disease prevalence in the population. However, it can be thought of as the probability that the animal is infected before any testing has been done. This is known as the *prior* probability that an animal is infected. The animal has been tested and we have a positive test result,

which represents new information about the animal. Using the equation on page 31, we can combine our prior knowledge with new information to produce a new estimate of the probability that the animal is infected (known as the *posterior*). The formulae for the positive predictive value on page 31 are an application of Bayes' theorem, which allows us to revise prior information with new information in order to give us an updated posterior probability.

# Chapter 3
# Diagnostic tests

*A. If reproducibility may be a problem, conduct the test only once.*
*B. If a straight line fit is required, obtain only two data points.*
Velilind's Laws of Experimentation

A test is broadly defined as any procedure that aims to divide a population into two groups: one with the characteristic of interest (disease, infection, presence of antibodies, etc.), and one without.

All tests may produce results that make errors in this classification. To qualify as a test, the procedure should succeed in classifying animals more accurately than a purely random procedure (such as tossing a coin).

The two types of **errors** that a test can make are:
- false positive – falsely identifying an animal that does not have the characteristic as having the characteristic;
- false negative – falsely identifying an animal that does have the characteristic as not having it.

In the following discussion, the terms *infected* and *uninfected* (or *diseased* and *not diseased*) are used as general terms to indicate the presence or absence of the characteristic of interest, while *positive* and *negative* are used to indicate test outcomes.

## SENSITIVITY AND SPECIFICITY

The **validity** of a test is the probability that it will get the classification correct. Validity is expressed in terms of **sensitivity** and **specificity**.

Sensitivity is the probability that a diseased animal will be identified as positive by the test (1 – false negative rate); this describes how well the test performs for truly infected animals.

Specificity is the probability that a non-diseased animal will be correctly identified as negative by the test (1 – proportion of false positives); this describes how well the test performs for truly uninfected (i.e. healthy) animals.

These ideas are illustrated in Figure 5 on page 34.

Sensitivity and specificity can be calculated using studies in which the test is applied to animals whose true status is known. The data are usually arranged in a two-by-two table as shown below.

|  |  | True status | | Total |
|---|---|---|---|---|
|  |  | Infected | Not infected |  |
| Test result | Positive | a | b | a + b |
|  | Negative | c | d | c + d |
|  | Total | a + c | b + d | a + b + c + d |

FIGURE 5



**Population**

**True state** → **Infected animals** / **Uninfected animals**

**Test result** → Positive / Negative / Positive / Negative

**Interpretation** →

| Sensitivity (true positive rate) | 1 - Se (false negative rate) | 1 - Sp (false positive rate) | Specificity (true negative rate) |

Se = sensitivity; Sp = specificity

In this table, the sensitivity is the number of correct positive results (true positives), *a*, divided by the total number of infected animals, *a* + *c*. The specificity is the number of correct negative results (true negatives), *d*, divided by the total number of uninfected animals, *b* + *d*.

**Example**

If a new test were applied to 100 animals, made up of 60 healthy animals and 40 infected animals, the results below might be obtained.

|  |  | True status | | Total |
|---|---|---|---|---|
|  |  | Infected | Not infected |  |
| Test result | Positive | 36 | 10 | 46 |
|  | Negative | 4 | 50 | 54 |
|  | Total | 40 | 60 | 100 |

In this example, the sensitivity of the test is 36/40 = 90% and the specificity of the test is 50/60 = 83.3%.

If the true status of animals is not known, this type of calculation cannot be used. New modelling techniques are available to estimate sensitivity and specificity when the true status of animals is not known. These techniques rely on the use of more than one test in a number of different populations. Detailed consideration of these techniques is beyond the scope of this manual.

## COMBINATION OF TESTS

A country would not be considered infected with an exotic disease just because a farmer found a sick animal and reported it. The first test (examination of the animal by the farmer) is quickly followed by a series of other tests, for example:

- clinical examination by a veterinarian
- laboratory tests for antibodies
- confirmatory laboratory tests for the disease agent.

Combinations of multiple tests allow us to avoid certain errors. In this case, we want to be sure that we are not falsely identifying an exotic disease, so we are trying to increase the specificity and decrease the chance of a false positive. The animal would only be considered positive if all of the following occurred:

- The farmer thought there was a problem.
- The veterinarian also thought there was a problem.
- The first (antibody) laboratory test gave a positive response.
- The confirmatory (agent) laboratory test gave a positive response.

If the results of all these tests are positive, we can be very certain that the animal is truly infected.

There is always a trade-off when combining tests. In the example on page 34:

- We increased specificity – with each extra test, the chance of making a false positive decreased.
- We decreased sensitivity – the animal would be considered negative if there was a negative result in any of the four tests, and, because each test has a chance of giving a false negative result, the chances of a false negative result increase with each extra test used.

In this case, in order to achieve high specificity, we need to sacrifice sensitivity. This is because with our interpretation of the results, the animal is only positive if it tests positive in all of the tests.

Using a different interpretation would change the overall test characteristics. If we consider that the animal is only negative if it is negative in all tests, the result would be to increase the sensitivity, but decrease the specificity.

## POPULATION SENSITIVITY

As already discussed, the sensitivity of a test applied to an individual is the probability that the test will produce a positive result, given that the individual is infected (or diseased). This concept can be extended to groupings of animals, or populations, at multiple levels. In this case, the definition of "infected" may be that the population is infected at a prevalence level equal to or greater than a specified threshold (design) prevalence. Similarly, the test may be a test of multiple units in the population using a test of defined performance characteristics (sensitivity and specificity), and a positive result might be defined as one or more of the sampled units producing a positive test result.

At the first level, this sort of population sensitivity is often termed herd or flock (or cluster-level) sensitivity. At higher levels, herds and flocks are aggregated into larger populations and the population sensitivity at this level is often termed the system or surveillance sensitivity.

The relationships between sensitivities at the various levels are discussed in more detail in later chapters. Specificity can also be extended to different population levels, although population specificity is generally less useful than population sensitivity.

## PREDICTIVE VALUES OF A TEST

The sensitivity and specificity of a test are important measures of the performance, or validity, of a test. However, sensitivity and specificity are of limited use when trying to interpret a test result, as they assume that you already know the status of the animal. Instead, what we would really like to know is the likelihood that an animal with a positive test result is truly infected, or conversely, the likelihood that an animal with a negative test result is truly uninfected. These concepts are termed the positive and negative predictive values of the test.

The positive predictive value was introduced in the previous chapter and is a function of the sensitivity and specificity of the test and the prior probability of being infected. The negative predictive value can be derived in similar fashion and is also a function of the sensitivity and specificity of the test and the prior probability of being infected. An important feature of predictive values is that if additional testing is undertaken, the result (predictive value) from one calculation can then be used as the prior for the subsequent calculation. This concept will become clearer in later chapters on combining surveillance system components, incorporating historical data and calculating confidence of population freedom from disease.

Just as the concepts of sensitivity and specificity can be extended to various cluster levels and population levels, so too can predictive values, with the negative predictive value at a population level providing a measure of the level of confidence that the population is "free" of disease (at the design prevalence). This is also discussed in more detail in later chapters.

# Chapter 4
# Concepts of freedom from disease

*"No amount of experimentation can ever prove me right;*
*a single experiment can prove me wrong."*
Albert Einstein (1879–1955)

## CONCEPTS AND PHILOSOPHY

When considering surveillance for diseases that are not known to be present, there may be two objectives:

- to demonstrate (or provide evidence) that the disease is not present, in order to support trade or stop unnecessary disease control activities, or
- to ensure that the disease would be able to be rapidly detected if it ever entered the country or region.

While it is common to talk about 'disease', this generally implies 'infection' and, essentially,

> **Freedom from disease implies freedom from infection.**

the objective is to prove that the pathogen is absent from the country or region. In order to design appropriate surveillance, we must therefore first ask the question: How can you prove that infection is not present?

Let us consider some possible approaches in a stepwise fashion:

- A simple approach may be to visit a farm and to look at some animals, while asking yourself: "Do any of these animals appear to be infected?" If the animals show no sign of the disease, then you may conclude that they are free from the infection.
  - This approach is quick and simple, but it does not prove that the country is free from infection, because:
    - It is not representative (only one farm was examined), and
    - The test used (clinical observation) is not good – there could be subclinically infected animals present.
- To address these problems, a structured serological survey is often used to support claims of freedom from infection. Consider a large survey, in which a random sample of 10 000 animals is selected from all parts of the country. Blood is collected from each animal and is tested in the laboratory for antibodies that would indicate any previous exposure to the pathogen. Clearly, this is much better than the first approach. However, if all the results from this survey are negative, does this prove that the country is free from infection?
  - If the population of susceptible species in the country is, for example, 10 million, then our sample, even though very large, is still indicative of only a small section of the population. While we have tested 10 000 animals, there are still 9 990 000 animals in the population that we have not tested. It is certainly possible that one or more of those animals is infected, but we have missed them in our survey.

This example shows that examining a small number of animals cannot prove that we are free from infection. Examining a larger number gives us a much better chance of finding the infection if it is present, but it still cannot prove that we are free. The more negative animals we observe, the more evidence we have that we may be free, and our *confidence* that we are free increases. However, there is still a chance that the infection is present, so we do not have absolute proof that we are free.

How then can we obtain this proof? We could try testing every single animal in the entire population. If all animals were negative, would this prove that we were free from disease? The problem here is that virtually all laboratory tests can make mistakes. When the sensitivity of the test is less than 100 percent, it means that there is a risk that a truly infected animal may give a negative test result. Even testing every single animal cannot prove that we are free from infection. However, even if there is not absolute proof, we would be very confident that the infection was extremely unlikely to be present.

If we had a perfect test that never made mistakes, by the time we had finished testing the animals, there would still be a chance that those animals that were tested first had become infected.

The simple conclusion is that it is impossible to prove that a population is free from

> It is **impossible** to prove that a population is free from infection.

infection. This problem has been examined extensively by philosophers concerned with science and knowledge. In this example, we have developed a theory – that the population is

free from infection. Each time we observe an animal that is not infected, it lends support to this theory, and as we see more and more uninfected animals, we become more and more confident that our theory is likely to be correct. However, no matter how many uninfected animals we see, we cannot prove that the theory is correct.

On the other hand, it is very simple to prove that the theory is *not* correct. All that is

> A single infected animal can prove that a population is not free.

needed is to find a single infected animal, and we have disproved the theory that the population is free from infection. Karl Popper established this principle of falsifiability as a criti-

cal foundation for science.

Thus, from a philosophical point of view, we cannot prove that a country is free from infection, no matter how many animals we test, but we *can* prove that a country *is* infected by finding a single infected animal. If our surveillance objective is to demonstrate that a country is free from infection, what can be done?

If absolute proof is not possible, then we must work with that which is possible. We cannot prove that we are free, but we can describe the level of confidence that we have, based on repeated observations of many non-infected animals. As in many areas of epidemiology where there is often uncertainty, instead of trying to achieve absolutes we are forced to work with probabilities.

<div style="color:#8b2942; border-left: 4px solid #8b2942; padding-left: 1em;">

Surveillance for disease freedom provides evidence that allows practical decisions to be made.

</div>

The reason for seeking to demonstrate freedom from infection is to support decision-making. For example:

- A trading partner may need to decide if it is safe to import animals from another country.
- Veterinary authorities need to decide if the vaccination programme can be stopped.

In order to support these decisions, absolute certainty of the disease status is not necessary. Having a high level of confidence about the disease status, so that the risk of being wrong is acceptably low, is normally adequate. The key requirement is that enough evidence is available to provide the confidence which allows practical decisions to be made. In order to do this, we have to work with probabilities.

## EXAMPLES OF SAMPLING

The easiest way to understand the factors that influence our confidence in freedom from disease is to look at some practical examples. The examples are based on classroom exercises that illustrate some of the concepts. The aim is to understand our intuitive feelings about how confident we are that a population may be infected or may be free from infection.

### Example 1: Disease free or high prevalence?

Consider a bag containing 1 000 coloured balls. Blue balls represent uninfected animals, and red balls represent infected animals. A number of bags have been prepared, some containing all blue balls (a disease-free population), and others have 200 red balls (a population that is infected with a prevalence of 20 percent). Without looking inside the bags, the task is to decide if the bag is one of the disease-free bags, or one of the infected bags.

→ One ball is drawn from the bag and it is blue.

   Q Are you able to guess if the population is the uninfected population, or the infected population?

If the ball had been red, then we would have proven that the population was infected, but as the ball is blue, we are not sure whether the population is infected or not, and a single ball does not provide much confidence.

→ Four more balls are drawn from the bag and they are all blue.

   Q Do you feel confident that the bag represents a disease-free population?

Your level of confidence has increased, as there is now more evidence available, but you cannot be sure.

   Q If the population is infected and the prevalence of disease (proportion of red balls) is 20 percent, how many red balls would you expect to have seen after drawing five balls from the bag?

The expected number of red balls is equal to the probability that each ball is red (20 percent) multiplied by the number of balls chosen (five) which equals one. So if the population is infected, you would expect to have seen one red ball after having drawn five balls.

   Q You have seen no red balls after drawing five balls. Does this mean that the population must be free?

Of course, the population could still be infected. The *expected* number of red balls is one, but this does not necessarily mean that you would get a red ball after five draws. The process is random, so even from an infected population with a prevalence of 20 percent, you may draw a red ball as the first ball, or you might not find a red ball until quite a few blue balls have been chosen. On *average*, you would expect to have one red ball after having drawn five balls.

→ Five more balls are drawn at random, and they are all blue.

**Q** How confident do you now feel that the population is free from disease?

After having drawn ten balls, you would expect to see two red balls if the population were infected. We have seen no red balls. By now, most people would feel pretty confident that the bag represented a disease-free population.

→ Ten more balls are drawn, and they are all blue.

There are now 20 balls, and we would have expected four red ones if the population were infected. Most people would be ready to conclude that the population is most likely free from disease. Note that it is possible, by chance, to draw 20 blue balls in a row from a population with 20 percent red balls, but it is very unlikely (the probability of this happening is only 1.15 percent).

## Example 2: Disease free or low prevalence?

Let us repeat this exercise, but this time the choice is different. Again, a number of bags have been prepared, but those that are infected have only 20 red balls out of 1 000, or a prevalence of 2 percent.

→ Five balls are drawn at random and they are all blue.

**Q** How confident are you that this population is free from disease?

This time, if the bag is infected, there are only a few red balls in the bag. It will therefore be much harder to find them. After drawing five balls, there is still a good chance that we would not have found one of the infected balls, so our level of confidence is very low.

→ Fifteen more balls are drawn and they are all blue.

**Q** How does your level of confidence this time compare with your level of confidence after drawing 20 balls in the previous example?

In the previous example, the expected number of red balls from an infected population after having selected 20 balls was four. In this example, the expected number is 2 percent × 20, or 0.4 balls. Even if the population is infected, it is still likely that we would not have detected an infected animal (red ball) yet. Our confidence is much lower this time, when compared to the same number of samples in the previous example.

→ Thirty more balls are drawn and they are all blue.

Now, 50 balls have been drawn and the expected number of red balls, if the population is infected, is one. After drawing 50 balls, our level of confidence is about the same as it was when we had drawn only five balls in the previous example – we are still unable to make a reasonable guess as to whether the population is infected or not.

## Example 3: Imperfect sensitivity and specificity

In the previous two examples, the colour of the ball has indicated the true disease status of the animals – if a ball is blue, the animal is not infected and if the ball is red, the animal is infected. The test (using our eyes to detect the colour) can be considered to be perfect – we never call a red ball blue, nor a blue ball red.

Consider an example where the test is not perfect. All the balls are the same colour, and they have to be tested by a machine to determine whether they are infected or not. However, the machine makes mistakes – the sensitivity is 95 percent and the specificity is 90 percent.

Using this machine, you could never be sure if a ball that gave a negative test result was truly negative, or maybe a false negative. A positive test result could indicate a true positive or a false positive.

Let us repeat the exercise in Example 1, where we were trying to distinguish between a population that is free and a population that had 20 percent infection.

→ After drawing five balls, they all test negative.

   **Q** What is our level of confidence after testing five balls, compared with the same stage in Example 1?

Although each ball tests negative, the sensitivity of the test is 95 percent. This means that an infected ball has a 5 percent chance of giving a false negative result. We may have already found an infected ball, but our test gave the wrong result. We are therefore a little less confident this time than we were in Example 1.

→ Five more balls are drawn and one tests positive.

   **Q** Is the population infected or not?

We now have ten balls with one positive test result. If the population is infected at 20 percent, we would expect to have two infected balls by now, although it is still quite likely that we would have only found one. However, the specificity of our test is 90 percent, which means that it would, on average, produce a false positive 10 percent of the time. If the population is not infected, we would expect to see one positive result after drawing ten balls. We are now not sure if the positive is a true positive or a false positive, or if the negatives are correct either. Our confidence level about the status of the population is lower than in Example 1, and it is very difficult to make any useful guess about the status of the population.

## Conclusion

These examples illustrate some important factors that influence our confidence as to whether a population is free from disease or not:

> Factors influencing confidence: sample size, assumed prevalence, sensitivity, specificity.

   1) Our confidence increases as the number of samples increases.

   2) Our confidence depends on the assumptions about the level of disease in the population. When we are trying to decide if the population is infected or not, if we assume that the disease would be common in an infected population, it would be easier to detect, and our confidence grows more quickly when we fail to detect it. On the other hand, if the disease is assumed to be rare in an infected population, more sampling is required in order to achieve the same levels of confidence.

   3) Our confidence depends on the sensitivity and specificity of the test we use.

These relationships can be expressed mathematically as follows:

Confidence ∝ sample size, assumed prevalence, sensitivity, specificity

This can be read as: "Confidence is proportional to sample size, assumed prevalence, sensitivity and specificity." It means that if any of these factors are increased, the confidence increases, and if any are decreased, the confidence decreases.

## PROBABILITIES, CONFIDENCE AND FREEDOM

The term 'confidence' has been used here in its general English meaning, to give an indication of how confident we feel about the surveillance and its ability to detect disease if it is present. When dealing in probabilities, it is important that the exact technical meaning of the various terms used is clearly understood.

When analysing surveillance, the aim is to determine the probability that the surveillance system would find at least one diseased animal based on the assumption that the population is infected at a specified prevalence. This may be written using probability notation:

Confidence in surveillance = $P(T+|D+)$

Where:
- T+ means getting a positive result from our surveillance. Here, surveillance is considered as a type of test of the entire population.
- D+ means that the population is infected (at the specified prevalence).

This is exactly the same concept as the sensitivity of a diagnostic test (the probability of getting a positive test result, given that the animal is infected):

> Sensitivity is a measure of the confidence that a surveillance system will detect disease if it is present.

Sensitivity = $P(T+|D+)$

Our measure of 'confidence' in our surveillance system is therefore a measure of the sensitivity of the surveillance system. The result of our analysis of a surveillance system is normally expressed in terms of sensitivity, but usually requires more detailed explanation. For instance:

"The sensitivity of the surveillance system is *x* percent, which means that the probability of finding at least one infected animal, assuming that the population is infected at a prevalence of *P*, is *x* percent."

Sensitivity is a useful measure of the quality of a surveillance system and its ability to detect disease. However, for decision-makers, a more intuitive measure is the probability that the country is free from disease. In probability notation, this may be expressed as:

Probability of freedom = $P(D-|T-)$

Or in words, the probability (or confidence) that the country is free from disease (D-), given that our surveillance did not detect any infected animals (T-). This is equivalent to the negative predictive value of a "country-level" test for the disease (our surveillance). Calculation of the probability of freedom from disease, based on the sensitivity of the surveillance system is discussed in Chapter 15.

## SPECIFICITY OF SURVEILLANCE

The performance of diagnostic tests on individual animals is described by the sensitivity and the specificity. The specificity is the probability that the test will give a negative result in an uninfected animal (the true negative rate). If we can also talk about the sensitivity of a surveillance system, then there must also be a specificity for a surveillance system – the probability that, if the country is free from disease, the surveillance system will give negative results.

When the purpose of surveillance is to demonstrate freedom from disease, imperfect specificity means that there is a possibility of false positives. A false positive means that we will conclude that the country is infected, when it is truly uninfected. This is a major error, as it may result in the implementation of costly emergency control activities and the loss of trade opportunities. For these reasons, steps are normally taken to ensure that the specificity of any diagnostic system in such surveillance is as good as possible. Normally, there is a series of confirmatory tests, and an animal is only considered positive if it gives a positive result in each of the confirmatory tests. This makes the specificity very high (but decreases the sensitivity).

Even with multiple tests, there is still a theoretical possibility that an animal in a surveillance system could give a false positive result. However, the specificity of the surveillance system is based not on the individual test results, but on the conclusions that are made about them.

If there is a positive test result that has been followed up with confirmatory tests, and it is still positive, the conclusion will be that it is a true positive and that the country is infected. Once this conclusion has been reached (even if it is occasionally incorrect), the question of freedom no longer arises – the country is deemed to be infected. If an animal that initially tested positive later tests negative on confirmatory tests, then it is assumed to have been a false positive, and the conclusion is that it is truly negative.

Based on this logic, the specificity of a surveillance system to detect or demonstrate freedom from infection is normally assumed to be 100 percent.

## DESIGN PREVALENCE

In the examples on page 40, you were asked to distinguish between two options: a bag representing a disease-free population, and a bag representing an infected population. The assumed level of disease (prevalence) in the infected bag influenced our confidence in the decision.

When analysing surveillance to demonstrate freedom from disease, this assumed prevalence value is important. If the value is high, the ability of the surveillance to detect disease (at that level) will be high. If the value is low, the ability to detect disease will be low.

The difficulty with this value is that it is not a real prevalence. We are dealing with a population that is free from disease, and therefore the real prevalence is zero. Instead, the value represents a hypothetical prevalence that is used to set the standard for our surveillance. To distinguish this value from a true prevalence, it is given the

> Design prevalence: a hypothetical prevalence that sets the standard for surveillance.

name *design prevalence*, as it is used to establish the design of our surveillance. In equations, prevalence is usually denoted by the letter P, but design prevalence is represented with the symbol P*.

If there is no disease, then it is not possible that the surveillance system would be able to detect disease. When we analyse surveillance, we are calculating the probability that the surveillance undertaken would be able to detect disease, *if disease were present at a specified level*. The design prevalence specifies the hypothetical level of disease that is used to measure the quality of our surveillance.

In order to account for disease clustering, it is often necessary to specify two levels of design prevalence: the proportion of infected herds in the population ($P^*_H$), and the proportion of infected animals in those infected herds ($P^*_A$). Clustering is discussed in detail in Chapter 13.

## How to decide on an appropriate design prevalence

The design prevalence sets the standard of proof for the surveillance. There is no right or wrong design prevalence. It is simply a value that has to be fixed in order to evaluate the surveillance. The main requirement of the design prevalence is that it is acceptable to those who need to make decisions on the basis of the surveillance.

---

**Example**

Surveillance has been undertaken in country A to demonstrate freedom from disease, in order to support animal exports. In analysing the surveillance, the exporting country uses a design prevalence of 20 percent. This results in an estimate of the sensitivity of the surveillance, which is very high (99.5 percent).

The country that wishes to import animals (country B) examines the analysis of the surveillance. They point out that this simply means that country A has a 99.5 percent chance of finding the disease if 20 percent or more of the population is infected. Failing to detect infection simply means that the population could be infected at anything less than 20 percent. Country B suggests instead that a design prevalence of 1 percent should be used.

Country A objects, because if a design prevalence of 1 percent is used, the estimated sensitivity of the surveillance decreases to 64 percent.

---

For a given surveillance system, increasing the design prevalence will increase the sensitivity and vice versa. The requirement in this example is that both countries agree on a design prevalence value, and assess country A's surveillance against this single fixed value.

Unfortunately, the process of agreeing on an appropriate design prevalence is not simple. There are a number of possible approaches, and these are listed below, in order of preference.

### *Global standards*

The OIE Terrestrial Animal Health Code and Aquatic Animal Health Code contain recommendations and standards for surveillance. For a small number of diseases, these standards include information on the required design prevalence, although these values may be

expressed in a number of different ways. Where such standards exist, these should be used. Examples include:

> ### Bovine tuberculosis
> "Regular and periodic testing of all cattle, water buffalo, and wood bison herds did not detect *M. bovis* infection in at least 99.8 percent of the herds and 99.9 percent of the animals in the country or zone for three consecutive years."
>
> **Terrestrial Animal Health Code, 2008, Article 11.7.2, section 3**

While it is not entirely clear, this requirement implies a herd-level design prevalence of 0.2 percent. Normally, an animal-level design prevalence indicates the prevalence of infected animals within infected herds. However, this text specifies an animal-level design prevalence of 0.1 percent across the entire population. If 0.2 percent of the population is infected, this implies that up to 50 percent of animals in each infected herd may be assumed to be infected (an unrealistically high number for bovine tuberculosis).

> ### Rinderpest
> "Annual sample sizes shall be sufficient to provide 95 percent probability of detecting evidence of rinderpest if present at a prevalence of 1 percent of herds or other sampling units and 5 percent within herds or other sampling units."
>
> **Terrestrial Animal Health Code, 2008, Article 8.13.22, section 5:a:ii**

> ### Bovine spongiform encephalopathy (BSE)
> "The application of Type A surveillance will allow the detection of BSE around a design prevalence of at least one case per 100,000 in the adult cattle population in the country, zone or compartment of concern, at a confidence level of 95 percent."
>
> **Terrestrial Animal Health Code, 2008, Article 11.6.22, section h:1**

### Regional standards
In cases where the OIE code does not specify appropriate information, regional standards may apply. The Council Directives of the European Economic Community (EU regulations) provide an example of regional standards.

### Trading partner requirements
Where no standards have been specified, and the purpose of the surveillance is to support international trade, the requirements of the importing country should be adhered to. This is appropriate when a country establishes clear standards for the importation of animals, but such is not always the case.

### *Acceptable level of protection (ALOP)*

A theoretical approach to determining the appropriate design prevalence is to calculate it based on the importing country's acceptable level of protection. The exposure assessment of import risk analysis normally starts with the prevalence of disease in the exporting country, and ends by calculating a probability of introduction of the disease. This may be compared to a national standard of acceptable risk (the ALOP) – if the risk is higher, imports are not permitted or risk mitigation strategies are required. If the risk is lower, trade is permitted. If a quantitative risk analysis is performed, and the ALOP is specified quantitatively, it is possible to do a risk analysis in reverse. This means that the risk of introduction is determined from the ALOP, and the prevalence in the country of origin that would result in this exact risk is calculated. This prevalence can then be used as the design prevalence, and if the level of disease is lower, the risk for importations will be acceptable.

In most cases, this approach is only theoretical, because:

- Virtually no countries have an explicit, quantitative ALOP. This is a concept that is embedded in the World Trade Organization (WTO) Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement) but is almost never translated into reality.
- Conducting a quantitative import risk analysis is complex and time-consuming.

### *Biology*

The most common way to determine suitable design prevalence values (as the previous options are frequently not possible) is to base the value on an understanding of the biology of the disease.

---

**Example**

In a naive population, FMD normally spreads rapidly and infects a high proportion of the exposed animals. Typically, 60 to 80 percent of a herd would become infected and would seroconvert. If the disease had been introduced to a susceptible herd, it would be extremely unusual for less than, say, 50 percent of the herd to be seropositive within a month or two of the infection.

In this case, if a design prevalence of 50 percent is used, the surveillance would be able to conclude, if no infection is found, that the disease, if present, has infected less than 50 percent of the animals. Normally, this would not be considered adequate proof, but for a highly contagious disease such as FMD, it is biologically implausible that the disease could be established and infect less than 50 percent of the animals. Demonstrating that the disease, if present, is present in less than 50 percent of the herd is logically equivalent to demonstrating that the disease is not present at all.

---

Normally, when this approach is used, the design prevalence is decreased somewhat to account for unusual circumstances where disease spread is slower. Even if it is biologically extremely unlikely that less than 50 percent of the herd would be infected, a design prevalence of 20 percent or even 10 percent is often used.

This approach is appropriate for highly contagious diseases. For less contagious and slowly developing diseases, it is often biologically plausible for an extremely small proportion of the herd to have been infected, without significant further spread. In these cases, the biology of the disease does not help guide the decision of design prevalence.

A particular problem arises in vaccinated populations, even when dealing with highly contagious diseases. For example, surveillance to demonstrate freedom from FMD infection is sometimes conducted in vaccinated populations using non-structural protein (NSP) ELISA tests that can distinguish between antibodies derived from vaccination and those derived from natural infection. Normally, for FMD, a reasonable design prevalence would be 10 percent or 20 percent. However, in a vaccinated population, the disease can no longer be considered highly contagious. It is biologically feasible for a small number of (non-immune) animals in a herd to have been exposed and seroconvert, without the rest of the herd being affected. The choice of the design prevalence therefore depends not just on the disease, but on the characteristics of the population being studied.

> Vaccinated populations require a lower design prevalence.

### *Practical considerations*
For slow-moving diseases, or diseases that are difficult to transmit (such as tuberculosis or BSE), the final choice of design prevalence is often dictated by practical considerations. As these diseases may affect a small proportion of the population, the design prevalence should be as low as possible. However, extremely low design prevalence values mean that very large sample sizes are required in order to achieve an acceptable level of sensitivity. In practice, the design prevalence is made as small as possible, while still being able to conduct affordable surveillance.

The judgement of what is practical depends on the nature of the disease, the resources of the countries involved, and the consequences of infection. Typically, the lowest design prevalence values that are used are 0.1 percent, as this is judged to be the lowest for which surveillance can be affordably run. Tuberculosis is one example of a disease for which a design prevalence of 0.1 percent has been used. The only disease for which a lower design prevalence has been used is BSE (0.001 percent), and this is due to the perceived high consequences of human exposure. Ideally, the design prevalence should not be influenced by such factors. Instead, the target confidence of freedom may be increased to take consequences into account.

### *Arbitrary choice based on commonly used values*
Finally, if none of the above considerations help define an appropriate value for design prevalence, the choice becomes arbitrary. It is more important to have a fixed, agreed design prevalence value than to worry too much about getting the 'right' value. The most commonly used values are 1 percent at the herd level and 1 percent, 5 percent or 10 percent at the animal level.

## Integer design prevalence values
Design prevalence, as with real prevalence values, is a proportion and is often expressed as a percentage. It describes the proportion of animals in a herd, or the proportion of herds in the population, that may be infected.

Consider a herd with 15 animals. If the design prevalence in this herd is 1 percent, it means that 1% × 15 = 0.15 animals are infected. It is not possible to have a fraction of an animal infected – the whole animal is either infected or not infected. The possible prevalence values for this herd are therefore 0 percent, 6.7 percent, 13.3 percent, 20 percent, 26.7 percent, and so on. It is not possible to have an infected herd with a prevalence lower than 6.7 percent.

When the number of animals in a herd is small and the design prevalence is also small, the effective design prevalence is determined by rounding up the target design prevalence to the nearest value that is possible, based on a whole number (integer) of infected animals. Herds of different sizes will have different effective design prevalence values, and may have different numbers of animals that are assumed infected.

One approach to simplify this situation is to express the design prevalence in terms of an integer number of infected animals, instead of as a proportion. For instance, a design prevalence that is sometimes used is one animal per herd or one herd in the population. For herds of different sizes, this represents a varying proportion, but it is still an acceptable and unambiguous definition of the design prevalence. A single infected animal per herd is the most commonly used integer design prevalence, but larger numbers may also be used.

> Design prevalence may be expressed as the integer number of infected animals per herd, or integer number of infected herds in the population.

One interesting side-effect of using a design prevalence of one animal per herd is the ability to assess confidence in absolute freedom, rather than freedom relative to a specified level of disease. The sensitivity measures the probability of detecting disease at the specified design prevalence, and if no disease is detected, we can conclude that, if present, the disease prevalence is lower than the design prevalence. When the design prevalence is one single infected animal, it is not possible to have a lower disease prevalence, so failing to find disease at this prevalence means that disease is not present.

## Design prevalence for early warning systems

Thus far, the discussion has focused on surveillance to demonstrate freedom from disease or infection, mainly for the purpose of supporting international trade. Another purpose of surveillance is to ensure that if disease enters the country, it can be detected as quickly as possible, so that an emergency response can be launched.

Consider a country with 100 000 herds. A herd-level design prevalence of 1 percent means that the surveillance has a good chance of detecting disease if at least 1 000 herds are infected. This represents a very large number of infected herds, and while it may be adequate for trade purposes, it is not adequate for early detection and emergency response. Surveillance systems for this purpose must be able to detect disease at much lower levels, preferably before the disease starts spreading and when only the first or second herd is infected. In this example, it would mean a design prevalence of, say, 2/100 000 or 0.002 percent.

Normally, we set the design prevalence, and then calculate the sensitivity of the surveillance system. For early warning systems, it is also possible to do it the other way around. Instead of asking: "What is the sensitivity of my surveillance if the design prevalence is 1 percent?", we could ask: "How many herds would have to be infected before my surveillance system could detect them with a sensitivity of 95 percent?"

## RELATIVE AND ABSOLUTE FREEDOM

The use of the design prevalence means that the sensitivity of our surveillance system is being measured against an agreed standard. If our surveillance has failed to detect disease, and if the sensitivity of the surveillance is good, we may conclude that the disease is not present at a level equal to or higher than the specified design prevalence. However, this still means that the disease could be present at a level lower than the design prevalence. This raises the question of what we actually mean when we talk about demonstrating freedom from disease and specify a certain design prevalence.

---

### Example

A surveillance programme is in place for bovine tuberculosis, following an eradication programme. Tuberculosis is a disease that spreads slowly, so it is biologically feasible that a very small proportion of a herd could be infected and a small number of herds in the country could be infected. The herd-level design prevalence ($P^*_h$) set by OIE for demonstration of freedom from tuberculosis is 0.2 percent.

The surveillance programme detects a small number of infected herds, representing 0.1 percent of the population.

**Q** Is the population free from tuberculosis?

---

There are two possible answers to this question. The first states that as the prevalence detected is less than the design prevalence, the population can be considered as 'officially free'. This reflects the concept of 'relative freedom'. Freedom is defined as a prevalence of disease less than the specified design prevalence.

The second approach is to recognize that any infected animals in the population means that the population is not free from infection. The best that can be claimed is that the prevalence of disease is very low. This reflects the concept of 'absolute freedom'.

---

### Example

A second country is also completing an eradication programme. Their surveillance (designed using a design prevalence of 0.2 percent) has failed to find any infected herds. Their conclusion is that they are free from disease.

---

Both countries have conducted detailed surveillance and have demonstrated that if the disease is present, the prevalence is lower than 0.2 percent. The difference between the two countries is that the first *knows* that there is still disease present, but in the second, they may be free, or disease may be present but remains undetected. Should the second country be considered to have a better disease status than the first?

This is a question that is difficult to answer. Many experts prefer to use a definition of freedom based on what we know. If we know that infection is present in a country, then the probability that the country is free is zero (it is known to be infected). If surveillance

has failed to find infection in the country (even though it is possible that some infected animals remain undetected), then the probability of freedom is greater than zero and may be calculated.

Chapter 5

# Representative surveys to demonstrate freedom from disease

*"Anyone who attempts to generate random numbers by deterministic means is, of course, living in a state of sin."*

**John von Neumann (1903–57)**

In the past, representative surveys were considered the best way to gather evidence to demonstrate freedom from disease. Representative surveys are based on random sampling and have two major advantages:

- All animals in the population are represented. This avoids bias and gives you confidence that you have not missed part of the population.
- Analysis of surveys based on random sampling is relatively simple.

The disadvantages of representative surveys are that they are often both very expensive and inefficient. More recently, approaches to the analysis of risk-based surveillance (i.e. non-representative surveillance) have been developed. While the purpose of this manual is to give the reader the skills required to design and analyse risk-based surveillance, it is important to first understand how representative surveys for freedom from disease are designed and analysed.

## SURVEY DESIGN

The simplest form of a representative survey to demonstrate freedom from disease involves single-stage, simple random sampling.

---

**Example**

A flock has 2 000 birds. The objective is to demonstrate that the flock is not infected with avian influenza.

A simple random sample of 50 birds is selected, meaning that each bird in the flock has the same probability of being selected as every other bird (50/2 000 or 2.5 percent).

---

In this type of survey, we assume no knowledge about the individual birds. Some birds may be older or younger; some birds stronger or weaker; birds in different parts of the house may have a different risk of exposure to wild birds carrying the disease. In representative surveys using random sampling, none of these factors are taken into account.

The beauty of simple random sampling is that it ensures that the sample selected will be as representative of the population as possible. If 10 percent of the population is made up of older birds, the sample will have approximately 10 percent older birds. If 5 percent of the population is exposed to a higher risk of infection than the rest, the sample will have approximately 5 percent at higher risk.

More complex representative designs are possible, including multistage surveys and those using different approaches to random sampling. These are discussed in more detail on page 58.

## CALCULATION OF SENSITIVITY

The sensitivity of a survey is the probability that, if the population is infected (it is disease positive, D+) at a given design prevalence (P*), at least one infected animal will be detected by the survey (the survey, as a test of the population, would have a positive test result, T+). In probability notation:

$$\text{Survey sensitivity} = P(T+|D+, P*)$$

In representative surveys where animals are chosen by simple random selection, the probability that each animal is infected is equal to P*, the design prevalence. This makes the calculation of sensitivity relatively simple.

### Simple example

Consider a flock of 2 000 birds. Let us use a design prevalence (P*) of 5 percent, meaning that our survey is aiming to detect disease if at least 5 percent of the population is infected. Our aim is to calculate the sensitivity of our survey or the probability that we would successfully detect disease if it were present.

The method of calculating survey sensitivity is based on simple application of probability rules in a step-by-step manner.

**Q** If a single animal is chosen at random, what is the probability that it would be infected?

If the population is infected, our survey design uses the design prevalence to specify the level of infection that would be present. In this case, our design prevalence is equal to 5 percent. If the prevalence of infection is 5 percent, then the probability that any animal chosen at random would be infected would also be 5 percent.

$$P(\text{infected}) = P* = 5\%$$

What is the probability that a single animal chosen at random would *not* be infected? This is an application of the NOT probability rule:

$$P(\text{not infected}) = 1 - P(\text{infected})$$
$$= 1 - P*$$
$$= 95\%$$

**Q** What is the probability that two animals chosen at random would not be infected?

This could be rephrased – what is the probability that the first animal chosen is not infected *and* that the second animal chosen is not infected. Using the AND probability rule:

$$P(2 \text{ animals not infected}) = P(1^{st} \text{ not infected}) \times P(2^{nd} \text{ not infected})$$
$$= (1 - P*) \times (1 - P*)$$
$$= (1 - P*)^2$$
$$= 95\%^2$$
$$= 90.25\%$$

**Q** What is the probability that 50 animals chosen at random would not be infected?

This is a simple extension of the previous example. The probabilities of each animal not being infected are multiplied together. If we use *n* to represent the sample size, then:

$$P(50 \text{ animals not infected}) = (1 - P*)^n$$
$$= 95\%^{50}$$
$$= 7.69\%$$

**Q** What is the probability that at least one animal out of those 50 *is* infected?

This again is an example of the NOT probability rule. If one or more animals is infected, it means that all 50 animals are not uninfected:

$$P(\text{at least one infected}) = 1 - P(\text{all uninfected})$$
$$= 1 - (1 - P*)^n$$
$$= 1 - 95\%^{50}$$
$$= 92.3\%$$

We have calculated the probability that we would get one or more positive test results using a perfect test (sensitivity and specificity both equal to 100 percent) if we sampled 50 animals from the population, with a design prevalence of 5 percent. This is therefore the sensitivity of our survey.

$$\text{Survey sensitivity} = Pr(T+ \mid D+, P*)$$
$$= 1 - (1 - P*)^n$$

## Imperfect sensitivity

The previous example gave the formula for selecting at least one infected animal in our survey. The problem is that once an infected animal has been selected, we need to test the animal to determine if it is infected or not. Unfortunately, our diagnostic tests are virtually never perfect.

Let us assume that the sensitivity (Se) of our test is 90 percent, but the specificity (Sp) is 100 percent. This assumption of perfect specificity was discussed on page 43. We are now interested not in the probability of selecting an infected animal, but the probability of getting a positive test result. If we have perfect specificity, we cannot have a false positive, so any positive result indicates that the population is truly infected.

> **Q** What is the probability that a bird chosen at random from the population will give a positive test result?

In order to give a positive test result, the bird must first be infected (D+), and then it has to test positive (T+) to our diagnostic test, given that it is infected (D+). These two events mean that we need to use the AND probability rule.

$$P(\text{animal tests positive}) = P(D+) \times P(T+ \mid D+)$$
$$= P^* \times Se$$
$$= 5\% \times 90\%$$
$$= 4.5\%$$

> **Q** What is the sensitivity of the survey, taking imperfect sensitivity into account?

The same logic from the previous example can be used, but this time, ($P^* \times Se$) replaces $P^*$. The final formula is:

$$\text{Survey sensitivity} = 1 - \left(1 - \left(P^* \times Se\right)\right)^n$$

This is an important result, as it is used as the basis for the analysis and design of both representative and risk-based surveillance. It is therefore worth writing it in a larger font, so that it is easier to remember.

$$\text{Survey sensitivity} = 1 - \left(1 - \left(P^* \times Se\right)\right)^n$$

## Small populations

The formula for survey sensitivity presented above is adequate for most situations, but it is based on two assumptions to simplify matters. This section and the one following discuss these assumptions. They are slightly more advanced topics and are not essential to the understanding of the fundamentals of risk-based surveillance. The calculations involved in these sections are normally left to computer software to implement.

The first assumption is that the probability of selecting an infected animal is independent of the result of other selections i.e. it is not affected by whether an infected animal was selected previously or not.

---

**Example**

Consider a small herd of 20 animals, with a design prevalence P* of 20 percent. This means that, if the herd is infected, 20 percent × 20, or four animals would be infected.

The probability of selecting an infected animal when the first animal is chosen would be 4/20 or 20 percent. However, if an uninfected animal were selected first, the probability of selecting an infected animal at the second draw would be 4/19 or 21 percent. On the other hand, if the first animal selected were infected, then the probability that the second animal chosen would be infected is 3/19 or 15.8 percent. The probabilities of selecting a positive or negative at each step for the first three animals sampled are shown in Figure 3 on page 30.

FIGURE 6

**Probabilities of selecting infected (+) or uninfected (-) animals from a population of 20 with a design prevalence of 20 percent**
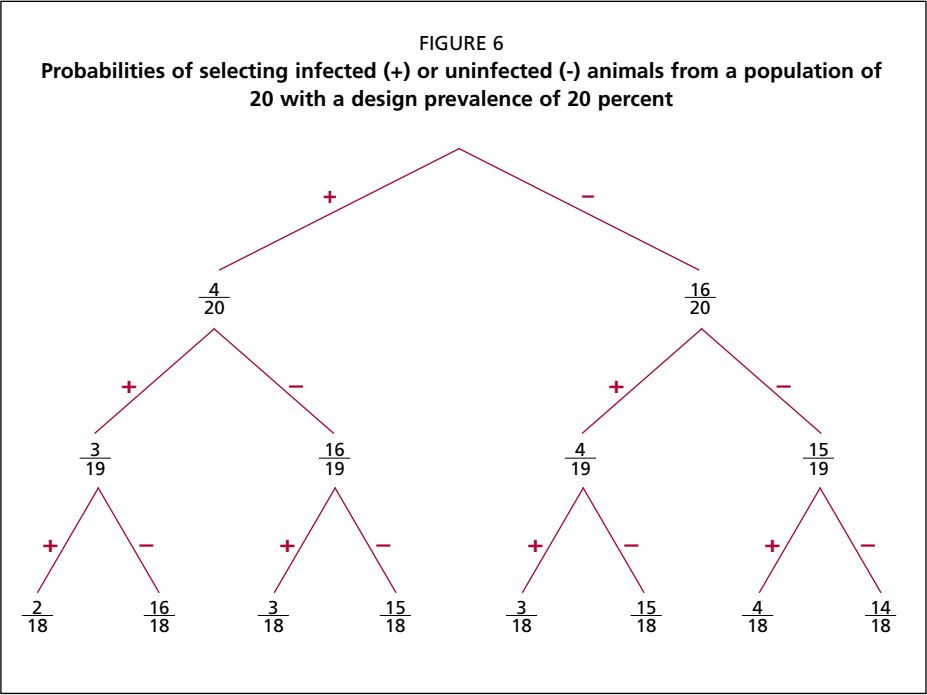
With each animal sampled, the probability of selecting an infected animal changes. The formula for survey sensitivity shown on page 54 is no longer valid, as that formula assumes a constant probability of selecting infected animals (P*). There is a more complex formula for survey sensitivity that takes this effect into account, based on the hypergeometric distribution. This and the other formulae mentioned are implemented in the EpiTools (http://epitools.ausvet.com.au), as discussed in Chapter 17.

When the sample size is small relative to the population size, the effect of changing probabilities is very small. For instance, in the example on page 54 with a population of 20 animals, the change in the probability of selecting an infected animal between the first and second selections (if an infected animal were chosen first) is from 4/20 (20 percent) to 3/19 (15.8 percent) or a decrease of 4.2 percent. If the population size was 2 000, the probabilities would be 400/2 000 (20 percent) to 399/1 999 (19.96 percent) or a decrease of 0.04 percent. In large populations, it is common to assume that the change in probabilities is so small that it can be ignored.

The other approach that is sometimes used is to carry out 'sampling with replacement'. This means that whenever an animal is chosen from the population, it is sampled and then returned to the population (which means that it has a chance of being chosen a second time). In other words, the probabilities do not change as more animals are selected and the formula remains valid.

## Imperfect specificity

If a test has imperfect specificity, it is possible for it to give a positive result when the animal is truly negative (a false positive). This means that when testing an animal, it is possible to get a positive result because the animal is truly infected and the test gives a true positive OR because the animal is uninfected and the test gives a false positive. The chance of selecting an uninfected animal is one minus the chance of selecting an infected animal (1- P*) and the chance of the test giving a false positive is one minus the chance of it giving a true negative (1-specificity). This can be expressed as:

$$\text{Pr(animal tests positive)} = \text{Pr(true positive)} + \text{Pr(false positive)}$$
$$= (P^* \times Se) + \left((1 - P^*) \times (1 - Sp)\right)$$

Using this approach, the formula for survey sensitivity becomes:

$$\text{Survey sensitivity} = 1 - \left[1 - \left((P^* \times Se) + \left((1 - P^*) \times (1 - Sp)\right)\right)\right]^n$$

This is the probability of getting at least one animal with a positive test result. With imperfect specificity, the animals with positive test results could well be false positives, so this does not necessarily mean that there are infected animals in the population. To overcome this problem, the survey design is often modified when using tests with imperfect specificity, so that a certain number of positive (assumed false positive) results are permitted before classifying the population as infected. Determining the acceptable number of positive results is discussed in the next section.

The formula is rapidly becoming more complex, so, where we need to take imperfect specificity into account, it is better to leave the calculations to computer software.

## CALCULATION OF SAMPLE SIZE WITH IMPERFECT SPECIFICITY

Most of the surveillance techniques in this manual are based on the assumption of perfect specificity as discussed on page 43. This section looks at a special case, which is the use of representative surveys with *imperfect* specificity. The calculation of sample size is based on a comparison of two sampling distributions: the distribution of the likely number of positive results if the population is free from infection (i.e. false positives), and the distribution of the number of positive results if the population is infected at the design prevalence. When these distributions overlap significantly, it means that it is quite possible to get the same number of positive results from either a free or infected population. Increasing the sample size spreads out the two distributions and decreases the overlap. The size of the overlap determines the probability of making an error. If we set the level of the acceptable error to 5 percent (equivalent to a confidence level of 95 percent) we can determine the sample size that achieves an overlap just less than 5 percent between the two curves.

**Example**

The probability distributions below are based on a design prevalence of 5 percent, a test sensitivity of 80 percent, a specificity of 90 percent and a sample size of 300. The overlap between the free and the positive distribution is large, so the sample size is not large enough.
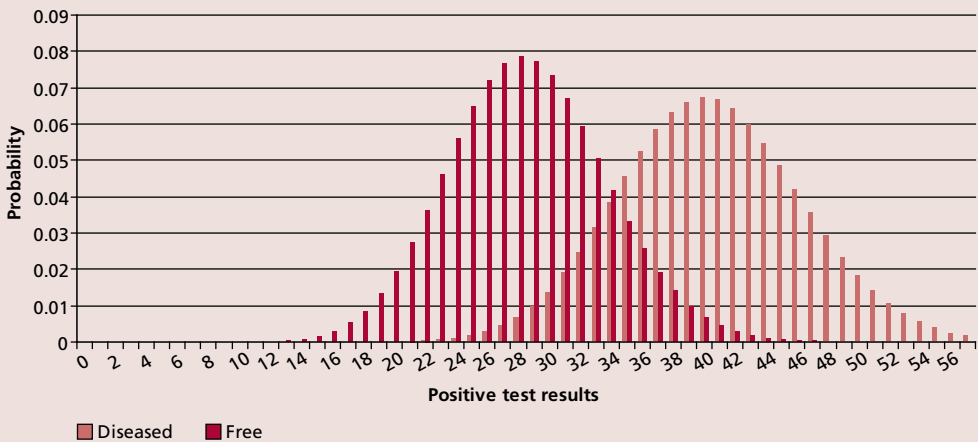


In the second example, the sample size has been increased to 700. Now, if we use a cut-off of 80 animals to define the population as positive or negative, the proportion of the free curve greater than 80 is less than 5 percent, and the proportion of the infected curve less than 80 is less than 5 percent. The sample size in this case needs to be 700 or more.
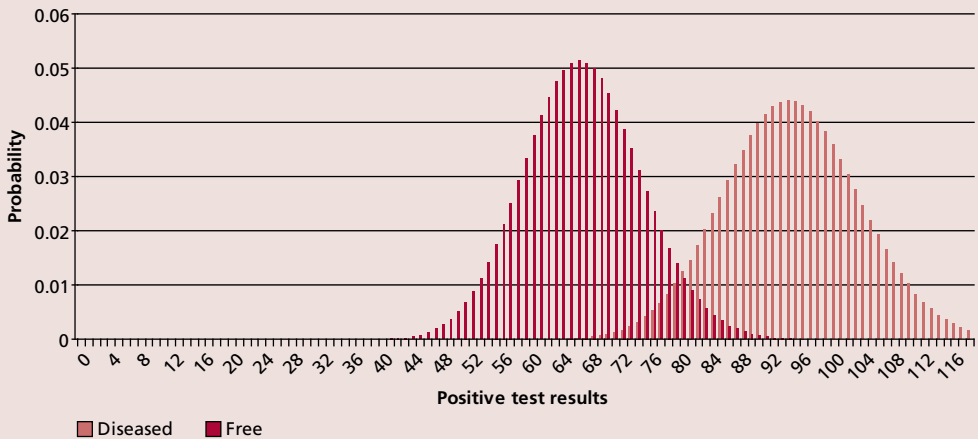


The calculations are complicated, so it is always better to use software tools to calculate sample sizes. A series of web-based tools for such calculations is available on the EpiTools web site: http://epitools.ausvet.com.au.

## TWO-STAGE SURVEY DESIGN

Surveys to demonstrate freedom from infection in small populations are relatively simple to design. However, when the population is large, new complications arise.

## Clustering of infection

Infection is rarely evenly distributed throughout a population. Normally, it forms clusters.

---

**Example**

Consider FMD. When the disease is present in a country, it is not spread evenly across the whole country. Rather, at any one time (even if the disease is endemic), a small proportion of herds is infected, but most herds are not infected. However, in those infected herds, a high proportion of animals may be infected.

The result is small patches of a high prevalence of infection, while most of the other herds have zero prevalence. This 'patchy' distribution of infection is known as clustering.

---

Disease may cluster due to a range of factors, but population groupings are the most common factors, for example, herds, villages with shared grazing, fish with a common water source and so on.

Where clustering occurs (which is almost always the case with large populations), the use of a single value for the design prevalence to describe the level of infection is not sufficient. The overall prevalence of FMD in the population may be only 0.1 percent, but in infected herds, 60 percent of animals may be infected. In these cases, the level of infection is described by using two different values for the design prevalence – one at the animal level, and one at the group (herd) level.

For FMD, the animal-level design prevalence in an infected herd (known as $P^*_A$) may be 20 percent, but the herd-level design prevalence (proportion of infected herds in the entire population, $P^*_H$) may be 1 percent. For surveys of large populations, it is normal to define these two levels of design prevalence.

Where the grouping structure of a population is more complex, there may be even more design prevalence levels. For instance, pigs may be grouped into pens or sheds and then into farms. This may require three levels of design prevalence in order to describe it accurately. Increasing the number of design prevalence levels complicates calculations considerably, so it should be avoided wherever possible. This explanation will be limited to two levels.

## First-stage calculations

Sample size calculations can be thought of as carrying out surveys at different levels. First, we carry out a survey of each herd to find out if it is infected or not. Our herd-level survey has a risk of error, and so it has a defined sensitivity and specificity (these are the probabilities that we set in order to define the cut-off values for the overlap of the free and positive probability distributions). Each herd has a result (positive or negative) with a certain

**TABLE 2**

| Error type | Probability value | Error description | Herd-level equivalent |
|---|---|---|---|
| Type I | Alpha | False positive rate | 1 – Specificity |
| Type II | Beta | False negative rate | 1 – Sensitivity |

sensitivity and specificity. We can treat these herd-level results as if they were just another test, and we can analyse the population of all the herds we have tested to determine if the entire population is free.

The first-stage calculations are therefore the same as those illustrated in Table 2. We use software to determine the sample size required, based on:

- the individual animal test sensitivity and specificity;
- the animal-level design prevalence;
- the error levels that we wish to set.

Because the error levels can be a bit confusing, some terminology is explained in the Table 2.

---

**Example**

If we choose a Type I error level (alpha) of 5 percent and a Type II error level (beta) of 1 percent, it means that our herd-level survey will have a sensitivity of 99 percent and a specificity of 95 percent.

---

Changing the error levels will change the sample size and the herd-level sensitivity and specificity.

## Second-stage calculations

Calculations at the second stage are very similar. The difference here is that we are calculating the number of herds that need to be sampled. The design prevalence is the herd-level design prevalence, and the sensitivity and specificity are not the animal-level diagnostic test values, but are instead those values that we defined by our Type I and Type II error rates at the first stage.

The error levels at the second stage determine the overall survey sensitivity and specificity.

## Optimizing the survey design

The relationship between error levels and sensitivity and specificity means that, in the survey design, we can set different first-stage error levels, but still achieve the same overall survey sensitivity and specificity. This simply means either testing more animals per herd and fewer herds, or fewer animals per herd and more herds.

This flexibility gives us the opportunity to optimize the survey design, based on cost. These calculations assume that there is a per-herd cost and a per-animal cost. By varying the parameters, the least cost combination of animals per herd and total herds can be calculated.

These calculations are also available on the EpiTools web site:

http://epitools.ausvet.com.au/content.php?page=2StageFreedomSS

# Chapter 6
# Risk-based surveillance

*"The obscure we see eventually.*
*The completely obvious, it seems, takes longer."*
**Edward R. Murrow (1908–65)**

*"The more original a discovery, the more obvious it seems afterwards."*
**Arthur Koestler (1905–83)**

Up until this point, the manual has discussed different aspects of surveillance, but has largely limited itself to trying to develop a good understanding of representative surveillance based on random sampling. Representative sampling is appropriate when we want to:
- measure the level of disease in a population and avoid bias;
- detect changes in the level of disease over time;
- describe the distribution of disease.

This type of surveillance asks the question: "How much disease or infection is there and where is it?" The question is answered by giving information on measures of the level of disease, such as prevalence.

However, the main focus of this manual is surveillance in **order** to demonstrate freedom from disease. The question being asked is: "Is disease or infection present?" The answer takes the form of a probability – the probability that our surveillance would have detected disease if it were present (the sensitivity of our surveillance). For this type of surveillance, representative sampling is often not the best approach.

---

**Example**

You arrive in a new city that you have never visited before. You start to feel ill and realize that you need to visit a doctor. What do you do?
  A) Select shops and houses at random, knock on the door and ask if there is a doctor available?
  B) Visit a cinema, a car repair shop, a computer shop and a bus station to see if you can find a doctor?
  C) Go to a hospital and ask to see a doctor?

---

Option A can be thought of as a representative survey. It is possible that, if the sample size were big enough, you would eventually find a doctor by visiting randomly selected houses and shops. However, the proportion of people in the population who are doctors is relatively small (there is a small design prevalence), so it could take a long time to find one.

Option B is a form of targeted surveillance. You decide to concentrate your search for a doctor in a number of different locations. However, the choice of locations is not good. It is possible that, by chance, you could find a doctor in a car repair shop, a computer shop or a bus station, but the probability may be even lower than if you had used random sampling.

Option C represents risk-based surveillance. In order to find a doctor, we go to a place

> **Risk-based surveillance means looking for something where it is mostly likely to be found.**

where we know that doctors are most likely to be found. There is a very small chance that there will not be any doctors on site (if the hospital is closed, or the doctors are on strike), but by choosing a hospital, we give ourselves the very best chance of finding a doctor quickly.

Risk-based surveillance involves looking for something where we think it is mostly likely to be found. The main thing that distinguishes risk-based surveillance from representative surveillance is our knowledge about the disease and the risk factors associated with the disease. In our example, we know that doctors often work in hospitals, so there is a higher probability of finding a doctor in a hospital than in most other locations.

---

**Example**

Consider another example. You have arrived in a city and feel perfectly healthy. A friend from home told you that you should meet somebody he once knew, called Ahmed, but he does not know where Ahmed lives or works, or even if he is still in the city.

---

In this example, we know almost nothing about Ahmed except his name. Looking in a hos-

> **To use risk-based sampling, you must know some risk factors.**

pital, a bus station, a car repair shop or a computer shop would all have some chance of finding him, but so would randomly selecting shops and houses. If we knew some risk factors, we would be able to search more efficiently (perhaps if we knew

that Ahmed likes Chinese food, we could search in Chinese restaurants), but if we do not know any risk factors, random sampling is a reasonable way to search (but not very efficient, as the prevalence of 'Ahmed' may be very low – just one individual in the entire city).

We could try to guess some risk factors. For instance, the friend who told you to meet Ahmed likes swimming and often goes to the pool. You could assume that because they are friends, Ahmed likes to go to the pool as well, so you should search at swimming pools. The value of this approach depends on whether your assumption is right or wrong. Searching at swimming pools if Ahmed does not like swimming (and therefore never goes to the pool) could be a much worse way to try to find him than using random sampling.

In conclusion:

- Risk-based surveillance involves using knowledge of risk factors to improve the probability that we will find disease or infection.
- Risk-based surveillance is more efficient at finding disease or infection than representative (random) sampling.
- If we do not know about the disease or any suitable risk factors, it is not possible to use risk-based surveillance.

- Surveillance that is based on some factor that is *not* a risk factor for the disease may be *less* efficient than representative sampling.

In order to benefit from the greater efficiencies of risk-based sampling, there are a number of requirements:

- We must have a reasonable understanding of risk factors influencing disease occurrence.
- We must have ready access to information about the population and the distribution of these risk factors. If accessing the necessary information is difficult and time-consuming, risk-based surveillance may no longer be more efficient than traditional approaches.

## FACTORS INFLUENCING SENSITIVITY

In Chapter 4 we presented an example of sampling from different bags to try to work out if the bag was 'infected' (had red balls present) or not. In the discussion about this example on page 42, the factors that influence our confidence about surveillance (or the sensitivity of the surveillance) were listed:

- the design prevalence ($P^*$)
- the sensitivity (Se) and specificity (Sp) of the test used
- the number of animals included in the surveillance (n)

Equation (1) on page 54 showed how these factors are related (when specificity is assumed to be 100 percent):

$$\text{Surveillance sensitivity} = 1 - \left(1 - \left(P^* \times Se\right)\right)^n$$

## POPULATION VARIATION

The equation above makes an important assumption. Remember that the middle term ($1 - (P^* \times Se)$)

> Animals vary in their probability of infection and the probability of detection.

represents the probability that an animal will *not* provide a true positive test result. This is raised to the power of *n* animals in the surveillance, which implies that all those animals are assumed to have the same values for $P^*$ and Se i.e. that all animals have the same probability of being infected, and that they all have the same probability of being detected.

This is clearly not true. If disease is present in a population, some animals are at a greater risk of becoming infected than others, depending on the nature of the disease. Some diseases affect young animals more than old animals, and some diseases affect females but do not affect males. There are many possible risk factors that describe differences in the risk of infection for different **groups within** the population. Similarly, infection may be easier to detect in some animals compared to others. For example, while both cattle and sheep can become infected with FMD, cattle often show clear clinical signs, whereas in sheep these signs may be absent or very subtle. Species is therefore a factor that influences the probability of clinical detection of FMD.

When random sampling is used to ensure a representative sample, the average probability of infection, $P^*$, of the animals sampled will be the same as the average for the population (that is the purpose of representative sampling). Similarly, the average sensitivity, Se, will be the same. Even though there may be significant variation in the probability of infection and sensitivity between individual animals, the average is the same as in the **entire** population, so the equation above can be used.

## RISK-BASED SURVEILLANCE

When the selection of animals for surveillance is not representative, Equation (1) can no longer be used in quite the same way. This is because the average P* and Se in the sample is no longer necessarily the same as the average for the entire population. While this may complicate things, it also provides an important opportunity for increasing the efficiency of surveillance.

Risk-based surveillance aims to take into account the differences in risk for animals in the population. By selecting animals with a higher probability of being infected (P*), or a higher probability of being detected if they are infected (Se), the sensitivity of the surveillance can be increased without increasing the total number of animals being tested.

Risk-based surveillance also aims to account for differences in P* and Se in different **groups within** the population. It does this by dividing the population into separate risk groups.

---

### Example

In a population, the P* for disease X is set at 5 percent, and the average sensitivity of the test being used is 90 percent. If we sampled 20 animals using random (representative) sampling, the sensitivity would be:

$$\text{Surveillance sensitivity} = 1 - (1 - (P^* \times Se))^n$$
$$= 1 - (1 - (0.05 \times 0.9))^{20}$$
$$= 60.2\%$$

However, if disease X is present in the population, it is three times more likely to affect young animals than older animals. The average probability of being infected is 5 percent, but as only 20 percent of animals are young, the probability in those young animals is 10.7 percent, while the probability in older animals is 3.6 percent (the way these figures are determined will be explained later). If we use risk-based surveillance, we would concentrate on the section of the population with the higher risk. By sampling only young animals, the sensitivity for a sample of 20 would be:

$$\text{Surveillance sensitivity} = 1 - (1 - (P^* \times Se))^n$$
$$= 1 - (1 - (0.107 \times 0.9))^{20}$$
$$= 86.8\%$$

By focusing our surveillance on the group at the higher risk, we were able to increase the sensitivity by about 26 percent without testing any additional animals.

# Chapter 7
# Analysis of complex surveillance systems

*"Things that are complex are not useful.*
*Things that are useful are simple."*
**Mikhail Kalashnikov (1919–13)**

## TRADITIONAL APPROACHES

Data collected by surveillance systems can be analysed and used for a variety of purposes. This chapter considers the situation where surveillance is used to provide evidence that a zone, country or region is free from disease, in order to support trade in animals or animal products. In the past, two distinct approaches have been available, sometimes used in combination.

### Structured surveys

The most common approach to demonstrating freedom from disease is to design and conduct a structured survey. The survey is designed in such a way as to achieve a specified sensitivity (e.g. 95 percent). This approach may be used at various levels, including a single herd, farm or pond, up to an entire country or group of countries.

The advantages of this system are that it is:

- Quantitative: as the survey is designed by those conducting the surveillance, it can be constructed in such a way that the results can be easily analysed using traditional probability theory, as discussed in Chapter 5. This usually means that random sampling is used to ensure a representative sample.

- Transparent: the method used to collect the sample and analyse the data can be easily documented, so that anybody using the results of the analysis (an importing country, for instance), can see exactly what was done.

- Repeatable: the results of the analysis are likely to be very similar (except for any random error present), **irrespective of who carries out the survey**.

- Objective: once the methodology is documented, the surveillance and analysis is completely objective. As the result is quantitative, there is no element of personal judgement involved. Normally, a target or standard is established, and if the surveillance meets this standard, it is considered to be adequate; otherwise it is not.

The disadvantage of structured surveys is that they are often very expensive and wasteful. Before conducting a structured survey, it is important to be reasonably confident that the area is already disease free; otherwise the expense involved in conducting the survey is wasted. This means that there is likely to be a great deal of surveillance data that have already been collected, and that provide a reasonably high level of confidence that the

> Structured surveys ignore other available surveillance data.

disease is not present. This prior evidence usually comes from complex surveillance activities that are hard to analyse, but nevertheless provides valuable evidence.

When a structured survey is used to provide evidence of freedom from disease, this normally implies that all previous surveillance has been ignored, and that a single survey is being used to provide new evidence.

## Expert panels

An alternative approach has been used in some cases, either in bilateral trade negotiations, or for a small number of diseases for which OIE grants official disease-free status. This involves the use of a small panel of experts who visit the area of interest, examine all the available surveillance data, laboratories, veterinary services and other infrastructure, and based on the overall picture, provide a judgement on whether the evidence is adequate to support a claim of freedom from disease.

This approach has some important advantages:

- It is able to take into account all the available surveillance data, including not only structured surveys, but also complex surveillance systems (such as farmer disease reporting systems) and historical surveillance data.
- It is able to use information on the quality of the veterinary services to assess the reliability of the surveillance.

The use of an expert panel, therefore, is less wasteful, as it takes into account a whole range of complex factors and weighs all available evidence. The problem is that this process goes on in the heads of the expert panel members, and the results are therefore:

- Qualitative. It is very difficult to decide whether the available evidence meets a specified standard or not.
- Subjective. The decision depends very much on the personal opinions, experiences and biases of the experts involved. For instance, a prominent virologist may feel that the diagnostic test they have developed is the best one to use, and they may be unreasonably biased against a country that is using a different test that was developed by a competing laboratory.
- Non-reproducible. Different expert panels may come up with different conclusions about the same situation.

## IDEAL SYSTEM

The ideal system for analysing surveillance to demonstrate freedom from disease would provide a combination of the advantages of these two approaches. Specifically, it should:

- be capable of incorporating all available evidence, including both structured surveys and complex non-structured surveillance activities, such as farmer disease reporting systems, abattoir surveillance and so on, as well as current and historical surveillance data;
- be capable of capturing information about the quality of surveillance and the quality of the veterinary services;
- be objective. The results of the analysis should not depend on who is doing the analysis.
- be repeatable. Repeated analysis, either by the same or different people, should provide the same result (allowing for random error).

- provide a quantitative outcome that allows simple evaluation of whether the evidence meets the required standard, and also allows comparison of the strength of evidence provided by different surveillance activities or by different countries.
- be easily communicable. The principles, methods, assumptions and results should be able to be clearly documented and relatively easily understood by those with an interest in the analysis.

## OVERVIEW – AN ANALOGY

This manual presents a collection of analytical techniques that meet most of the requirements of the ideal system. The approach used may be easiest to grasp with the use of an analogy.

Consider a set of old-fashioned grocer's weighing scales. Imagine that these scales are designed to help weigh evidence of freedom from infection. The numbers at the top indicate the probability that the country is free from infection, starting at 0 on the left (infected) **and progressing** to 1 on the right (definitely free). Evidence (in the form of weights) is placed on the tray on the right of the scales and makes the needle rise, indicating an increasing probability of freedom, depending on the strength of the evidence (size of the weight).

In this analogy, the weights represent the evidence gathered from surveillance, and the size of the weight represents the sensitivity of that surveillance. It is possible to achieve a high probability of freedom by putting a single heavy weight on the tray (i.e. by doing surveillance that has very high sensitivity). However, it is also possible to achieve the same probability of freedom by putting several smaller weights on the tray (i.e. by combining the evidence from several different surveillance activities). This shows how different types of surveillance, even surveillance with relatively low sensitivity, can be combined with others to produce a high probability of freedom from infection.

All of the weights do not have to be placed on the tray simultaneously. It is possible to place relatively small weights (use surveillance with poor sensitivity), but continue to place additional small weights on the tray over a period of time. Eventually, given enough time, it is possible to accumulate **sufficient numbers** of small weights to provide a high probability of freedom. This demonstrates how surveillance evidence can accumulate over time.

However, there is also a tray on the left of the scales. When weights are placed on this tray, the needle is pushed back towards the left, decreasing the probability that the country is free from disease. The left tray represents the risk that new infection may be introduced, because of poor biosecurity. If the biosecurity is perfect, and there is no chance of introducing new disease, then no weights will be placed on the left tray. It is relatively easy to build up enough surveillance evidence on the right tray to give a high probability that disease is not present. But if the level of biosecurity is poor, then there is an ongoing risk that new infection will be introduced. This means that some weights are steadily being placed on the left tray. If there is no new surveillance, then the needle will gradually move to the left, decreasing the probability of freedom.

The probability of freedom from infection is therefore a balance between two factors – on one side, the evidence of freedom from infection based on surveillance (which may be made up of multiple different types of surveillance with different sensitivities), and on the other side, the probability that infection may be introduced.

To complete the analogy, the designer of the scales unfortunately decided to put a spring in the mechanism that tries to pull the needle back towards zero. As the needle moves towards 1, it gets harder and harder to move it further. This means that the first weight that is placed on the right tray gives us a relatively high probability of freedom. But, if an equal weight is added again, it results in a smaller increase in the probability of freedom than the first weight. Evidence of freedom is therefore not additive.

## METHODOLOGICAL REQUIREMENTS

In order to understand and measure the balance between new surveillance evidence from different surveillance activities and the risk of introducing infection (and therefore to describe the probability that the country is free from infection), a number of distinct methodological tools are required.

- A method to quantify the sensitivity of a component of a surveillance system. In our analogy, this is a method for determining the weight of evidence that a surveillance activity contributes. Analytical methods for structured random surveillance already exist, but a new method is needed for complex surveillance.
- A method to combine the evidence provided by different surveillance components. This is the equivalent of placing multiple different weights on the right tray of the scales.
- A method that allows us to calculate the probability of freedom from infection, based on the combined sensitivity of the surveillance. This represents the internal mechanism of the scale that determines where the needle points to.
- A method to account for the balance between the risk of introduction of new disease, and the progressive accumulation of evidence from surveillance over time. This will allow us to determine the true value of historical surveillance data.

### Quantifying the sensitivity of complex surveillance

Chapter 5 discussed the analysis of structured surveys to demonstrate freedom from disease. Some relatively simple formulae were developed to analyse surveillance data and calculate the sensitivity of the surveillance.

Complex surveillance systems cannot be analysed in the same way, because there is a whole range of different biases, which means that some animals are more likely to be infected than others, and some animals are more likely to be selected than others.

One methodology that may be used to estimate the sensitivity of complex surveillance systems is scenario tree modelling, which is described in detail from Chapter 8 to Chapter 13. This method uses a tree structure to describe the population and surveillance structures, and to explicitly capture the probability that any given animal might be infected with the disease or that it might be detected. Scenario trees are the tool we use for quantifying risk-based surveillance, as they provide a quantitative measure of the sensitivity of surveillance components.

### Combination of evidence from multiple surveillance components

Chapter 14 looks at possible approaches to combine evidence from different surveillance components. The basic technique of calculating the combined sensitivity of two or more components is very simple, but becomes much more complex when there is overlap

between the coverage of the surveillance components, as this means that we need to avoid 'double counting' the same evidence.

## Calculation of the probability of freedom from infection

The quality of surveillance is most commonly described in terms of sensitivity. However, the probability of freedom from infection is a more intuitive and often more useful measure of the quality of surveillance. Chapter 15 discusses how this is calculated, based on the combined sensitivity of all the available surveillance.

## Incorporating historical data

As illustrated with the analogy of the scales, evidence can accumulate over time, but its value is decreased if there is an ongoing risk of introduction of new disease. Chapter 16 describes how to incorporate historical surveillance data and quantitatively measure the balance between evidence and risk of introduction of infection.

# Chapter 8
# Introduction to scenario tree modelling

*"Absence of proof is not proof of absence."*
**William Cowper (1731–1800)**

## A SIMPLE EXAMPLE

In Chapter 5 we derived the basic formula that is used to analyse representative surveys to demonstrate freedom from infection. In that formula, the probability that an animal tests positive is indicated by $P^* \times Se$, or, in words, the probability that the animal is infected times the probability that it tests positive, given that it is infected. This formula was based on the assumption that the specificity of the test was perfect.

This can be represented by the simple tree shown in Figure 7 on page 72. The diagram shows that there are two ways to obtain a positive test result (T+): an infected animal with a true positive test result, or an uninfected animal with a false positive test result. To calculate the probability of obtaining a positive test result, we use our two probability rules: the AND rule to multiply probabilities down the branches of the tree, and the OR rule to add the resultant probabilities. Thus, the result is:

P(Infected AND true positive) = $P^* \times Se$
P(Uninfected AND false positive) = $(1- P^*) \times (1-Sp)$
P(either one OR other of the above) = $(P^* \times Se) + [(1- P^*) \times (1-Sp)]$

If specificity is equal to 1, this simplifies to $P^* \times Se$.

This example assumes that all animals have the same probability of being infected, and also have the same sensitivity. If all animals are not the same, we can add some new nodes to the tree to describe the differences between animals. While the tree shown in Figure 8 on page 73 looks more complicated, but it is actually very simple. Two new nodes have been added: AGE and SPECIES. For this example disease, age influences the probability that an animal will be infected (young animals are at higher risk of becoming infected than older animals). Species influences the probability that an infected animal will be detected (sensitivity). This tree may represent a clinical surveillance system – cattle show typical clinical signs, while sheep often show only mild clinical signs.

The first tree in Figure 7 divided the population into four groups, based on infection status and test result. The tree in Figure 8 has divided the population into 16 different groups. Within each group, animals are similar, but each group is different with respect to the factors included in the tree.

FIGURE 7
**Simple scenario tree**
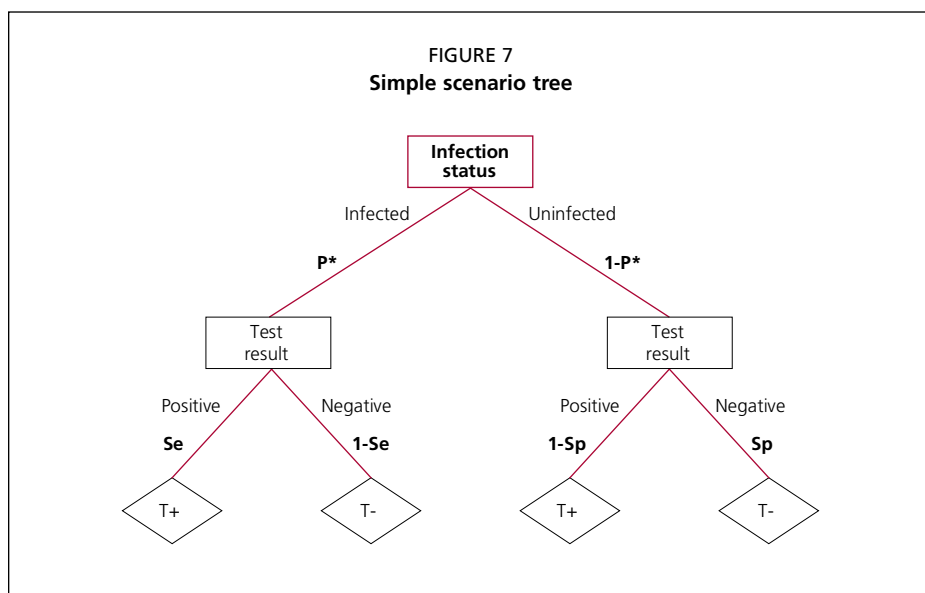
Including different factors in the tree allows us to assign different probabilities. If young animals have a greater risk of being infected, we can use a different probability of infection for young animals (on the left side of the tree) compared to old animals. If infected cattle are easier to detect than infected sheep, the sensitivity (probability of a positive test result in infected animals) will be different for cattle than for sheep.

## PURPOSE OF THE SCENARIO TREE
A scenario tree is a tool to assist in the calculation of the sensitivity of a component of a surveillance system. In contrast to the simple analysis of representative surveys, the purpose of a scenario tree is to take into account the fact that not all animals in the population:
- have the same probability of being infected (some are at greater risk than others);
- nor do they have the same probability of being detected (the sensitivity of detection is greater in some animals than in others).

Remember our formula for surveillance sensitivity:

$$\text{Surveillance sensitivity} = 1 - \left(1 - \left(P* \times Se\right)\right)^n$$

In this formula, all *n* animals in the population are assumed to have the same values for P*

Scenario trees divide
the population
into homogenous
subpopulations. Animals in
a subpopulation have the
same risk of infection and
probability of detection.

(probability of being infected) and Se (probability of being detected). The scenario tree divides the population into smaller subpopulations, based on risk factors and detection probabilities. The reason this is so valuable is because it allows us to analyse non-representative surveillance. If surveillance is targeted towards a group of animals that are at higher risk of being infect-ed, a scenario tree allows us to calculate the sensitivity that we achieve for that particular group.

FIGURE 8

**Scenario tree with one factor affecting probability of infection (age) and one factor affecting the probability of detection (species).**

A scenario tree represents a series of different limbs, or paths from the beginning to the end, and each limb defines a subpopulation. For instance, in Figure 8, the first of the 16 subpopulations is made up of those animals that are young, infected, cattle and test positive, while the last is made up of animals that are mature, uninfected, sheep and test negative. The probability that a randomly selected animal falls into one of these groups can be calculated by multiplying the probabilities at each step down the limb (using our AND probability rule). For instance, the probability that an animal is in the last group is the probability that an animal is mature multiplied by the probability that it is uninfected multiplied by the probability that it is a sheep multiplied by the probability that it tests negative.

## TERMINOLOGY

A number of terms are commonly used when discussing scenario trees, and these terms need to be defined. As follows:

NODE    A node represents a factor that is used to divide a population into a number of groups. In Figure 8, nodes are depicted as square boxes. A good way to think about nodes is that they are asking a question about the animal or group of animals. For example, the AGE node is asking: "What is the age of the animals?" and the SPECIES node is asking: "What species are the animals?"

*BRANCH*  A branch represents the answer to the question, dividing the population into different groups on the basis of the node from which the branches come. Branches are depicted in Figure 8 as lines coming out of the bottom of each node box. The branches from the AGE node answer the question: "Is the animal *YOUNG* or *OLD*?", thereby dividing the population into two groups.

Branches have probabilities associated with them, indicating the probability that an animal or group of animals will belong to that branch.

Outcome    The outcome (sometimes called a 'leaf') is at the end of the last branches of the scenario tree. In Figure 8 on page 73, these are depicted as diamonds. The outcome represents the final conclusion about animals in that group. The two possible outcomes are normally *TEST POSITIVE* or *TEST NEGATIVE*.

Limb       This describes the path through a particular series of nodes and branches, starting at the beginning and continuing to the end, which results in a single outcome.

## BRANCH PROBABILITIES

Branch probabilities are often proportions.

Each branch is associated with the probability that an animal or group of animals will fall into that branch category. In Figure 8, the AGE node has two branches: *YOUNG* and *MATURE*.
The probability that an animal falls into the *YOUNG* branch category is determined by the proportion of young animals in the population. Therefore, branch probabilities are often proportions.

Branch probabilities can also be sensitivities or specificities.

Another example is the TEST RESULT node. Here, the proportion of animals that have a positive test result depends on whether the animals are infected. If they are infected, the probability is the individual animal test sensitivity. If the animal is not infected, it is one minus the test specificity.

In the example of the TEST RESULT, the probability depends on the previous branch – is the animal infected or not. In fact, probabilities in a scenario tree are always conditional, which means that all probabilities depend on all the previous branches in the tree.

Probabilities are conditional on all previous branches in the tree.

For instance, the SPECIES node has two branches – *CATTLE* and *SHEEP*. However, there are four different *CATTLE* branches. The probability for the first *CATTLE* branch is based on the subpopulation of *YOUNG* and *INFECTED* animals. Some young infected animals are cattle, and some are sheep. The proportion of young infected animals that are cattle is the correct branch probability to use in this case.

The second *CATTLE* branch is for *YOUNG UNINFECTED* animals, the third is for *MATURE INFECTED* animals, and the last is for *MATURE UNINFECTED* animals. Each of the probabilities for the four cattle branches may therefore be different, and depends on the previous branches.

Changing the order of nodes in a tree changes the conditional probabilities. In the previous example, it was necessary to estimate the proportion of young animals that are cattle. This is a non-intuitive value and may be difficult to estimate. Changing the node order so that SPECIES comes first, and AGE comes second, would mean that the probability to estimate is the proportion of *CATTLE* that are *YOUNG*.

## NODE TYPES

There are three main types of nodes in a scenario tree, each of which serves a different purpose. When developing a scenario tree model, it is important to ensure that you are clear what type each node is. The three main types of nodes are infection nodes, detection nodes and category nodes.

## Infection node

An infection node represents the question: "Is the animal or group of animals infected?"

Infection nodes always have two branches: *INFECTED* and *NOT INFECTED*. Remember that analysis of surveillance for freedom from infection is based on estimating the surveillance system sensitivity, which is the probability of detecting disease if the disease is present at a defined level. An infection node defines the level of disease that is present. The probability associated with the *INFECTED* branch of an infection node is the design prevalence (P*).

Scenario trees must always have at least one infection node. Often, for surveillance in large populations, there will be two (or more) levels of design prevalence specified in order to take clustering of infection into account (see Clustering of infection on page 58). In this case, there will be more than one infection node (one for each level of clustering).

The probability for an infection node at a given level never changes. For example, in Figure 8 on page 73, there are two infection nodes (one for *YOUNG* and one for *OLD* animals). The value for the infected branch for both is the same – the design prevalence.

If a tree describes national surveillance for a disease that clusters, there are likely to be two infection nodes, one at the herd level and one at the animal level. The *INFECTED* branches for each of the herd-level nodes will all have the same probability, $P^*_H$, the herd-level design prevalence (which, for example, may be 1 percent). The value for the branches for the animal-level nodes will be the animal-level design prevalence, $P^*_A$ (which, for example, could be 20 percent for a highly infectious disease).

However, like other nodes, infection nodes are conditional on the previous branches in the limb. The probability that an animal is infected if the herd is not infected is equal to zero. Where there are more than one infection nodes, the probability for the *INFECTED* branch will be zero if any of the previous infection node branches were *NOT INFECTED*.

## Detection node

A detection node describes the probability that an animal will be detected as being infected. Each scenario tree must have at least one detection node, but some have many. The last node in a tree is always a detection node. Moreover, detection nodes always have two branches: *DETECTED* (the 'yes' answer to the question), or *NOT DETECTED* (the 'no' answer). The sensitivity of a surveillance system is the probability that infection will be detected if it is present at a defined level. If there are no detection nodes, there is no way to describe how the disease can be detected.

Typically, scenario trees can have two different detection structures based on laboratory testing or clinical surveillance.

For surveillance based on laboratory testing (e.g. a structured survey where sampled animals are all tested with a specified test), the detection node corresponds to the test used.

---

**Example**

For instance, if an ELISA is used to detect antibodies, then:
  – the detection node would be the ELISA result;
  – the question would be: "Is the sample positive to the ELISA test or not?";
  – the branches would be "Yes (*POSITIVE*)" and "No (*NEGATIVE*)";
  – for animals that are infected, the *POSITIVE* branch probability would be the sensitivity of the ELISA test;
  – for animals that are not infected, the *NEGATIVE* branch probability would be the specificity of the ELISA test.

---

Where follow-up tests are used, these can be added as a series of further detection nodes, one for each test.

For clinical surveillance, the detection nodes describe steps in the detection process. Typically, the sequence of events that must occur in order for clinical detection of an infected animal to take place are as follows:

**Clinical surveillance.**

- Infected animal shows clinical signs.
- Owner notices clinical signs.
- Owner contacts veterinarian.
- Veterinarian examines animal.
- Veterinarian takes appropriate samples.
- Samples tested for the disease.

Each of these steps is represented by a separate detection node, and has a probability of occurring. This is then normally followed by one or more laboratory tests to detect the infection and possibly confirm the diagnosis.

## Category node

Infection and detection nodes describe the probability of being infected and of being detected. However, on their own, they are capable of describing differences between subpopulations. Category nodes are used to divide the population into groups according to relevant factors. Category nodes allow us to explicitly account for the impact of a wide range of factors in our surveillance.

**Category nodes are used to describe factors influencing infection or detection.**

There are three types of category nodes: risk category nodes (which influence the risk of infection), detection category nodes (which influence the probability of detection), and group category nodes (which are used to describe the coverage of the surveillance).

While every scenario tree must contain at least one infection node and one detection node, category nodes are optional. However, without at least one category node, the tree is, in effect, assuming that all animals in the population have equal probability of being infected and detected. The results of analysis of the scenario tree will therefore be identical to a simple analysis assuming representative sampling. The aim of scenario trees is to take into account differences in subpopulations, and category nodes are the way to achieve that.

Category nodes must have at least two branches, but they may have more than two,
depending on the nature of categories being considered. For instance, the risk of infection may vary geographically. If there are seven geographical regions, each with a different risk, there would be seven branches to each REGION category node.

Category nodes have two or more branches.

The probability associated with the branches of a category node represents the proportion in each of the groups. For instance, if 20 percent of animals are young, then the *YOUNG* branch of the AGE category node would have a probability of 20 percent, and the *OLD* branch would have a probability of 80 percent.

Two different proportions are used, referring to two different populations. The first is the population proportion (*PrP*), which uses the entire population as the reference (conditional on earlier nodes). The second is the surveillance proportion (*PrSSC*), which only uses those animals included in the surveillance system component as the reference. The use of population and surveillance system component proportions is explained in Chapter 9.

A simple way to remember the difference between the node types is to consider the values used for probabilities with the node branches:

- infection nodes: design prevalence
- detection nodes: sensitivity
- category nodes: proportions

### Risk category nodes

A risk category node is used to describe the effect of a risk factor for infection. In Figure 8 on page 73, AGE is a risk category node, as it describes a risk factor that influences the probability of infection. Young animals are more susceptible to infection than older animals.

**Risk category nodes:** factors influencing risk of infection.

Risk category nodes are more complex than the other types of nodes, as they need to describe the difference in risk of infection between the two categories. The way in which this is achieved is discussed in detail in Chapter 9.

### Detection category nodes

The sensitivity of a test or other component of a detection system may vary. For instance, some tests are more sensitive in earlier stages of infection than later stages; the probability that an animal will show clinical signs may depend on the serotype infecting the animal. These factors may not influence the probability of infection, but they can influence the probability of detection. Detection nodes allow these factors to be included in the scenario tree.

**Detection category nodes:** factors influencing probability of detection.

### Group category nodes

Group category nodes describe factors that have no direct impact on the probability of infection or the probability of detection. They therefore do not influence the results of the scenario tree. The only reason for including these nodes is to allow the scenario tree to be analysed separately for different populations of interest.

**Group category nodes:** factors to describe coverage.

For example, there may be no geographical difference in the risk of infection. However, it may be useful to include a group category node for region. This will allow the surveillance system sensitivity to be analysed region by region (instead of using just a single summary measure for the whole country), thus enabling the quality of surveillance to be compared between regions.

## BUILDING A SCENARIO TREE

A scenario tree describes:

- the risk that an animal or group of animals might be infected (based on the risk factors for infection and the structure of the population); and
- the way in which an infected animal may be detected, based on the structure of the surveillance system.

### Step 1: Infection nodes

The best way to build a scenario tree is to consider these two areas separately. Let us use as an example a simple scenario tree to describe surveillance for bovine brucellosis. Rather than drawing a tree as shown in Figure 7 and Figure 8 on pages 72 and 73 respectively, it is usually much simpler to just write a list of nodes.

Start with the infection nodes, as every tree must have at least one infection node. Surveillance systems for diseases that cluster normally have two nodes. Because brucellosis clusters at the herd level, we will start with two infection nodes, so our list looks like this:

REGION
HERD TYPE
HERD SIZE
**HERD INFECTED**
RECENTLY ABORTED
VACCINATED
**ANIMAL INFECTED**
AREA
RBT RESULT
SNT RESULT

Note that we start from the largest unit (herd) and progress down to the smallest unit (animal).

### Step 2: Risk factors at the herd level

The next step is to identify all the possible risk factors (factors that influence the probability of infection). This can be done in two parts: factors operating at the herd level, and factors operating at the animal level.

If brucellosis is present in the population, there may be factors which mean that one herd is more likely than another to become infected. For instance, brucellosis is often more common in dairy herds than in beef herds, so the type of herd should be included as a risk

factor. Because small herds buy in fewer animals, they may have a low risk of being infected than large herds; therefore, herd size could be included as a risk factor.

Risk factors are added to our list *before* the infection node that they influence. These two risk factors are therefore added *above* the herd infection node.

Our list now looks like this:

> **REGION**
> **HERD TYPE**
> **HERD SIZE**
> **HERD INFECTED**
> **RECENTLY ABORTED**
> **VACCINATED**
> **ANIMAL INFECTED**
> **AREA**
> **RBT RESULT**
> **SNT RESULT**

### Step 3: Animal-level risk factors

The next step is to list those risk factors operating at the animal level. This could include recently aborted cows. Vaccination is sometimes used, and this makes it less likely that an animal will be infected (but it is still not completely protective). These factors can be included in our list before the animal infection node:

> **REGION**
> **HERD TYPE**
> **HERD SIZE**
> **HERD INFECTED**
> **RECENTLY ABORTED**
> **VACCINATED**
> **ANIMAL INFECTED**
> **AREA**
> **RBT RESULT**
> **SNT RESULT**

### Step 4: Detection nodes

We now have a description of the risk structure of the population. Next, we need to describe how our surveillance system detects disease. For this example, we are doing a targeted survey. Specimens are collected from selected animals, and tested using first the Rose Bengal test (RBT). This is then confirmed with the serum neutralization test (SNT). Animals are considered positive if they test positive to both of these tests. There are therefore two detection nodes:

REGION
**HERD TYPE**
**HERD SIZE**
**HERD INFECTED**
**RECENTLY ABORTED**
**VACCINATED**
**ANIMAL INFECTED**
AREA
**RBT RESULT**
SNT RESULT

## Step 5: Detection category nodes

The next step is to identify any factors that influence the probability of detection. Our surveillance is being conducted in a country where transport is difficult, and samples are tested at the country's central laboratory. Lengthy transport may result in poor sample quality, thus decreasing the sensitivity of the test. We could therefore divide the country into areas (*REMOTE* and *NOT REMOTE*) as an indicator of sample quality, and use different sensitivity values for our tests (lower sensitivity for remote areas). Our list now looks like this:

REGION
**HERD TYPE**
**HERD SIZE**
**HERD INFECTED**
**RECENTLY ABORTED**
**VACCINATED**
**ANIMAL INFECTED**
**AREA**
**RBT RESULT**
**SNT RESULT**

## Step 6: Group category nodes

The final step is to consider if we wish to add a group category node, thus allowing us to analyse the results of the tree for different parts of the population. For instance, we could add a REGION node so that we have surveillance sensitivity estimates for each different region in the country. Group category nodes, if used, are normally added to the top of the tree.

TABLE 3

| Node | Type | Branches |
| --- | --- | --- |
| Region | Group category | Region 1, 2, 3, etc. |
| Herd type | Risk category | Beef, dairy |
| Herd size | Risk category | Small, medium large |
| Herd infected | Infection | Infected, not infected |
| Recently aborted | Risk category | Aborted, not aborted |
| Vaccinated | Risk category | Vaccinated, not vaccinated |
| Animal infected | Infection | Infected, not infected |
| Area | Detection category | Remote, not remote |
| RBT result | Detection | Positive, negative |
| SNT result | Detection | Positive, negative |

**REGION**
**HERD TYPE**
**HERD SIZE**
**HERD INFECTED**
**RECENTLY ABORTED**
**VACCINATED**
**ANIMAL INFECTED**
**AREA**
**RBT RESULT**
**SNT RESULT**

## Step 7: Node types and branches

At this stage, it is a good idea to review the list. We may have forgotten some factors, or included some factors which, on consideration, are unlikely to play a significant role. For each node, you should then identify the type of node, and specify the branches, as shown in Step 8.

## Step 8: Branch probabilities

Once the nodes and branches are defined, the next task is to calculate or estimate branch probabilities and other model parameters. This is discussed in Chapter 11.

### Step 9: Implement the tree

The scenario tree then needs to be put into a format that allows it to be analysed. There are several options for doing this: on paper (for very simple trees); use a spreadsheet (while suitable for simple and complex trees, this requires a great deal of time and care); or use specially designed software for scenario tree analysis (much simpler and suitable for simple to moderately complex trees – highly complex trees may need the flexibility of a spreadsheet).

### Step 10: Analysis

Analysing the tree model involves using the AND and OR probability rules. The probability of an animal being in each of the subpopulations is calculated by multiplying the probabilities for each of the branches in the limb. The probabilities for each limb that gives a positive outcome (infection detected) are added. This provides the *unit sensitivity* (*USe*),which is the average probability that a single animal that passes through the surveillance system will give a positive result.

In risk category nodes, there are special rules for handling the multiplication of probabilities. These are explained in Chapter 9. This manual assumes that most analysis will be done using specialized software, which simplifies these calculations. The software is described in Chapter 17.

## TREE BUILDING RULES

The example above described the general process of building a scenario tree. The following rules may be useful when building your own trees:

- Infection and detection nodes have two branches.
- Category nodes have two or more branches.
- A tree should be symmetrical – the nodes encountered are the same along every limb of the tree. While this is not absolutely essential, it makes calculation easier.
- There must be at least one infection node. There are often two, but rarely more than two such nodes.
- There must be at least one detection node, but there are often more than that number.
- The last node in the tree must be a detection node.
- Risk category nodes must **be placed** before the infection node that they are influencing.
- Detection category nodes must **must be placed** before the detection node that they are influencing.
- For each node, the probabilities of all branches must add up to one.
- Probabilities are conditional on all previous branches in the limb.

## NODE ORDER

The general order for nodes in a scenario tree is:

1. group category node (if required)
2. nodes relating to infection
   a. zero or more risk category nodes describing risk factors operating at the herd level
   b. zero or one herd infection node
   c. zero or more risk category nodes describing risk factors operating at the animal level
   d. zero or one animal infection node (as some trees may stop at the herd level)

3. nodes relating to detection
   a. zero or more detection category nodes
   b. one or more detection nodes describing the surveillance system

The order of risk category nodes when there are multiple nodes relating to the same infection node is not important. However, as the lower nodes are conditional on the higher ones, these conditional probabilities are often easier to estimate with one node order compared to another.

---

**Example**

Sex (*MALE/FEMALE*) and region (*1/2/3/4*) are risk category nodes in a scenario tree. It is possible to include them in either order.

If sex comes before REGION, the probability for the *MALE* branch of the *SEX* node is the proportion of males in the population. The probability for the REGION 1 branch of the REGION node is the proportion of all males that are in Region 1. While there is clearly a correct value for this proportion, it is not intuitively easy to grasp and may be difficult to calculate.

If REGION comes before SEX, the REGION proportions will be the proportion of all animals in each region. The proportion for the MALE branch of the SEX node under REGION 1 will now be the proportion of males in Region 1. This is conceptually easier to understand and the figures are likely to be more readily available in this form.

# Chapter 9
# Incorporating risk into a scenario tree

*"There are sadistic scientists who hurry to hunt down errors instead of establishing the truth."*
Marie Curie (1867–1934)

The purpose of a scenario tree is to describe how different parts of the population have different probabilities of being infected and being detected. Scenario trees allow us to analyse risk-based surveillance, targeted at groups that are more likely to be infected.

The previous chapter provided an overview of building a scenario tree and it also introduced the different types of nodes. This chapter focuses on the use of risk category nodes to incorporate risk into a scenario tree.

In common usage, 'risk' is defined as the likelihood of an adverse event occurring. However, in risk analysis, risk is a combination of both the likelihood and the consequences of an adverse event. In scenario tree modelling, the term risk is used to describe only the likelihood of an event.

## QUANTIFYING TARGETING IN RISK-BASED SURVEILLANCE

Risk-based surveillance is an approach to disease surveillance that involves looking for disease where it is most likely to be present. Instead of representative surveillance (where we assume we know nothing about the risk of different subpopulations), we use our understanding of the disease to determine those animals that are most likely to be infected, and concentrate our surveillance effort there. Clearly, this is more efficient – by examining the high-risk groups, you have a greater chance of finding the disease (if it is present) than by examining animals that are at lower risk.

To capture the benefit of this type of surveillance in the scenario tree, we need to understand exactly what we are doing. First, we are talking about identifying groups of animals that are at higher risk of disease. To do this, we need to answer the questions:

- Which animals? How do we define the group?
- How do these animals differ from the rest of the population? What is the difference in the risk?

Second, we need to understand how our surveillance is targeting this risk group. We need to be able to determine if animals in the risk group have a higher probability of being included in our surveillance than other animals in the population.

## Describing differences in risk

Consider a risk factor for bovine tuberculosis – the presence of infected wildlife in the area. To use this risk factor in a scenario tree, we need two things:

1) To define the high-risk and low-risk populations
2) To describe the differences in risk between them.

The first has almost been done – our high-risk population comprises those farms that are in an area where wildlife are infected with tuberculosis. To use this definition, we would need maps of the distribution of infected wildlife, and maps of the cattle farms. Using these maps, we could divide the population of cattle farms into those in the areas with infected wildlife and those in areas where there are no infected wildli.

Quantifying the difference in risk is done using the relative risk (sometimes called the risk ratio, and abbreviated as *RR*). The relative risk describes the risk of being infected in one group, relative to the risk of being infected in the other group. Typically, this can be estimated by observational studies that measure the prevalence or incidence of disease.

---

**Example**

A study of herd infection rates has been conducted in a country infected with bovine tuberculosis. The study found that the incidence of new herd breakdowns in the areas with wildlife was 3.6 breakdowns per 100 herds per year. In the areas without wildlife infection, the incidence was 1.2 breakdowns per 100 herds per year. The relative risk is the incidence in the high-risk group relative to (or divided by) the incidence in the low-risk group:

$$\mathrm{RR} = \frac{3.6}{1.2}$$
$$= 3$$

This can be interpreted as meaning that herds in the high-risk area are three times more likely to become infected than herds in the low-risk area.

---

Relative risk is a ratio, and it can range between 0 and infinity. If the relative risk equals one, it means that the risk in the two populations is the same (i.e. the 'risk factor' is actually having no effect).

We use the relative risk to describe the difference in risk between different **sections** of the population, and to identify our high-risk groups.

We can calculate a relative risk for risk factors that comprise more than two groups.

**Example**

We want to divide the areas with infected wildlife into heavily and lightly infected areas, which means that we will have three risk groups: no wildlife infection, low levels of wildlife infection and high levels of wildlife infection. The incidence of herd break-downs has been calculated for each of these areas:

No wildlife infection: 1.2 breakdowns per 100 herds per year

Low levels of wildlife infection: 2.7 breakdowns per 100 herds per year

High levels of wildlife infection: 4.3 breakdowns per 100 herds per year

To calculate the relative risk, we compare each of the risk groups to the group with the lowest risk (no wildlife infection). This gives relative risks of:

$$RR_{\text{no infection}} = \frac{1.2}{1.2} = 1$$

$$RR_{\text{low infection}} = \frac{2.7}{1.2} = 2.25$$

$$RR_{\text{high infection}} = \frac{4.3}{1.2} = 3.58$$

In many cases, suitable studies measuring the different incidence or prevalence of disease associated with the risk factor are not available. In these cases, it is necessary to estimate the relative risk, as discussed in Chapter 11.

### *Putting relative risk into the tree*

As we now know the different relative risks for the different groups in the population, we can use this in our risk category node. Consider a small section of a scenario tree:



FIGURE 9
**Example of a small section of a scenario tree showing a risk category node and three infection nodes**
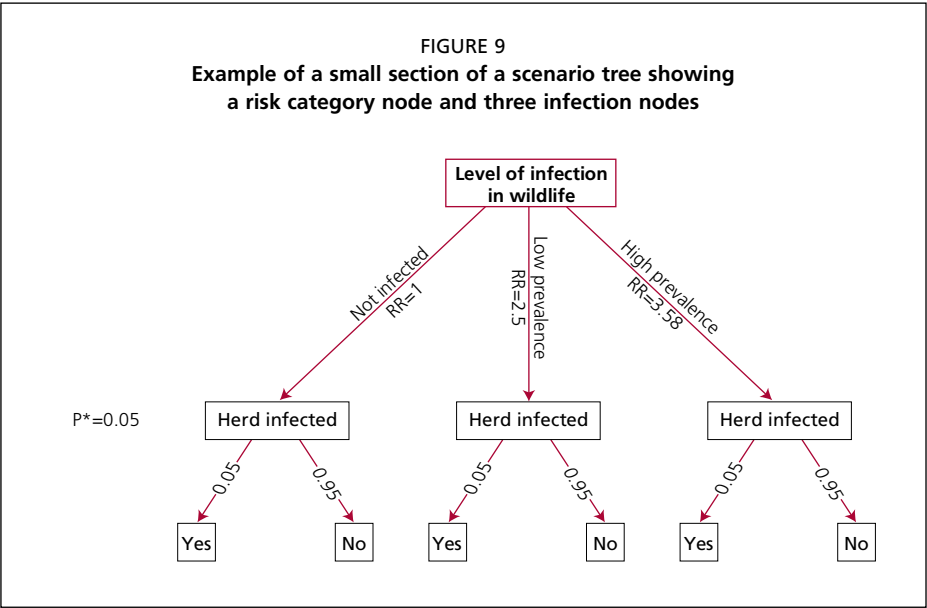
When analysing a scenario tree, we multiply the figures down each of the branches to determine the probability that an animal will be in a specific group. In this example, the probability that a herd will be infected if it is in an area with no wildlife infection is 1 (the *RR*) times 0.05 (the design prevalence), which is 0.05. However, if the herd is in a high-prevalence area, the probability is 3.58 times 0.05, which is 0.18.

In effect, by using the relative risks, we have changed the design prevalence (or probability that a herd or animal will be infected), and made it higher for our high-risk groups. In this example, we are now saying that the chance of a herd being infected in the high-risk area is 18 percent, whereas it is only 5 percent in a low-risk area. If the level of infection is that high, it will be much easier to find disease with our surveillance, so our surveillance will be much more sensitive in the high-risk area.

This is the result we want. The relative risk is used to adjust the standard design prevalence to show how much more likely it is that high-risk herds or animals would be infected.

Unfortunately, there is a problem with this approach. The design prevalence sets the standard for our surveillance and is a measure of the average probability that an animal or herd will be infected. In this example, we said that the design prevalence is 0.05, or that, on average, herds have a 5 percent probability of being infected. After adjusting for the risk of the different herds, the average design prevalence (assuming that each type of herd is equally represented in the population) is (0.05 + 0.11 + 0.18)/3 = 0.11 or 11 percent. Using relative risk in this way has meant that the average probability of infection for the entire population has increased. We are no longer using our standard design prevalence.

While we have described the difference in risk, the values we use for risk in the model are changing our average design prevalence. These numbers need to be adjusted so that the average risk does not change. The approach to adjusting the numbers involves understanding population proportions and targeting. The relative risk describes the difference in the risk of the different populations, but we have not yet described how our surveillance is targeting these populations.

## Describing targeting

Risk-based surveillance targets high-risk populations. Targeting means that some herds or animals are selected based on their risk. To understand targeting, it is necessary to compare it to representative sampling. In representative sampling (for instance, using random sampling), the proportions in the sample are the same as the proportions in the population. If 20 percent of the population has a high risk of being infected, a representative sample would have 20 percent of animals in the high-risk group. In contrast, targeted sampling would have more than 20 percent of the sample from the high-risk group.

*Targeting is described by comparing the population proportion to the surveillance proportion.*

To describe targeting, it is therefore necessary to compare two proportions: first to identify what proportion of the entire population is in the risk group, and then to identify what proportion of the sample is in the risk group. The difference between these two proportions describes the level of targeting.

---

**Example**

Ten percent of cattle herds are in areas with high levels of wildlife tuberculosis infection (our *population proportion* or *PrP* is 10 percent).

If 10 percent of animals in our surveillance (our *surveillance system component proportion* or *PrSSC*) are from the high-risk group, there is no targeting.

If 50 percent of animals in our surveillance are from the high-risk group, it is clear that we have concentrated our surveillance in this area and we are employing targeted surveillance. This approach will give us a higher chance of detecting the disease than if we employed representative sampling.

If only 5 percent of animals in our surveillance come from the high-risk group, we have under-represented this group of the population. This is called *negative targeting* or negatively biased surveillance. Using this approach, we are less likely to find disease than if we used representative sampling.

---

## Implementing risk in a scenario tree

Branches for most nodes have just one number associated with them – a probability or proportion for the surveillance system component. Out of all the farms or animals that are included in the surveillance system component, this is the proportion that falls into each branch.

However, it is now clear that for a full understanding of risk and targeting, the branches for our risk category nodes in the scenario tree require three pieces of information:

1. the relative risk for that branch (*RR*)
2. the proportion of the *population* in that branch (*PrP*)
3. the proportion of the *surveillance component* in that branch (*PrSSC*)

How are these three numbers used when calculating a scenario tree? This is explained in detail in the next section, but **briefly summarized as follows**:

1. The *PrSSC* is used in the same way as the proportions in other nodes.
2. The population proportion and relative risk are used to adjust the design prevalence that the risk category node refers to.
   a. The population proportion and relative risk are combined into an *adjusted risk*, which ensures that the average design prevalence is constant across the population.
   b. The adjusted risk is multiplied by the design prevalence to calculate the *effective probability of infection* (*EPI*), which takes the place of the design prevalence.

## WHAT YOU NEED TO KNOW

If you are implementing a scenario tree using the software described in Chapter 17, you can skip the next section, as you already know everything you need to know. To capture the effect of risk-based surveillance in a scenario tree, you need to provide three different numbers for each branch of a risk category node:

- the relative risk of herds or animals in that group being infected (relative to the group with the lowest risk)
- the proportion of the population in that group, and
- the proportion of the surveillance system component in that group

If you are implementing a scenario tree using a spreadsheet, or if you just want to understand how the tree works, then you should read the following section.

## CALCULATION OF ADJUSTED RISK

A relative risk is a ratio of the incidence or prevalence of infection in one group compared

This section is not required if you intend to use the freedom software.

to another. If the prevalence in one group is 20 percent and in the other group is 10 percent, then the relative risk is 20/10 or 2:1. This ratio can be expressed in different ways and still have the same value: 20/10 = 2/1 = 4/2 = 1/0.5

The adjusted risk for a branch is a measure of the risk in that branch. The ratio of adjusted risks has the same value as the relative risk, but it has adjusted risks expressed in different figures. This adjustment does not change the measure of risk, but it is done so that the weighted average probability of infection across all groups remains the same as the design prevalence.

### The constraints

To calculate the adjusted risk, we must consider two constraints.

**Constraint 1:** The ratio of the adjusted risks must remain the same as the original relative risk.

This can be expressed by the formula:

$$\frac{AR_1}{AR_2} = \frac{R_1}{R_2}$$

Where:

*AR* is the adjusted risk,

*R* is the risk, and

$\frac{R_1}{R_2}$ is the relative risk

This means that if the relative risk, as shown above, is 2:1, then the ratio of the adjusted risks must be in the same ratio (e.g. 4/2, 6/3, 1.5/0.75)

**Constraint 2:** The average risk across the population is, by definition, equal to one. This average risk must not change.

As a formula this can be expressed as:

$$\left(AR_1 \times PrP_1\right) + \left(AR_2 \times PrP_2\right) = 1$$

FIGURE 10

Relative risk

High-risk group

Population proportion
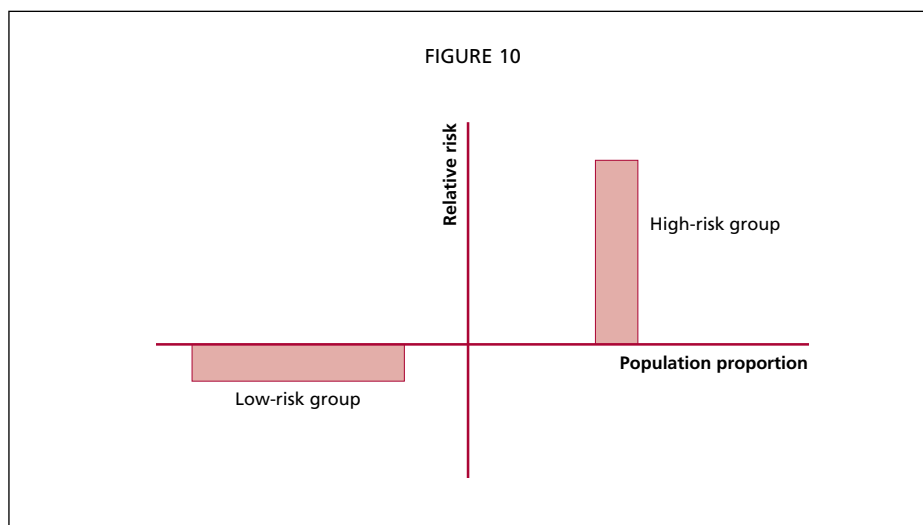
Low-risk group

Or more generally (when there are more than two risk groups):

$$\sum_{i=1}^{I}\left(AR_i \times PrP_i\right)=1$$

where *PrP* is the population proportion.

The objective is to find the numbers for the adjusted risk that meet both of these constraints. The solution is presented graphically in Figure 10. The x-axis shows the population proportion and cuts the y-axis (relative risk) at the value of 1.

The correct values for adjusted risk occur when the areas of the two rectangles are equal. In this example, there is a large proportion of animals in the low-risk group. The risk in this group must be lower than one, to balance the risk in the other group, but as there are many animals in this group, the adjusted risk can be just a little below one. In the high-risk group, there are very few animals. In order to balance the many animals in the low-risk group, the adjusted risk has to be much greater than one.

## The solution

Mathematically, the solution is found using simultaneous equations, to solve for $AR_2$ as shown. Constraint 1 on page 90 gives us:

$$\frac{AR_1}{AR_2}=\frac{R_1}{R_2}$$

Rearranging, we can equate this to $AR_1$.

$$AR_1=\frac{AR_2 \times R_1}{R_2}$$

Our second constraint is:

$$\left( AR_1 \times PrP_1 \right) + \left( AR_2 \times PrP_2 \right) = 1$$

If we replace $AR_1$ in this formula with the expression for $AR_1$ from the previous formula, we get:

$$\left( \frac{AR_2 \times R_1}{R_2} \times PrP_1 \right) + \left( AR_2 \times PrP_2 \right) = 1$$

To solve this for $AR_2$, we first extract $AR_2$ from the terms on the left:

$$AR_2 \left[ \left( \frac{R_1}{R_2} \times PrP_1 \right) + PrP_2 \right] = 1$$

Reorganize the complex middle term to use the same denominator by multiplying $PrP_2$ by $R_2/R_2$:

$$AR_2 \left[ \left( \frac{R_1 \times PrP_1}{R_2} \right) + \left( \frac{R_2 \times PrP_2}{R_2} \right) \right] = 1$$

Now divide both sides by the middle term and further simplify:

$$AR_2 = \frac{R_2}{\left( R_1 \times PrP_1 \right) + \left( R_2 \times PrP_2 \right)}$$

The value of $R_2$ is normally 1 as this is the low-risk (or reference) category. The formula can therefore be simplified to:

$$AR_2 = \frac{1}{\left( R_1 \times PrP_1 \right) + PrP_2}$$

Finally, once $AR_2$ is known, $AR_1$ can be calculated:

$$AR_1 = AR_2 \times \frac{R_1}{R_2}$$

Or, if $R_2$ is equal to 1,

$$AR_1 = AR_2 \times R_1$$

**Example**

Ten percent of farms are from an area with a high prevalence of tuberculosis in wildlife (the population proportion for the high-risk group, $PrP_1$), and 90 percent ($PrP_2$) are in areas that have no infection. The risk in the high-prevalence area is three times greater than in the area with no infection ($R_1=3$, $R_2=1$).

The adjusted risk for the low-risk group is:

$$AR_2 = \frac{1}{(R_1 \times PrP_1) + PrP_2}$$

$$= \frac{1}{(3 \times 0.1) + 0.9}$$

$$= \frac{1}{1.2}$$

$$= 0.83$$

The adjusted risk in the high-risk group is:

$$AR_1 = AR_2 \times R_1$$

$$AR_1 = \frac{1}{1.2} \times 3$$

$$= 2.5$$

Note how, in order to average to one, the adjusted risk in the low-risk group is always less than one and the adjusted risk in the high-risk group is always more than one.

# Chapter 10
# Calculating sensitivity with a scenario tree

> *"Things should be made as simple as possible, but not any simpler."*
> Albert Einstein (1879–1955)

At this point, we have seen an example of how to build a scenario tree, and how to capture risk within the tree. We are now able to use the tree to calculate the sensitivity of our surveillance system component. This is done in two stages:

- Calculate the average sensitivity of the surveillance system component for a *single animal*, or the component unit sensitivity (C*SeU*).
- Estimate the sensitivity for the entire surveillance system component (*CSe*) based on the total number of animals that are processed by the system.

## CALCULATION OF UNIT SENSITIVITY

Let us consider a simple example of a tree with one risk factor (age), one infection node and one detection node (an ELISA test). This may describe a surveillance system that involves collection of blood from animals, where age is a known risk factor (young animals at higher risk), and where most of the blood samples come from younger animals.

### Building the scenario tree

In this case, the structure of the scenario tree is very simple, with just the three nodes.

### Organizing the model parameters

It helps at this stage to carefully organize the parameters that will be used in the model. The first parameters will be those for the AGE risk category node. The risk in young animals is five times greater than in older animals. Young animals make up 20 percent of the population but 90 percent of the animals under surveillance. These figures are summarized on page 96, and the adjusted risk is calculated using the formulae shown in Chapter 9.

The infection and detection nodes also require parameters as well. The design prevalence is 0.05 and the sensitivity of the ELISA is 95 percent.

### Drawing the tree and adding parameters

In the examples we have used thus far, we have drawn the tree, starting from the top, with the branches leading down. This is a useful way to illustrate how the process works. Calculating the tree requires operations on many different numbers, and a spreadsheet is useful when performing such operations. As the downward branching structure of the tree is difficult to represent in a spreadsheet, it is more common to draw the tree starting from the left and moving across the page. An example of how our simple tree can be

TABLE 4

| Node | Type | Branches |
|------|------|----------|
| AGE | Risk category | *YOUNG, OLD* |
| ANIMAL INFECTED | Infection | *INFECTED, NOT INFECTED* |
| ELISA RESULT | Detection | *POSITIVE, NEGATIVE* |

TABLE 5

| Risk factor | Branches | Relative risk | Population proportion | Surveillance proportion | Adjusted risk |
|-------------|----------|---------------|-----------------------|-------------------------|---------------|
| AGE | *Young* | 5 | 0.2 | 0.9 | 2.78 |
|  | *OLD* | 1 | 0.8 | 0.1 | 0.56 |

TABLE 6

| Node | Branches | Probability type | Value |
|------|----------|------------------|-------|
| INFECTION | *INFECTED* | Design prevalence | 0.05 |
|  | *NOT INFECTED* |  | 0.95 |
| ELISA RESULT | *POSITIVE* | Sensitivity | 0.95 |
|  | *NEGATIVE* |  | 0.05 |

represented in a spreadsheet is shown on page 97. The columns represent nodes, and each row represents one possible limb (path through the tree).

Once the tree structure is prepared, the parameter values can be entered. This is easier if the parameters have been well organized in the same spreadsheet, as illustrated in Table 7.

For the AGE node, *PrSSC* represents the proportion in each group in the surveillance system.

For the INFECTED node, *EPI* represents the *effective probability of infection*. This is the adjusted risk (from Table 7) multiplied by the design prevalence (0.05).

For the ELISA RESULT node, *Se* represents the sensitivity. Note that the sensitivity (probability of getting a positive test result) is zero where the animal is not infected (the *NO* branch of the infection node), as we are assuming perfect specificity.

The tree with the required parameters is shown in Table 8.

## Calculating the tree

We are now ready to calculate the probability of each branch. This is done by multiplying each of the values across the table for each branch. For instance, the probability for the first branch (young, infected, positive test result) is $0.9 \times 0.139 \times 0.95 = 0.119$, as shown on page 98.

TABLE 7

| Age | | Infected | | ELISA result | | |
|---|---|---|---|---|---|---|
| **Branch** | **PrSSC** | **Branch** | **EPI** | **Branch** | **Se** | **Outcome** |
| Young | | Yes | | Pos | | Pos |
| | | | | Neg | | Neg |
| | | No | | Pos | | Pos |
| | | | | Neg | | Neg |
| Old | | Yes | | Pos | | Pos |
| | | | | Neg | | Neg |
| | | No | | Pos | | Pos |
| | | | | Neg | | Neg |

TABLE 8

| Age | | Infected | | ELISA result | | |
|---|---|---|---|---|---|---|
| **Branch** | **PrSSC** | **Branch** | **EPI** | **Branch** | **Se** | **Outcome** |
| Young | 0.9 | Yes | 0.139 | Pos | 0.95 | Pos |
| | | | | Neg | 0.05 | Neg |
| | | No | 0.861 | Pos | 0 | Pos |
| | | | | Neg | 1 | Neg |
| Old | 0.1 | Yes | 0.028 | Pos | 0.95 | Pos |
| | | | | Neg | 0.05 | Neg |
| | | No | 0.972 | Pos | 0 | Pos |
| | | | | Neg | 1 | Neg |

The probability for the fourth branch would be $0.9 \times 0.861 \times 1 = 0.775$.

The completed calculations are shown on page 99.

Before going any further, it is always a good idea to check for errors. The tree represents the possible outcomes for a single animal in the surveillance system. When the probabilities of all the possible outcomes (i.e. the last column) **are added up**, the answer should always come to one.

## Calculating the component unit sensitivity (CSeU)

The component unit sensitivity is the probability that a single unit (animal) passing through the surveillance system would be detected. This means that we are only interested in those animals that have a positive outcome. (If the outcome is negative, the animal has not been detected.)

The component unit sensitivity is therefore the sum of the probabilities of all the different limbs that can lead to detection. These are highlighted in Table 12.

Only two limbs have a non-zero probability; therefore, the component unit sensitivity is $0.119 + 0.003 = 0.122$. This means that the probability that the surveillance system will detect disease by examining one animal is (on average) 12.2 percent.

TABLE 9

| Age | | Infected | | ELISA result | | | |
|---|---|---|---|---|---|---|---|
| Branch | PrSSC | Branch | EPI | Branch | Se | Outcome | Probability |
| Young | **0.9** | Yes | **0.139** | Pos | **0.95** | Pos | **0.119** |
| | | | | Neg | 0.05 | Neg | |
| | | No | 0.861 | Pos | 0 | Pos | |
| | | | | Neg | 1 | Neg | |
| Old | 0.1 | Yes | 0.028 | Pos | 0.95 | Pos | |
| | | | | Neg | 0.05 | Neg | |
| | | No | 0.972 | Pos | 0 | Pos | |
| | | | | Neg | 1 | Neg | |

TABLE 10

| Age | | Infected | | ELISA result | | | |
|---|---|---|---|---|---|---|---|
| Branch | PrSSC | Branch | EPI | Branch | Se | Outcome | Probability |
| Young | **0.9** | Yes | 0.139 | Pos | 0.95 | Pos | 0.119 |
| | | | | Neg | 0.05 | Neg | 0.006 |
| | | No | **0.861** | Pos | 0 | Pos | 0.000 |
| | | | | Neg | **1** | Neg | **0.775** |
| Old | 0.1 | Yes | 0.028 | Pos | 0.95 | Pos | |
| | | | | Neg | 0.05 | Neg | |
| | | No | 0.972 | Pos | 0 | Pos | |
| | | | | Neg | 1 | Neg | |

TABLE 11

| Age | | Infected | | ELISA result | | | |
|---|---|---|---|---|---|---|---|
| Branch | PrSSC | Branch | EPI | Branch | Se | Outcome | Probability |
| Young | 0.9 | Yes | 0.139 | Pos | 0.95 | Pos | 0.119 |
| | | | | Neg | 0.05 | Neg | 0.006 |
| | | No | 0.861 | Pos | 0 | Pos | 0.000 |
| | | | | Neg | 1 | Neg | 0.775 |
| Old | 0.1 | Yes | 0.028 | Pos | 0.95 | Pos | 0.003 |
| | | | | Neg | 0.05 | Neg | 0.000 |
| | | No | 0.972 | Pos | 0 | Pos | 0.000 |
| | | | | Neg | 1 | Neg | 0.097 |

TABLE 12

| Age | | Infected | | ELISA result | | | |
|---|---|---|---|---|---|---|---|
| **Branch** | **PrSSC** | **Branch** | **EPI** | **Branch** | **Se** | **Outcome** | **Probability** |
| Young | 0.9 | Yes | 0.139 | Pos | 0.95 | **Pos** | **0.119** |
| | | | | Neg | 0.05 | Neg | 0.006 |
| | | No | 0.861 | Pos | 0 | **Pos** | **0.000** |
| | | | | Neg | 1 | Neg | 0.775 |
| Old | 0.1 | Yes | 0.028 | Pos | 0.95 | **Pos** | **0.003** |
| | | | | Neg | 0.05 | Neg | 0.000 |
| | | No | 0.972 | Pos | 0 | **Pos** | **0.000** |
| | | | | Neg | 1 | Neg | 0.097 |

## Comparison with representative sampling

It is interesting to note at this point what the effect of our risk-based surveillance has been on the sensitivity. If we had used representative sampling – so that the average probability of infection was the same for all animals – the probability of getting a positive result from one animal (the unit sensitivity) would be:

$$CSeU = P^* \times Se$$
$$= 0.05 \times 0.95$$
$$= 0.0475$$

By focusing our sampling on young animals (the group with the highest risk) we have almost tripled the unit sensitivity from 4.75 percent to 12.2 percent. This increase in sensitivity is the reason we use risk-based surveillance.

## CALCULATION OF COMPONENT SENSITIVITY (CSE)

The component unit sensitivity is interesting, but it is not the value we are looking for. We want to know how good our surveillance is at detecting the disease, not by examining a single animal, but for the whole surveillance system component, which is examining many animals.

This step is simple, as we can use a formula that we discussed earlier. The CSe (surveillance system component sensitivity) is the probability that all animals in the surveillance system do not give a negative result:

$$CSe = 1 - (1 - CSeU)^n$$

If our surveillance system component (sampling blood from mostly young animals) involved the collection of 20 blood samples then:

$$CSe = 1 - (1 - 0.122)^{20}$$
$$= 0.925$$

We can compare the result of 92.5 percent to the sensitivity we would have got if we had used representative rather than risk-based sampling: 62.2 percent. Clearly, targeting the high-risk animals has provided a major advantage in this case.

## WHAT NEXT?

This chapter has presented an example of a very simple scenario tree, and has used it to calculate the sensitivity of a component of a surveillance system. This is the first complete example of the use of scenario trees presented in this manual, and it shows that they can be a useful and relatively simple tool. However, the example shown here needs to be refined in a number of ways in order to address the following:

- Complexity
  - Most surveillance systems are more complex than the example shown in this chapter, with many more risk factors, two levels of infection nodes, and multiple detection nodes. The same basic principles apply to complex trees, but they are harder to implement.
- Model parameters
  - Scenario trees use many parameters to describe risk and proportions. These figures are often difficult to ascertain, and a solution must be found when you do not know what the real values are.
- Uncertainty and variability
- If we are not sure about some of the inputs, we need a way to express our uncertainty in our estimate of sensitivity.
- Clustering
- This simple example assumed that all animals were at equal risk of being infected and each animal contributed the same amount of evidence that the population was free from infection. However, when disease clusters, testing many animals from the same herd means that there is less chance of finding the infection than would be the case if testing the same number of animals were spread over a large number of herds.

The next chapters will examine these issues in order to make the sensitivity estimates from the scenario tree more accurate and reliable.

Chapter 11
# Probability estimates in a scenario tree

> *"An expert is a person who has made all the mistakes*
> *that can be made in a very narrow field."*
> Niels Bohr (1885–1962)

Scenario trees require many numbers. Each branch has a probability associated with it, and some nodes (risk category nodes) have three numbers for each branch. As we work along each limb in the tree, each probability is conditional on all the previous probabilities in that limb. Where do all these numbers come from and how do we make sure that they are correct? This chapter provides some guidance in finding the right parameters for a scenario tree.

## SUMMARY OF REQUIRED VALUES

The purpose of a scenario tree is to calculate the sensitivity of a component of a surveillance system. Sensitivity is the probability that the surveillance system component will detect at least one infected animal, if the population is infected at the design prevalence. The result of the scenario tree analysis (sensitivity) is a probability; therefore all the branch parameters are also probabilities as well.

The probability values are different, depending on the type of node.

### Infection node

An infection node has two branches: *INFECTED* and *NOT INFECTED*. The probability associated

Infection node: design
prevalence.

with the *INFECTED* branch is the design prevalence ($P^*$). Prevalence is defined as the proportion of the population with a defined characteristic (in this case, the proportion that is infected). It also can be thought of as a probability: if an animal is drawn from the population at random, the probability that it will be infected is equal to the prevalence.

The approach to selecting an appropriate design prevalence was discussed in the section on page 44. To summarize, the way to choose a **design prevalence**, in order of preference, is:

- Use global standards (e.g. from the OIE code).
- Use regional standards (e.g. EU regulations).
- Check the requirements of your trading partners.
- Calculate based on your trading partner's stated acceptable level of protection (ALOP) – this approach is rarely feasible.
- Determine based on the biology of the disease (e.g. the minimum expected prevalence if infection is established).

- Determine based on practical considerations (what level of surveillance is affordable).
- Make an arbitrary choice based on common values (1 percent, 5 percent or 10 percent).

### Detection node

A detection node has two branches, representing *DETECTED* and *NOT DETECTED*. The probability

<div style="background:pink">
Detection node:
sensitivity.
</div>

associated with the *DETECTED* branch is a sensitivity. This is evident when the detection node refers to something like a laboratory test (e.g. a complement fixation test (CFT)). However, detection nodes are also often used to describe other complex components of a surveillance system, such as the probability that a farmer will call the veterinarian if they notice that an animal is sick. This may also be thought of as a sensitivity.

### Risk category node

A risk category node is the most complex type of node in the scenario tree, as discussed in

<div style="background:pink">
Risk category node:
relative risk, population
proportion and surveillance
system component
proportion.
</div>

Chapter 9, as each branch has three figures associated with it.

#### Relative risk

The relative risk describes how some parts of the population are at higher risk than others. This is the only figure used in a scenario tree that is not a probability. It is a ratio that can take a value from zero to infinity. When adjusted (using the population proportion) to create the *adjusted risk* (*AR*), it is multiplied by the design prevalence to give the *effective probability of infection* (*EPI*) which, again, is a probability.

### Population proportion (PrP)

The population proportion is used to change the relative risk to the adjusted risk. It represents the proportion of the entire population that is in the branch category. This is important as it allows the tree to take targeting into account.

### Surveillance system component proportion (PrSSC)

This is the proportion of animals in the surveillance system that fall into the branch category. Targeting is expressed by the difference between the population proportion and the SSC proportion.

With some surveillance systems, the *PrP* and the *PrSSC* are the same.

---

**Example**

Consider surveillance based on a survey using random representative sampling. Twenty percent of the population are in a high-risk group. The representative sampling will ensure that 20 percent of the sample (the PrSSC) is also in the high-risk group. Representative sampling does not target high-risk groups, so scenario tree analysis will give the same result as simpler methods of analysis.

> **Example**
>
> A passive farmer reporting system may have coverage of the entire population, as every animal is owned by a farmer, and therefore every animal has a chance of being reported and detected if it becomes infected.
>
> In this case, the PrSSC is the same as the entire population, so the PrSSC is the same as the population proportion. Surveillance systems that have complete coverage of the population therefore do not take different risk groups into account. However, they do take differences in the probability of detection into account, and thus scenario tree analysis is very useful for these situations.

### *Detection category node*

A detection category node is used to divide the population into groups that have different

> Detection category node: SSC proportion (and population proportion).

probabilities of being detected. The branches of a detection category node are associated with the proportions of the surveillance system component that fall into that category.

Often, we will also need to note the population proportion for detection category nodes, so as to determine how good the sensitivity of our surveillance is compared to representative (non-risk-based) sampling.

### *Group category node*

The group category node is similar to a detection category node in that it uses the

> Group category node: SSC proportion (and population proportion).

surveillance system component proportion, but will often have the population proportion recorded for comparison purposes.

In summary, the values that may be required for node branches include:

- relative risk
- sensitivity
- surveillance system component proportion
- population proportion

## SOURCES OF ESTIMATES

Sometimes figures for the probabilities are already available. Sometimes they are not, but data that allow the figures to be calculated or estimated are available. And sometimes, nothing is available. How do we deal with these different situations?

## Sensitivity

Sensitivity is the probability of getting a positive test result if the animal tested truly is infected. Sensitivity is used in detection nodes. Where a detection node refers to the use of a laboratory test, sensitivity estimates may be available. For more complex detection nodes (for instance, the probability that a sick animal will be noticed by the farmer, or the probability that a veterinarian will take samples for laboratory analysis), there are unlikely to be existing estimates.

### *Laboratory tests*

### *Existing estimates for validated tests*

In some cases, there are published studies in which a laboratory test has been validated, and the sensitivity and specificity have been calculated.

Even when this information has not been published, internal validation studies may have been carried out by a laboratory, and the figures may be available directly from that source. When these figures are available, it is reasonable to use them in the scenario tree model. However, there are a couple of considerations that need to be taken into account, and these are as follows:

- Sensitivity varies due to a variety of factors, including laboratory techniques and factors associated with the population under study. If validation studies have been carried out in another part of the world and on different populations, the values for sensitivity may not be directly applicable to the local population.
- Where the key factors that influence sensitivity are known, these should appear in the model as detection category nodes. In this case, different values for sensitivity should be used for the different categories of animals. It is relatively rare for studies aimed at estimating sensitivity to calculate different values according to a range of influencing factors – instead they tend to estimate an average sensitivity across the population studied.
- Often, sensitivity estimates are published as point values (for instance, 93.5 percent). However, these estimates were calculated using sampling approaches, and there is therefore some element of random error associated with the estimates. Ideally, this should be reported along with the estimates, in the form of a 95 percent confidence interval (for instance, in the form 93.5 percent [87.2 percent – 98.5 percent]). Where confidence intervals are not available, it may be possible to calculate them based on the sample size used to make the sensitivity estimate. The confidence interval describes the uncertainty around an estimate, which can be incorporated into the model, as described in Chapter 12.

### *Generating new estimates*

Where no published information or internal estimates for a laboratory test are available, an alternative solution is to undertake studies to generate the required estimates. The traditional study would involve identifying a number of truly positive animals (based on the use of a 'gold standard' test), and **subjecting** them to the test that needs to be validated. The sensitivity is the proportion of these animals that have a positive test result.

A newer, alternative study design (latent class analysis) makes it possible to estimate sensitivity and specificity without necessitating the use of a gold standard test. While this design requires the use of at least two different tests and two populations with different disease prevalence levels, it does not require the true status of individual animals to be known. Occasionally, large volumes of historical laboratory records are available to meet these requirements. **It may be possible to analyse** this data relatively quickly and cheaply, in order to produce good estimates of the test performance.

Both of these approaches pose a number of particular problems. Both the high cost and the time involved in conducting a study often make it impractical. However, more importantly,

the reason for the sensitivity estimate is to support surveillance to demonstrate that the infection is not present in the country. If the country is free from the infection, there are no infected animals to test (and artificially infecting animals would be very dangerous). In fact, this creates a paradox: in order to estimate the sensitivity of the test correctly, it should be evaluated on the animals of interest (the local population), but when the population is free from infection, the test cannot be evaluated.

Normally, we are forced to use estimates from areas where the disease is present, either historical data in the country of interest (if the disease has been eradicated), or from other countries with roughly similar populations.

### Expert opinion

If suitable estimates of test sensitivity are not available from any source, it may still be **possible to obtain** appropriate figures for use in the scenario tree. Even if the test has not been formally validated, it is likely that many scientists have used it in different situations for some time. These people are likely to have a reasonably good understanding of the performance of diagnostic tests. Formal approaches to gathering and analysing expert opinion offer a method for collecting sensitivity estimates when no other information is available. These approaches are discussed in detail in the next section.

### Other detection probabilities

Detection probabilities that are not associated with a laboratory test are most commonly found in trees using some aspect of passive reporting. A typical 'detection cascade' in a passive farmer reporting system may look like this:
- Infected animal shows clinical signs.
- Farmer notices animal with clinical signs.
- Farmer contacts veterinary services.
- Veterinarian examines animal.
- Samples taken for analysis.
- Samples tested for disease in question.

This is normally followed by one or more laboratory tests to detect and then confirm the infection.

Other non-laboratory detection probabilities may be associated with activities such as abattoir meat inspection.

### Existing estimates

The probabilities listed above have rarely been explicitly studied or quantified. The exception is perhaps the sensitivity of abattoir meat inspection for the detection of various diseases. It is unlikely that other useful information will be available in the published literature.

However, some figures could be available. For example, the first in the list, the probability that an infected animal shows clinical signs, is likely to be included in general descriptions of the epidemiology of the disease, and could be included in both textbooks and published papers.

*Generating new estimates*

Unlike the evaluation of laboratory tests, estimating the probabilities associated with these non-laboratory detection nodes may be feasible, even where the disease does not exist. One of the key advantages of a scenario tree is that it explicitly identifies the various probabilities involved in the detection system, and each of these can be studied separately.

Some of the probabilities listed on page 105 may be calculated from existing records. For instance, veterinary visit records may indicate how often a veterinarian collects samples for analysis from cases with a certain collection of presenting signs (consistent with the disease in question). Similarly, it may be possible to determine the probability that a sample submitted from a possible case is tested by examining laboratory testing records. While accessing and analysing these data sources may be difficult, they offer an approach to getting realistic probability estimates for some of the required parameters.

However, for others (e.g. the probability that a farmer would notice clinical signs, or that they would contact the veterinary services), no records are likely to exist. It may be possible to conduct small studies to directly measure these probabilities. One approach would be to convene a number of farmer meetings in different areas. At each meeting, farmers would be shown a series of photographs or videos, and asked to answer a number of questions. The videos could include a combination of scenes in which all the animals are healthy, and one or more animals are showing signs of the disease. The questions posed may be:

- Do you notice anything unusual about this group of animals?
- If yes:
  - What is unusual?
  - Would you take any action as a result of this observation?
  - What action would you take?

Such meetings should be conducted without giving any prior information to the participants (for instance, do not invite them to a meeting using a form of words which indicates that the purpose of the meeting is to study the detection of classical swine fever).

Clearly, while responses to questions that are posed in a formal setting, **such as a public meeting**, may not accurately reflect people's actual behaviour, such responses nevertheless provide some indication of behaviour, and this information may be usefully applied in the scenario tree.

*Expert opinion*

The previous example showed how a study of farmer behaviour could be analysed in the same way as other surveys, with the precision of the estimate being related to sample size (the number of farmers included in the study). If a structured study such as that described is not feasible, expert opinion may provide an alternative source of data. While similar in some ways, gathering expert opinion is fundamentally different to a study such as that described. In a traditional study, each participant provides a single observation. **On the other hand, however, when expert opinion is solicited, individual experts are likely to present quite different observations**, based on their experience.

In many cases, the "expert" who is gathering expert opinion is assumed to be some respected, well-educated person – a laboratory scientist, university professor etc. The explicit probabilities required by a scenario tree model show that the appropriate experts may be

TABLE 13

| Node | Type | Branches | Proportion |
|------|------|----------|------------|
| Region | Group category | *Region 1, 2, 3* | Herd |
| Herd type | Risk category | *Beef, dairy* | Herd |
| Herd size | Risk category | *Small, medium large* | Herd |
| **Herd infected** | **Infection** | ***Infected, not infected*** | |
| Recently aborted | Risk category | *Aborted, not aborted* | Animal |
| Vaccinated | Risk category | *Vaccinated, not vaccinated* | Animal |
| **Animal infected** | **Infection** | ***Infected, not infected*** | |
| Area | Detection category | *Remote, not remote* | Animal |
| RBT result | Detection | *Positive, negative* | |
| SNT result | Detection | *Positive, negative* | |

very different for the different questions. For instance, the experts considering a farmer's ability to detect animals with disease may be either a number of farmers (with experience of the behaviour of their peers), or, **alternatively**, may be just one individual who has extensive daily contact with farmers, sees their animals and hears about their observations of disease; such an individual may be a field veterinarian, a paraveterinary worker, or a trader. In order to address each question, a different expert is likely to be required.

Details on the use of expert opinion are discussed in the next section.

## Proportions

In a scenario tree, branch probabilities for category nodes are based on proportions. They all require the proportion of animals in the surveillance system component that fall into the category represented by the branch, but it is generally useful to know the population proportion as well (and this is also required for risk category nodes).

### *Proportion of herds or proportion of animals?*

It is important to note that the proportions used in category nodes refer to the units in the infection node immediately following. For instance, consider the following example list of nodes for a scenario tree used to analyse brucellosis surveillance:

The two infection nodes (HERD and ANIMAL) have been highlighted. The first three nodes (REGION, HERD TYPE and HERD SIZE) are all category nodes; therefore, their branches require proportions. As these three nodes are placed before the herd infection node, the proportions refer to the proportion of *herds* that fall into each group. For instance, if there are 10 000 herds in the country, and 4 000 of these herds are in Region 1, 3 500 in Region 2 and 2 500 in Region 3, then the probability for the *REGION 1* branch of the REGION node is 40 percent. In the *REGION 2* branch it will be 35 percent, and in *REGION 3* it will be 25 percent.

The herd size node also refers to the proportion of *herds* because it is placed before the herd infection node.

The RECENTLY ABORTED node, however, is placed before the animal infection node, and therefore refers to the proportion of *animals* that have recently aborted. The VACCINATION node refers to the proportion of animals vaccinated.

After the animal-level infection node, any detection category nodes refer to the animal level as well. Thus, the AREA node (*REMOTE* or *NOT REMOTE*) in the node list on page 107 refers to the proportion of animals that are in remote areas, or are in not remote areas. It would also have been possible (and maybe simpler) to place the AREA detection category node ahead of the HERD node, as it is the herd that is located in a remote or not remote area. Even though the node relates to detection of individual animals, it can be placed at higher points in the tree if it is logical to do so.

### *Conditional proportions*
Remember too that all of these proportions are conditional on the previous nodes, depending on which limb the node appears on. For instance, the branch probabilities for the AREA node (*REMOTE* and *NOT REMOTE*) are conditional on REGION, HERD TYPE, HERD SIZE, RECENTLY ABORTED and VACCINATED. In most cases, however, some of the nodes may be considered independent of previous nodes. For instance, the probability that an animal is in a remote area possibly depends on the region (some regions have more remote areas than other regions), but it may not be related to the animal's vaccination or abortion status, as these are independent of geographic remoteness.

In practice, to determine the population proportions for the AREA node, one could:
1. Obtain a map of the country.
2. Identify the three regions.
3. Calculate the number of herds in each region (this provides the data for the branch probabilities for the REGION node).
4. Identify on the map those areas that are remote and those that are not. (This may be simply done by marking a line at a given distance from the diagnostic laboratories (buffering). A more sophisticated approach would be to calculate 'travel time contours' based on road distance and speed – these are lines joining points that take the same time to travel to/from the laboratory).
5. Based on region and area, divide the country into six areas (*REMOTE* and *NOT REMOTE* in each of the three regions).
6. In each of these six areas, calculate the total number of animals.
7. The category proportions can then be calculated.
   a. For *REGION 1, REMOTE*, this is the number of animals in the remote area of Region 1 divided by the total number of animals in Region 1.
   b. For *REGION 1, NOT REMOTE*, it is one minus the proportion of animals in the remote area.

This example shows how the AREA node was only conditional on one of the previous levels and could be considered independent of the others. For each node, it is necessary to determine on which of the previous nodes the node is conditional, and for which it is independent.

For example, RECENTLY ABORTED is probably conditional on HERD TYPE (*BEEF* or *DAIRY*) and may also be conditional on REGION. Laboratory abortion investigation records may provide some

information about whether there are more abortions in one region than another – although the difference may be due to bias because of different reporting rates between regions; therefore, such data must be interpreted with caution.

The vaccinated node is likely to be conditional on region and herd type, and it should be possible to estimate the population proportion by using veterinary service vaccination records or vaccine sales records. It may also be conditional on area (remoteness), in which case area **should be placed** higher in the tree.

### *Population and surveillance system component proportions*

The previous examples focused on the use of official statistics or other records to provide information about the population. SSC proportions are often simpler to calculate, as we normally have data about the animals that are included in our surveillance.

Ideally, surveillance data should contain a record for each animal in the surveillance system component, and each record should contain information on each of the nodes included in our tree. For instance, for each animal, the dataset should contain the following information:

- a herd identifier linked to herd information, including:
  - the region the herd is in
  - the type of the herd (beef/dairy)
  - the number of animals in the herd
  - whether the herd is in a remote area or not
- whether the animal has recently aborted or not
- whether the animal has been vaccinated or not
- the results for the RBT and SNT

If this information is available, the dataset can be quickly summarized to provide all the SSC proportions required. Another approach to using this type of complete dataset will be discussed in Chapter 13.

Often, this type of complete data is not collected as part of the surveillance, and therefore some values have to be estimated. One approach to estimating the SSC proportions is to assess whether there was any targeting or bias related to that factor (normally in discussion with the surveillance designers or field teams). For instance, did they attempt to preferentially collect samples from animals that had recently aborted, or did they try to avoid vaccinated animals? If not, then the population proportion can be used, on the assumption that without targeting for these factors, the animals would be roughly representative. If there was targeting, an estimate of the level of targeting will be required to calculate the SSC proportion.

## Relative risk

Risk category nodes require estimates of the relative risk for each of the categories. These probably represent the most difficult values that are needed for a scenario tree.

For well-known risk factors, specific risk factor studies may have been carried out, thus providing reliable estimates of the relative risk for different categories. As with published sensitivity estimates, it is important to consider confidence intervals when using published estimates of relative risk.

In most cases, however, little information will be available. It is very difficult and expensive to carry out risk factor studies to measure the relative risk (such studies can only be conducted when the disease is present). Expert opinion is usually necessary in order to estimate relative risks.

## EXPERT OPINION

The use of expert opinion as a method of estimating parameters (such as the sensitivity of a test, or a relative risk) has often been viewed as undesirable and unreliable. This may be due to the common misconception that the approach is based on asking an expert (a wise person who is respected in their field) what they think the answer is, and the expert then makes a guess at the answer.

The appropriate use of expert opinion is very different. Philosophically, it is based on the same type of approach as scenario tree modelling to demonstrate freedom from infection. To demonstrate freedom, we aim to use all available sources of evidence, even if they are complex and even if alone they do not contribute very much evidence. The principle is not to say that the evidence is imperfect, and therefore to reject it, but to carefully understand the limitations of the different data sources, and to use that which is good (risk-based surveillance), and take into account that which is bad (the presence of bias).

Expert opinion is appropriate when no data based on direct structured observation are available. It may also be used to supplement information from small or potentially biased studies. The principle of the use of expert opinion is to acknowledge that, even if no study exists, there are usually a number of people who have a great deal of experience with the question of interest. Rather than say this experience is not as well structured or as easily captured as an objective study, we try to capture the experience, while at the same time making sure that any limitations are clearly taken into account.

The three main rules for expert opinion are:
- Ask the right experts.
- Ask them specific questions in a way that enables them to provide specific answers.
- Always capture uncertainty.

No matter how experienced experts are, there is a significant chance that they will give an incorrect answer. This does not matter too much if the answer is close to the correct answer, but it could be important if it is a long way from the correct answer. Capturing uncertainty involves asking the experts not only to say what they think the answer is, but also to indicate how confident they are about their answer. This is normally done by asking them to provide a confidence interval. If their estimate is wrong, the confidence interval indicates the range in which they are very sure the true estimate lies.

This approach is the same as that used with other parameters of the scenario tree that may be uncertain. For instance, a sensitivity estimate from a validation study is based on a sample of animals and therefore has a confidence interval that is related to the sample size. When we explicitly include confidence intervals in our scenario tree model, we accept that the result may be wrong, but we can offer a range in which we are very confident that the correct result lies. Chapter 12 provides a detailed description of how to incorporate uncertainty into a scenario tree model; therefore, for the moment, we just need to ensure that experts provide confidence intervals.

A great deal of research into collecting and using expert opinion has been carried out, but this discussion will only address a small number of common approaches that are generally practical and suitable for eliciting the required parameters for scenario trees.

## Gathering expert opinion
### *Working with experts*
### *Choosing experts*
As discussed previously, the right experts are not **necessarily** scientists and professors, and they are almost certainly not the person who is building the scenario tree. The right experts are individuals who have direct and significant experience in an area related to the specific question being asked, which means that the right experts **will differ depending on the issue at hand**.

Experts' estimates are more likely to be applicable to the population of interest if they have a broad **knowledge** of that population. This means that it is much better to have access to a group of experts, as opposed to just one or two experts. While such a group could comprise as few as five individuals, it could equally comprise hundreds – although there is a point when the exercise stops being the solicitation of expert opinion and **instead becomes** a survey (for example, a survey of farmer behaviour).

### *Interaction between experts*
Group dynamics can play an important role in the estimates provided by experts. Two main approaches are commonly used.

The first involves working with each expert independently. This may be achieved by a one-on-one interview (conducted by telephone or by e-mail), or by asking experts in a group setting to consider the question and provide written answers without conferring with the other experts. The main advantage of using this approach is that individuals' opinions are not influenced by the opinions of others in the group, thus allowing the full range of experience to be captured. As well as avoiding the danger of having one or two dominant personalities in a group exerting too much influence on the opinions of others, the 'independent' approach provides one response per expert, which can be used to assess the variation in responses, and thus as a measure of uncertainty.

The other option is to work in a group setting. The group is asked to discuss the question together. At the end of the discussion, you may ask the group to produce a single estimate based on consensus (with a confidence interval), or ask each expert to record their own estimate, in the light of the group discussions. Some of the advantages of the group approach include:
- Discussion **that takes place in a** group setting ensures that there is a common understanding of the question at hand, whereas when questions are answered by individuals separately, there is a risk that some respondents will interpret the question slightly differently.
- Discussion also often prompts the memory of others within the group, and allows the question to be considered from a variety of points of view.
- It is an efficient way to gather estimates rapidly.

The best approach may vary for different situations and different expert groups.

### Questions

The way in which a question is asked has an important impact on the answers that may be received.

### Ensuring that the question is understood

It is very important that the experts fully understand the question that is being asked. This is not always as simple as it seems. For instance, when asking experts to estimate relative risk, it is possible or even likely that not all experts will be familiar with the concept of relative risk. This problem can be addressed in two ways.

First, there may be a need for training and explanation. This is generally a good idea in any case, but it is even more important when the questions **incorporate** technical concepts. Relative risk is a common concept among epidemiologists, but it may be poorly understood by many others involved in animal health. A brief explanation of what a relative risk is, how it is calculated, and how it is used will improve the quality of responses. However, it may also be valuable to provide a list of examples of relative risks for known risk factors, which may be used as a comparison. Those who are inexperienced in the calculation of relative risks may think that values of 10 or 50 appear reasonable, without realizing that most risk factors for many diseases have relative risks which are much lower than these (often in the range of 1.2 to 3).

The second approach is to ask questions in terms that are already understood by the experts. If the experts are not familiar with relative risks, then ask them for estimates that would enable a relative risk to be calculated. For instance:

"Imagine two groups, each comprising 100 animals. Group 1 has the risk factor, and Group 2 does not have the risk factor. If the disease is present in the area, how many animals in Group 1 would you expect to have the disease? How many animals in Group 2 would you expect to have the disease?"

The two prevalence estimates from the previous question can be used to calculate an estimate of the relative risk.
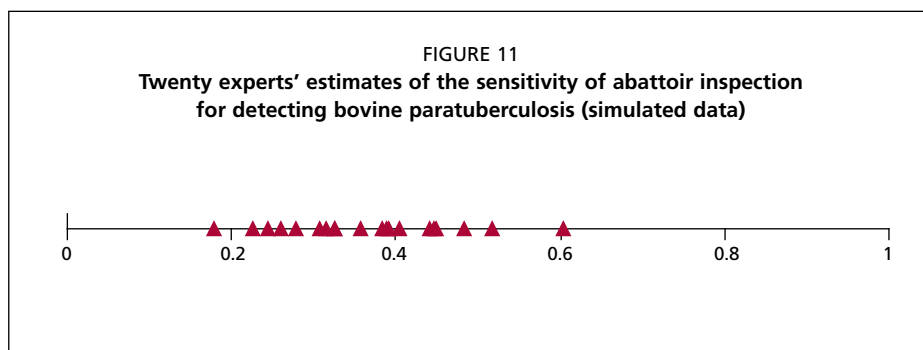
### Confidence intervals

For each question, it is important to capture the experts' uncertainty. This is normally done in two stages:

1. "What do you think is the correct value?"
2. "If you are wrong, what is the lowest possible value that could be correct, and what is the highest?"

This approach will provide a range in terms of the minimum and maximum possible values, with the most likely value somewhere in between. This is the most common approach used in expert opinion. In statistics, a 95 percent confidence interval is usually used, but this tends to be more difficult for many experts to imagine and to estimate.

## Combining expert opinion

When a group approach is used to collect expert opinion, and the group is able to provide a single consensus estimate of the value (and of the confidence interval), these values can be used directly in the scenario tree. However, when experts provide values independently,

FIGURE 11
**Twenty experts' estimates of the sensitivity of abattoir inspection
for detecting bovine paratuberculosis (simulated data)**

there will be a number of different estimates and confidence intervals. We need an approach to combine or summarize these, in order to determine what should be used in the model.

### Uncertainty and variability

There are two reasons why an expert may not be able to give a precise single estimate of a value (for instance, a prevalence).

- They may not know the correct answer (they are uncertain).
- There may not be a single correct answer, as the prevalence may be different in different situations (there is variability in the answer).

The approach to combining estimates from different experts differs depending on whether the main reason for the confidence interval is uncertainty or whether it is variability. Of course, in many cases, both will be present.
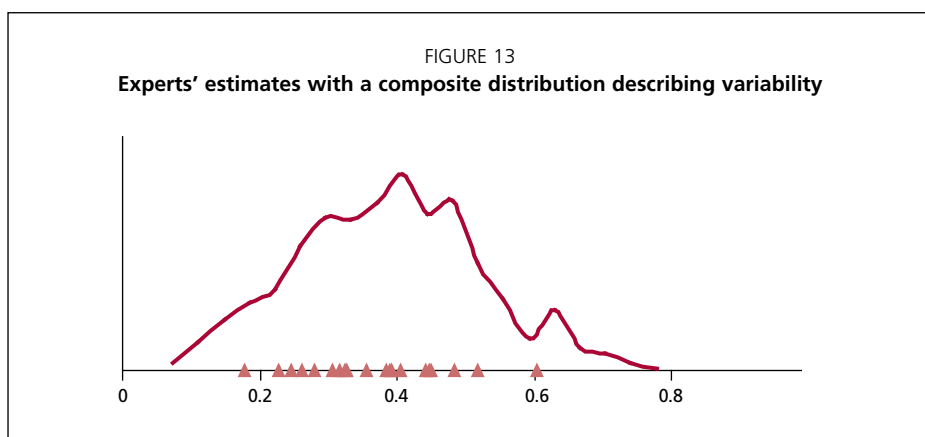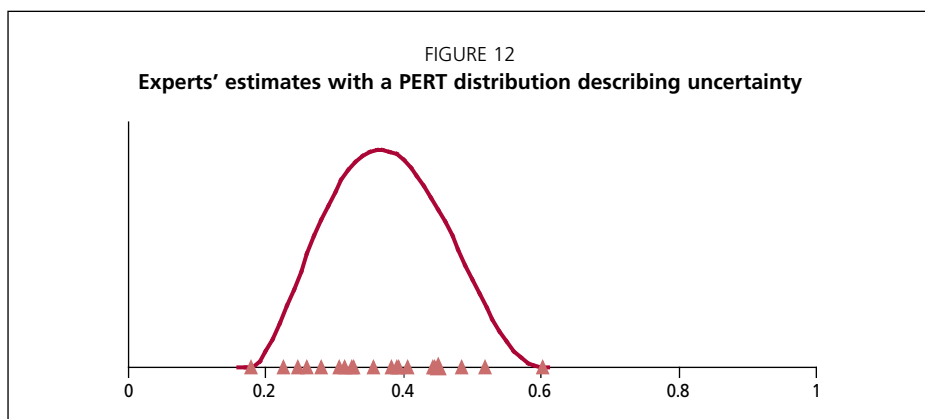
### Uncertainty

When the variability in estimates is due to uncertainty, it implies that there is a single correct answer. Estimates from experts can be thought of as samples from a population of experts, each with some random error, but distributed around the true value.

Figure 11 shows an example of the results that 20 experts may provide, estimating the sensitivity of abattoir meat inspection for identifying paratuberculosis.

If it is considered that there is one true value for the sensitivity, and it is assumed that the estimates of the experts are not biased, then we could summarize the results by taking the average of the estimates as an estimate of the true value. In this case, the average is 0.37.

There are two methods that we could use to describe the uncertainty. The first is to consider that any of the estimates could be correct, and to use the lowest and highest estimates as the minimum and maximum possible values. This is the most conservative approach and would provide the widest range. The confidence interval is shown on page 114. This is based on a beta-PERT probability distribution, as discussed in the next chapter.

The second method would be to consider the confidence intervals provided by the experts, and to take the mean of the lower ends and the mean of the upper ends of these intervals. Depending on the width of the experts' separate confidence intervals, this may produce an overall confidence interval that is wider or narrower than the one shown on page 114.

FIGURE 12
**Experts' estimates with a PERT distribution describing uncertainty**



FIGURE 13
**Experts' estimates with a composite distribution describing variability**



*Variability*

If differences in expert opinion are considered to represent varying correct values for the parameter under different conditions, then the summary of the experts' views should retain these differences. Using the average is no longer appropriate. A common approach is to consider each expert's estimate and confidence interval as a distribution, and build up a composite distribution based on the views of all the experts. The details of how this is done will be discussed in the next chapter, but Figure 13 illustrates an example of the output. Each expert's opinion has contributed to the shape of the final curve.

## Rinderpest example

In a scenario tree for analysing livestock and wildlife surveillance for rinderpest, vaccination was considered an important factor. Vaccination had ceased at different times in the areas being considered, but, for the purpose of this example, the last vaccination was administered six years before the surveillance. Animals less than six years old were certain not to have been vaccinated, while animals older than six years may have been vaccinated.

TABLE 14

| Years | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cattle | 0.7 | 0.69 | 0.65 | 0.5 | 0.48 | 0.46 | 0.44 | 0.4 | 0.38 | 0.35 | 0.3 | 0.23 | 0.2 | 0.16 | 0.05 |
| Sheep | 0.65 | 0.4 | 0.2 | 0.18 | 0.1 | 0.04 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Goats | 0.65 | 0.4 | 0.2 | 0.18 | 0.1 | 0.04 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Buffaloes | 0.85 | 0.8 | 0.83 | 0.8 | 0.75 | 0.6 | 0.57 | 0.45 | 0.4 | 0.34 | 0.33 | 0.29 | 0.25 | 0.2 | 0.15 |
| Warthogs | 0.6 | 0.4 | 0.2 | 0.18 | 0.1 | 0.04 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Kudus | 0.55 | 0.5 | 0.4 | 0.18 | 0.1 | 0.04 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Giraffes | 0.88 | 0.85 | 0.83 | 0.8 | 0.75 | 0.6 | 0.57 | 0.45 | 0.4 | 0.34 | 0.33 | 0.29 | 0.25 | 0.2 | 0.15 |
| Eland | 0.88 | 0.85 | 0.83 | 0.8 | 0.75 | 0.6 | 0.57 | 0.45 | 0.4 | 0.34 | 0.33 | 0.29 | 0.25 | 0.2 | 0.15 |
| Gerenuks | 0.65 | 0.59 | 0.4 | 0.35 | 0.3 | 0.21 | 0.12 | 0.07 | 0.04 | 0.02 | 0 | 0 | 0 | 0 | 0 |
| Camels | 0.88 | 0.85 | 0.83 | 0.8 | 0.75 | 0.6 | 0.57 | 0.45 | 0.4 | 0.34 | 0.33 | 0.29 | 0.25 | 0.2 | 0.15 |

Age was used as a risk category node with two branches: less than six years, and greater than or equal to six years. The surveillance targeted younger animals; therefore, it was necessary to estimate the population proportion and the SSC proportion of animals less than six years old.

After discussion with local experts (13 field veterinarians with extensive experience of each of the relevant species), it was agreed that it would be very difficult to directly estimate the proportion of each species less than six years old. Instead, an indirect approach was used.

Experts were asked to describe the age structure of each species in terms of a survival curve. For each one-year age bracket, they were asked to estimate the proportion of animals born that survived to that age group. Table 14 is an example of the information gathered from one expert.

As it was assumed that there was a single correct survival curve for each species, the estimates from each expert were averaged. This produced the following survival curves.

The uncertainty around each of these curves was calculated based on the standard deviation of the estimates. The proportion of animals less than six years old is the area under the curve left of the six-year point on the x-axis, as a proportion of the total area under the curve for that species. Figure 15 on page 116 shows the estimated proportions.

This example illustrates how it may be possible to gather the required estimates for a scenario tree by asking questions in a form that is easier for the experts to understand. Although the true survival curves for the different species in the study area are not known, the results in Figure 14 are certainly biologically believable and are consistent with expectations.

FIGURE 14
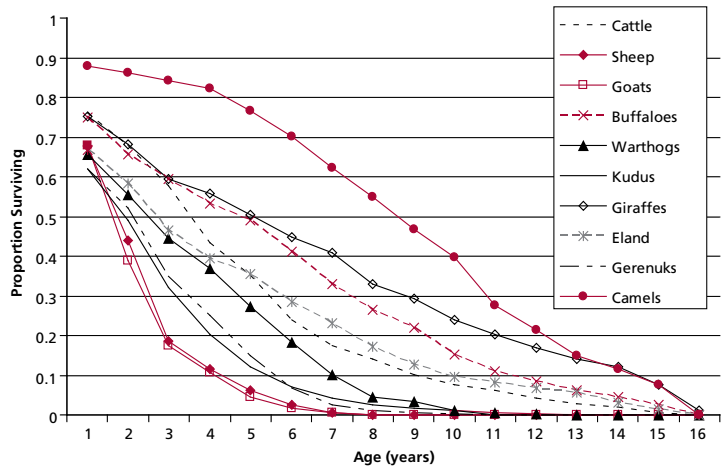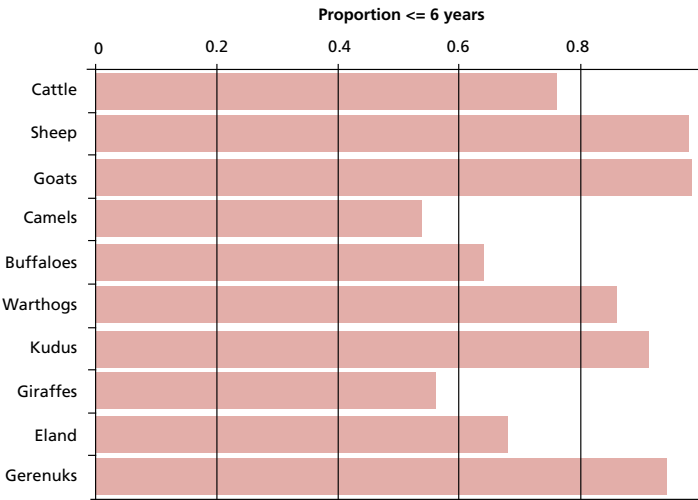**Estimated survival curves based on the means of 13 experts' estimates**



FIGURE 15
**Estimated proportion of animals of different species less than six year of age, based on expert opinion**

## Chapter 12
# Incorporating uncertainty

*"Do not expect to arrive at certainty in every subject which you pursue.*
*There are a hundred things wherein we mortals...*
*must be content with probability, where our best light*
*and reasoning will reach no farther."*
Isaac Watts (1674–1748)

### CAPTURING UNCERTAINTY AND VARIABILITY IN A MODEL

A model is a simplification of reality. Models are never completely perfect, and the objective is to create a model that provides answers which are good enough for effective decision-making. If a model is too simple, it may miss important factors and **therefore may not be** capable of providing useful information. If it is too complex, it may be too difficult to find all the parameters required.

Scenario trees are models of a surveillance system component. They attempt to capture the effect of all the major factors involved in the distribution of infection and the operation of the surveillance system, but they cannot and should not include every small detail. As discussed in the previous chapter, it is often not possible to find the exact information for a parameter, as the data are not available. Instead, it is necessary to estimate the value required, recognizing that it may be incorrect. If it is incorrect and we pretend that it is correct, then the model results will be wrong. However, if it is incorrect and we use confidence intervals to describe how close to the right answer we think we are, the model can take this uncertainty into account.

Each time we run the model using a different input, we will get a different output. This example shows how the three different inputs (the bottom, middle and top of our confidence interval) can produce three different results that indicate the bottom, middle and top of the range of possible values for the sensitivity of our surveillance system.

This approach is good when there is just one parameter that has uncertainty, but normally there are many parameters. If two parameters have confidence intervals (the minimum for parameter 1 [$min_1$], the most likely for parameter 1 [$ml_1$], and the maximum for parameter 1 [$max_1$] and the same for parameter 2), we could run the model multiple times to see what results we would get for the following combinations:

- $ml_1$ and $ml_2$
- $min_1$ and $ml_2$
- $max_1$ and $ml_2$
- $ml_1$ and $min_2$
- $ml_1$ and $max_2$
- $max_1$ and $min_2$

### Example

Consider a scenario tree model where all the parameters are known perfectly, except for one: the sensitivity of the laboratory test. A small study has been carried out to estimate the sensitivity; it has produced an estimate of 96.5 percent, but the 95 percent confidence interval is from 81.2 percent to 98.6 percent.

If we analyse the scenario tree using the best estimate (96.5 percent), we will get a result for the sensitivity of our surveillance system component, say 88 percent. This may be close to the right answer, but it could also be wrong. We need to be able to communicate to those who are using the results of our analysis that the answer could be wrong, and describe how wrong it could be. We have a confidence interval for our input, so it would be useful to have a confidence interval for our output as well.

One way to do this would be to run the model again, but this time, instead of using the best estimate, we could use the lower limit of the confidence interval (81.2 percent). On this occasion, our scenario tree gives a different result, 83 percent. We could then run the model a third time, using the upper limit of the confidence interval (98.6 percent) and we would get a another different result (91 percent).

As the number of uncertain parameter **increases, so too** would the number of different combinations, which renders this approach somewhat impractical.

The other problem with this approach is that all values between the minimum and the maximum are not equally likely to be correct. An expert's best estimate, or the point estimate of a survey, indicates the result that is most likely to be correct. The upper and lower limits could each, conceivably, be correct, but they are much less likely to be so. Simply testing the upper and lower limits does not give an indication of which value is most likely to be the correct value.

Instead of just using three points to measure uncertainty, it is more effective to use a probability distribution. A probability distribution describes how likely each value is to be correct, over a given range. In the example from the previous chapter, we summarized expert opinion by using the average value as the most likely correct result, and the lowest and highest of the experts' estimates as the minimum and maximum possible values. This is shown again in Figure 12. The line is a probability distribution which shows that the value of 0.37 is the most likely. Values just above or just below 0.37 are also very likely, but as you get further away, the results are possible, but increasingly less likely.

If we use probability distributions as the input to our scenario tree model, then we can get a probability distribution as the output, describing how likely a range of different values are to be correct.

The example in Chapter 10 showed how to calculate sensitivity of the surveillance system with a scenario tree. At each step, we did calculations with numbers, multiplying them together to give another number as a result. If we want to use a distribution as the input to the scenario tree, instead of using a number, we need to use a tool that lets us do the same calculations on distributions. The tool is known as stochastic modelling.

FIGURE 16
**Expert estimates and a PERT probability distribution**
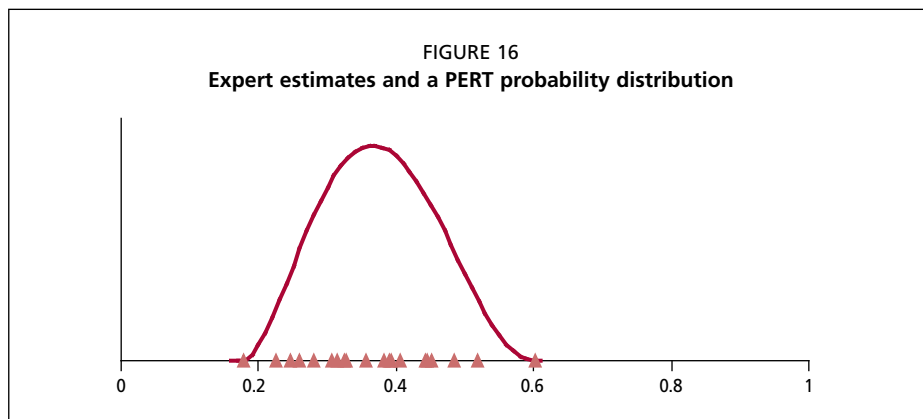
## STOCHASTIC MODELLING

The principle behind stochastic modelling is very simple. Just as we described earlier, if we run the scenario tree model using different inputs, we will get different results.

Stochastic modelling uses computer software to analyse the model repeatedly, each time using different inputs and recording the result of each analysis. Typically, a model may be run 1 000 or 10 000 times, and each time the answer will be different.

But what values are used for the inputs? In our example we chose the minimum, most likely and maximum value from the distribution. This gave us the limits, but it wasn't able to show which values were more common and which were less common. In stochastic modelling, the input values are chosen *at random* from the input distributions. This approach gives the technique its other common name – Monte Carlo simulation, named after the famous casinos in Monte Carlo, where all activities are based on random chance.

The steps in running a stochastic model are:

1. Build your scenario tree.
2. Describe every parameter for which there is uncertainty or variability in terms of a distribution.
3. Tell the software to analyse the model for a set number of times (or iterations).
4. For each iteration, the model will randomly select a single number from each of the parameter distributions. For any other values in the model that use a fixed number, this number will be used for every iteration.
5. The result of each iteration is stored.
6. When the iterations are finished, you can use the output values to draw a histogram that describes the output distribution.

### Describing distributions

Probability distributions are an important part of a stochastic model because they allow inputs to be described not as points, but as ranges of possible values, each with a specified likelihood of being correct. Rather than specifying the probability of every value in the range, distributions are usually described in terms of a number of parameters.

*Normal*

The most well known is the normal distribution, which is described in terms of the mean and the standard deviation, as shown in Figure 17.

The normal distribution is used to describe many biological measurements, such as weight or production. It is rarely used in scenario tree models.

*Beta*

A more common distribution for scenario tree models is the beta distribution. This is described by two parameters, alpha and beta, and is bounded in the range from zero to one. It is therefore very useful for modelling probabilities and proportions, such as prevalence, population proportions and sensitivity. An example of the beta distribution is shown in Figure 18.

A beta distribution is also extremely useful for representing uncertainty about proportions generated from count data. The alpha and beta parameters can be calculated directly from the data:

alpha = x+1,

beta = n-x+1

where

- x is the number of successes and
- n is the sample size.

They can also be calculated from the mode of the distribution, showing the highest point or most likely value, and a percentile (such as the 95 percent percentile) of the distribution, indicating the spread.
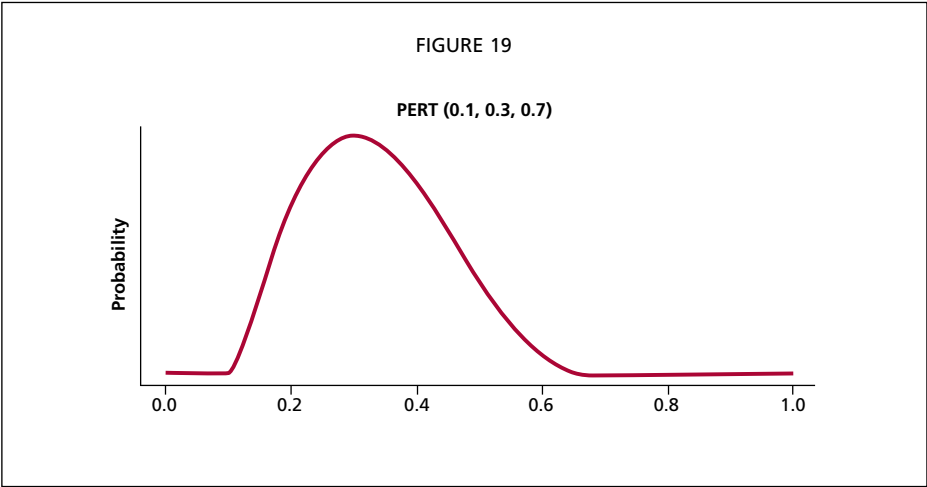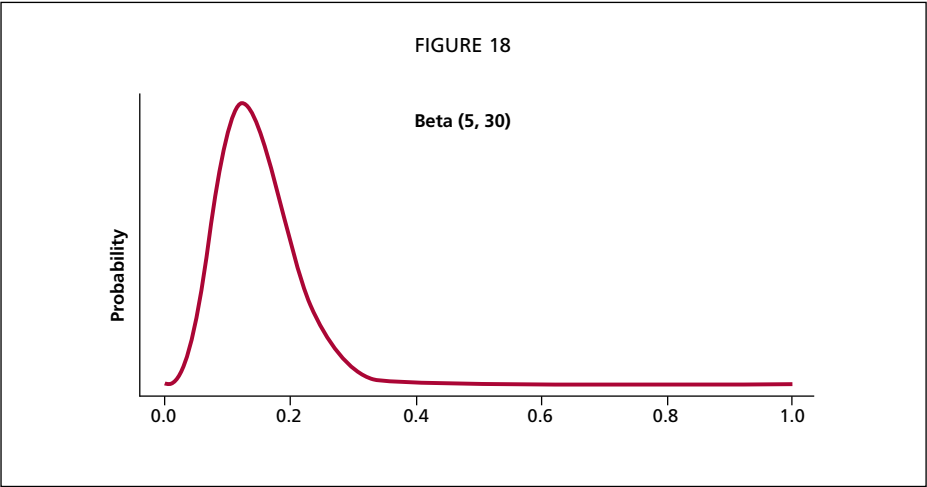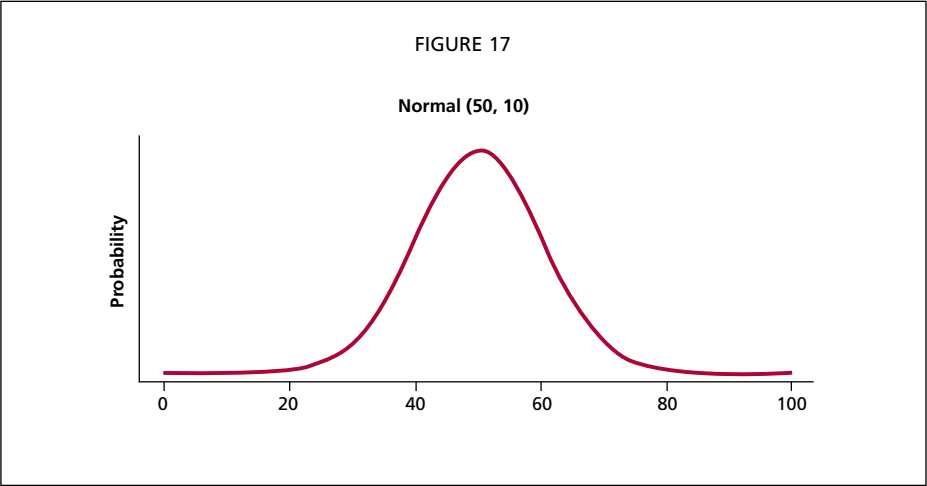
*PERT*

A special form of the beta distribution has been developed to deal specifically with expert opinion. This is called the beta-PERT distribution, and it has the parameters: minimum, most likely, and maximum. This is the most commonly used distribution in scenario tree modelling when expert opinion is used as shown in Figure 19.

*Other distributions*

A wide range of other distributions are available for particular purposes. **These distributions are as follows**:

- Binomial – for a binary event (such as tossing a coin or becoming infected), this shows the likely number of successes in a given number of events.
- Discrete – this is an arbitrary distribution normally described by a dataset or a histogram summarizing a dataset.
- Hypergeometric – this is used to model sampling from a population without replacement.
- Lognormal – this is a logarithmic transformation of the normal distribution and is used to describe skewed data such as herd or flock size, or the incubation period of disease.
- Triangular – this takes the same three parameters as a PERT distribution but joins the three with straight lines. The PERT is generally thought to provide a more realistic description of probabilities associated with expert opinion, as it provides greater weight on values close to the mode and reduced weight in the tails.

FIGURE 17

**Normal (50, 10)**



FIGURE 18

**Beta (5, 30)**



FIGURE 19

**PERT (0.1, 0.3, 0.7)**

*Rules of thumb*
- If you are dealing with expert opinion, use the PERT distribution.
- If you are dealing with other probabilities (sensitivity, proportions, prevalence), use the beta distribution. There are a number of tools available to calculate the alpha and beta parameters of a beta distribution if you know the mode and the 5th or 95th percentile values (see http://epitools.ausvet.com.au), or they can be calculated from the data, as described on page 121.
- For other distributions, only use them if you know they are appropriate for the data that you are analysing. If in doubt, consult a statistician.

## SOFTWARE FOR STOCHASTIC MODELLING

Stochastic modelling requires specialized software that is able to randomly select values from defined input distributions and analyse the model repeated over many iterations. There are a number of software packages available for this purpose, but only three will be **described** here. The first two are add-ins for Microsoft Excel®, providing new formulae and menu items. The third is the R statistical software which includes probability distributions and other functions required for stochastic simulation modelling.

### Palisade @Risk

@Risk is a well-known, powerful commercial software package available from http://www.palisade.com/risk. It is widely used and is capable of performing all the functions required for scenario tree modelling. It is available for recent MS Windows operating systems and is accompanied by an extensive help system, and therefore it is not described here any further.

### PopTools

PopTools provides an effective, free alternative to @Risk. This software works in a very similar way, as an Excel plug-in with new formulae and menus. It was originally developed for ecological modelling, but it has a range of stochastic modelling tools that are more than capable of supporting the needs of scenario tree modelling. It also has an impressive array of other analytical tools and utilities. It is available for free download from http://www.poptools.org. The package includes a large collection of example spreadsheets, designed to illustrate the different functions; in addition, it has an inbuilt help system. More comprehensive help is available for separate purchase at a nominal fee.

The examples later in this chapter are based on the use of PopTools, but they can be adapted for use with @Risk by making minor modifications.

### R (or other statistical software)

Many statistical software packages, such as R (http://cran.r-project.org/) include probability distribution functions and are well suited to simulation modelling. Because R is a dedicated statistical and programming environment, it has much greater flexibility than Excel add-ins and it can be used for complex models. However, this increased flexibility and power comes at the cost of users having to learn a **relatively** complex language and programming structure. In summary, R is a very powerful and useful package for simulation modelling, but users need programming skills and experience in order to make the most of its capabilities.

TABLE 15

|  | Min | Most likely | Max | Weight |
|---|---|---|---|---|
| **Expert 1** | 0.2 | 0.3 | 0.4 | 5 |
| **Expert 2** | 0.25 | 0.3 | 0.55 | 5 |
| **Expert 3** | 0.65 | 0.75 | 0.9 | 1 |
| **Expert 4** | 0.4 | 0.5 | 0.7 | 3 |
| **Expert 5** | 0.3 | 0.55 | 0.75 | 2 |

## EXAMPLE EXERCISES
### Exercise 1: Combination of expert opinion

In Chapter 11 we discussed the use of expert opinion. This exercise will show how expert opinion can be combined, based on the assumption that variation is primarily caused by variability rather than uncertainty. This is a small stochastic model, but it is not a scenario tree model.

Example: Five different experts have been asked to estimate the probability that a veterinarian would collect specimens from a case showing signs consistent with the disease of interest. Each expert has been asked to provide their most likely estimate, as well as the minimum and maximum values. The data from the experts are shown in Table 15. In addition, the experts have been asked to evaluate their own level of expertise in relation to the question, on a scale of 1 to 5, with 5 indicating a high level of expertise and 1 a very low level. This is shown in the *weight* column.

This data are entered into Excel, as shown in Table 16. Colour coding is used to distinguish the values: **black** for labels, **blue** for input data, **orange** for random variables and **red** for output.

The spreadsheet is now ready to be converted into a stochastic model. The first step is to enter formulae for the random variables. We will use a PERT distribution to describe the estimates for each of the experts. There are two ways to enter the formula:

Using the menu (this is easier when you are not familiar with the formulae)

1. Place the cursor in the cell where the formula should be entered (E2).
2. Click on the **PopTools** menu.
3. Select **Random variable.**
4. In the dialog select PERT as the distribution.
5. Leave Length blank.
6. The output cell should already be set to E2.
7. Set the Min value to B2.
8. Set the Likely value to C2.
9. Set the Max value to D2.
10. Leave the weight as 4.
11. Click Go.

TABLE 16

|   | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 |  | Min | Mode | Max |  | Weight |
| 2 | Expert 1 | 0.2 | 0.3 | 0.4 |  | 5 |
| 3 | Expert 2 | 0.25 | 0.3 | 0.55 |  | 5 |
| 4 | Expert 3 | 0.65 | 0.75 | 0.9 |  | 1 |
| 5 | Expert 4 | 0.4 | 0.5 | 0.7 |  | 3 |
| 6 | Expert 5 | 0.3 | 0.55 | 0.75 |  | 2 |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |

TABLE 17

|   | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 |  | Min | Mode | Max | Random | Weight |
| 2 | Expert 1 | 0.2 | 0.3 | 0.4 | 0.264064 | 5 |
| 3 | Expert 2 | 0.25 | 0.3 | 0.55 | 0.391712 | 5 |
| 4 | Expert 3 | 0.65 | 0.75 | 0.9 | 0.730146 | 1 |
| 5 | Expert 4 | 0.4 | 0.5 | 0.7 | 0.586328 | 3 |
| 6 | Expert 5 | 0.3 | 0.55 | 0.75 | 0.618053 | 2 |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |

Type the formula yourself (this is faster when you are familiar with the formulae).

1. Place the cursor in the cell where the formula should be entered (E2).

2. Type: **=dPertDev(B2,C2,D2,4)**

3. Press enter.

If PopTools is loaded and the formula has been entered correctly, you should now see a number in the cell E2. The number is a random value drawn from the PERT distribution, so it will be different from the numbers in Table 17.

Copy the formula down to cells E3 to E6, so that your spreadsheet looks like the layout in Table 18 on page 125.

The random numbers will be changed every time you recalculate the spreadsheet. This can be easily demonstrated by pressing the F9 key. Each time you press F9 key, the numbers in column E will change, representing new random numbers from the defined PERT distributions.

We have now described the input to our model in terms of distributions, and we have entered the formulae to select random numbers from those distributions to be used for each iteration of the model. The next step is to do the model calculations to produce the output. The method we used to combine expert opinion is to randomly select a value from one expert at each iteration.

TABLE 18

|   | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 |  | **Min** | **Mode** | **Max** | **Random** | **Weight** |
| 2 | **Expert 1** | 0.2 | 0.3 | 0.4 | 0.264064 | 5 |
| 3 | **Expert 2** | 0.25 | 0.3 | 0.55 | 0.391712 | 5 |
| 4 | **Expert 3** | 0.65 | 0.75 | 0.9 | 0.730146 | 1 |
| 5 | **Expert 4** | 0.4 | 0.5 | 0.7 | 0.586328 | 3 |
| 6 | **Expert 5** | 0.3 | 0.55 | 0.75 | 0.618053 | 2 |
| 7 |  |  |  |  |  |  |
| 8 |  |  |  | **Result** | 0.391712 |  |

The experts gave themselves a weight between 1 and 5 to indicate their level of exper-
tise. We will use these weights when we choose which expert's opinion to select. Experts
with a weight of 5 will be five times more likely to be selected than experts with a weight
of 1. In this way, the opinion of our 'strong' experts will contribute more to our output than
will the opinion of our 'weak' experts, but all will have some contribution.

Enter the formula into cell E8 using either the **Random variable** menu option (Discrete
distribution) or type in:

  **=DiscreteDev(E2:E6, F2:F6)**

Note that the parameters are both ranges.

The spreadsheet should now look something like the one Table 18. Press the F9 key
again a few times to see how the numbers change when the sheet is recalculated.

Our model is now complete, with inputs expressed in the form of random variables from
defined distributions, and an output value. The next step is the Monte Carlo simulation –
analysing the model many times and collecting the result of each iteration.

1. Click on the **PopTools** menu, select **Simulation tools** and then select **Monte Carlo
   analysis**.
2. In the Monte Carlo analysis dialog you need to enter details of the simulation:
   a. Dependent range: This is the output cell, E8. You are able to select multiple cells if
      you want to analyse a number of different outputs, but they must all be set out in
      a single column and there must be no blank spaces.
   b. Test values: This is used to test whether your result is greater or less than some fixed
      value. Normally, this is left blank.
   c. Lower percentile and upper percentile: these are used to provide statistics for the
      test values. Again, these can be left blank.
   d. Number of replicates. This determines how many times the model will be run. The
      default is 100, but for simple models, it is normally possible to run 1 000 iterations
      relatively quickly.
   e. Output: This is where a brief summary of the results of the analysis will be pro-
      duced. Specify a blank cell.
   f. Test criterion: These are only used when test values are specified. Normally, they
      can be ignored.

g. Random seed: This can normally be left as zero. If you set a fixed random seed (other than zero), every time you run the model with the same random seed, you will get exactly the same result (as the same sequence of random numbers is used). When zero is applied, the computer's clock is used to generate a new random seed for each analysis.

h. Keep results: This allows you to store the result of each analysis on a new sheet in your spreadsheet. You should always make sure this box is selected, so that you can look at the output distribution.

i. Colour code for demonstration: This is for training purposes only – leave this blank.

j. Click Go and watch the counter indicate the iterations.

When the analysis is completed (it should be very quick for such a simple model), the summary results will be displayed, and a new page, called **Monte Carlo simulation results 1**, will have been added to the spreadsheet. The summary results are not very revealing, but we can obtain more information by analysing the new spreadsheet, as we can use it to see the output distribution.

To show the output distribution:

1. Open the new page, and highlight all the numbers in column B under Var 1.
2. Click on the PopTools menu, then Simulation Tools and then click Summary stats.
3. The input range should already be set to show the column of numbers.
4. The test value can be left blank.
5. For the output range, enter a blank cell.
6. Check the Sort range for percentiles box.
7. Select 20 bins for histogram.
8. Click Go.

PopTools produces some more detailed summary statistics and then draws a histogram of the output results, which should be similar to the one shown on page 127.

This histogram represents the combined estimates of the five experts. Note that it is not a smooth, regular curve. The experts had quite different views, and each of these views is reflected in the output. This approach indicates that the probability that a veterinarian would submit samples varies considerably – in some circumstances it is reasonably high, but most of the time it is quite low.

The great value of this output distribution is that it does not claim that there is a single correct value. It describes the variability, as well as the experts' uncertainty.

## Exercise 2: Analysis of a simple scenario tree

For this exercise, we will use the simple three-node scenario tree that we introduced in Chapter 10. This tree includes one risk category node (age), as young animals are at higher risk than older animals; an infection node, and a detection node. The calculations are the same as those used previously, but for this exercise, we will introduce uncertainty in some of the parameters:

- the relative risk for younger animals compared with older animals;
- the population proportion of younger and older animals;
- the SSC proportion of younger and older animals;
- the sensitivity of the ELISA.

FIGURE 20
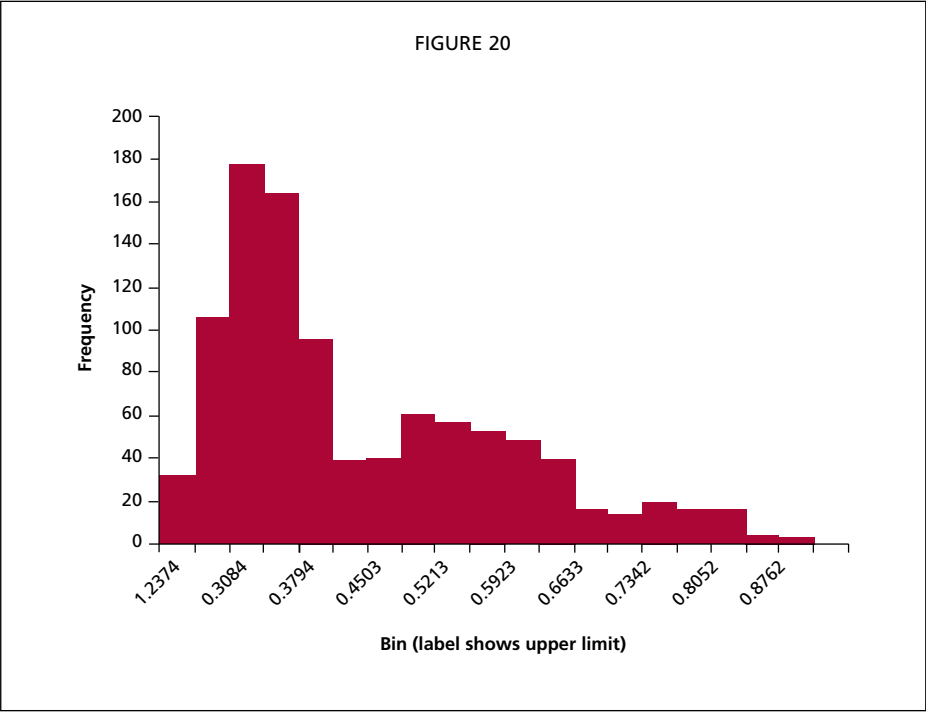
The layout of the spreadsheet is shown on Table 19; it includes model parameters, the scenario tree and the results. For more complex models, these three sections are often divided between three pages. Remember the meaning of the colour coding, which makes it easier to understand the model: **input values, random variables, normal formulae,** and **outputs**.

The figures displayed represent a single iteration of the model. Table 20 on page 129 shows the same spreadsheet layout, but cells containing formulae have been displayed with the formula rather than the result. Most of the formulae use well-known functions but some functions merit special mention. As follows:

- The PopTools functions for a random variable from the PERT distribution in F3, F5, F7 and F12.
- The formula for calculating the adjusted risk in cells C10 and C9.
- The **SUMIF()** function in C30. This calculates the sum of a column of numbers if the value in another column matches certain criteria. The formula used **=SUM-IF(G19:G26,"Pos",H19:H26)** adds each value in H19 to H26 if the corresponding value in G19:G26 is equal to "Pos".

Try setting up this spreadsheet in Excel yourself and make sure that it is working properly.

Once the model has been constructed, it can be analysed stochastically. Use the Pop-Tools, Simulation tools, Monte Carlo analysis menu to run the model 1 000 times, using the values in C30 to C34 as the dependent range, and making sure that you check the 'Keep results' box. This will create a new page of Monte Carlo results, containing one column of results for each of the five output values, and 1 000 rows, one for each iteration.

TABLE 19

### Parameters

| Age | Branch | Min | Most likely | Max | Value |
|---|---|---|---|---|---|
| Relative risk | Young | 1.5 | 5 | 7 | 4.04579626 |
| | Old | | | | 1 |
| Population proportion | Young | 0.18 | 0.2 | 0.24 | 0.19426404 |
| | Old | | | | 0.80573596 |
| SSC proportion | Young | 0.85 | 0.9 | 0.92 | 0.88348941 |
| | Old | | | | 0.11651059 |
| Adjusted risk | Young | 2.541826 | | | |
| | Old | 0.628264 | | | |
| Design prevalence | | 0.05 | | | |
| ELISA | Sensitivity | 0.85 | 0.95 | 0.99 | 0.91843519 |
| Animals in SSC | | 20 | | | |

### Scenario tree

| Age | | Infected | | ELISA result | | | |
|---|---|---|---|---|---|---|---|
| Branch | PrSSC | Branch | EPI | Branch | Se | Outcome | Probability |
| Old | 0.1165 | Yes | 0.0314 | Pos | 0.9184 | Pos | 0.0034 |
| | | | | Neg | 0.0816 | Neg | 0.0003 |
| | | No | 0.9686 | Pos | 0 | Pos | 0.0000 |
| | | | | Neg | 1.0000 | Neg | 0.1129 |
| Young | 0.8835 | Yes | 0.1271 | Pos | 0.9184 | Pos | 0.1031 |
| | | | | Neg | 0.0816 | Neg | 0.0092 |
| | | No | 0.8729 | Pos | 0 | Pos | 0.0000 |
| | | | | Neg | 1.0000 | Neg | 0.7712 |
| | | | | | | Check Sum: | 1.0000 |

### Results

| | |
|---|---|
| Unit sensitivity (actual) | 0.106487 |
| Unit sensitivity (representative) | 0.045922 |
| Component sensitivity (actual) | 0.894799 |
| Component sensitivity (representative) | 0.609447 |
| Sensitivity ratio | 1.468215 |

TABLE 20

|  | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | **Parameters** | | | | | | | |
| 2 | **Age** | **Branch** | **Min** | **Most likely** | **Max** | **Value** | | |
| 3 | Relative risk | Young | 1.5 | 5 | 7 | =dPertDev(C3,D3,E3,4) | | |
| 4 | | Old | | | | 1 | | |
| 5 | Population proportion | Young | 0.18 | 0.2 | 0.24 | =dPertDev(C5,D5,E5,4) | | |
| 6 | | Old | | | | =1-F5 | | |
| 7 | SSC proportion | Young | 0.85 | 0.9 | 0.92 | =dPertDev(C7,D7,E7,4) | | |
| 8 | | Old | | | | =1-F7 | | |
| 9 | Adjusted risk | Young | =F3*C10 | | | | | |
| 10 | | Old | =1/((F3*F5)+(F4*F6)) | | | | | |
| 11 | Design prevalence | | 0.05 | | | | | |
| 12 | ELISA | Sensitivity | 0.85 | 0.95 | 0.99 | =dPertDev(C12,D12,E12,4) | | |
| 13 | Animals in SSC | | 20 | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |
| 16 | **Scenario tree** | | | | | | | |
| 17 | **Age** | | **Infected** | | **ELISA result** | | | |
| 18 | **Branch** | **PrSSC** | **Branch** | **EPI** | **Branch** | **Se** | **Outcome** | **Probability** |
| 19 | Old | =F8 | Yes | =C11*C10 | Pos | =$F$12 | Pos | =F19*D19*B19 |
| 20 | | | | | Neg | =1-F19 | Neg | =F20*D19*B19 |
| 21 | | | No | =1-D19 | Pos | 0 | Pos | =F21*D21*B19 |
| 22 | | | | | Neg | =1-F21 | Neg | =F22*D21*B19 |
| 23 | Young | =F7 | Yes | =C11*C9 | Pos | =$F$12 | Pos | =F23*D23*B23 |
| 24 | | | | | Neg | =1-F23 | Neg | =F24*D23*B23 |
| 25 | | | No | =1-D23 | Pos | 0 | Pos | =F25*D25*B23 |
| 26 | | | | | Neg | =1-F25 | Neg | =F26*D25*B23 |
| 27 | | | | | | | Check Sum: | =SUM(H19:H26) |
| 28 | | | | | | | | |
| 29 | **Results** | | | | | | | |
| 30 | Unit sensitivity (actual) | | =SUMIF(G19:G26,"Pos",H19:H26) | | | | | |
| 31 | Unit sensitivity (representative) | | =C11*F12 | | | | | |
| 32 | Component sensitivity (actual) | | =1-(1-C30)^C13 | | | | | |
| 33 | Component sensitivity (representative) | | =1-(1-C31)^C13 | | | | | |
| 34 | Sensitivity ratio | | =C32/C33 | | | | | |

FIGURE 21
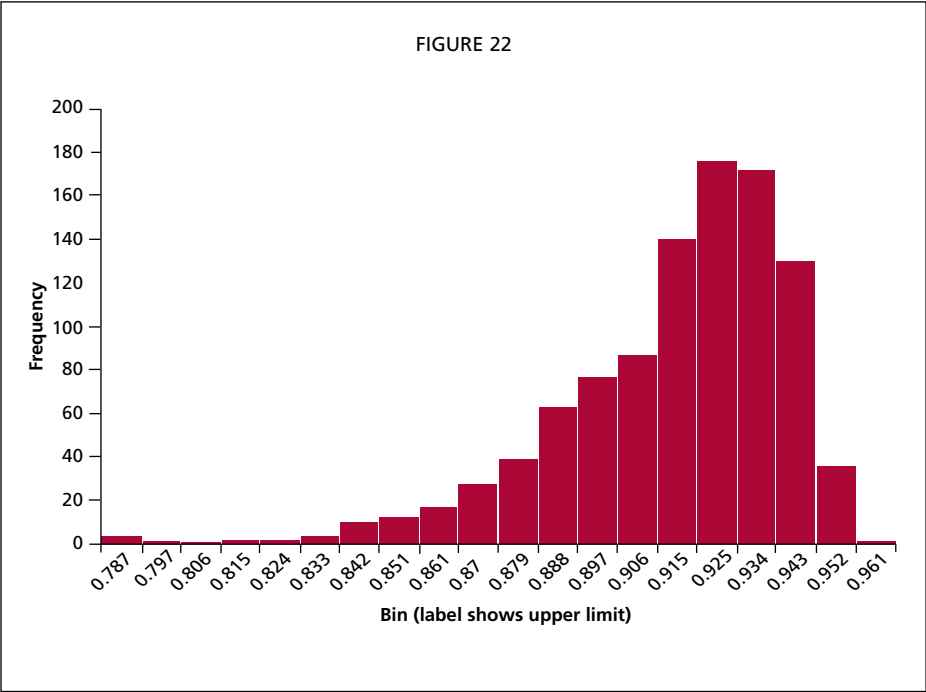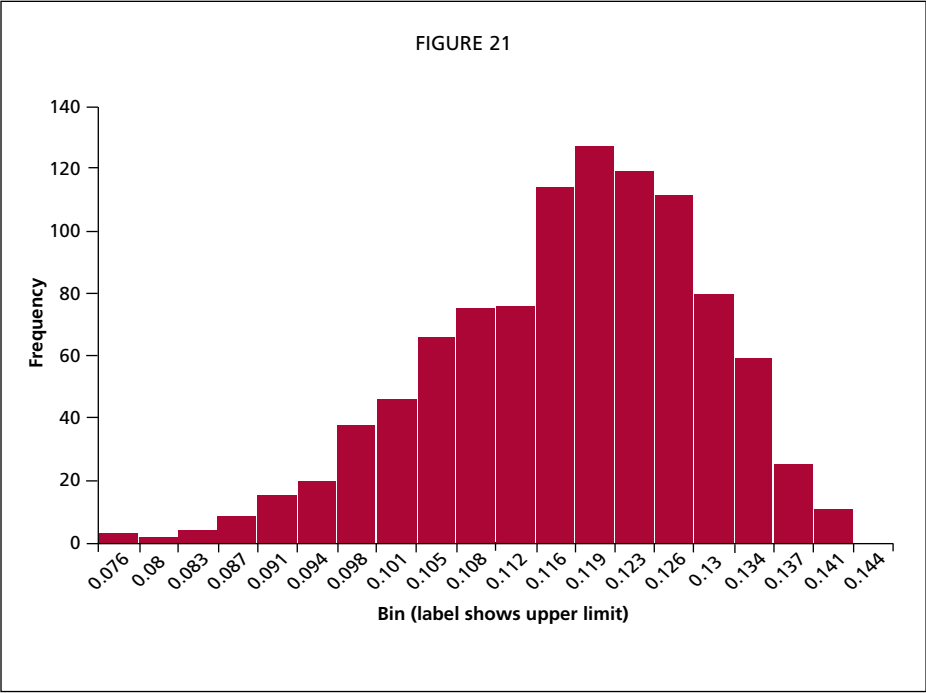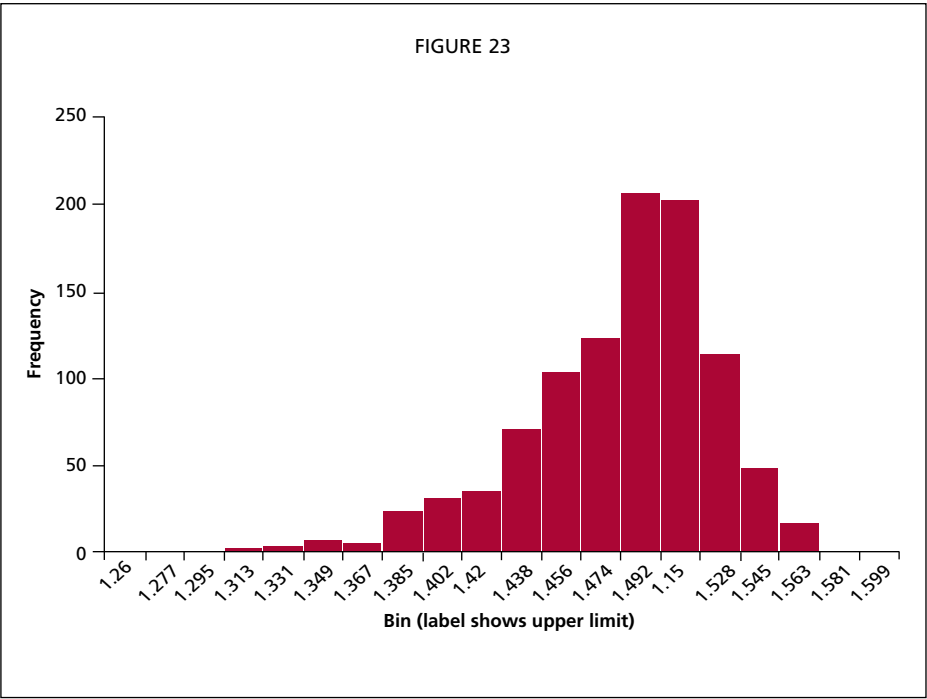


FIGURE 22

FIGURE 23

Use the PopTools, Simulation Tools and Summary stats tool to summarize each of these columns and **create a** graph the output. You should get a result something similar to the histograms shown in Figures 21 to 23.

The first histogram (Figure 21) shows the distribution of the unit sensitivity, i.e. the probability of detecting disease if just a single animal were in our surveillance system component. It ranges between 7 percent and 14 percent, with a mode of around 12 percent.

Now let us consider at the third variable: the surveillance system component sensitivity. This is normally the answer that we are most interested in. It shows that the sensitivity is about 92 percent, but most of the values are in the range from 84 percent to 95 percent.

The fifth value has not yet been discussed in detail. It is often interesting to know how well surveillance is targeted. This can be done by comparing the sensitivity of the actual surveillance with a hypothetical surveillance system based on purely representative sampling. This value is known as the sensitivity ratio (*SR*)

$$\text{Sensitivity Ratio} = \frac{\text{CSe}_{\text{Actual}}}{\text{CSe}_{\text{Representative}}}$$

If the *SR* is equal to one, it means that the surveillance is as efficient as representative (e.g. random) sampling (i.e it is well targeted). If the SR is greater than one, the surveillance is more efficient than representative sampling. If the SR is less than one, the surveillance is less efficient because it is poorly targeted or negatively biased.

The distribution of the *SR* in our model is shown in the Figure 23 on page 131. It is about 1.50, ranging from 1.3 to 1.56. This means that the sensitivity of the surveillance is about 1.5 times greater than representative surveillance using the same number of animals.

This example has been simplified so as to make it easier to understand. In reality, there are usually more risk factors to be taken into account, and certainly more follow-up tests would be required before one could conclude that the country or zone is infected. These tests normally result in indicating a lower unit sensitivity. In addition, most surveillance activities focus on many more than 20 animals, which means that the SSC sensitivity is often much higher.

## POPTOOLS REFERENCE
These examples show that it can be quite simple to design a stochastic model in a spreadsheet using PopTools, although larger models rapidly become more complex. This section provides some brief notes on using PopTools. See the help system and example spreadsheets for more information.

### Installation
To install PopTools, download the executable installation file from the PopTools web site at http://www.poptools.org/download and save it to your hard disk. Double click to start the installation process. In Windows Vista you may need to right click on the file and select 'Run as Administrator' for successful installation.

Once installed, Excel will open with a spreadsheet containing a welcome message. To confirm that the system has been correctly installed, check to see if there is a new menu item: PopTools.

### Random variable functions
Some of the more commonly used distributions available in PopTools are listed in Table 21.

Other less common distributions are also available including:
- Cauchy
- Correlated random variable
- Gamma
- Geometric
- Negative binomial
- Normal (integer)
- Pareto
- Weibull

TABLE 21

| Function | Parameters | Distribution and notes |
|---|---|---|
| dBetaDev | Alpha, beta | Beta distribution (defined by alpha and beta) |
| dBetaMSDev | Mean, standard deviation | Beta distribution (defined by the mean and standard deviation) |
| dBinomialDev | Trials, probability | Binomial |
| dExpDev | Mean | Exponential |
| dHyperDev | Samples, affected, population | Hypergeometric |
| dLogNormalDev | Mean, standard deviation | Log normal |
| dNormalDev | Mean, standard deviation | Normal |
| dPertDev | Min, most likely, max, weight | PERT. Weight should always be set to 4 for consistency with other implementations of the PERT |
| dPoissonDev | Mean | Poisson |
| dRandInt | Lower, upper | A uniform random integer between the lower and upper bounds |
| dRandReal | Lower, upper | A uniform random real number between the lower and upper bounds |
| dTRand | Min, most likely, max | Triangular |
| DiscreteDev | Numbers, frequencies | Discrete distribution – selects a number at random from the list of numbers with a probability proportional to the frequencies |

## Other useful functions

In addition to the random variable functions, PopTools makes a large number of other specialized functions available. Many of these are designed to assist with matrix operations or statistical analysis. Some of the more useful general functions include:

- **FormulaText(ref)** – returns the formula in the referenced cell, displayed as text. This is useful for displaying how a spreadsheet works (it was used to produce the second diagram of the spreadsheet in Table 21), but it can also be useful for documenting a spreadsheet.
- **F(value)** – displays the value if it is not an error, otherwise it displays a blank cell. This is useful for hiding errors.
- **QSort(range)** – sorts a range of data. This is an array formula. Normal formulae only work on a single cell, but array formulae work on an array of cells. To enter an array formula:
  - Select a range of cells that the formula will occupy.
  - Type the formula.
  - Instead of pressing enter, press Shift-Control-Enter together.
- **RandFix(true/false)** – When this is true, every random formula returns the expected value rather than a random value. This stops the model from behaving stochastically, and can be useful when checking for errors. This can also be achieved from the menus (PopTools, Fix random generator)

## Selected menus and dialogs

PopTools contains many features that are not directly relevant to scenario tree modelling. Some of these features may be useful when calculating probability inputs for a model, or for other related purposes. This section of the manual deals with the features that are most likely to be of value.

### *Simulation tools menu*

We have already used two of the items here – Monte Carlo simulation and summary statistics.

### *Monte Carlo simulation*



Inputs for the dialog have already been described on page 127.

*Summary statistics*
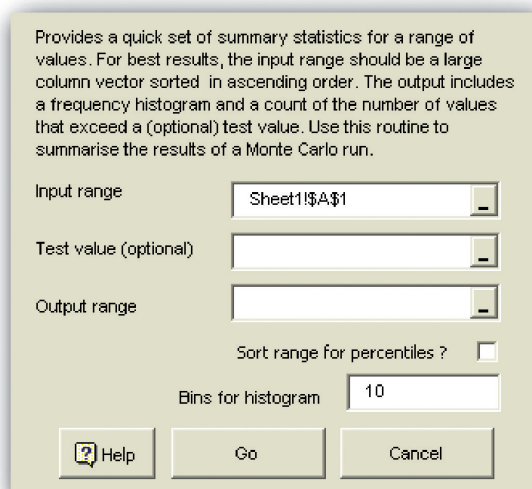


Inputs for the summary statistics dialog have been described on pages 126 and 127.

### Extra stats menu

PopTools provides access to a range of statistical tools that are not readily available in Excel. The most commonly used ones for those working in disease surveillance are likely to be:



- ANOVA: Analysis of variance
- Chi square
- Regression
- Relative risk
- Odds ratio

### Sampling menu

This menu has tools for simulation of sampling and sample size calculation.



- Simple random sampling: this selects a simple random sample of values from a range of data in the spreadsheet.
- PPS sample: this selects a sample of data from the spreadsheet using probability proportional to size sampling.
- Cross-sectional sample size: calculates the sample size for a prevalence survey.
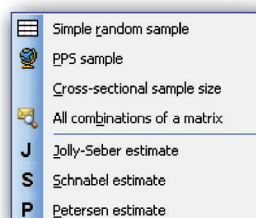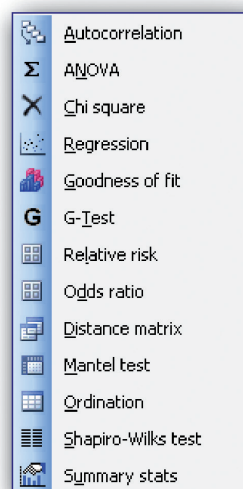
***Auditing menu***

This contains a very useful list of functions for checking a spreadsheet and understanding how values are related and calculated.

***Random variable***

This dialog is a convenient way to select a random variable with a specified distribution for use in a stochastic model.

An example of entering data into this dialog was provided on pages 123 and 124.
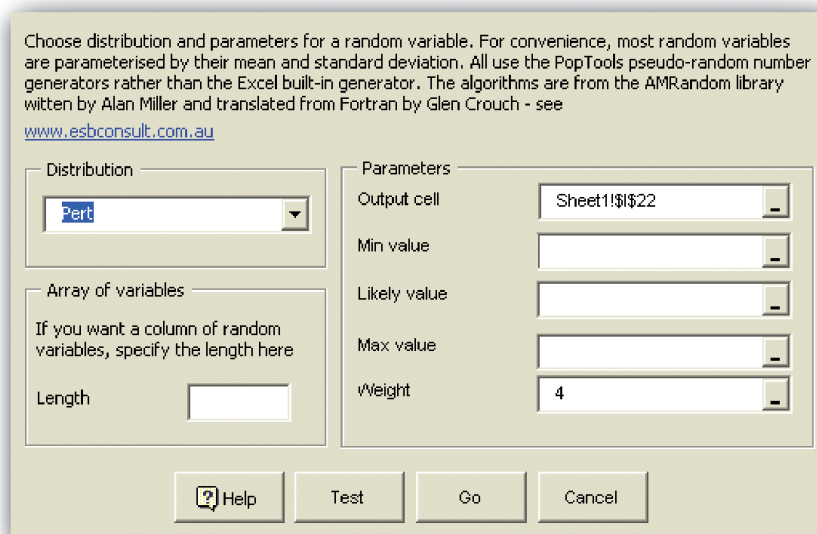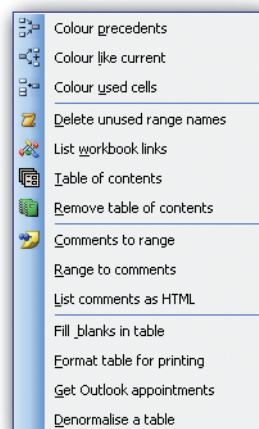
# Chapter 13
# Clustering

*"Freedom and independence form my character."*
**Mustafa Kemal Ataturk (1881–1938)**

## CLUSTERING OF DISEASE AND POPULATIONS

Directly transmitted diseases usually form clusters. If an infected animal is brought into a population, other animals in the same herd are more likely to become infected than animals in other herds. Livestock movements or other factors may move the disease to other herds, and animals in those new herds will become infected, thus generating new clusters.

While most of the major livestock diseases are directly transmitted **between animals within a herd**, some diseases do not cluster at the herd level. Vector-borne diseases such as Bluetongue are not constrained by herd structures and fences, but are distributed wherever the vector is able to find a suitable habitat.

### Lack of independence between animals

The examples in previous chapters have shown how to calculate the unit sensitivity (*CSeU*, probability of detecting disease with a single animal in the surveillance system). This is then used to calculate the SSC sensitivity using the formula:

$$CSe = 1 - (1 - CSeU)^n$$

This approach assumes that animals are independent, which means that the probability of one animal being infected is not related to the probability that another animal in the same herd is infected. However, this is clearly not always true. When animals are grouped into herds, and disease clusters, if one animal is infected, there are likely to be others that are also infected. The infection status of animals is therefore unlikely to be independent.

As the number of animals that test negative increases, the amount of new information that each new animal reduces **significantly**, because we are already reasonably sure of the status of the herd.

In probability terms, the independent probability that an animal will test positive (when we have no prior test results from the herd) is different to the conditional probability that an animal will test positive, given that other animals in the same herd have already tested negative.

The formula we used to calculate the SSC sensitivity:

$$CSe = 1 - (1 - CSeU)^n$$

assumes that the new information provided by each animal is the same. However, we have just seen that if the animals come from the same herd, subsequent animals provide less new evidence. If they come from different herds, they provide more evidence.

**Example**

Consider one herd of 40 animals examined as part of a surveillance programme that uses an animal-level design prevalence of 20 percent and a herd-level design prevalence of 5 percent. Before any surveillance is carried out in the herd, we do not know if the herd is infected or not, but the herd-level design prevalence tells us that the probability that the herd is infected is 5 percent.

If we sample one animal and it tests negative, does this change the probability that the herd is infected?

The answer to this is 'yes', because after one negative test, we still do not know the herd status. However, the herd is less likely to be infected. If we then test a second and a third animal, each time we get a negative result, we become more and more confident that the herd is uninfected.

After testing 20 animals from a herd of 40 animals, and all have tested negative, we can already be very confident that the disease is not present. Testing one more animal makes us more confident, but does not provide as much new information as the first animal tested, because we are already very confident.

For diseases that cluster, we need to find a different approach to calculating the SSC sensitivity, which is able to take the lack of independence between animals in the same herd into account.

## Step-wise calculation of sensitivity

We achieve this by calculating the sensitivity of the SSC in steps. First, we calculate the sensitivity for each herd. Then, based on the herd-level sensitivity, we can calculate the SSC sensitivity. By treating each herd separately, we are able to take into account clustering and lack of independence between animals at the herd level.

To do this, we must know something about the herds. In the previous calculations based on the assumption of independence, the only figure that was used was *n*, the total number of animals in the SSC. When we take lack of independence into account, we should ideally know:

- which herds in the population are part of the SSC;
- which animals in the SSC belong to which herd;
- the size of each herd;
- the risk characteristics of each of the animals;
- the risk characteristics of each of the herds.

This allows us to accurately calculate the herd-level sensitivity of every herd in the SSC. Where all these details are not known, a general description of the population (estimated number of herds, distribution of herd sizes, and estimated number of animals tested per herd) can be used to estimate the separate herd-level sensitivity for each herd in the SSC.

If you are using the web-based Freedom software, the details of these calculations will be handled automatically. If you are implementing the analysis in a spreadsheet, you should read the next section in order to better understand the approach to calculation of herd sensitivity.

TABLE 22

| Node | Type | Branches |
|------|------|----------|
| Adjacent to wetlands | Risk category | Yes, No |
| Flock type | Risk category | Commercial, Backyard |
| Flock infected | Infection | Infected, uninfected |
| Animal infected | Infection | Infected, uninfected |
| Initial test | Detection | Test positive, negative |
| Confirmatory test | Detection | Test positive, negative |

## HERD-LEVEL SENSITIVITY CALCULATION

The examples of scenario trees shown previously are unable to take into account the lack of independence between animals within herds. A different approach is required in order to analyse information at the herd level.

### SPREADSHEET LAYOUT EXAMPLE

Let us look at an example of how a spreadsheet could be organized to help with these herd-level calculations. Consider a surveillance system component for avian influenza in domestic chickens. A node list for a simplified scenario tree is provided in Table 22.

As with the earlier spreadsheet example, it helps to lay out the parameters clearly, so that you can access them for later calculations. For simplicity, this example will not be a stochastic model. The parameters are shown in Table 23.

In addition to the calculated figures for adjusted risk (*AR*, discussed in Chapter 9), two other figures have been calculated for convenience:

- the combined sensitivity of the two tests used, ($Se_1 \times Se_2$ or, in the spreadsheet **=B9 * B10)** and
- the probability that an animal will provide a positive test result ($P^*_A \times Se$, or **=B8 * B9 * B10)**

Instead of the model being displayed as a tree, we can use a table with one row per flock. The information on the line should include everything we need to calculate the flock sensitivity, and the probability of getting a positive test result from the flock. Our example surveillance system component only has 20 flocks, which are shown In Table 24.

The data columns in this table are:

A) The flock ID is a unique number to identify each flock.

B) n is the number of animals tested from the flock.

C) N is the total number of birds in the flock.

D) Wetland: 1 means that the flock is near a wetland, and 0 means that it is not. Most flocks are near a wetland, as the surveillance targeted these flocks.

TABLE 23

|   | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1 | Parameters | Branch | RR | PrP | PrSSC | AR |
| 2 | Adjacent to wetlands | Yes | 1.5 | 0.05 | 0.8 | 1.463415 |
| 3 |  | No | 1 | 0.95 | 0.2 | 0.97561 |
| 4 | Flock type | Backyard | 4 | 0.9 | 0.95 | 1.081081 |
| 5 |  | Commercial | 1 | 0.1 | 0.05 | 0.27027 |
| 6 |  |  |  |  |  |  |
| 7 | Flock infected | 0.01 |  |  |  |  |
| 8 | Animal infected | 0.05 |  |  |  |  |
| 9 | Initial test | 0.98 |  |  |  |  |
| 10 | Confirmatory test | 0.9 |  |  |  |  |
| 11 | Combined test Se | 0.882 |  |  |  |  |
| 12 | P(Animal Pos) | 0.0441 |  |  |  |  |

TABLE 24

|   | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 20 | Flock ID | n | N | Wetland | Backyard | SeH | EPI | P(neg result) |
| 21 | 1 | 23 | 33 000 | 1 | 1 | 0.645606 | 0.015821 | 0.989786 |
| 22 | 2 | 25 | 45 000 | 1 | 1 | 0.676174 | 0.015821 | 0.989302 |
| 23 | 3 | 16 | 28 000 | 1 | 1 | 0.514041 | 0.015821 | 0.991868 |
| 24 | 4 | 12 | 27 000 | 1 | 0 | 0.417964 | 0.003955 | 0.998347 |
| 25 | 5 | 10 | 43 000 | 1 | 1 | 0.363022 | 0.015821 | 0.994257 |
| 26 | 6 | 28 | 22 000 | 0 | 1 | 0.717155 | 0.010547 | 0.992436 |
| 27 | 7 | 6 | 33 000 | 0 | 1 | 0.237087 | 0.010547 | 0.997499 |
| 28 | 8 | 17 | 39 000 | 0 | 0 | 0.535472 | 0.002637 | 0.998588 |
| 29 | 9 | 10 | 18 000 | 1 | 1 | 0.363022 | 0.015821 | 0.994257 |
| 30 | 10 | 28 | 11 000 | 1 | 1 | 0.717155 | 0.015821 | 0.988654 |
| 31 | 11 | 15 | 47 000 | 1 | 1 | 0.491622 | 0.015821 | 0.992222 |
| 32 | 12 | 14 | 42 000 | 1 | 1 | 0.468168 | 0.015821 | 0.992593 |
| 33 | 13 | 25 | 16 000 | 1 | 1 | 0.676174 | 0.015821 | 0.989302 |
| 34 | 14 | 24 | 41 000 | 1 | 1 | 0.661235 | 0.015821 | 0.989539 |
| 35 | 15 | 10 | 48 000 | 1 | 1 | 0.363022 | 0.015821 | 0.994257 |
| 36 | 16 | 14 | 36 000 | 0 | 1 | 0.468168 | 0.010547 | 0.995062 |
| 37 | 17 | 22 | 39 000 | 1 | 1 | 0.629256 | 0.015821 | 0.990045 |
| 38 | 18 | 18 | 28 000 | 1 | 1 | 0.555958 | 0.015821 | 0.991204 |
| 39 | 19 | 11 | 14 000 | 1 | 1 | 0.391112 | 0.015821 | 0.993812 |
| 40 | 20 | 30 | 31 000 | 0 | 1 | 0.741552 | 0.010547 | 0.992179 |

E) Backyard: 1 means that the flock is a backyard flock, and 0 means that it is a commercial flock.

In general terms, these columns should include a herd or flock identifier, one column for each herd-level risk factor, indicating the branch that the herd falls into, the total number of animals in the herd and the number of animals tested. This structure can be further extended to include different risk groups for animals within a herd, with an extra row for each risk group.

In addition to the data columns, Table 24 also includes three calculated columns:

F) SeH is the flock-level sensitivity that has been achieved by testing the *n* animals in the flock. The different ways of calculating the herd-level or flock-level sensitivity are discussed in detail in the next section. This example uses the binomial formula that we are already familiar with:

$$SeH = 1 - (1 - P_A^* \times SeA)^{n_h}$$

This is included in the spreadsheet as:

**=1-(1-B12)^B21**

G) EPI or the effective probability of infection. This captures the adjusted risk values for the two risk category nodes, and uses them to adjust the flock-level design prevalence.

$$EPI = P_H^* \times AR_1 \times AR_2$$

This can be extended for as many risk category nodes as are required. The spreadsheet formula is:

**=IF(D21,$F$2,$F$3)*IF(E21,$F$4,$F$5)*$B$7**

In this formula, **IF()** statements are used to select the correct adjusted risk values for the flock's risk group. The first checks the value of D21 to see if the flock is near a wetland or not, and chooses the appropriate adjusted risk.

H) P(Neg result) is the probability that the flock will have all negative results from the testing. This is one minus the probability of getting at least one positive result. This, in turn, is the probability that the flock is infected (the EPIH) times the probability of it being detected if it is infected (the flock-level sensitivity, SeH).

$$\Pr(\text{negative herd result}) = 1 - (EPIH \times SeH)$$

Or, in the spreadsheet:

**=1 - (F21 * G21)**

Once the probability that each flock will produce a negative result in the surveillance has been calculated, we can calculate the SSC sensitivity. This is the probability that at least one flock will give a positive result, or one minus the probability that all will give negative results. To calculate the probability that all flocks will give negative results, we simply multiply together the probabilities of negative results for each flock.

The general formula is:

$$SSCSe = 1 - \prod_{i=1}^{I}(1 - EPIH_i \times SeH_i)$$

and in the spreadsheet

**=1-(PRODUCT(H21:H40))**

In our example, this gives us an estimated component sensitivity of 13.5 percent. For comparison purposes, we can analyse the same component using the scenario tree and assuming independence between animals. The spreadsheet implementation of the scenario tree is shown in Table 25.

As can be seen from this analysis, the component sensitivity, when assuming independence, is 20.1 percent, compared with 13.5 percent when we take clustering into account. Failing to account for lack of independence among animals will mean that the sensitivity is overestimated. This is because we assume that we are getting the same amount of new information from every animal that is tested. However, if we test many animals from the same herd, we are getting less and less information from each new animal, so our overall sensitivity is lower.

The size of the difference between sensitivity when we take lack of independence into account, and when we do not, depends on whether the collection of the surveillance data is also clustered. If there have been many animals taken from a small number of herds, it will make a big difference. If there have only been a few animals taken from many different herds, then accounting for lack of independence may make almost no difference at all.

## HERD-LEVEL SENSITIVITY FORMULAE

The approach to calculating the sensitivity at the herd level varies slightly depending on the size of the herd and the number of animals tested from each herd. This section discusses the various options. The principles described here can be extended to higher grouping levels if they exist (for example, there may be three levels of infection nodes for intensive animal production – animal, house and farm), or to the entire surveillance system component.

### *Small proportion of herd tested*

When the proportion of the herd that is tested is small, it is reasonable to assume that sampling without replacement does not significantly change the probability that the next animal selected will be an infected animal. In this case, we can use the simpler binomial formula to estimate the herd-level sensitivity. If there are no animal-level risk nodes, the formula is:

$$SeH_h = 1 - (1 - P_A^* \times SeA)^{n_h}$$

TABLE 25

| Wetland | | Type | | Flock infected | | Animal infected | | Test results | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Branch | PrSSC | Branch | PrSSC | Branch | EPI | Branch | P*A | Branch | Se | Outcome | Probability |
| Yes | 0.8000 | Backyard | 0.9500 | Yes | 0.0158 | Yes | 0.0500 | Pos | 0.8820 | Pos | 0.0005 |
| | | | | | | | | Neg | 0.1180 | Neg | 0.0001 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0114 |
| | | | | No | 0.9842 | Yes | 0.0500 | Pos | 0.0000 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0374 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.7106 |
| | | Commercial | 0.0500 | Yes | 0.0040 | Yes | 0.0500 | Pos | 0.8820 | Pos | 0.0000 |
| | | | | | | | | Neg | 0.1180 | Neg | 0.0000 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0002 |
| | | | | No | 0.9960 | Yes | 0.0500 | Pos | 0.0000 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0020 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0378 |
| No | 0.2000 | Backyard | 0.9500 | Yes | 0.0105 | Yes | 0.0500 | Pos | 0.8820 | Pos | 0.0001 |
| | | | | | | | | Neg | 0.1180 | Neg | 0.0000 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0019 |
| | | | | No | 0.9895 | Yes | 0.0500 | Pos | 0.0000 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0094 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.1786 |
| | | Commercial | 0.0500 | Yes | 0.0026 | Yes | 0.0500 | Pos | 0.8820 | Pos | 0.0000 |
| | | | | | | | | Neg | 0.1180 | Neg | 0.0000 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0000 |
| | | | | No | 0.9974 | Yes | 0.0500 | Pos | 0.0000 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0005 |
| | | | | | | No | 0.9500 | Pos | 0 | Pos | 0.0000 |
| | | | | | | | | Neg | 1.0000 | Neg | 0.0095 |
| | | | | | | | | | | Check | 1.0000 |
| | | | | | | | | | | CSeU | 0.000627 |
| | | | | | | | | | | CSe (independent) | 0.201045 |

Where:
- $SeH_h$ is the herd-level sensitivity for the $h^{th}$ herd
- $P^*_A$ is the design prevalence at the animal level
- $SeA$ is the animal-level sensitivity
- $n_h$ is the number of animals tested from the $h^{th}$ herd

If there is one or more risk category node related to the animal-level infection node, there will be a number of different groups of animals within the herd with different risks of being infected. In this case, the group-level sensitivities are separately calculated and multiplied together to give the herd-level sensitivity, using the following formula which assumes *J* different risk groups of animals within the herd.

$$SeH_h = 1 - \prod_{j=1}^{J}(1 - EPIA_j \times SeA_j)^{n_j}$$

### Large proportion of herd tested

When the proportion of the herd that is tested is large, the assumptions of the binomial formula used above are no longer valid. Instead, it is more appropriate to use the binomial approximation to the hypergeometric distribution to calculate sensitivity. The formula for the calculation is:

$$SeH_h = 1 - (1 - SeA_{Av} \times \frac{n_h}{N_h})^{EPIA \times N_h}$$

Where:
- $SeA_{Av}$ is the average animal-level sensitivity for the herd
- $EPIA$ is the effective probability of infection of animals in the herd
- $N_h$ is the total number of animals in the herd, and
- $n_h$ is the number of animals tested from the herd

### All animals tested

When the entire herd is tested, if there are infected animals in the herd, we can guarantee that those animals will be tested. The herd-level sensitivity is therefore based on the animal-level sensitivity and the number of infected animals in the herd (the animal-level design prevalence). The formula is:

$$SeH_h = 1 - (1 - SeA_{Av})^{d_h}$$

Where:
- $d_h$ is the number of infected animals in the herd, $EPIA \times N_h$ rounded up to an integer.

## Chapter 14

# Combining multiple surveillance components

> *"There is only one way in which a person acquires a new idea;*
> *by combination or association of two or more ideas he already has into*
> *a new juxtaposition in such a manner as to discover a relationship*
> *among them of which he was not previously aware."*
>
> Francis A. Carter

Scenario trees give us the capacity to analyse complex risk-based surveillance in order to estimate its sensitivity. Normally, a surveillance system comprises a number of different components which provide different types of evidence that the disease is not present, or different approaches for the early detection of a disease.

---

**Example**

Consider the possible sources of evidence for bovine tuberculosis status:

- routine herd tuberculin tests
- movement or export testing
- abattoir meat inspection for granulomas
- passive clinical surveillance
- human health surveillance detecting *M. bovis*

Each of these systems is a component of an overall surveillance system, and each has a different capacity to detect the presence of the disease. Using scenario trees, we are able to estimate the sensitivity of each of the components, which is useful. However, we are also interested in the system as a whole – what is the sensitivity of all our surveillance, considering each of the components together?

---

This chapter presents techniques to combine different components of a surveillance system to answer this question.

### SIMPLE EXAMPLE

Sensitivities are probabilities, and by this stage by this stage **the users of this manual will already be expert** at combining probabilities. If we have two surveillance components, each with their own sensitivity, we can use probability theory to combine them into a single sensitivity.

**Example**

Consider a surveillance system with two components – component 1 is a structured survey and component 2 is abattoir meat inspection. We have analysed each component and calculated the sensitivity:

$CSe_1 = 82\%$

$CSe_2 = 45\%$

Component sensitivity is the probability that we would detect the disease using that component. The combined surveillance system sensitivity (*SSe*) is the probability that we would detect disease in at least one of the components. Using logic that is now familiar, this can be calculated as one minus the probability of not detecting disease in any of the components, giving us the simple formula for the sensitivity of the surveillance system:

$$SSe = 1 - \big((1 - CSe_1) \times (1 - CSe_2)\big)$$

If we are combining *I* different components, this can be generalized to:

$$SSe = 1 - \prod_{i=1}^{I} (1 - CSe_i)$$

For this example, the combined sensitivity would be:

$$SSe = 1 - (1 - 0.82) \times (1 - 0.45)$$
$$= 0.901$$

This approach is straightforward, and is appropriate for combining some components in a surveillance system. Unfortunately, there are many situations where this approach overestimates the combined sensitivity, due to overlapping of surveillance components.

## OVERLAPPING SURVEILLANCE COMPONENTS

In the previous example we had two components – structured surveillance and abattoir meat inspection. Some herds are beef herds and, as the primary purpose of production is to sell animals for slaughter, many animals are sent to the abattoir. Other farms are dairy farms, or may be breeding farms, and therefore send very few animals to the abattoir. Surveillance sensitivity for the abattoir component is therefore better in some farms than others.

For the structured survey, some farms are selected and some farms are not – the sensitivity in those farms that are not selected is zero, but we collect useful information from those farms that are selected.

The problem occurs when farms are present in both surveillance components. If a farm is a beef farm, and sends a lot of animals to the abattoir, then this factor will make a significant impact on the overall sensitivity of the abattoir surveillance system. However, if this particular farm is also selected for inclusion in the structured survey, it is also contributing information to the sensitivity of that component. The difficulty is that the information the farm provides is not new information. If the farm has been sending animals to the abattoir throughout the year, and they all test negative, this provides quite a bit of evidence that the

farm is not infected. Testing the farm again as part of the structured surveillance operation does not provide as much new information as if we had tested a farm that does not send any animals to the abattoir.

Where herds are included in more than one component of a surveillance system, the components are not independent. It is necessary to take this lack of independence into account when analysing the data; otherwise, we will overestimate the combined sensitivity of the system.

## ACCOUNTING FOR THE OVERLAP

In Chapter 13 we looked at analysing a scenario tree herd by herd, in order to take clustering into account. This approach gives us an opportunity to account for the overlap between surveillance system components at the herd level as well.

Normally, when we analyse a single herd in an SSC, we have a number of pieces of information:

- the probability that the herd is infected (from the design prevalence), which is the same for all herds;
- the risk factors applying to that herd, expressed in terms of the adjusted risk;
- the sensitivity for that herd, based on the number of animals tested and the animal-level sensitivity.

The result of the analysis is an estimation of the probability that the herd will give a negative result. We can also estimate the probability, after testing, that the herd is infected. This can be done using Bayes' theorem, as discussed in Chapter 2. The prior probability of being infected is given by the design prevalence; the new information is the surveillance that has been carried out, and the posterior probability says how likely the herd is to be infected, based on the surveillance. As all our surveillance results are negative, the posterior probability will be lower than the prior (the design prevalence).

This posterior is a description of our state of knowledge about the herd after the surveillance has been carried out. We may have assumed that all herds had the same prior probability of being infected, but after testing some of the herds, we are more confident that they are free. In contrast, we have no information about the untested herds, so we must still assume that the probability that they are infected is equal to the design prevalence.

When it comes to the next component of the surveillance system, normally it is analysed in the same way. However, instead of starting with the assumption that we know nothing about the state of the herds (and therefore using the design prevalence as an estimate of the probability that each herd is infected), if we have already done some surveillance, we now have better information about the state of some herds. Those herds that have already been tested with negative results have a lower probability of being infected than the untested herds. We can start our analysis of the second component by using the updated information about herd status, based on the results of the first component. In Bayesian terms, this means that instead of using the design prevalence as our prior for the probability that each herd is infected, we use the posterior estimate from the first surveillance component.

If three or more components are analysed, this chain can continue. For each component, the prior probability that each herd is infected is the posterior probability from the analysis of the previous component.

When we use this approach, herds that have been tested in another component will have a lower prior probability of infection, which means that we will be less likely to find any infection in that herd, and the contribution that that herd makes to the component sensitivity will be lower. Herds that have not been previously tested will continue to have the design prevalence as their prior probability of infection, and so they will contribute the same amount to the component sensitivity as if we had analysed the component independently.

To calculate the system sensitivity, we use the same approach presented in the simple example at the start of this chapter to combine the sensitivity of each component. This now gives a more accurate measure of system sensitivity because the contribution of the overlapping herds has already been removed, and therefore the components may now be considered independent.

### Spreadsheet example

If you use the web-based Freedom software, and provide herd-level data that allow the herds to be matched between surveillance system components, the software is able to take overlap between components into account and calculate the combined sensitivity.

If you are using a spreadsheet, the easiest approach is to include all the surveillance system components on the same sheet, one next to the other, as shown in the example on page 150.

The process (referenced by columns) involves the following:

A) Identify every herd uniquely. Every herd that appears in any component of the surveillance system should be listed.

For each component:

B) The prior probability that the herd is infected. For the first component analysed, this is the herd-level design prevalence (B2). For the subsequent components, this is the posterior probability that the herd is infected from the previous component (see columns H and N).

C) The number of animals tested. If the herd was not included in the component, then the number of animals tested is zero.

D) The herd-level sensitivity, as calculated in previous examples. This example has been simplified, and thus no animal or herd-level risk factors have been included. The herd-level sensitivity in D10 is therefore:

=1-(1-$B$5)^C10

E) The effective probability of infection, as calculated in previous examples. This simple spreadsheet has no risk factors, and therefore this is equal to the prior probability that the herd is infected (B10).

F) The probability that the herd will have a negative result in the surveillance, as calculated in previous examples. This is one minus the probability that it is infected (the EPI) times the probability that the infection will be detected (the herd-level sensitivity):

=1-(D10 * E10)

G) The posterior probability that the herd is infected. This is an application of Bayes' theorem, analogous to the negative predictive value in clinical testing. Bayes' theorem was discussed on pages 30 to 32. This is calculated as:

$$\Pr(D+\,|\,T-) = 1 - \frac{(1-P)Sp}{(1-P)Sp + P(1-Se)}$$

When Sp is equal to 1 this simplifies to:

$$\Pr(D+\,|\,T-) = 1 - \frac{1-P}{1-P\times Se}$$

Or in the spreadsheet (G10):

=1-(1-E10)/(1-E10*D10)

The component sensitivities are calculated in cells F17, L17 and R17 as:

=1-PRODUCT(F10:F15)

The surveillance system (combined) sensitivity is calculated in cell F2 as:

=1-(1-F17)*(1-L17)*(1-R17)

The system sensitivity is 25.05 percent. For comparison purposes, the second sheet shows the same calculations performed without taking into account the overlap between the components. This is done by setting the prior probability that each herd is infected to be the design prevalence for all three components (rather than the posterior from the previous component). The system sensitivity in this case is 35 percent.

There was significant overlap between the components; therefore, the analysis resulted in a significant decrease in the estimate of system sensitivity. If there were less overlap, the difference would be less.

Compare the component sensitivities between the two approaches. For component 1, there is no difference, as in both cases the analysis used the herd-level design prevalence as the prior. However, the sensitivities for components 2 and 3 are progressively lower. Using this approach, we sometimes find that those components analysed last contribute almost nothing, as there is almost complete overlap with previously analysed components.

Changing the order of components in the analysis will not change the final estimate, but will change the sensitivity of the individual components.

SPREADSHEET 1
**Accounting for overlap between surveillance system components using Bayesian revision**

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Parameters | | | | Result | | | | | | | | | | | | | | |
| 2 | P*H | 0.05 | | | SSe | 0.2505 | | | | | | | | | | | | | |
| 3 | P*A | 0.2 | | | | | | | | | | | | | | | | | |
| 4 | SeA | 0.56 | | | | | | | | | | | | | | | | | |
| 5 | CSeU | 0.112 | | | | | | | | | | | | | | | | | |
| 6 | | | | | | | | | | | | | | | | | | | |
| 7 | Calculations | | | | | | | | | | | | | | | | | | |
| 8 | | Component 1 | | | | | | Component 2 | | | | | | Component 3 | | | | | |
| 9 | ID | Prior P(inf) | n | SeH | EPI | P(neg result) | Post P(inf) | Prior P(inf) | n | SeH | EPI | P(neg result) | Post P(inf) | Prior P(inf) | n | SeH | EPI | P(neg result) | Post P(inf) |
| 10 | 1 | 0.05 | 25 | 0.949 | 0.05 | 0.953 | 0.003 | 0.003 | 10 | 0.695 | 0.003 | 0.998 | 0.001 | 0.001 | 0 | 0.000 | 0.001 | 1.000 | 0.001 |
| 11 | 2 | 0.05 | 0 | 0 | 0.05 | 1.000 | 0.050 | 0.050 | 14 | 0.810 | 0.050 | 0.959 | 0.010 | 0.010 | 19 | 0.895 | 0.010 | 0.991 | 0.001 |
| 12 | 3 | 0.05 | 0 | 0 | 0.05 | 1.000 | 0.050 | 0.050 | 0 | 0.000 | 0.050 | 1.000 | 0.050 | 0.050 | 20 | 0.907 | 0.050 | 0.955 | 0.005 |
| 13 | 4 | 0.05 | 10 | 0.695 | 0.05 | 0.965 | 0.016 | 0.016 | 3 | 0.300 | 0.016 | 0.995 | 0.011 | 0.011 | 23 | 0.935 | 0.011 | 0.990 | 0.001 |
| 14 | 5 | 0.05 | 28 | 0.964 | 0.05 | 0.952 | 0.002 | 0.002 | 8 | 0.613 | 0.002 | 0.999 | 0.001 | 0.001 | 0 | 0.000 | 0.001 | 1.000 | 0.001 |
| 15 | 6 | 0.05 | 0 | 0 | 0.05 | 1.000 | 0.050 | 0.050 | 12 | 0.760 | 0.050 | 0.962 | 0.012 | 0.012 | 1 | 0.112 | 0.012 | 0.999 | 0.011 |
| 16 | | | | | | | | | | | | | | | | | | | |
| 17 | | | | | CSe1 | 0.1249 | | | | | CSe2 | 0.084114 | | | | | CSe3 | 0.0649 | |

This is a 

SPREADSHEET 2
**The same calculations without accounting for overlap, using a constant prior probability of infection**

|   | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | Calculations | | | | | | | | | | | | | | | | | | |
| 8 | | Component 1 | | | | | | Component 2 | | | | | | Component 3 | | | | | |
| 9 | ID | Prior P(inf) | n | SeH | EPI | P(neg result) | Post P(inf) | Prior P(inf) | n | SeH | EPI | P(neg result) | Post P(inf) | Prior P(inf) | n | SeH | EPI | P(neg result) | Post P(inf) |
| 10 | 1 | 0.05 | 25 | 0.949 | 0.05 | 0.953 | 0.003 | 0.003 | 10 | 0.695 | 0.050 | 0.965 | 0.016 | 0.016 | 0 | 0.000 | 0.050 | 1.000 | 0.050 |
| 11 | 2 | 0.05 | 0 | 0 | 0.05 | 1.000 | 0.050 | 0.050 | 14 | 0.810 | 0.050 | 0.959 | 0.010 | 0.010 | 19 | 0.895 | 0.050 | 0.955 | 0.005 |
| 12 | 3 | 0.05 | 0 | 0 | 0.05 | 1.000 | 0.050 | 0.050 | 0 | 0.000 | 0.050 | 1.000 | 0.050 | 0.050 | 20 | 0.907 | 0.050 | 0.955 | 0.005 |
| 13 | 4 | 0.05 | 10 | 0.695 | 0.05 | 0.965 | 0.016 | 0.016 | 3 | 0.300 | 0.050 | 0.985 | 0.036 | 0.036 | 23 | 0.935 | 0.050 | 0.953 | 0.003 |
| 14 | 5 | 0.05 | 28 | 0.964 | 0.05 | 0.952 | 0.002 | 0.002 | 8 | 0.613 | 0.050 | 0.969 | 0.020 | 0.020 | 0 | 0.000 | 0.050 | 1.000 | 0.050 |
| 15 | 6 | 0.05 | 0 | 0 | 0.05 | 1.000 | 0.050 | 0.050 | 12 | 0.760 | 0.050 | 0.962 | 0.012 | 0.012 | 1 | 0.112 | 0.050 | 0.994 | 0.045 |
| 16 | | | | | | | | | | | | | | | | | | | |
| 17 | | | | | CSe1 | 0.1249 | | | | | CSe2 | 0.149311 | | | | | CSe3 | 0.1356 | |

# Chapter 15
# Probability of freedom

*"If there be two subsequent events,*
*the probability of the 2ᵈ b/N and the probability of both together P/N,*
*and it being 1ˢᵗ discovered that the 2ᵈ event has also happened,*
*the probability I am right is P/b."*
Reverend Thomas Bayes (1702–61)
[first formulation of Bayes' theorem in
*Essay towards solving a Problem in the Doctrine of Chances*]

## SENSITIVITY VERSUS FREEDOM

The main measure of the quality of surveillance to demonstrate freedom, or for early detection of disease, is the sensitivity of the surveillance. Scenario tree analysis allows us to estimate the sensitivity of complex surveillance systems, and Chapter 14 introduced approaches that enable us to combine the sensitivity of multiple surveillance components.

Surveillance sensitivity and design prevalence are commonly used as standards for surveillance. Sensitivity is therefore a useful measure, but the concepts of sensitivity are sometimes difficult to communicate, especially to people with a non-technical background.

When dealing with freedom from infection, the first question that may be asked is: "Is

**Sensitivity is hard to understand for non-technical people.**

the country free from infection?" or "How confident are we that we are free?" The analyses that have been used thus far in this manual have attempted to answer that question by estimating sensitivity. To put the answer into words: "The probability that the surveillance would be able to detect infection, assuming that the population is infected at a level specified by the design prevalence, is **x** percent." For a non-technical person, this answer may be very hard to interpret. There seems to be a paradox: the question is about the country being free, and the answer assumes that the country is infected.

Most people would find it much easier to understand an answer expressed in terms of

**Probability of freedom is easier to understand.**

the probability of freedom. For instance: "The probability that the country is free from infection is **x** percent." Communicating the results of analysis in these terms makes it much easier for people to understand what is being said, and **and it also delivers** a number of other benefits as well.

## CALCULATION OF THE PROBABILITY OF FREEDOM FROM INFECTION

Surveillance sensitivity is a conditional probability – the probability that the surveillance system would find the disease, given that the country is infected at a specified design prevalence. In probability terms, this can be written as:

P(T+|D+)

Where:

- T+ stands for test positive, or the surveillance produces a positive outcome, and
- D+ stands for disease positive, or the country is truly infected at a specified level.

The probability of freedom can also be expressed in these terms:

P(D-|T-)

or the probability that the country is free from infection (D-) given that the surveillance has not produced a positive result (T-).

When you compare these two probability statements, there are a couple of important things to notice:

- They are not the same. Sensitivity cannot be interpreted as probability of freedom.
- The conditionality is reversed. Sensitivity is conditional on the population being infected, while probability of freedom is conditional on negative surveillance results.
- The probability of freedom at the country level looks rather like a negative predictive value at the animal level.

The negative predictive value of a diagnostic test is the probability that an animal is truly

> Probability of freedom is analogous to the negative predictive value of a test.

negative, given that we got a negative test result. At the country level, we are interested in the probability that the country is truly negative, given that we got negative results from our surveillance.

Just as with predictive values, we can use Bayes' theorem as shown on pages 30-32 to calculate the probability of freedom:

$$P(\textit{free}) = \frac{\text{True Negative}}{\text{True Negative} + \text{False Negative}}$$

$$= \frac{(1 - P) \times Sp}{(1 - P) \times Sp + P \times (1 - Se)}$$

Where:

- Sp and Se are the sensitivity and specificity of the surveillance system, and
- P is the *prior* probability that the country was infected.

If the specificity of the surveillance system is 100 percent, this is simplified to:

$$P(\textit{free}) = \frac{1 - P}{1 - (P \times Se)}$$

## SELECTING A PRIOR

In the formula for negative predictive value used for individual animal diagnostic testing, the prior probability of infection is estimated by the prevalence of the disease. However, when we are working at the country level, how do we know the value for the prior?

One possibility would be the prevalence at a country level. Consider the population to be all countries in the world, and the prevalence is the proportion of countries that are

infected. Most would agree that this is not a reasonable estimate of the probability that any particular country is infected, especially one that is claiming to be free, as geographic region and biosecurity play an important role in a country's disease status.

Instead, the prior should reflect the country's particular situation. If one believes that the country is free, and then undertakes surveillance to help support this claim, then the prior probability that the country is infected should be quite low.

---

### Example

We have undertaken surveillance that has a sensitivity of 75 percent. If we think that the prior probability of infection was 10 percent (which means that the prior probability of freedom was 90 percent), our posterior probability of freedom will be 97.3 percent.

---

If we had chosen a prior probability of infection of 80 percent (only 20 percent probability of being free), our posterior (with the same surveillance sensitivity) would have been 50 percent instead of 97.3 percent.

When reporting the probability that a country is free from infection based on surveillance, the choice of the prior can make a very big difference. If a country wishes to indicate that it is very likely to be free, it will choose a high prior probability of being free. However, trading partners may not agree with such an optimistic view of things, and may challenge the prior.

In this way, the prior is similar to the design prevalence. It has a big impact on the result, but is difficult to choose objectively. In order to avoid disagreements about the interpretation of the analysis of surveillance, there needs to be an objective and mutually agreeable method of selecting a prior.

One useful approach is to base the prior on previous information. The prior describes the probability of being free *before* the current surveillance was carried out, but it can be based on earlier surveillance. This means that the value for the prior for the current year can be justified by analysis of the previous year's surveillance data. This solution looks good until one asks what the prior should be for the analysis of the previous year's surveillance. The obvious answer is to base it on an analysis of the previous year's data. And so on.

Regardless of how many years of surveillance data are available, there will always be a starting point where there is either no previous surveillance data, or where it was known that the country was infected. In both cases, we are still left with the question of what value we should use for the prior at the beginning of our surveillance to demonstrate freedom from infection.

If the surveillance has started after an eradication programme, it means that in the previous year, cases were detected, so the probability of infection was 100 percent and the probability of freedom was 0 percent. In the following year, no cases were detected, and therefore the country may be free. The simplest approach to address the concerns of the stakeholders is to use a standard 'compromise' value for the prior probability of freedom before the first round of surveillance. This value is 50 percent.

The advantage of this approach is that, while the prior has a big impact on the posterior in any one year, for a series of analyses in which the posterior of one is used as the prior of

the next, the starting prior quickly loses its influence on the result. This means that, as long as the surveillance is reasonably sensitive, the probability of freedom after, say, five years of analysis is reasonably independent of the starting prior probability of infection being present. For example, with the surveillance described on page 155, the posterior probability of freedom would be greater than 99.5 percent after five years, regardless of whether a prior probability of infection of 10 percent, 50 percent or 80 percent had been used.

While a prior probability of freedom of 50 percent is commonly used, it is also important to recognize that a 50 percent value implies that infection and freedom were equally likely when we started the surveillance for freedom. In some cases this may be correct. However, in many cases, such as many years of history of absence of disease (although not formally quantified), or completion of a rigorous eradication programme, may provide a relatively high level of confidence that the disease is truly absent. In these situations it would be reasonable to argue for a higher value for the prior probability of freedom (for example 75 percent). Despite this, in the end it will often be a negotiated value agreed between trading partners.

# Chapter 16
# Incorporating historical surveillance data

*"History is the witness that testifies to the passing of time;*
*it illumines reality, vitalizes memory, provides guidance in daily life*
*and brings us tidings of antiquity."*

Cicero (106 BC–43 BC)

The previous chapter suggested that the best way to choose a suitable prior probability of freedom when calculating the current probability of freedom is to use the posterior probability from the previous year or years. This approach opens up the possibility of incorporating historical data into our analysis of surveillance.

## VALUE OF HISTORICAL DATA

For many diseases there is often a certain amount of historical surveillance data that can provide evidence for freedom from infection. Passive clinical surveillance with an absence of reports consistent with the infection is often available, but other types of surveillance may also exist.

Clearly, current surveillance data are useful. Surveillance data from the previous year are

> Surveillance data loses value as it ages.

probably useful as well. But is surveillance that was carried out 20 years ago relevant to the current disease situation? Most experts would consider such old information to be irrelevant. This demonstrates the principle that surveillance information loses its value as it gets older. It would not be valid to analyse surveillance data carried out 20 years ago and claim that the data have the same value in demonstrating current freedom.

On the other hand, we would probably be more confident about the status of a country that has undertaken surveillance every year for the last 20 years, with consistently negative results, compared to one that has only carried out surveillance for the first time during the current year. Historical information has some value, which can accumulate over time, but this value decreases with age.

The reason for the decrease in value of historical data is the risk of introduction of new disease that would change the disease-free status of the population. Where the risk of introduction of disease is small, historical information retains more of its value. Where the risk is great, the value quickly vanishes.

Remember the analogy of the scales that was used on page 67. This showed that the probability that a population is free from infection is a balance between the surveillance evidence that accumulates over time and the risk of introduction of new disease into the herd. When the surveillance evidence outweighs the risk of introduction, the evidence for freedom accumulates and the probability rises towards 100 percent. When the risk of introduction outweighs the surveillance evidence, the probability of freedom decreases towards 0 percent.

**Example**

What is it that makes old surveillance data less valuable? Consider the example of a single farm and a hypothetical disease that can only be introduced into the farm via live animals. The farm carried out detailed surveillance of all of its animals 20 years ago and demonstrated, to a very high level of confidence, that the disease was not present.

If the farm is a closed herd, breeds its own replacements and never introduces animals from outside the herd, there has been no opportunity to introduce the infection from outside. The confidence in freedom is the same today as it was 20 years ago, as there is no way the farm could have lost its free status.

On the other hand, if the farm sells 20 percent of its animals every year, and buys in a further 20 percent to replace them, there has been a constant risk of introducing new disease. The surveillance from 20 years ago tells us nothing about the current status.

## RISK OF INTRODUCTION

The risk of introduction of disease is measured as the probability that disease will enter the herd during a certain time period (the time period of analysis).

The good news is that there is a well-defined methodology that enables us to estimate

Risk analysis is used to estimate the probability of introduction of infection.

the risk of introduction, and that many of the techniques used are very similar to the techniques that have been discussed in this manual. The methodology is quantitative risk analysis.

The bad news is that, as with the creation of a scenario tree model, performing a thorough quantitative risk analysis can be a very challenging and time-consuming task. In the ideal situation, a risk analysis has already been undertaken, and the results can be used directly. Often, this is not the case and either a detailed risk analysis is required, or a quick, simple risk analysis can be undertaken.

Risk analysis methodology has been extensively described elsewhere and is beyond the scope of this manual. We will limit ourselves to the following comments:

- A full risk analysis involves a number of steps, including hazard identification, risk assessment and risk mitigation. For the purposes of assessing the probability of introduction of infection, only a small part of this overall process is required – the release and exposure assessment.
- This is based on the use of risk pathway diagrams (similar to a scenario tree) which describe the events that must occur in order for infection to move from the population of origin and become established in the target population.
- Quantitative risk assessment is based on the multiplication of probabilities for each pathway, as has been discussed for scenario trees.
- Uncertainty can be incorporated into risk pathway models using the same approach as that taken for scenario tree models. Input probabilities are described as distributions, and stochastic modelling generates an output distribution. The output distribution can then be used in the scenario tree model. For quick, simple risk pathway analyses, it is important to be realistic about uncertainty and to use stochastic modelling.

## CALCULATION OF POSTERIOR PROBABILITY OF FREEDOM

If you are analysing historical surveillance data using the web-based software, the system is able to perform the calculations automatically. You will need to provide the following information:

- the date of each surveillance observation (so that the data can be divided into multiple periods);
- the length and number of periods to be analysed;
- the probability of introduction of infection during each period.

Calculations using a spreadsheet are illustrated on page 160.

### Time period of analysis

Any analysis of surveillance data requires a definition of the time period over which data are analysed. If the surveillance is based on a time-limited activity (such as a structured survey), then the time period is the duration of the activity. However, surveillance generally tends to be ongoing or sporadic. For instance, abattoir meat inspection surveillance is carried out every day, and so there is a constant stream of data available.

One of the most important factors influencing the sensitivity of surveillance is the number of animals that pass through the surveillance system. When there is an ongoing stream of surveillance information, it must be divided into discrete time periods in order to analyse it. The longer the time period for analysis, the greater the number of animals that will be included in the analysis. Choosing a long time period therefore increases the apparent sensitivity, whereas short time periods have a lower sensitivity.

While it may be tempting to use a long time period to give a higher sensitivity, this is rarely the best approach. Consider analysis of FMD surveillance data for a period of five years. It could be analysed as a single dataset for five years; five time periods of one year each, or for a series of one-month time periods. The sensitivity for a one-month period would be very much lower than the five-year sensitivity. However, as this chapter has shown, it is possible to combine the monthly data together to generate an overall probability of freedom. Normally, an estimate based on combined short time periods will be a little lower than an estimate based on analysis of all the data as a single time period, but the **two estimates** will be broadly similar.

The reason for the difference is that analysing the data as short time periods allows the value of the older data to be discounted according to the risk of introduction of infection. When analysing the data as a single time period, the data relating to five years ago and the data relating to yesterday are both treated as if they have the same value.

It is therefore better to analyse the data as multiple relatively short time periods rather than a single long time period. But how long should these short time periods be? The answer largely depends on the nature of the disease. For rapidly spreading diseases with short incubation periods, such as FMD, newly introduced disease can spread and reach the design prevalence very quickly. A short period of analysis is appropriate – normally one month – but it could be as short as one week. For slow-developing diseases, such as tuberculosis or BSE, a one-year period of analysis is usually used.

TABLE 26

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Time period | SSSe | P(intro) | Prior P(inf) | Prior P(free) | Post P(free) | Post P(inf) | Post P(inf) adjusted | Post P(free) adjusted |
| 2 | 0 | 0.459 | 0.043 | 0.5 | 0.5 | 0.648929 | 0.351071 | 0.378975 | 0.621025 |
| 3 | 1 | 0.597 | 0.03 | 0.378975 | 0.621025 | 0.802615 | 0.197385 | 0.221463 | 0.778537 |
| 4 | 2 | 0.634 | 0.044 | 0.221463 | 0.778537 | 0.905705 | 0.094295 | 0.134146 | 0.865854 |
| 5 | 3 | 0.699 | 0.018 | 0.134146 | 0.865854 | 0.955444 | 0.044556 | 0.061754 | 0.938246 |
| 6 | 4 | 0.761 | 0.023 | 0.061754 | 0.938246 | 0.984513 | 0.015487 | 0.038131 | 0.961869 |
| 7 | 5 | 0.411 | 0.023 | 0.038131 | 0.961869 | 0.977183 | 0.022817 | 0.045292 | 0.954708 |
| 8 | 6 | 0.638 | 0.014 | 0.045292 | 0.954708 | 0.983116 | 0.016884 | 0.030647 | 0.969353 |
| 9 | 7 | 0.485 | 0.01 | 0.030647 | 0.969353 | 0.983979 | 0.016021 | 0.025861 | 0.974139 |
| 10 | 8 | 0.616 | 0.02 | 0.025861 | 0.974139 | 0.989909 | 0.010091 | 0.02989 | 0.97011 |
| 11 | 9 | 0.405 | 0.012 | 0.02989 | 0.97011 | 0.981998 | 0.018002 | 0.029786 | 0.970214 |
| 12 | 10 | 0.509 | 0.043 | 0.029786 | 0.970214 | 0.98515 | 0.01485 | 0.057212 | 0.942788 |
| 13 | 11 | 0.451 | 0.017 | 0.057212 | 0.942788 | 0.967759 | 0.032241 | 0.048693 | 0.951307 |
| 14 | 12 | 0.553 | 0.032 | 0.048693 | 0.951307 | 0.977632 | 0.022368 | 0.053652 | 0.946348 |
| 15 | 13 | 0.536 | 0.049 | 0.053652 | 0.946348 | 0.974368 | 0.025632 | 0.073376 | 0.926624 |
| 16 | 14 | 0.417 | 0.032 | 0.073376 | 0.926624 | 0.955872 | 0.044128 | 0.074716 | 0.925284 |
| 17 | 15 | 0.513 | 0.011 | 0.074716 | 0.925284 | 0.962163 | 0.037837 | 0.048421 | 0.951579 |

## Spreadsheet implementation

Incorporation of historical surveillance data involves the repeated calculation of the probability of freedom, starting with the earliest time period, using Bayes' theorem. For each period, the posterior probability of infection for the previous period is used as the prior probability of infection for the current period. To account for the risk of introduction of disease, the posterior probability of infection is adjusted to account for the possibility that disease may have been introduced during that period.

The calculation is based on the surveillance system sensitivity and the risk of introduction for each time period. The spreadsheet in Table 26 illustrates how the calculations can be set up.

The columns and formulae are as follows:

A. The period number. Periods can be any length, but the values in B and C must be calculated based on the selected time period.

B. The surveillance system sensitivity, based on the analysis of one or more components.

C. The probability of introduction of disease, based on published data or risk analysis.

D. The prior probability that the country is infected. For the first time period, this is set at a standard starting value of 0.5. For subsequent time periods, it is the adjusted posterior probability of infection for the previous time period. For example, cell D3 contains **=H2**

E. The prior probability of freedom. This is one minus the prior probability of infection. Cell E2 contains **=1-D2**

F. The posterior probability of freedom calculated using Bayes' theorem. This is the key result that we are after for each time period. The formula (assuming perfect specificity) and its simplification are:

$$\text{Post P(free)} = \frac{1-P}{(1-P)+P \times (1-SSe)}$$

$$= \frac{1-P}{1-P \times SSe}$$

The implementation in cell F2 is **= (1-D2)/(1-(D2*B2))**

G. The posterior probability of being infected. This is simply one minus the posterior probability of being free. Cell G2 contains =1-F2

H. The posterior probability of infection at the beginning of the *next* time period adjusted for the probability of introduction of infection. This value is used as the prior for the next time period. This calculation is based on the generalization of the OR probability rule for non-exclusive outcomes. The country could be infected because:

- it was not free to begin with, or
- it became infected during the time period, or
- both occurred.

This is calculated as the sum of the two possibilities minus the overlap, or the probability that both outcomes have happened. This can be expressed as:

$$P(A \text{ or } B) = P(A) + P(B) - P(A) \times P(B)$$

In the spreadsheet cell H2 contains **=G2+C2-(G2*C2)**

I. The adjusted posterior probability of being free at the start of the next time period is 1 minus the adjusted probability of being infected. From the table, I2 contains **=1-H2**.

The data from the spreadsheet have been plotted in the figures on page 162.

## EXAMPLES

The ability to combine historical surveillance data provides a great deal of flexibility when seeking to demonstrate freedom from infection. If the risk of introduction is low, then inexpensive surveillance with relatively low sensitivity can generate a high probability of freedom if it is continued over a long enough period.

The figures show surveillance with 20 percent sensitivity and a 0.1 percent probability of introduction of disease.

If the risk of introduction of disease is high, even very sensitive surveillance may be inadequate to achieve a high probability of freedom. Figure 26 on page 162 shows surveillance with a sensitivity of 80 percent but a risk of introduction of 70 percent.

FIGURE 24



FIGURE 25



FIGURE 26
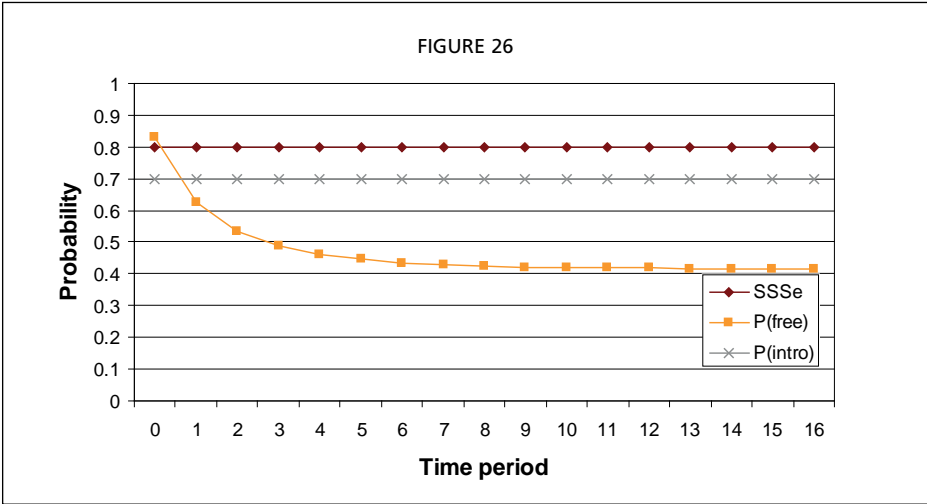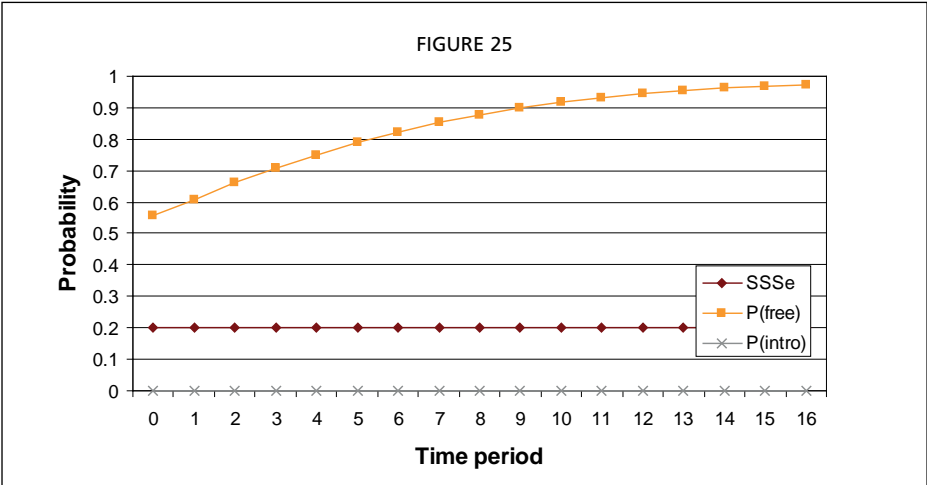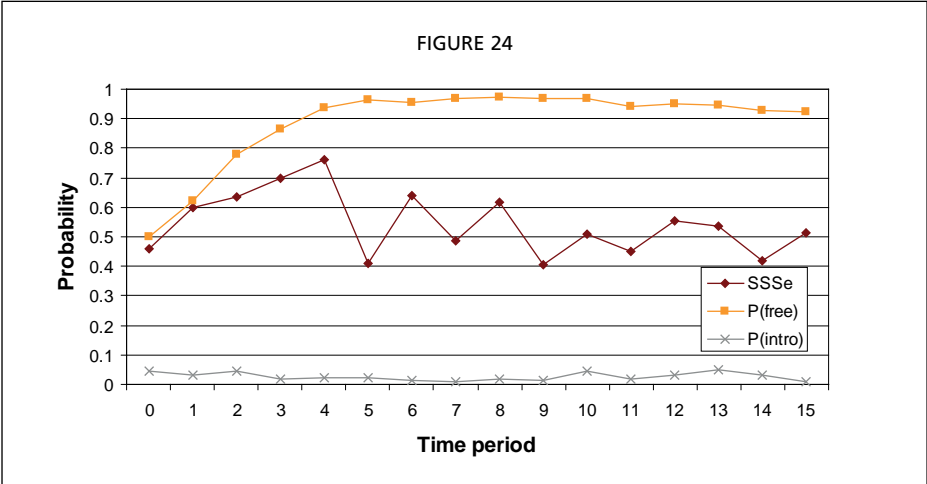
# Chapter 17
# EpiTools software

*"We've heard that a million monkeys banging on a million typewriters will eventually reproduce the entire works of Shakespeare. Now, thanks to the Internet, we know this is not true."*

**Robert Wilensky (1951-2013)**

Analysing surveillance to demonstrate freedom from infection using scenario trees is possible using a spreadsheet with a Monte Carlo simulation add-in. However, the process is complicated and it is easy to make mistakes, especially with a large-scale model.

EpiTools (http://epitools.ausvet.com.au) is a web-based suite of epidemiological calculators and utilities that can be used for many of the calculations required for the analysis of representative surveillance or risk-based surveillance. In order to undertake a valid analysis of surveillance, it is necessary to have a good understanding of the principles of scenario tree modelling. In addition, it should be noted that while EpiTools can manage complex calculations related to scenario tree modelling, it is important that those analysing surveillance data should have a good understanding of the principles of scenario tree modelling. Furthermore, they should develop an appropriate model structure and provide the correct input parameters.

## OVERVIEW

EpiTools is a comprehensive system of epidemiological calculators and other utilities. As such, it serves many purposes, in addition to being capable of producing relevant calculations for freedom surveillance. The front page (or main menu) is divided into the following sections: *Surveillance utilities*, *Epidemiological studies*, *Application of diagnostic tests* and *New additions*, as shown below. In this manual we will focus only on *Surveillance utilities* and, within surveillance, only on those sections that are relevant to surveillance for disease freedom.

## SURVEILLANCE UTILITIES

The surveillance section includes options for a variety of surveillance-related activities, as shown below. Each of these options is described in more detail in the rest of this chapter.



**EpiTools epidemiological calculators**

This site has been developed by AusVet Animal Health Services, with funding from the Australia estimating disease prevalence or demonstrating freedom from disease through structured surve

**Surveillance utilities**

- **1-Stage representative freedom surveys** New menu
- **2-Stage representative freedom surveys** New menu
- **Risk-based freedom surveys** New menu
- **Random Sampling from a population** New menu
- **Estimating true prevalence**
- **Pooled prevalence calculator**
- **Survey Toolbox for livestock diseases and freedom in finite populations**
- **HerdPlus module for herd-sensitivity and freedom in finite populations**

## 1-stage representative freedom surveys

This section provides utilities for designing and analysing 1-stage representative freedom surveys (see Chapter 5). Options include surveys assuming perfect test specificity (for example, using follow-up of positives with a definitive test) or allowing for imperfect test specificity, if necessary. All of the formulae discussed in Chapter 5 are implemented here, including formulae for large or small population sizes. The specific formulae are listed in the appendix on page 190.



**1-Stage Freedom analysis**

Use these utilities to assist with planing and analysis of simple 1-stage surveys to demonstrate disease freedom. Methods are appropriate for sampling units (animals, fish, plants, etc) to demonstrate cluster (herd, flock, farm, tank, etc) freedom, or for other individuals or clusters to demonstrate population freedom. All methods assume representative sampling of units from the population and methods are included for large (or unknown) and finite (known) population size. Most of the options (except where noted) assume perfect (100%) test specificity.

If you prefer the old menu click here.

Click here for Some important formulae for freedom analyses.

**Assuming perfect test specificity (100%)**

| *Sample size calculation* | *Population sensitivity calculation (SeH or SSe)* |
|---|---|
| Use these options to estimate required sample size assuming perfect test specificity: | Use these options to calculate cluster or population-level sensitivity assuming perfect test specificity: |
| • For known or unknown population size (replaces previous options for large and finite populations) | • For perfect specificity and known or unknown population size (replaces previous options for large and finite populations) |
| • For pooled sampling in a large population | • For varying unit sensitivity |
| • To achieve target confidence of freedom | |

**Allowing for imperfect test specificity (<100%)**

| *Sample size calculation* | *Population sensitivity with imperfect test specificity* |
|---|---|
| Use these options to estimate required sample size assuming imperfect test specificity: | Use these options to calculate cluster or population-level sensitivity assuming imperfect test specificity: |
| • FreeCalc sample size for imperfect tests | • For a large population size and variable cut-point number of reactors |
| | • FreeCalc calculator for imperfect tests |
| | • Using several different methods |

**Some other useful options**

| *Design prevalence* | *Confidence of population freedom from disease* |
|---|---|
| • Design prevalence required to achieve target population sensitivity for given sample size | • For a single time period |
| | • For multiple time periods |

### *Assuming perfect specificity*

The options presented in this section allow for calculation of sample size and population sensitivity for 1-stage surveys where there is follow-up testing to confirm any positive results to screening tests as either true positives (population is infected) or false positives (no evidence of infection). These calculations use the formulae described in Chapter 5 and Chapter 6.

*Sample size for known or unknown population size*

Use this option to calculate the sample size required to achieve a target population or cluster-level sensitivity for a 1-stage survey. For cluster (herd, flock, etc.) calculations, enter test sensitivity, unit-level design prevalence and required cluster-level sensitivity. Alternatively, for population-level calculations enter cluster-level (herd) sensitivity, cluster-level design prevalence and required population-level sensitivity. Design prevalence can be entered either as a proportion or as a specified number of units. Calculations use the hypergeometric approximation if population size is specified, or the binomial method if population size is not specified.

**Sample size to achiev**

**Input Values**

| Design prevalence (proportion or units): | 0.01 |
| | ◉ Proportion |
| | ○ Unit(s) |
| Unit (test or cluster) sensitivity: | 0.9 |
| Required population sensitivity: | 0.95 |
| Population size (if known): | |

[ Submit ]    [ Reset ]

*Sample size for pooled testing for freedom*

Use this option to calculate the number of pools that need to be tested in order to provide a target population sensitivity at the specified design prevalence, for a large (infinite) population, using pooled testing, and assuming a fixed pool size and a test of known sensitivity and 100 percent specificity. Inputs are the pool size used, pool-level test sensitivity, desired population sensitivity and design prevalence. Test sensitivity is measured at the pool level (i.e. the probability that a pool will test positively if it represents one or more infected animals).

**Sample size for**

**Input Values**

| Size of pools: | 10 |
| Test Sensitivity: | 0.9 |
| Desired herd-sensitivity: | 0.95 |
| Design (target) prevalence: | 0.02 |

[ Submit ]

*Sample size to achieve target confidence of freedom*

Use this option to calculate the sample size required to achieve a desired level of confidence of population freedom at either the herd level or the population level. *Note: this option is for calculating confidence of population freedom, not confidence of detection (population sensitivity).* If applied at the herd level, sensitivity and design prevalence are unit-level (animal-level) values, whereas if applied at the population level, sensitivity and design prevalence are cluster-level (herd-level/flock-level) values. Design prevalence can be entered either as a proportion or as a specified number of units. Prior confidence of freedom is the level of confidence of population freedom *before* undertaking the surveillance being analysed. Probability of introduction is the probability that disease will be introduced into the population during the period when the surveillance is being undertaken.

**Sample size required t**

**Input Values**

| Test (or Herd) Sensitivity: | 0.9 |
| Prior confidence of freedom: | 0.7 |
| Probability of introduction during period: | 0.01 |
| Required confidence of freedom: | 0.95 |
| Population size (leave blank or 0 if not known): | |
| Design prevalence (proportion or units): | 0.02 |
| | ◉ Proportion |
| | ○ Unit(s) |

[ Submit ]    [ Reset ]

### Calculate population sensitivity for known or unknown population size

Use this option to calculate population-level or cluster-level sensitivity (confidence of detection) for a 1-stage survey. For these calculations, unit specificity is assumed to be 100 percent. For calculating cluster (herd, flock, etc.) sensitivity, enter test sensitivity and unit-level design prevalence, or for population-level sensitivity enter cluster-level (herd) sensitivity and cluster-level (herd) design prevalence. Design prevalence can be entered either as a proportion or as a specified number of units. Calculations use the hypergeometric approximation if population size is provided, or the binomial method if population size is not specified.

### Calculate population sensitivity when unit sensitivity varies

Use this option to calculate population-level or cluster-level sensitivity (confidence of detection) and confidence of freedom when individual unit sensitivity varies. This is useful, for example, in circumstance where, for some reason, different units have been subjected to tests with different sensitivities, or in cases where herd sensitivity has been shown to vary in calculations of overall population-level sensitivity. For calculating cluster (herd, flock, etc.) sensitivity, enter test sensitivity and unit-level design prevalence, or for population-level sensitivity enter cluster-level (herd) sensitivity and cluster-level (herd) design prevalence. Design prevalence can be entered either as a proportion or as a specified number of units. If the population size is not specified, the binomial method is used in all cases; in all other cases, the hypergeometric approximation is used. This option requires input of data on test or cluster sensitivity for each unit sampled.

### Assuming imperfect specificity

Options under this section allow for calculation of sample size and population sensitivity for 1-stage surveys where the screening test has imperfect (< 100 percent) specificity and there is no follow-up of positive results. To account for imperfect specificity, a non-zero number of positive test results can occur and the overall result may still be considered negative. See Chapter 5 for descriptions of these methods.

*FreeCalc calculator for sample size for known or unknown population size*

Use this option to calculate the sample size required to achieve a target population or cluster-level sensitivity for a 1-stage survey when test specificity is < 100 percent. For cluster-level (herd-level, flock-level etc.) calculations, enter test sensitivity, unit-level design prevalence and required cluster-level sensitivity. Alternatively, for population-level calculations enter cluster-level (herd) sensitivity, cluster-level design prevalence and required population-level sensitivity. Design prevalence can be entered either as a proportion or as a specified number of units. Maximum allowable Type I and Type II errors can be specified (1 minus minimum population sensitivity and 1 minus minimum population specificity, respectively). Calculations can be specified to use either the hypergeometric approximation or the binomial (large population) method.

*Calculate population sensitivity for known or unknown population size*

Use this option to calculate population-level or cluster-level sensitivity (confidence of detection) for a 1-stage survey where unit specificity is assumed to be < 100 percent. For calculating cluster (herd, flock, etc.) sensitivity, enter test sensitivity and unit-level design prevalence, and for population-level sensitivity, enter cluster-level (herd) sensitivity and cluster-level (herd) design prevalence. Design prevalence can be entered either as a proportion or as a specified number of units. The cut-point number of reactors is the maximum number of positive results that can occur and for the population still to be considered free from infection. Calculations assume a large population size.

*FreeCalc calculator for population sensitivity*

This is another approach to calculate population or cluster-level sensitivity (confidence of detection) for a 1-stage survey where unit specificity is assumed to be < 100 percent. For calculating cluster (herd, flock, etc.) sensitivity, enter test sensitivity and unit-level design prevalence; for population-level sensitivity, enter cluster-level (herd) sensitivity and cluster-level (herd) design prevalence. Design prevalence can be entered either as a proportion or as a specified number of units. The cut-point number of reactors is the maximum number of positive results that can occur and for the population still to be considered free from infection. Calculations can be specified to use either the hypergeometric approximation or the binomial (large population) method.

### *Other options for 1-stage surveys*
### *Calculate design prevalence*

Use this option to calculate the design prevalence necessary to provide a specified population sensitivity (confidence of detection) for a given sample size. For example, assuming that you have sampled 140 units (instead of the target of 150), this utility can calculate the design prevalence that you would need to use in order to achieve 95 percent population sensitivity for that sample size and specified unit sensitivity. Enter test sensitivity if the target is a herd-level (cluster-level) sensitivity, or herd sensitivity if the target is a system-level or population-level sensitivity. Calculations use the hypergeometric approximation if population size is provided, or the binomial method if population size is not specified.

**Design prevalence re**

**Input Values**

| | |
|---|---|
| Test (or Herd) Sensitivity: | 0.9 |
| Sample Size: | 30 |
| Population size (leave blank or 0 if not known): | |
| Required herd (system or population) sensitivity: | 0.95 |

Submit          Reset

### *Calculate confidence of population freedom for a single time period*

Use this to calculate confidence of population freedom from disease for a single time period or one-off survey (see Chapter 15 for details). Inputs are the population-level sensitivity from the surveillance programme and the prior confidence of freedom status (i.e. your opinion of the situation that prevailed before the surveillance was carried out).

**Confidence of populati**

**Input Values**

| | |
|---|---|
| Population-level Sensitivity: | 0.9 |
| Prior confidence of freedom | 0.7 |

Submit

### *Calculate confidence of population freedom for multiple time periods*

Use this to calculate confidence of population freedom from disease for multiple time periods as part of an ongoing surveillance programme (see Chapters 15 and 16 for details). Inputs are the prior confidence of freedom for the first time period (before the surveillance started) and a spreadsheet table of the population-level sensitivity from the surveillance programme for each time period and the probability of disease introduction during each time period.

**Freedom Calculator**

**Confidence of population fr**

**Input Values**

*Submit Data*

| | |
|---|---|
| Initial prior confidence of freedom | 0.5 |

Data: Download example data

## 2-stage representative freedom surveys

This section provides utilities for designing and analysing 2-stage representative freedom surveys (see Chapter 5). These options all assume perfect (100 percent) test specificity after follow-up of positives with a definitive test. This section uses the same formulae as described for 1-stage surveys with perfect specificity. It includes a link to a file which provides details of all of the formulae used (http://epitools.ausvet.com.au/docs/Important-formulae-for-surveillance.pdf).

**Epi Tools - 2-Stage surveys for demonstration of freedom**

These utilities can be used to estimate sample sizes and to analyse results of 2-stage surveys for demonstrating disease freedom. For the first stage a sample of herds is selected while the second stage is to select a sample of animals from each selected herd. Current versions assume perfect test specificity but allow for imperfect sensitivity. Methods are also available for situations where a complete list of herd sizes is available and also where no information is available on the target population.

Click here for Some important formulae for freedom analyses.

**New**

• Calculate sample sizes (clusters and units per cluster) for a 2-stage freedom survey with specified herd sensitivity.

*Sample size calculation*

Use these options to calculate required sample sizes (clusters and units per cluster) assuming perfect test specificity:

   • **Least-cost sample sizes where herd sizes are known (and select clusters for testing).**
   • **Least-cost sample sizes where herd sizes are NOT known.**
   • **Sample sizes for specified herd (cluster) sensitivity.**

*Population sensitivity calculation (SSe)*

Use these options to calculate population-level sensitivity assuming **perfect** test specificity:

   • Analyse results of 2-stage freedom survey.
   • Stochastic analysis of 2-stage freedom survey.

## Sample size calculation

Options in this section allow for sample size calculation (numbers of clusters and numbers of units per cluster) for 2-stage representative surveys for disease freedom, assuming perfect test specificity (see Chapter 5 for details).

## Least-cost sample sizes where herd sizes are known

Use this option to calculate the number of herds and the number of animals within each herd to be tested, in order to provide a specified system sensitivity (probability of detecting disease) for the given animal-level and herd-level design prevalence and test sensitivity. Calculations are based on actual herd sizes provided and a list of randomly selected herds; in addition, the number of animals to sample for each selected herd is included in the outputs. Test specificity is assumed to be 100 percent (or follow-up testing of any positive will be undertaken to confirm or exclude disease). Sample sizes are optimized to minimize overall cost for given relative herd-level and animal-level testing costs. A maximum sample size per herd can be specified, if desired, and calculations can be specified to ensure either a fixed sample size per herd or a fixed (minimum) herd sensitivity.

**Calculate sample sizes for**

**Input Values**

| | |
|---|---|
| Animal-level design prevalence: | 0.1 |
| ⦿ Proportion | |
| ○ Animal(s) | |
| Herd-level design prevalence: | 0.01 |
| ⦿ Proportion | |
| ○ Herd(s) | |
| Test sensitivity: | 0.9 |
| Relative testing cost per herd: | 1000 |
| Relative testing cost per animal: | 20 |
| Target System Sensitivity (confidence): | 0.95 |
| Maximum sample size per herd: | 30 |
| ⦿ Aim for constant sample size per herd | |
| ○ Aim for constant (minimum) SeH | |

Sample sizes are calculated using the hypergeometric probability distribution (assuming sampling without replacement). Design prevalence (specified level of disease to be detected) must be specified at both animal and herd levels, and can be specified either as a proportion or as a number of units. A spreadsheet list of the number of animals in each herd must be entered.

### *Least-cost sample sizes where herd sizes are NOT known*

Use this option to calculate the number of herds and the number of animals within each herd to be tested in order to provide a specified system sensitivity (probability of detecting disease) where herd sizes are unknown. Test specificity is assumed to be 100 percent (or follow-up testing of any positive will be undertaken to confirm or exclude disease). Sample sizes are optimized to minimize the overall cost for given relative herd-level and animal-level testing costs. A maximum sample size per herd can be specified, if desired, and calculations can be specified to ensure either a fixed sample size per herd or a fixed (minimum) herd sensitivity.

**Calculate sample**

**Input Values**

Animal-level design prevalence: `0.1`
- ◉ Proportion
- ○ Animal(s)

Herd-level design prevalence: `0.01`
- ◉ Proportion
- ○ Herd(s)

Test sensitivity: `0.9`
Target Herd Sensitivity: `0.5`
Target System Sensitivity (confidence): `0.95`
Number of herds in the population (optional): `[    ]`

[ Submit ]

Sample sizes are calculated using the binomial probability distribution (assuming large population sizes relative to the number of animals sampled). Design prevalence (specified level of disease to be detected) must be specified at both animal and herd levels, and can be specified either as a proportion or as a number of units.

### *Least-cost sample sizes for specified herd (cluster) sensitivity*

Use this option to calculate sample sizes for 2-stage surveys for demonstrating disease freedom using a specified cluster-level sensitivity. This analysis calculates the number of herds and the number of animals within each herd that need to be tested in order to provide specified herd/cluster and system sensitivities (probability of detecting disease) for the given animal-level and herd-level design prevalence and test sensitivity. Test specificity is assumed to be 100 percent (or follow-up testing of any positive will be undertaken to confirm or exclude disease). Design prevalence (specified level of disease to be detected) must be specified at both animal and herd levels, and can be specified either as a proportion or as a number of units.

Numbers of herds to test are calculated using the hypergeometric approximation, if the number of herds in the population is specified, as well as using the binomial formula assuming an unknown (large) number of herds in the population. If the population size is not specified, only the binomial results are presented. Numbers of animals to test in each herd are calculated for a range of herd sizes using the hypergeometric approximation and, for unknown (large) herd sizes, using the binomial calculation.

### *Calculate population sensitivity*

Options in this section allow for calculation of population (system) sensitivity for 2-stage representative surveys for disease freedom, assuming perfect test specificity (see Chapter 5 for details). These options require input of testing results for each herd or cluster sampled.

*Analysis of 2-stage freedom survey*

Use this option to analyse herd testing data for 2-stage surveys for demonstrating disease freedom. This analysis calculates the overall system sensitivity for the survey and the resulting probability of population freedom from disease. It assumes that a random sample of herds (or all herds) has been selected for testing from the population and that a random sample of animals (or all animals) has been tested within each selected herd. It also assumes that the test system has a specificity of 100 percent (any positive results are further investigated to exclude false positives) and that no positive results were recorded. The analysis adjusts for imperfect sensitivity of the test used.

The analysis calculates both herd-level and system-level sensitivity estimates, using either binomial or hypergeometric methods, depending on the available data. Design prevalence (specified level of disease to be detected) must be specified at both animal and herd levels, and can be specified either as a proportion or as a number of units. A spreadsheet table listing numbers of units tested per cluster and, optionally, population size for each cluster is required.

*Stochastic analysis of 2-stage freedom survey*

This option is similar to the previous option and can be used to analyse herd testing data for 2-stage surveys for demonstrating disease freedom. The main difference is that this option allows for a stochastic analysis, with test sensitivity and prior confidence of freedom entered as probability distributions, and with outputs similarly presented as probability distributions. The analysis calculates the overall system sensitivity for the survey, and the resulting probability of population freedom from disease, in a similar manner and with similar inputs to the previous (deterministic) method.

## Risk-based surveillance

This section provides a number of utilities to help design and analyse surveillance using risk-based sampling (Chapters 6 to 13). The tools provided here are suitable for relatively simple systems using either 1-stage or 2-stage sampling with a single risk factor at each of the unit and cluster levels. All methods assume 100 percent test specificity, and also assume that populations are large relative to the sample size (binomial sampling). Scenario trees may be useful to help improve understanding and visualization of the surveillance systems being analysed, but they are not essential for these analyses. All analyses are deterministic and use mean or average values rather than probability distributions.

## 1-stage risk-based surveys

Options in this section are for simple 1-stage surveys (for example, animals in a herd or herds in a larger population) with a single risk factor and/or variation in sensitivity.

### Sample size for single risk factor and single level of sensitivity

Use this option to calculate the sample size for simple risk-based surveillance, for example, a survey in which a high-risk population is targeted. This analysis assumes that there is no clustering of disease (for instance, that we are working at the herd level), and that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered; in order to do so, the following information is required:



- The relative risk is the risk of animals being infected in the high-risk group, relative to the risk of animals being infected in the low-risk group. The population proportion is the proportion of animals from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of animals from the surveillance **exercise** that are in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion), the individual unit (animal or herd) sensitivity and the target surveillance sensitivity required.

## Sample size for single risk factor and two levels of test sensitivity

Use this option to calculate the sample size for simple risk-based surveillance with a single risk factor and where two different values for test sensitivity are possible (for example, using two different tests). This analysis assumes that there is no clustering of disease (for instance, that we are working at the herd level), and that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered; in order to do so, the following information is required:

**Simple risk-based surv⟨**
**Input Values**

| | |
|---|---|
| Relative risk : | 3 |
| Population proportion in high risk group : | 0.2 |
| Surveillance proportion in high risk group : | 0.7 |
| Surveillance proportion with high sensitivity in high risk group : | 0.8 |
| Surveillance proportion with high sensitivity in low risk group : | 0.6 |
| Design prevalence : | 0.05 |
| Sensitivity in high-sensitivity group : | 0.95 |
| Sensitivity in low-sensitivity group : | 0.8 |
| Target surveillance sensitivity : | 0.95 |

Submit

- The relative risk is the risk of animals being infected in the high-risk group, relative to the risk of animals being infected in the low-risk group.
- The population proportion is the proportion of animals from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of animals from the surveillance **exercise** that are in the high-risk group.

Other parameters that are required include: the design prevalence (as a proportion); the individual unit (animal or herd) sensitivity for both sensitivity levels; target proportions for the sample in high-risk/high-sensitivity and low-risk/high-sensitivity groups, and the target surveillance sensitivity required.

## Population sensitivity for single risk factor and single level of sensitivity

Use this option to calculate the population sensitivity and confidence of freedom for simple risk-based surveillance, for example, a survey in which a high-risk population is targeted. This analysis assumes that there is no clustering of disease (for instance, that we are working at the herd level), and that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered; in order to do so, the following information is required:

**Simple risk-based surv**
**Input Values**

| | |
|---|---|
| Relative risk : | 3 |
| Population proportion in high risk group : | 0.2 |
| Surveillance proportion in high risk group : | 0.7 |
| Design prevalence : | 0.05 |
| Test sensitivity : | 0.95 |
| Number tested : | 50 |
| Prior confidence of freedom : | 0.5 |

Submit

- The relative risk is the risk of animals being infected in the high-risk group, relative to the risk of animals being infected in the low-risk group.
- The population proportion is the proportion of animals from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of animals from the surveillance **exercise** that are in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion), the individual unit (animal or herd) sensitivity, the number of units tested, the target surveillance sensitivity required and the prior confidence of freedom before undertaking the surveillance.

### Sample size for single risk factor and two levels of test sensitivity

Use this option to calculate the population sensitivity and confidence of freedom for simple risk-based surveillance with a single risk factor and where two different values for test sensitivity are possible (for example, using two different tests). This analysis assumes that there is no clustering of disease (for instance, that we are working at the herd level), and that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered; in order to do so, the following information is required:

| Simple risk-based su |  |
| --- | --- |
| **Input Values** | |
| Relative risk : | 3 |
| Population proportion in high risk group : | 0.2 |
| Design prevalence : | 0.05 |
| Sensitivity in low-sensitivity group : | 0.8 |
| Sensitivity in high-sensitivity group : | 0.95 |
| Prior confidence of freedom : | 0.5 |
| Number tested in high-risk high-sensitivity group : | 20 |
| Number tested in high-risk low-sensitivity group : | 10 |
| Number tested in low-risk low-sensitivity group : | 10 |
| Number tested in low-risk high-sensitivity group : | 5 |
| Submit | |

- The relative risk is the risk of animals being infected in the high-risk group, relative to the risk of animals being infected in the low-risk group.
- The population proportion is the proportion of animals from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of animals from the surveillance operation that are in the high-risk group.

Other parameters that are required include: the design prevalence (as a proportion); the individual unit (animal or herd) sensitivity for both sensitivity levels; the numbers of units sampled in each of the high-risk/high-sensitivity, high-risk/low-sensitivity, low-risk/high-sensitivity and low-risk/low-sensitivity groups; the target surveillance sensitivity required and the prior confidence of freedom before undertaking the surveillance.

### Calculate minimum detectable prevalence for single risk factor and single level of sensitivity

Use this option to calculate the minimum detectable prevalence (the design prevalence required to give a specified level of population sensitivity) for simple risk-based surveillance. This analysis assumes that there is no clustering of disease (for instance, that we are working at the herd level), and that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered, for which the following information is required:

| Simple risk-based surve |  |
| --- | --- |
| **Input Values** | |
| Relative risk : | 3 |
| Population proportion in high risk group : | 0.2 |
| Surveillance proportion in high risk group : | 0.7 |
| Target Surveillance (population) sensitivity : | 0.95 |
| Test sensitivity : | 0.95 |
| Number tested : | 50 |
| Submit | |

- The relative risk is the risk of animals being infected in the high-risk group, relative to the risk of animals being infected in the low-risk group.
- The population proportion is the proportion of animals from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of animals from the surveillance that are in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion), the individual unit (animal or herd) sensitivity, the number of units tested and the target surveillance sensitivity required.

### 2-stage risk-based surveys with single risk factor

Options in this section are for simple 2-stage surveys with a single risk factor acting at the cluster (herd/flock) level only.

### Sample size for single risk factor at the cluster level

Use this option to calculate the sample sizes for simple, 2-stage risk-based surveillance, for example, a survey in which high-risk herds are targeted, but sampling within herds is representative. This analysis accounts for clustering of disease in herds or flocks, and assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered at the cluster level only; for this, the following information is required:

- The relative risk is the risk of *clusters* being infected in the high-risk group, relative to the risk of *clusters* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of *clusters* from the surveillance operation that are in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion), the individual unit (animal) test sensitivity, the desired cluster-level (herd) sensitivity and the target surveillance sensitivity required. Increasing the cluster sensitivity will increase the number of units to be tested per cluster, but will decrease the numbers of clusters; decreasing target cluster sensitivity will do the reverse.

### Population sensitivity and confidence of freedom for single risk factor at the cluster level

Use this option to calculate population sensitivity and confidence of freedom for simple 2-stage risk-based surveillance, for example, a survey in which high-risk herds are targeted, but where sampling within herds is representative. This analysis accounts for clustering of disease in herds or flocks, and assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives). One risk factor is considered at the cluster level only; for this, the following information is required:

- The relative risk is the risk of *clusters* being infected in the high-risk group, relative to the risk of *clusters* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of *clusters* from the surveillance that are in the high-risk group.

Other parameters that are required include the design prevalence at both unit and cluster levels (as proportions), the individual unit (animal) test sensitivity, the numbers of high-risk and low-risk clusters tested, the number of units tested per cluster (assumed to be constant) and the prior confidence of freedom before undertaking the surveillance.

*Population sensitivity and confidence of freedom for single risk factor at the cluster level using actual testing data*

This option is similar to the previous one, except that instead of assuming that within-cluster sample size is constant, it allows for entry of actual data on risk status (high or low), number of units tested and (optionally) cluster or herd size for each cluster sampled. One risk factor is considered at the cluster level only; for this, the following information is required:

- The relative risk is the risk of *clusters* being infected in the high-risk group, relative to the risk of *clusters* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* from the entire population that are in the high-risk group.
- The surveillance proportion is the proportion of *clusters* from the surveillance exercise that are in the high-risk group.

Other parameters that are required include the design prevalence at both unit and cluster levels (as proportions), the individual unit (animal) test sensitivity, the prior confidence of freedom before undertaking the surveillance and a spreadsheet table of data for each cluster sampled.

**2-stage surveys with risk factors at both cluster and unit levels**

Options in this section allow for more complex 2-stage surveys with two risk factors acting at both cluster (herd/flock) and unit (animal) levels. Options are also provided that allow for two levels of sensitivity varying between risk groups.

*Sample size for 2-stage surveys with risk factors at cluster and unit levels and varying sensitivity*

Use this option to calculate the sample sizes for more complex 2-stage risk-based surveillance, for example, a survey in which sampling at the herd level is targeted at high-risk herds, and sampling within herds is targeted at high-risk animals. This analysis assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives) and that test sensitivity can have two different values, independent of risk factors. Risk factors are considered at both cluster and unit levels and are assumed to be independent. In order to consider risk factors, the following information is required:

- The relative risk is the risk of *clusters* or *units* being infected in the high-risk group, relative to the risk of *clusters* or *units* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* (or *units*) from the population (or cluster) that is in the high-risk group.

- The surveillance proportion is the proportion of *clusters* (or units) from the surveillance exercise that is in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion) at unit and cluster levels, the individual unit (animal) test sensitivity for high-sensitivity and low-sensitivity groups, the surveillance proportions in high-sensitivity and low-sensitivity groups and the target surveillance sensitivity required.

### Sample size for 2-stage surveys with risk factors at cluster and unit levels, and sensitivity correlated with risk group

This option is very similar to the previous one, except that in this case, sensitivity varies between high-risk and low-risk groups of animals, rather than being independent of risk. The analysis assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives) and that test sensitivity can have two different values, for high-risk and low-risk units (animals). Risk factors are considered at both cluster and unit levels and are assumed to be independent. The following information is required for risk factors:

- The relative risk is the risk of *clusters* or *units* being infected in the high-risk group, relative to the risk of *clusters* or *units* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* (or units) from the population (or cluster) that is in the high-risk group.
- The surveillance proportion is the proportion of *clusters* (or units) from the surveillance that is in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion) at unit and cluster levels, the individual unit (animal) test sensitivity for high-risk and low-risk groups, the desired cluster-level (herd) sensitivity and the target surveillance sensitivity required.

**Complex risk-based su**

**Input Values**

*Herd-level risk factor*

| | |
|---|---|
| Relative risk : | 3 |
| Population proportion in high risk group : | 0.2 |
| Surveillance proportion in high risk group : | 0.7 |
| Design prevalence : | 0.1 |

*Animal-level risk factor*

| | |
|---|---|
| Relative risk : | 2 |
| Population proportion in high risk group : | 0.5 |
| Surveillance proportion in high risk group : | 0.9 |
| Design prevalence : | 0.1 |

*Diagnostic test*

| | |
|---|---|
| Population proportion in high sensitivity group : | 0.4 |
| Surveillance proportion in high sensitivity group : | 0.8 |
| Sensitivity in high sensitivity group : | 0.95 |
| Sensitivity in low sensitivity group : | 0.9 |
| Number tested : | 20 |

Submit

### Population sensitivity and confidence of freedom for 2-stage surveys with risk factors at cluster and unit levels and varying sensitivity

Use this option to calculate population sensitivity and confidence of freedom for more complex 2-stage risk-based surveillance, for example, a survey in which sampling at the herd level is targeted at high-risk herds and sampling within herds is targeted at high-risk animals. This analysis assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives) and that test sensitivity can have two different values, independent of risk factors. Risk factors are considered at both cluster and unit levels, and are assumed to be independent. The following information is required for risk factors:

- The relative risk is the risk of *clusters* or *units* being infected in the high-risk group, relative to the risk of *clusters* or *units* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* (or units) from the population (or cluster) that is in the high-risk group.

**Complex 2-stage**

**Input Values**

*Herd-level risk factor*

| | |
|---|---|
| Relative risk : | 3 |
| Population proportion in high risk group : | 0.2 |
| Design prevalence : | 0.05 |

*Animal-level risk factor*

| | |
|---|---|
| Relative risk : | 2 |
| Population proportion in high risk group : | 0.3 |
| Design prevalence : | 0.2 |

*Diagnostic test*

| | |
|---|---|
| Sensitivity in high risk group : | 0.95 |
| Sensitivity in low risk group : | 0.9 |
| High-risk herds tested : | 20 |
| Low-risk herds tested : | 10 |
| High-risk animals tested per herd : | 6 |
| Low-risk animals tested per herd : | 4 |
| Prior confidence of freedom : | 0.5 |

Submit

- The surveillance proportion is the proportion of *clusters* (or units) from the surveillance that is in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion) at unit and cluster levels, the individual unit (animal) test sensitivity for high-sensitivity and low-sensitivity groups, the population and surveillance proportions in high-sensitivity and low-sensitivity groups, and the number of units tested.

### Population sensitivity and confidence of freedom for 2-stage surveys with risk factors at cluster and unit levels and sensitivity correlated with risk group

This option is very similar to the previous one, except that in this case, sensitivity varies between high-risk and low-risk groups of animals, rather than being independent of risk. The analysis assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives) and that test sensitivity can have two different values, for high-risk and low-risk units (animals). Risk factors are considered at both cluster and unit levels, and are assumed to be independent. In order to consider risk factors, the following information is required:

**Complex 2-stage**

**Input Values**

*Herd-level risk factor*

Relative risk : 3

Population proportion in high risk group : 0.2

Design prevalence : 0.02

*Animal-level risk factor*

Relative risk : 2

Population proportion in high risk group : 0.4

Design prevalence : 0.1

*Diagnostic test*

Sensitivity in low-risk animals : 0.8

Sensitivity in high-risk animals : 0.95

Prior confidence of freedom : 0.5

Submit

Paste herd testing data in the *'LRTested')* and herd risk-grou be included.
Download example data

- The relative risk is the risk of *clusters* or *units* being infected in the high-risk group, relative to the risk of *clusters* or *units* being infected in the low-risk group.
- The population proportion is the proportion of *clusters* (or units) from the population (or cluster) that is in the high-risk group.
- The surveillance proportion is the proportion of *clusters* (or units) from the surveillance that is in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion) at unit and cluster levels; the individual unit (animal) test sensitivity for high-risk and low-risk groups; the numbers of high-risk and low-risk herds sampled, and the numbers of high-risk and low-risk animals sampled in each herd (assumed to be constant).

### Population sensitivity and confidence of freedom for 2-stage surveys with risk factors at cluster and unit levels and actual testing data

This option is similar to the previous options, except that in this case the analysis is based on actual testing data for each herd. The analysis assumes that the effective specificity of the surveillance system is equal to 100 percent (all positives are followed up to ensure that they are not false positives) and that test sensitivity can have two different values, for high-risk and low-risk units (animals). Risk factors are considered at both cluster and unit levels, and are assumed to be independent. The following information is required for risk factors:

- The relative risk is the risk of *clusters* or *units* being infected in the high-risk group, relative to the risk of *clusters* or *units* being infected in the low-risk group.

- The population proportion is the proportion of *clusters* (or units) from the population (or cluster) that is in the high-risk group.
- The surveillance proportion is the proportion of *clusters* (or units) from the surveillance that is in the high-risk group.

Other parameters that are required include the design prevalence (as a proportion) at unit and cluster levels, the individual unit (animal) test sensitivity for high-risk and low-risk groups and prior confidence of freedom. Actual testing data are entered as a spreadsheet table, with one row for each herd tested, including herd risk group and numbers of high-risk and low-risk animals tested.

## Random sampling from populations

This section provides four utilities that can be used to assist in generating random samples from a population. For representative freedom surveys, random sampling is the best way to ensure that the resulting sample is representative of the population. Even for risk-based surveys, it is important that sampling within risk groups is representative of the group, so, again, random sampling is important.

### Random sampling from a sampling frame

Use this option if you have a sampling frame from which to select the sample: for example a list of farms or cattle owners. The list (sampling frame) needs to be in electronic spreadsheet format, with one row per unit to be sampled. Options for sampling include:

- simple random sampling or probability proportional to size;
- sampling with or without replacement;
- stratification (proportional to stratum size) by a stratifying factor (must be included as an additional column in the data);
- subsetting of the list according to an additional factor (for example only those with > 20 animals).

The required sample size must be specified and the sampling frame must be entered in spreadsheet format.

### Random geographic coordinates sampling

Use this option to generate the specified number of random geographic coordinates from a rectangular space defined by minimum and maximum X (longitude) and Y (latitude) coordinates. Input and output coordinates are in decimal degrees, with a specified precision of output values.

### *Random sampling of animals*

Use this option to generate a random sample of individuals from a sampling frame of animal owners in a village, or farms in a district/region. A full sampling frame of all animals is generated, with animals for each owner numbered sequentially. Simple random sampling is then carried out to generate a sample, with animals identified first by owner, and then by sequential number for each owner.

Sampling can be based on either a specified number of animals to be sampled or a specified percentage of the total number of animals in the sampling frame.

A list of all animal owners and the number of animals owned by each owner is required for the generation of the sampling frame. This can be copied and pasted from Excel or other software, and it should be formatted as two columns: the first column is the unique owner identification number and the second column is the number of animals for each owner.

### *Generate a list of random numbers*

Generate the specified number of random numbers from either a specified range (defined by minimum and maximum values) or from a submitted list of numbers. To choose numbers from a list, the list of numbers should be entered as a single column of values; alternatively, it can be pasted from a spreadsheet. Sampling can be carried out with or without replacement.

# Appendices

## GLOSSARY

**Acceptable level of protection**
In risk analysis, the level of protection that a country sets, and against which the results of risk analyses are judged.

**Active surveillance**
Surveillance in which the primary users of the surveillance data (usually the veterinary authorities) initiate and design the data collection.

**Adjusted risk**
This is a measure of risk in a specified branch of a risk category node. The ratio of the adjusted risks between two branches has the same value as the relative risk, but the adjusted risks are adjusted to ensure that the average risk in the population remains equal to one.

**Bayes' theorem**
A probability formula that enables prior knowledge to be combined with new information to give an updated (posterior) probability estimate. In probability notation, it is expressed as:

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \times \Pr(A)}{\Pr(B)}$$

**Bias**
If a survey procedure were repeated using the same methodology many times, bias is the difference between the true value and the mean of the estimated values.

**Branch**
In a scenario tree, 'branch' is one of several outcomes from a node. For instance, a node *age* may have two branches: *old* and *young*.

**Category node**
A node in a scenario tree that divides the population into two or more categories according to some criterion. The main types of category nodes are *risk category nodes* (which divide the population into groups that have different probabilities of being infected) and detection category nodes (which divide the population into groups with different probabilities of being detected).

**Census**
A survey or other surveillance activity that examines every member of the population (as opposed to examining just a sample of the population).

| | |
|---|---|
| **Clustering** | The phenomenon whereby disease is not evenly distributed in a population but instead forms pockets of high prevalence among areas of no disease. |
| **Component sensitivity** | The sensitivity of a component of a surveillance system – the probability that the component will detect disease if the population is infected at or above the design prevalence. Also called surveillance system component sensitivity (SSCSe). |
| **Conditional probability** | The probability of an event occurring, given that another event is known to have occurred. |
| **Design prevalence** | A hypothetical prevalence of disease, against which the surveillance system is evaluated. |
| **Detection category node** | A node in a scenario tree that divides the population into groups with different probabilities of being detected. |
| **Detection node** | A node in a scenario tree relating to a step in a surveillance procedure that is necessary in order for a case of disease to be detected. The probability of success at this step is the sensitivity. |
| **Effective probability of infection** | The design prevalence multiplied by the adjusted risk, giving the probability that a particular group in a scenario tree will be infected. |
| **General surveillance** | Surveillance that is able to detect many diseases or any disease (in contrast to surveillance that is targeted at detecting only one disease). |
| **Herd sensitivity, group sensitivity** | The probability of detecting at least one infected animal when a herd is examined. |
| **Independence** | In probability, two events are independent if the outcome of one is not influenced by the outcome of the other. |
| **Infection node** | In a scenario tree, a node which describes the probability that an animal or group of animals will be infected. The probability in the *infected* branch of an infected node is the design prevalence. |
| **Monte Carlo simulation** | An analytical technique involving repeating an analysis many times, using different input parameters drawn randomly from defined distributions. |

| | |
|---|---|
| **Node** | In a scenario tree, a node represents a factor that may take two or more values, each with assigned probabilities. |
| **Passive surveillance** | An activity in which the primary purpose of the collection of the data is *not* surveillance. |
| **Population proportion** | The proportion of the entire population that has some defined characteristic of interest (in contrast to the surveillance system component proportion). |
| **Posterior probability** | An estimate of the probability of an event, calculated using Bayes' theorem, based on prior knowledge and new information. |
| **Prior probability** | When using Bayes' theorem, an estimate of the probability of an event occurring, before new information about the event has been collected. |
| **Random error** | In sampling, error due to the random effect of selecting one animal or another. Random error leads to lack of precision that can be minimized by using a large sample size. |
| **Relative risk** | The probability of an event occurring in one part of the population, divided by the probability of it occurring in another. Also known as risk ratio. |
| **Risk** | The probability of an adverse event occurring. In this manual, risk is defined simply as a probability, in contrast to its use in risk analysis, where it is likelihood combined with consequences. |
| **Risk category node** | In a scenario tree, a node that classifies the population into two or more groups, each with a different risk of being infected. |
| **Scenario tree** | A branching quantitative model used for the analysis of surveillance system components. |
| **Sensitivity** | The probability of getting the right answer from a test on an *infected* population. The true positive rate. |
| **Specificity** | The probability of getting the right answer from a test on an *uninfected* population. The true negative rate. |
| **Stochastic** | Describes a process involving chance. |
| **Surveillance system** | When applied to surveillance for a particular disease, the collection of activities that produce data which contribute to our understanding about the status of that disease. |

| | |
|---|---|
| **Surveillance system component** | A component of a surveillance system. A single activity that produced data about disease status. Abbreviated as SSC. |
| **Surveillance system component proportion** | The proportion of animals or herds in a surveillance system component that have a characteristic of interest, in contrast to the population proportion. |
| **Surveillance system sensitivity** | The probability that the surveillance system would detect disease if the population is infected at or above the design prevalence. |
| **Syndrome** | A defined collection of clinical signs possible with other epidemiological information. |
| **Systematic error** | An error in surveys or surveillance that results in the expected value (mean value of many repetitions of the activity) being different from the true population value. Systematic error causes bias or lack of accuracy, and may be caused by sampling bias, measurement bias, analysis bias or confounding. |
| **Targeted surveillance** | 1. Surveillance aimed at detecting a specific disease, as opposed to general surveillance.<br>2. Surveillance targeted at a portion of the population. Risk-based surveillance.<br>The two different usages are unfortunate, but context usually makes it possible to determine which meaning is intended. |
| **Unit sensitivity** | A general term for the probability that a single unit passing through the surveillance system would be detected as being infected, assuming that the population is infected at or above the design prevalence. |

# Abbreviations and symbols

**AR**   Adjusted risk. This is a measure of risk in a specified branch of a risk category node. The ratio of the adjusted risks between two branches has the same value as the relative risk, but ARs are adjusted to ensure that the average risk in the population remains equal to one.

**CSe**   Surveillance system component sensitivity (SSCSe)

**CSeU**   Component unit sensitivity. This is a general term for the probability that a single unit passing through the surveillance system would be detected as being infected, assuming that the population is infected at the design prevalence.

**EPI**   The effective probability of infection. This is the design prevalence multiplied by the adjusted risk.

**GSe**   Group sensitivity. The probability that disease would be detected in a group of animals.

**P(x)**   The probability of event *x*

**P\***   The design prevalence. This is a hypothetical prevalence of disease against which the surveillance system is evaluated.

**P\*$_A$**   Animal-level design prevalence. This is the proportion of animals infected within an infected herd.

**P\*$_H$**   Herd-level design prevalence. This is the proportion of herds infected within the population.

**PrP**   The population proportion. This is the proportion of animals in the study population that fall into a specified group as defined by a category node.

**PrSSC**   The surveillance system component proportion. This is the proportion of animals in the surveillance system component (SSC) that fall into a specified group as defined by a category node.

**RR**   Relative risk (also known as risk ratio).

**Se**   Sensitivity (true positive rate)

**SeA**          The animal-level sensitivity. This is the same as the unit sensitivity when the animal is the unit of analysis (which is the most common case in livestock applications).

**SeH**          Herd sensitivity. This is the same as group sensitivity when the herd is the grouping level (which is the most common case in livestock applications).

**Sp**           Specificity (true negative rate)

**SS**           Surveillance system

**SSC**          Surveillance system component

**SSe**          Surveillance system sensitivity. The probability that the surveillance system would detect disease if the population is infected at or above the design prevalence.

# Important formulae for surveillance

This summary of key formulae for surveillance was compiled by Dr Evan Sergeant, AusVet Animal Health Services, and has been reproduced here with his permission.

## REPRESENTATIVE SURVEILLANCE FOR DISEASE FREEDOM

### Terminology

In this manual some specific terminology relating to unit, cluster and population values have been used in order to try and simplify the formulae presented. These terms are explained here:

A **unit** is either an individual (animal, plant, etc.) as part of a cluster, or a cluster as part of a larger population.

A **cluster** is a grouping level of individual units (animals, plants, fish, etc.) at a higher level, such as a herd, flock, tank, pen, farm, etc. Usually, clusters are only considered at one level, but can occur at multiple levels (for example, pens within farms within districts).

**Unit sensitivity** is the sensitivity at the unit level for a particular analysis. For cluster-level analyses, unit sensitivity is the sensitivity of the test (or combination of tests) used, whereas for population-level analyses, unit sensitivity is the cluster-level sensitivity for clusters sampled.

**Population sensitivity** is a sensitivity calculated at some population or grouping level. Depending on the context, the population can be either a cluster of multiple individuals or a larger population comprising multiple clusters.

**Component sensitivity** is a population-level sensitivity, usually at a country or regional level, calculated for one part (component) of a surveillance system which comprises multiple separate components or activities.

**System sensitivity** is a population-level sensitivity, usually at a country or regional level, calculated from one or more components

### Population sensitivity and sample size

***Binomial***
*Population sensitivity*

$$SeP = 1 - (1 - SeU \times P^*U)^n$$

$$= 1 - \prod (1 - SeU_i \times P^*U)$$

where SeU varies among units

*Sample size:*

$$n = \log(1 - SeP)/\log(1 - SeU \times P^*U)$$

*Assumes:*
- sampling with replacement or sample small (< 10 percent) relative to population
- specificity = 100%

**Hypergeometric approximation**
*Population sensitivity:*

$$SeP = 1 - (1 - SeU \times n/N)^d$$

$$SeP = 1 - (1 - SeUavg \times n/N)^d$$

where SeU varies among units

*Sample size:*

$$n = (N/SeU)^*(1 - (1 - SeP)^{1/(P^* \times N)})$$

*Assumes:*
- sampling without replacement or where sample size is large relative to population
- specificity = 100%

**Exact**
*Population sensitivity:*

$$SeP = 1 - (1 - SeU)^d$$

$$SeP = 1 - (1 - SeUavg)^d$$

where SeU varies among units

*Assumes:*
- sampling of the entire population
- specificity = 100%

## Negative predictive value (confidence of population freedom: PFree)

$$PFree = (1 - PrInf)/(1 - PrInf \times SeP) \text{ or}$$

$$= PriorPFree/(1 - SeP \times (1 - PriorPFree))$$

assuming specificity = 100%

## Revising confidence of freedom in successive time periods

$$PFree_t = 1 - [1 - PFree_{t-1} + PIntro_t - ((1 - PFree_{t-1}) \times PIntro_t)]$$

## Equilibrium PFree

Maximum or minimum stable value for PFree for given combinations of SeP and PIntro

$$PFree_{equ} = (1 - (PIntro / SeP)) / (1 - PIntro)$$

Maximum or minimum value for PriorPFree (after discounting) for given combinations of SeP and PIntro

$$PriorPFree_{equ} = 1 - (PIntro / SeP)$$

## Design prevalence to achieve specified population sensitivity

Where cluster size is unknown (binomial):

$$P^*U = (1 - \exp((\log(1 - SeP))/n))/SeU$$

Where cluster size is known (hypergeometric approximation):

$$P^*U = \log(1 - SeP)/\log(1 - SeU \times n/N)/N$$

## Population sensitivity required to achieve desired PFree

$$SeP = (1 - PriorPFree/PFree)/(1 - PriorPFree)$$

where PFree is the target value and PriorPFree is the current prior value.

## Population sensitivity required to stay above specified threshold PFree

$$SeP = PIntro/(1 - \text{Target PFree})$$

## Combining test sensitivities in series
(For example, in a diagnostic process with multiple steps)

$$SeUcombined = \prod Se_i$$

## Combining component sensitivities in parallel, assuming independence
Calculates system sensitivity from multiple components, assuming independence (no overlap between units sampled) between components, for example, different compartments or different clusters represented in the surveillance system.

$$SeP = 1 - \prod (1 - CSe_i)$$

## Updating cluster sensitivities between components where there is overlap
This assumes no independence between components, for example, where the same clusters (herds or flocks, etc.) are represented in multiple surveillance system components. The probability of infection for each cluster is adjusted between components and resulting component sensitivities are then combined as for assuming independence. For this example, binomial calculations are used, but hypergeometric or exact **measurements** could also be used, if appropriate:

### *Method 1: Adjusting effective probability of infection between components*
1. Calculate SeC for each cluster [$SeC = 1 - (1 - SeU \times P^*)^n$] for each component.

2. Calculate posterior confidence of freedom and hence posterior probability of infection for each cluster for the first component (component order is a matter of convenience):

$$PFree_c = (1 - P^*)/(1 - P^*_c \times SeC)$$

where $P^*_c$ is the cluster-level design prevalence

$$PostPInf_c = (1 - PFree_c)$$

3. Calculate probability that each cluster has a negative test result and hence component sensitivity (CSe) for first component:

$$P(Neg) = 1 - P^*_c \times SeC$$

$$CSe = 1 - \prod (P(Neg))$$

4. Calculate P(Neg) for each cluster and CSe for the second component after substituting $PostPInf_h$ instead of P* in formula:

$$P(Neg) = 1 - PostPInf_h \times SeC_2$$

$$CSe = 1 - \prod (P(Neg))$$

5. Repeat for as many components as necessary.

6. Clusters start with P* at the first component in which they appear and then get updated as necessary.

7. When all component sensitivities have been calculated, calculate overall system sensitivity (probability that one or more components will yield a positive result if the population is infected at the design prevalence), using independence formula.

$$SSe = 1 - \prod (1 - CSe_i)$$

### *Method 2: Aggregating data between components*
An alternative (often simpler) approach is to aggregate the data for each cluster in order to calculate single SeC values, and then combine these values to calculate overall system sensitivity:

$$SeC = 1 - \prod((1 - P\text{*} \times SeU_i)^{n_i})$$

For where $SeU_i$ and $n_i$ are test sensitivity and sample size for each of the i components in the surveillance system.

| Abbreviation/symbol | Meaning |
|---|---|
| **n, N** | Sample size and corresponding population size |
| **d** | Number of diseased elements in a population |
| **t** | Time period |
| **P*U** | Unit-level design prevalence (individual or cluster) |
| **Se** | Test sensitivity |
| **SeU** | Unit-level sensitivity (test sensitivity when calculating cluster/herd-level sensitivity or cluster/herd-level sensitivity when calculating population or component sensitivity) |
| **SeP** | Population sensitivity (can be cluster level or overall) |
| **SeC** | Cluster sensitivity |
| **SeCi** | Cluster sensitivity for the i-th cluster |
| **SeUavg** | Average unit sensitivity across all units (individuals or clusters) |
| **CSe$_i$** | Component sensitivity for the i-th surveillance system component |
| **SSe** | System sensitivity |
| **PFree** | Confidence of population freedom (= negative predictive) value) |

| **PriorPFree** | Confidence of population freedom before undertaking current surveillance |
|---|---|
| **PrInf** | Prior probability of being infected = 1 – prior confidence of freedom |
| **PostPInf** | Posterior probability of being infected = 1 – posterior confidence of freedom freedom (NPV) |

## RISK-BASED FREEDOM SURVEILLANCE

### Adjusted risk and effective probability of infection

$$ARL = 1/(RR \times PPrH + PPrL)$$

$$ARH = RR \times ARL$$

or for multiple risk levels:

$$ARi = RRi / \Sigma( RR \times PPr)$$

$$EPI = P\text{*} \times AR \text{ (for respective risk categories)}$$

EPI > 1 is invalid – design prevalence and/or relative risk should be revised to ensure EPI < 1.

### Population sensitivity for simple, 1-stage, no risk factors, one factor affecting sensitivity

Assuming large population relative to sample size (binomial) and only two unit sensitivity values:

$$SeP = 1 - (1 - P\text{*} \times SeUH)^{n(h)} \times (1 - P\text{*} \times SeUL)^{n(l)}$$

n(h) and n(l) are sample sizes for high-sensitivity and low-sensitivity groups, respectively; or assuming small population:

$$SeP = 1 - (1 - SeU_{avg} \times n/N)^d$$

### Sample size for simple, 1-stage, one risk factor (two levels), constant sensitivity

$$USe = EPI_H \times SeU_H \times SPr_H + EPI_L \times SeU_L \times SPr_L$$

$$n = \log(1 - SeP)/\log(1 - USe)$$

SeUH and SeUL are the mean values for SeU for high-risk and low-risk groups respectively.

## Population sensitivity for simple, 1-stage, one risk factor, one factor affecting sensitivity

$$SeP = 1 - (1 - EPI_H \times SeU_H)^{n(hh)} \times (1 - EPI_H \times SeU_L)^{n(hl)} \times (1 - EPI_L \times SeU_H)^{n(lh)} \times (1 - EPI_L \times SeU_L)^{n(ll)}$$

n(hh), n(hl),n(lh) and n(ll) are sample sizes for high-risk and high-sensitivity, high-risk and low-sensitivity, low-risk and high-sensitivity and low-risk and low-sensitivity groups, respectively.

## Sample size for simple, 1-stage, one risk factor, one factor affecting sensitivity

$$LRSe = SPr_{LH} \times SeU_H + (1 - SPr_{LH}) \times SeU_L$$

$$HRSe = SPr_{HH} \times SeU_H + (1 - SPr_{HH}) \times SeU_L$$

$$USe = EPI_H \times HRSe \times SPr_H + EPI_L \times LRSe \times SPr_L$$

$$n = \log(1 - SeP)/\log(1 - USe)$$

LRSe, HRSe are weighted average sensitivity in low-risk and high-risk samples, respectively. USe is the probability of a single randomly selected animal from the sample being positive, given that the population is infected at the design prevalence.

SPrH, SPrL, SPrLH, SPrHH are proposed sample proportions from the high-risk subpopulation, low-risk subpopulation, high-sensitivity group in high-risk subpopulation and high-sensitivity group in low-risk subpopulation, respectively.

### See also key for representative freedom surveys

| Abbreviation/symbol | Meaning |
|---|---|
| **RR** | Relative risk |
| **AR** | Adjusted risk |
| **PPrH, PPrL** | Population proportions in high-risk and low-risk-groups groups, respectively |
| **SPrH, SPrL** | The proportion of the surveillance sample from the respective risk group |
| **EPI, EPIH, EPIL** | Effective probability of infection and EPI in high-risk and low-risk groups. Probabilities of infection after adjusting design prevalence for group relative risks |

| **SeUH, SeUL** | Sensitivity in high-risk and low-risk groups, respectively. May be test (animal) sensitivity or herd sensitivity, depending on level at which being calculated. |
| **USe** | The probability of a single randomly selected animal from the surveillance sample being positive, given that the population is infected at the  design prevalence. |

## PREVALENCE ESTIMATION

### Apparent or seroprevalence
(assumes perfect test sensitivity and specificity)

#### *Estimated prevalence:*

$$P = x/n$$

Asymptotic (normal approximation) confidence intervals:

$$CI = P \pm Z\sqrt{((P(1-P)/n)}$$

Alternative (binomial, Wilson binomial) CI methods usually better, particularly as P approaches 0 or 100 percent.

#### *Sample size:*

$$n = (Z^2 \times P(1-P))/e^2$$

Assumes a large population. Where expected sample size is large (10 percent) relative to populations size, use the following adjustment:

$$n_{adj} = (N \times n)/(N + n)$$

### Estimated true prevalence
(allows adjustment for imperfect sensitivity and specificity)

$$TP = (AP + SP - 1)/(Se + Sp - 1)$$

**Note:** Method fails when Se + Sp = 1 due to division by 0.
TP may be negative if AP + Sp < 1 (Sp estimate is lower than suggested by the results).

Asymptotic (normal approximation) confidence intervals assuming known sensitivity and specificity:

$$CI = TP \pm Z\sqrt{[AP(1 - AP)/(n \times (Se + Sp - 1)2)]}$$

Assumes Se and Sp known exactly (no uncertainty).
Lower CI may be < 0 if TP is close to 0.

### *Sample size:*

$$n = (Z/e)^2 \times (Se \times TP + (1 - Sp) \times (1 - TP)) \times (1 - Se \times TP - (1 - Sp) \times (1 - TP))/(Se + Sp - 1)^2$$

Asymptotic (normal approximation) confidence intervals assuming uncertain sensitivity and specificity:

$$CI = TP \pm Z\sqrt{[AP \times (1-AP)/(n \times (Se + Sp - 1)^2) + (Se \times (1-Se) \times TP^2)/(M \times (Se + Sp - 1)^2)}$$
$$+ (Sp \times (1-Sp)*(1-TP)^2)/(R*(Se + Sp - 1)^2)]$$

| Abbreviation/symbol | Meaning |
|---|---|
| **n, N** | Sample size and corresponding population size (animal level) |
| **P** | Observed or expected prevalence (proportion) |
| **x** | Number of units with the characteristics of interest |
| **Z** | Z distribution value corresponding to desired confidence level Z = 1.96 for 95%, 2.58 for 99% and 1.64 for 90% |
| **e** | Desired precision of estimate (± relative to estimate). Confidence interval width = 2e |
| **nadj** | Sample size adjusted for small population |
| **TP** | True prevalence estimate |
| **AP** | Apparent prevalence estimate |
| **Se, Sp** | Sensitivity and specificity of the test used |
| **CI** | Confidence interval |
| **M** | Sample size for estimating test sensitivity |
| **R** | Sample size for estimating test specificity |

# References

**Cameron, A.R.** 2009. Surveillance. *Animal Health Management Essentials.* (to be published by the OIE Regional Coordination Unit, Bangkok).

**OIE.** 2014. Terrestrial Animal Health Code (available at http://www.oie.int/international-standard-setting/terrestrial-code/)

**Cameron, A.R.** 1999. Survey toolbox: a practical manual and software package for active surveillance of livestock diseases in developing countries. Canberra, Australian Centre for International Agricultural Research (ACIAR) (also available at http://aciar.gov.au/publication/mn054).

**Palisage@Risk software.** Risk Analysis using Monte Carlo. Website (available at http://www.palisade.com/risk).

**EpiTools.** 2012. Website (available at http://epitools.ausvet.com.au/).

**PopTools.** 2011. Website (available at http://www.poptools.org/).

**Richard A., Becker & Allan R. W.** 1993. Website (available at http://cran.r-project.org/).

**FAO ANIMAL PRODUCTION AND HEALTH MANUALS**

1.  Small-scale poultry production, 2004 (E, F, Ar)
2.  Good practices for the meat industry, 2006 (E, F, S, Ar)
3.  Preparing for highly pathogenic avian influenza, 2006 (E, Ar, S$^e$, F$^e$, Mk$^e$)
3.  Revised version, 2009 (E)
4.  Wild bird HPAI surveillance – a manual for sample collection from healthy, sick and dead birds, 2006 (E, F, R, Id, Ar, Ba, Mn, S$^e$, C$^e$)
5.  Wild birds and avian influenza – an introduction to applied field research and disease sampling techniques, 2007 (E, F, R, Ar, Id, Ba, S**)
6.  Compensation programs for the sanitary emergence of HPAI-H5N1 in Latin American and the Caribbean, 2008 (E$^e$, S$^e$)
7.  The AVE systems of geographic information for the assistance in the epidemiological surveillance of the avian influenza, based on risk, 2009 (E$^e$, S$^e$)
8.  Preparation of African swine fever contingency plans, 2009 (E, F, R, Hy, Ka, S$^e$)
9.  Good practices for the feed industry – implementing the Codex Alimentarius Code of Practice on good animal feeding, 2009 (E, C, F, S, Ar**, P**)
10. Epidemiología Participativa – Métodos para la recolección de acciones y datos orientados a la inteligencia epidemiológica, 2011 (S$^e$)
11. Good Emergency Management Practices: The essentials, 2011 (E, F, S, Ar, C**)
12. Investigating the role of bats in emerging zoonosese – Balancing ecology, conservation and public health interests, 2011 (E)
13. Rearing young ruminants on milk replacers and starter feeds, 2011 (E)
14. Quality assurance for animal feed analysis laboratories, 2011 (E, F**, R$^e$)
15. Conducting national feed assessments, 2012 (E, F)
16. Quality assurance for microbiology in feed analysis laboratories, 2013 (E)

Availability: November 2014

| | | | | |
|---|---|---|---|---|
| Ar | – | Arabic | Multil – | Multilingual |
| C | – | Chinese | * | Out of print |
| E | – | English | ** | In preparation |
| F | – | French | $^e$ | E-publication |
| P | – | Portuguese | | |
| R | – | Russian | Mk – | Macedonian |
| S | – | Spanish | Ba – | Bangla |
| Mn | – | Mongolian | Hy – | Armenian |
| Id | – | Bahasa | Ka – | Georgian |

The *FAO Animal Production and Health Manuals* are available through the authorized FAO Sales Agents or directly from Sales and Marketing Group, FAO, Viale delle Terme di Caracalla, 00153 Rome, Italy.



Find more publications at
http://www.fao.org/ag/againfo/resources/en/publications.html

Increasing global population and improvements in the standard of living mean that there is a rapidly increasing demand for animal protein with intensified animal production. The international movement of animals and animal products has been made cheaper and faster through improved transport infrastructure. Increasing human and livestock population has placed pressure on wildlife habitats, resulting in closer contact between wildlife, domestic animal populations and humans with spreading and re-emergence of diseases as consequences of these risk factors.

Managing these disease threats poses enormous challenges and requires good quality information: what diseases exist; where they are found; what impact they are having; which populations are at risk; how we can prevent, control or eradicate these diseases. Animal disease surveillance plays a central role in providing this information.

Risk-based surveillance is not a particular technique; rather, it describes a general approach to undertaking disease surveillance. The principle is simple and self-evident: the most efficient way to find disease is to survey the animal populations that are most likely to be affected. This is in contrast to the more traditional statistically-based approach of taking representative samples from a population. While the idea of risk-based surveillance is simple, the implications are complex. The approach can be much more cost-effective for some purposes, but if misused, it can lead to serious errors or it can be more expensive than traditional approaches.