Principles for the Design and Conduct of Surveys to show Presence or Absence of Infectious Disease in Aquatic Animals

Angus Cameron

AusVet Animal Health Services Pty Ltd
140 Falls Road
Wentworth Falls, NSW 2782
angus@ausvet.com.au

_____

SUMMARY

This document provides standards and guidelines for implementing surveys for the purpose of demonstrating freedom from aquatic animal diseases.

Due to the wide variety of species cultured, the pathogens and management systems, the guiding principle applied in this document is that surveillance systems should be designed to meet the needs of a specific situation. No one surveillance system with a prescribed sampling strategy and sample size can address the range of situations experienced in aquaculture. Instead, guidelines are provided as to standard of proof required to demonstrate freedom of aquatic animals from infectious diseases, and the issues that must be considered in the design of a surveillance system.

Factors that must be taken into account when designing a surveillance system include:
- the definition of the population, including any sub-populations that should be targeted to improve the probability of detecting disease,
- clustering of disease,
- documentation of the methodology used, survey design and data analysis procedures,
- the test or test system being used,
- the design prevalence or minimum expected prevalence in the presence of disease,
- sampling approaches, and
- quality assurance systems.

Three examples describe possible surveillance systems for use when aiming to declare freedom from three different hypothetical infections. They illustrate a number of the principles discussed, and are based on 1) a farm-accreditation scheme for freshwater fish culture, 2) a national survey to demonstrate freedom from disease, and 3) a survey of molluscs using spatial sampling.

Introduction

Surveillance and Monitoring

The terms surveillance and monitoring have been used with a variety of meanings. In this document, *surveillance* is the structured collection and analysis of data for the purpose of detecting incursions of new or emerging disease or infection in an area, or for demonstrating freedom from a disease or infection. *Monitoring*, in contrast, is conducted for the purpose of assessing changes in the level or distribution of disease in an area. The main distinction is that surveillance is concerned with exotic disease, while monitoring is concerned with endemic disease.

Surveillance can be divided into two main objectives: the first is early detection of incursions of disease, to ensure that rapid responses can be implemented to control or eradicate the disease; the second is demonstration of freedom from disease, either to support trade, or as part of a disease control programme, for example, the establishment of disease free zones.

Scope of this document

This document is concerned with the establishment of standards for the conduct of disease surveillance, and specifically for surveillance with the objective of demonstrating freedom from disease, by using structured disease surveys.

Evidence for freedom from disease may be gathered by many mechanisms, including use of laboratory databases and field reporting systems, structured negative reporting systems, and the analysis of farm production records. However, the most common and analytically convenient approach is to use a structured survey designed with a known probability of detecting disease, if it is present.

This document does not provide a step-by-step guide to the use of such a survey; other texts are available for this purpose. Instead, it provides a set of standards and guidelines for the conduct of a survey to demonstrate freedom from disease, aimed to ensure that such surveys are carried out to a consistent level of quality, and will meet current and future international standards. Guidance on methods of sampling will be found in Appendix 3.

Similarly, this document is not concerned with surveys designed for the purposes of monitoring endemic disease, for example, prevalence surveys. Although such surveys share some of the same principles discussed in these standards, the theoretical framework in which they are conducted is fundamentally different: they aim to estimate population parameters, while surveys to demonstrate disease are based on hypothesis testing.

Finally, this document deals with the conduct of surveys, rather than the design of surveillance systems. A surveillance system typically involves a number of data collection approaches, and also incorporates data management, analysis and reporting systems. Structured surveys may be one component of a surveillance system, but the others are beyond the scope of this document.

There is a certain amount of assumed knowledge required for the effective application of these standards, including a basic understanding of aquatic animal health and production systems, and a familiarity with commonly used epidemiological terms. Some of these are defined in Appendix 1, while others may be found in epidemiological or survey texts.[1,2]

This document includes three examples of surveys conducted at different levels for different purposes, and illustrating the application of the standards.

International Framework

While one of the reasons for undertaking a survey to demonstrate freedom from disease is to support domestic programmes to control disease, for example, through the use of disease-free zones, these surveys are also often used as an aid to trade. As such, it is important that national standards for surveys comply with international standards set by OIE.

The chapter, Requirements for Surveillance for International Recognition of Freedom from Infection, in the *Diagnostic Manual for Aquatic Animal Diseases*[3], represents a significant change in the approach to setting standards for aquatic animal disease surveillance.

This change can be summarised as:

- The development of a single set of generic surveillance guidelines, applicable to all diseases, rather than inclusion of separate surveillance guidelines in each disease chapter. The latter approach resulted in the risk of duplication, unnecessary inconsistency between different diseases, and delays in the development of appropriate guidelines for many diseases.

- Inclusion of limited amounts of the required disease-specific surveillance information in each disease chapter, so that the generic guidelines can be refined to apply to the specific disease. Examples include:
  - guidance on the appropriate sub-populations to target for the highest chance of detecting disease, if it is present; and
  - biologically determined survey design parameters, such as the design prevalence.

- A move to output-based standards, rather than the existing input-based approach. The draft standards require that demonstration of freedom from disease meet a defined level of confidence (95%). The way in which this level of confidence is met is not prescribed. As a result, tables of sample sizes, definitions of sampling intervals and so on have been removed from the guidelines, leaving countries with the flexibility to apply the most appropriate surveillance to meet their own particular circumstances - as long as the final confidence achieved by that surveillance meets the international standard.

- This increased flexibility is qualified by the inclusion of requirements that surveillance approaches need to be scientifically based, and be consistent with the latest advances in the field. In this way, as new and improved methodologies for surveillance are developed, they can be immediately adopted - in contrast to the current system which risks prescribing the use of out-dated methodologies.

- Introducing a clear requirement for transparency and quality management. When applying for recognition of disease freedom, countries must fully document all aspects of their evidence, methodology and analysis, as well as demonstrating that a quality assurance system is in place to document that the planned or described surveillance activities were actually undertaken as intended.

One consequence of these changes is that the draft chapter does not lay out a recipe for surveillance, suitable for application aquatic animal health staff in any situation. Its flexibility and non-prescriptive nature mean that different countries have the freedom to develop surveillance systems that are individually designed and appropriate for their particular circumstances.

Australian and New Zealand Standards

Introduction

The standards for surveillance in this section are applicable to all diseases of all species. In order to be applied, some information about a particular species/disease combination is required for the design of the survey and the analysis of the results. Examples include the *design prevalence* (the expected level of disease if the agent was present) and the most appropriate population in which to look for disease, for example, at a particular stage of growth or time of year.

It is anticipated that, in the future, this type of information (listed in more detail later in this document) will be specified in each disease chapter of the OIE Aquatic Animal Code.[4] In the meantime it is necessary to provide some guidance on the selection of appropriate values. These standards therefore list either generally applicable recommended values, or provide guidelines for the choice of suitable parameters required when planning surveillance activities.

General Principles

Demonstrating freedom from infection involves providing sufficient evidence to demonstrate that infection with a specified agent is not present in a specified population. In practice, it is not possible to definitively prove that a population is free from infection unless every member of the population is examined simultaneously with a perfect test with both sensitivity and specificity equal to 100%. Instead, the aim is to provide adequate evidence to an acceptable level of confidence, that infection, if present, is present in less than a specified proportion of the population.

In the presence of uncertainty, a probabilistic approach must be taken, rather than providing a definitive statement of freedom. The accepted approach is to calculate the probability that a particular surveillance plan will detect disease, if disease is present. If the surveillance system finds no disease, it is not proof that disease is not present. It is saying only that we are confident that, if disease was present, it was present at a level lower than that specified. See the definition of *design prevalence* below for more information about this specified level of disease.

Methods to demonstrate freedom from infection should be adaptable to deal with the complexity of real life situations. No single method is applicable in all cases. Methods must be able to accommodate the variety of aquatic animal species, the multiple diseases of relevance, varying production and *surveillance systems*, and types and amounts of data and information available.

Significant progress is being made in the area of surveillance and in particular in the development of survey methods and approaches to the analysis of surveillance data. As a result, any particular method that is recommended or prescribed today is likely to be superseded by newer and better techniques in the near future. The preferred approach is therefore for those working in the field to select the most appropriate surveillance tools that are currently available, as long as they are technically valid and in accord with current scientific thinking. The methodology should be well-documented and supported with references to scientific publications and other sources, including expert opinion.

Requirements for Demonstration of Freedom from Infection

These requirements take the form of seven key concepts, which must be considered when conducting surveys to demonstrate freedom from infection. These concepts are explained, and any requirements specified. The concepts are:
- Population
- Statistical Methodology, Survey Design and Analysis
- Clustering
- Design Prevalence
- Test Characteristics
- Sampling
- Quality Assurance

Population

Careful consideration of the population is required to ensure that the results of surveillance are appropriately interpreted. Similarly, a consideration of alternative populations provides authorities with the flexibility to design well-targeted surveillance systems. The target population to which the demonstration of freedom from infection applies is all individuals of all species susceptible to the infection in the country or zone.

Whenever the study population (the individuals selected to participate in the study) is different from the target population (the population of interest, for example, all farmed prawns in Australia), there is a risk that the findings from the study population may not represent the true situation in the target population. Despite this, it is rarely feasible to conduct surveillance in which the study and target populations are exactly the same, due to practical constraints. However, if, based on our knowledge of a particular disease, we can make reasonable assumptions about the differences between target and study populations, we may be able to interpret results from the study population with confidence.

For instance, a particular pathogen may infect a range of species but cause clinical disease in a few species only, and have a particular seasonal pattern of manifesting itself, for instance, only in late summer. The target population is all susceptible species, and they should be tested at all times of the year to ensure full representation. Instead, using a study population defined as only those species that show clinical disease, and only in late summer provides a much greater chance of detecting the disease if it is present. A study based on this population would provide greater confidence in freedom from disease than would a fully representative study of the whole target population.

On the other hand, a study of a disease in wild stocks may be based on sampling from the commercial catch. This study population may have a lower chance of containing diseased fish, as the fishing techniques used may be designed to capture only well-grown healthy fish. Fish affected by the disease may die before they are large enough to be caught, or may otherwise be less likely to be caught. This type of study would provide less confidence than a representative study.

Whatever the study population used, it is most important to document it, to consider how it differs from the target population, and to account for any effects these differences may have on the overall assessment of confidence in freedom from infection.

The study population should be (in order of preference):
- The appropriate study population as defined in the relevant disease chapter of the OIE Aquatic Animal Health Code, where such a definition exists,
- A subset of the target population that defines a group of animals, which, if infection was present, would be most likely to have a higher prevalence of infection than the target population. This subset should be defined in terms of:
  - species;
  - time, for example, season or month of year;
  - stage of life cycle or growth period;
  - production system and/or management characteristics;
  - location;
  - readily identifiable physical or behavioural characteristics.
- The same as the target population,
- A subset of the target population with the same or lower probability of infection. The nature and impact of any biases on the results of the analysis must be considered, documented and taken into account in the analysis.

Statistical Methodology, Survey Design and Analysis

While the intention of these standards is to allow a range of adaptable methods to be used, it is a necessary requirement that the output of any method used must be the same: a measure of the confidence that the survey would have detected disease if it were present at specified levels. This limitation ensures that, despite the use of different statistical methods, the results will be fully comparable between different surveys.

A statistical hypothesis-testing framework is used for analysis, an approach that should be familiar to most scientists. It is based on the establishment of a *null hypothesis* and the examination of evidence supporting this null hypothesis. If there is inadequate evidence to support the null, then it is rejected and an alternate hypothesis is accepted.

Alpha is the probability of rejecting a null hypothesis that is true. Therefore $1 - \alpha$ is a measure of supporting evidence for the null hypothesis. The common standard for all analytical techniques is that the result of analysis should be an estimate of α.

An important difference between this approach and that used in much statistical hypothesis testing is the choice of the null. Normally the null means 'no effect', so that, for instance, treatment A has no different effect to treatment B. In this context, the null is that *disease is present*. This means that if inadequate evidence is available to reject the null, that is, α is large, due to, for instance, a poor surveillance system that collects only a small number of samples, then we fail to reject the null. The result is that the country or zone is considered as having disease, until enough evidence is presented to demonstrate otherwise.

The definition of the null (*disease is present*) must be qualified: how much disease? The level of disease under the null is specified by the *design prevalence*, described below. A full statement of the null hypothesis is therefore that disease is present at a level equal to or greater than that specified by the design prevalence.

The alternate hypothesis, the one we accept if we reject the null, is that disease is *not* present at a level equal to or greater than that specified by the design prevalence. This means that disease is either not present at all, or, if it is present, the level of disease is less than that specified by the design prevalence.

The required level of confidence in the surveillance system, which is the probability that the system would detect infection if infection were present at the specified level, must be greater than or equal to 95%. This is the most important fundamental standard in this document. It is not important what surveillance system or evidence is used, the nature of the production system, which tests were used, or how the information was analysed, provided it is scientifically valid and takes all these other considerations into account). What matters is that it is able to meet this standard level of confidence: 95%.

Demanding this high level of confidence means that there is only a small risk that a country may claim freedom when it is in fact not free. The requirement for a high level of

confidence recognises the potentially serious consequences of such a mistake (known as a type I error): the spread of infection between countries.

The other type of error that can result from an analysis of evidence (type II error) is that a country determines that it is infected, when it is in fact free from disease. While the consequences of this error may be significant for the country itself, for example, loss of trade opportunity, such an error would not result in an increased risk of spread of disease.

The *power* of the analysis, in this context, is the probability of avoiding a type II error. As this does not affect the risk of international spread of disease, no international standard value has been set. Instead, it is up to the country claiming freedom to determine the best value for their own purposes. A high power would minimise the risk of falsely concluding that disease is present, but is usually more expensive. By convention, this is often set to 80%. Where the test system, which is defined below, has an imperfect specificity, it is important to ensure that power is set at a reasonably high level, despite the increase in sample size that this results in. In practice, many test systems involve one confirmatory test that is considered for all intents and purposes to have a specificity of 100%. If this is the case, the effective power of the surveillance system is also 100%. Power is therefore only important when a probabilistic interpretation of positive reactors is used with tests with imperfect specificity.

The method used to calculate the *confidence* in the *surveillance system* must be scientifically based and clearly documented, including references to published work describing the method.

The design of the survey will depend on the size and structure of the population being studied. If the population is relatively small and can be considered to be homogenous with regards to risk of infection, a single stage survey can be used.

In larger populations where a sampling frame is not available, or when there is a likelihood of clustering of disease, multi-stage sampling is required. In two-stage sampling, at the first stage of sampling, groups of animals, for example, ponds, farms or villages are selected. At the second stage of sampling, animals are selected for testing from each of the selected groups.

If testing of individual animals is involved, single stage surveys are relatively unusual. They may be used when the aim is to demonstrate that a single aquaculture establishment is free from infection, but even then, only if the establishment does not divide animals into groups, for example, multiple ponds or cages. Another example would be certification of batches of animals for export.

Stratification may be useful for both practical purposes, for example, ease of conducting the survey, and to make the sampling design more flexible. For instance, stratifying by administrative subdivision may mean that different survey teams can be used in each area. Stratifying by species or production system may mean that different sampling strategies can be used, in order to account for varying risk in different sub-populations. For instance, stratifying by species may allow most resources to be dedicated to the species known to be most at risk. However, other species can also be included in the survey to provide supportive evidence that the assumptions about species susceptibility and risk were valid.

Analysis of test results from a survey should take at least the following considerations into account:

- The survey design;
- The sensitivity and specificity of the test, or test system; and
- The design prevalence, or prevalences where a multi-stage design is used;

Software to assist with the analysis of survey results, and which meets the listed criteria is freely available.[5]

Clustering of Infection

That almost all diseases cluster at multiple levels complicates the design and analysis of surveillance systems. For example, a 10% animal-level prevalence of infection may mean that one in ten animals, evenly spread through the population, are infected (unclustered). Alternatively, it may mean that there are a small number of groups of animals in which almost every animal is infected, while the rest of the population is uninfected (clustered infection). While the overall prevalence is the same in both cases, detecting disease in the latter case may be more difficult, as it is necessary to identify one of the few infected groups before disease can be detected. Design prevalence (below) must often be expressed in terms of prevalence at different levels of

clustering, in recognition of the clustered nature of infection and populations.

A number of methods exist that can take clustering of infection into account, but these complicate design and analysis. There may be many levels at which infection clusters. Necessarily, the mathematical models upon which analysis of the real world are based are gross simplifications. It may—be generally impractical to incorporate all possible levels of infection clustering into such a model. As a result, it is recommended to incorporate only a single level of infection clustering, unless dealing with small homogenous populations. The judgement as to which is most significant depends on the details of the production system and disease, but will usually represent the group of aquatic animals, for example, cage, pond etc or the management unit, for example, farm.

Design Prevalence

Calculation of the confidence of a surveillance system is based on the null hypothesis that infection is present in the population. The level of infection is specified by the design prevalence, which is distinct from prevalence. Design prevalence forms part of the definition of the null hypothesis. As part of a *hypothesis*, it is an abstract statement of what may be present in nature. It is *not* a statement of the proportion of infected animals or other units, because, when demonstrating freedom from infection, we are assuming that the prevalence of infection is zero.

A number of other terms have been used to describe design prevalence, including *minimum expected prevalence*, *maximum acceptable prevalence*, and *minimum detectable prevalence*. These terms describe aspects of the ways in which the values are used, but as their main role is to define the hypothesis, and therefore the design of surveillance or the analytical approach, the more general term *design prevalence* is preferred.

The choice of values for the design prevalence is problematic. While any value could be chosen, the value chosen influences our interpretation of the results of analysis. For instance, if a design prevalence of 50% is used when assessing batches of shrimp post-larvae for white spot syndrome virus (WSSV) infection, a negative result would mean that we are 95% confident that the prevalence of infection in the batch is less than 50%. This

does not constitute very strong evidence for freedom. On the other hand, if a value of 1% is used, a negative result would mean that it is still possible that infected post-larvae exist, but at a prevalence of no more than 1%. Considering the biology of WSSV, it is unlikely that only 1% of an infected batch would be infected. Vertical transmission of the virus means that if any are infected, the brood stock must have been infected, and with the expectation that a reasonably high proportion of infected post-larvae would be infected. As a result of this argument, it may be considered that evidence that the prevalence is less than 1% is evidence for freedom.

A further point that should be borne in mind is that these standards deal with freedom from infection. Freedom means that there is not one infected animal in the country or zone. If it is known that infected animals are present in the country, even at a prevalence well below the specified design prevalence, then it is not possible to claim freedom. In other words, freedom means zero prevalence, not low prevalence. While we may not be able to prove freedom with 100% confidence, we can prove that a country is *not* free from infection, by identifying one infected animal.

Published values for design prevalence used in calculations should be those specified in the relevant disease chapter of the OIE Manual[3]. If not published for the particular disease, justification for the selection of design prevalence values must be provided, and should be based on the following guidelines:

- At the individual animal level, the design prevalence is based on the biology of the infection in the *population*. It is equal to the minimum expected prevalence of infection in the *study population*, if the infection had become established in that *population*. It is dependent on the dynamics of infection in the *population* and the definition of the *study population*, which may be defined to maximise the expected prevalence in the presence of infection.

  The design prevalence at the individual animal level applies to the specific study population specified. For instance, while it may be possible for infection to persist in carriers at low levels throughout the year, if the study population is a highly susceptible (indicator) species at the normal time of year for outbreaks, then it may be reasonable to assume that, if disease is present, its prevalence *in that study population* would

be much higher than in the target population at other times of the year. This means that a higher value could be used for design prevalence, making it easier to provide adequate evidence.

A suitable design prevalence value at the animal level, for example, prevalence of infected animals in a cage, may be:
- between 1% and 5% for infections that are transmitted slowly; and
- more than 5% for more contagious infections.

These suggested ranges for design prevalence are intended as a broad guide only. Analyses based on high design prevalence values of more than 5%, are, in general, less convincing, and must therefore be supported by sound arguments as to why failing to find infection at the specified level may be considered equivalent to complete freedom from infection.

- Unlike at the individual animal level, where most infections may be expected to spread among in-contact animals, at higher levels, such as the farm, it is possible that disease may be maintained within a single farm. The biology of the disease and nature of the farming system determine whether spread between farms is likely, but in many cases, very low farm-level prevalence values are possible. Demanding a 'zero-risk' approach is not permissible, so it is necessary to decide what level of risk we are willing to tolerate, and balance this with the likely behaviour of the disease.

On the basis of these principles, the design prevalence at the level of a farm, pond, cage or village should be set as low as possible, while still making it practicable for countries to meet this level of proof. Extremely low values, for example, 0.1%, require massive surveys or the accumulation of huge amounts of evidence. Higher values, greater than 2%, run the risk of failing to detect low levels of premises with infected animals. The final figure should be a compromise between what is practical and the level of security that is desired. In the absence of compelling reasons for a shift in one direction or another, a value of 1% is recommended as a generally acceptable figure.

A suitable design prevalence value for the first level of clustering, for example,

proportion of infected farms in a zone, may be up to 2%.

Test Characteristics

Surveillance involves performing one or more *tests* for evidence of the presence of current or past infection, ranging from detailed laboratory examinations to observations by farmers. The level of performance of a test at the population level is described in terms of its sensitivity and specificity. Imperfect sensitivity and/or specificity affect the interpretation of surveillance results and must be taken into account in the analysis of surveillance data. In the past, generic tables of sample size have been used as the basis for surveillance planning, without regard to which tests may be used. Even small differences between tests, for example, the difference between 99% and 98% specificity, may require a marked increase in sample size.

All calculations must take the sensitivity and specificity of any tests used into account. The values of sensitivity and specificity used for calculations must be specified, and the method used to determine or estimate these values must be documented.

This means that whenever a test is used to provide evidence for freedom, its sensitivity and specificity must be known or estimated. This is a challenging requirement, as there are many tests in use for which there are either only poor or no estimates of their sensitivity and specificity. Ideally, test performance should be evaluated using the same population as that in which the test is to be used. This is impossible in the case of demonstrating freedom from infection, as sensitivity cannot be calculated in an infection-free population.

If suitable studies have not been done to provide the necessary information about tests that are being used, there are a number of options. The first and preferred option is to perform and publish the necessary studies in the best way possible. This means either by using an infected population in another zone or country that is considered to be similar to the disease-free population, or to use experimentally infected animals. While the conduct of such studies is time consuming and expensive, the potential benefits, such as the ability to open up new export markets due to disease-free status, will normally far outweigh the costs.

The second option is to use data from existing studies performed in other populations, which may be dissimilar to the local population. The third option, which is by far the least preferred, but nevertheless acceptable, is the use of expert opinion to estimate the values of sensitivity and specificity. If expert opinion is used, values of 100% for sensitivity or specificity are unlikely to be accepted without stringent questioning.

The considerable flexibility offered by these guidelines in the approach to obtaining estimates of sensitivity and specificity is in recognition of 1) the great importance these values have in interpretation of surveillance data, and 2) the current scarcity of good estimates for these values. Regardless of the manner in which an estimate of the values was obtained, the basis for this estimate must be clearly documented, so that its validity can be judged. Where it is difficult or impossible to conduct the necessary field studies, those with the responsibility of evaluating evidence for freedom are likely to consider estimates obtained through another means as reasonable.

The other implication of these provisions is that any test can be used, as long as figures for sensitivity and specificity can be provided. For instance, clinical observations by farmers or fish health workers, or analysis of farm records may both be interpreted as tests for the presence of a particular disease. The sensitivity of both may be very low, and specificity moderately high. However, if a valid approach is used to estimate the sensitivity and specificity of the test, and these values are incorporated into the analysis, it is possible that evidence based on clinical observations or farm records could contribute significantly to a claim of freedom from infection.

Most surveillance systems using laboratory or other tests are based on the use of more than one test. For instance, an initial screening test may be applied, followed by the use of a confirmatory test for any specimens that give a positive screening test result. Often, the testing protocol may be quite complicated and involve the use of more than two tests.

The combination of multiple tests always has an effect on the overall test system sensitivity and specificity. In general, if the combination aims to increase specificity, as in the example above, the result will be a decrease in sensitivity, and vice versa.

Where more than one test is used in a surveillance system (sometimes called using tests in series or parallel), the overall test system sensitivity and specificity must be calculated using a valid method. Simple formulae are available in epidemiology texts.[1,2] and illustrated in the examples included in this document, for the calculation of the combined sensitivity and specificity of multi-test systems. While these formulae are currently acceptable, they suffer from the invalid assumption that the tests being combined are independent. If newer approaches to calculating the combined performance of multiple tests that take this lack of independence into account are developed in the future, these should be used in preference to the current methods.

In the absence of such new methods, and as part of good design, the tests that are used in combination should be as biologically independent as possible. For instance, the use of one test that detects pathognomonic tissue changes and another that detects antigen is far preferable to two tests that both detect antigen by slightly different methods.

Pooled testing involves the pooling of specimens from multiple individuals and performing a single test on the pool. Pooled testing is an acceptable approach. Where pooled testing is used, the results of testing must be interpreted using sensitivity and specificity values that have been determined or estimated for that particular pooled testing procedure and for the applicable pool sizes being used. Analysis of the results of pooled testing must, where possible, be performed using accepted, statistically based methods, which must be fully documented including published references.

Sampling

1. Sampling Principles
The objective of sampling from a population is to select a subset of units that is representative of the population with respect to the characteristic of interest, in this case, the presence or absence of infection. Sampling should be carried out in such a way as to provide the greatest likelihood that the sample will be representative of the population, within the practical constraints imposed by different environments and production systems. Representative sampling is sometimes difficult, so the aim is to do the best possible job under the circumstances.

The intended output of a sampling exercise is a representative sample. This is usually best achieved by random sampling. However, random sampling is the means, not the objective.

Biased or targeted sampling in this context involves sampling from a defined study population that has a higher probability of infection than the target population of which it is a sub-population. Once the study population has been identified, the objective is still to select a representative sample from this sub-population. In other words, all sampling should aim to select a representative sample; any targeting or bias is achieved not by sampling, but by defining the population from which the sample is selected.

2. Sampling methods

The survey design may involve sampling at several levels.

For sampling at the level of the *epidemiological units* or higher *units*, a formal *probability sampling*, for example, simple random sampling, method must be used. This is because probability-sampling techniques are the best available techniques to reliably generate a representative sample. One of the common requirements for probability sampling is the existence of a *sampling frame* listing every member of the population, although there are approaches that surmount this requirement, such as systematic sampling and spatial sampling. It should be possible to generate a sampling frame of epidemiological or higher units, for example, ponds, farms, or villages, in most circumstances, and therefore probability-sampling techniques can be applied. In those circumstances where a sampling frame is not available or cannot be generated, alternative approaches, for example, ponds systematic or spatial sampling, can be applied.

In contrast to sampling at higher levels, sampling of lower units, and particularly individual animals, where required, is often difficult. It is rarely possible to generate a sampling frame, and systematic or spatial sampling techniques can be used only in particular cases.

In recognition of the problems facing sampling at this level, these standards do not prescribe any particular approach. Instead, what is required is that the desired output, a representative sample, be kept in mind, that every effort is made to achieve this, and that the approach used be documented in detail and justified. Sampling strategies that are based on convenience sampling are very unlikely to be considered acceptable in any circumstances.

3. Sample size

The number of *units* to be sampled from a *population* should be calculated using a statistically valid technique, which takes at least the following factors into account:
- The sensitivity and specificity of the *diagnostic test*, or *test system*;
- The design prevalence, or prevalences where a multi-stage design is used;
- The level of *confidence* that is desired of the survey results.

Additionally, other factors may be considered in sample size calculations, including, but not limited to:
- The size of the *population*, but it is acceptable to assume that the *population* is infinitely large;
- The desired power of the survey;
- Uncertainty or variability in estimates of sensitivity and specificity.

Software packages are widely available to assist with sample size calculations, along with texts describing their application for surveys to demonstrate freedom from infection. These are listed in the examples.

Quality Assurance

Surveys should include a documented quality assurance system, to ensure that field and other procedures conform to the specified survey design. Acceptable systems may be quite simple, as long as they provide verifiable documentation:

- That those involved in field and laboratory aspects of the survey actually did what they where meant to do, for instance, that sampling protocols were adhered to; and
- Any discrepancies between the planned and achieved activities, for instance, if the actual sample size was less than the target sample size.

Appendix 1: Definitions

Epidemiological Unit
This unit is a group of animals that share approximately the same risk of exposure to a disease agent with a defined location. This may be because they share a common aquatic environment, for example, fish in a pond, caged fish in a lake, or because management practices make it likely that a disease agent in one group of animals would quickly spread to other animals, for example, all the ponds on a farm, all the ponds in a village system.

Test
A procedure used to classify a *unit* as either positive or negative with respect to an infection or disease. Tests may be classified as:
• diagnostic, when applied to clinically diseased individuals;
• screening, when applied to apparently healthy individuals; or
• confirmatory, when applied to confirm the result of a previous test.

Test system
A combination of multiple *tests* and rules of interpretation that are used for the same purpose as a *test*.

Probability sampling
A sampling strategy in which every *unit* has a known non-zero probability of inclusion in the sample.

Targeted surveillance
Surveillance targeted at a specific disease or infection.

Early detection system
A system for the timely detection and identification of the incursion or emergence of infection or disease in a country, zone or aquaculture establishment. An early detection system must be under the control of the competent authorities and must include the following characteristics:
• representative coverage of target animal *populations* by field services;
• ability to undertake effective disease investigation and reporting;
• access to laboratories capable of diagnosing and differentiating relevant diseases;

• a training programme for veterinarians or fish health specialists for detecting and reporting unusual disease occurrence.

Prescribed Biosecurity Conditions
A set of conditions applying to a particular disease or infection, and a particular zone or country, required to ensure adequate biosecurity to maintain freedom from infection, namely:
• the disease is legally notifiable;
• an *early detection system* is in place;
• no vaccination against the disease is carried out;
• infection is not known to be established in wild *populations*;
• import requirements to prevent the introduction of disease or infection into the country or zone, as outlined in the OIE Aquatic Animal Health Code.[4]

Units
Individually identifiable elements. This is a generic concept used to describe, for example, the members of a *population*, or the elements selected when sampling. In these contexts, examples of units include individual animals, to ponds, nets, cages, farms, villages, and districts.

Population
A group of *units* sharing a common defined characteristic.

Target population
For the purposes of demonstrating freedom from infection, the *population* of interest, usually made up of all aquatic animals of species susceptible to a specified infectious agent in a defined country, zone or aquaculture establishment.

Study population
The *population* from which evidence of freedom from infection is derived. This may be the same as the *target population* or a subset of it.

Surveillance System
A method of surveillance that generates a source of information on the animal health status of *populations*.

Confidence

In the context of demonstrating freedom from infection (in which the null hypothesis is that infection is present), the confidence is the probability that a *surveillance system* or combination of *surveillance systems* would detect the presence of infection if the *population* were infected. The confidence depends on the design prevalence, or the assumed level of infection in an infected *population*. Confidence therefore refers to our confidence in the ability of a *surveillance system* to detect infection, and is equal to the sensitivity of the system. This is distinct from, but may be used to calculate, the probability that a given *population* is free from infection, based on the results of one or more *surveillance systems*.

Appendix 2: Example Surveillance Systems

The following examples describe structure surveys that are able to meet the requirements of standards. The purpose of these examples is:

- to illustrate the range of approaches that may be acceptable;
- to provide practical guidance and models, which may be used for the design of surveys; and
- to provide references to available resources that are useful in the development and analysis of surveys.

While these examples demonstrate ways in which freedom from infection may be successfully demonstrated, they are not intended to be prescriptive. Those applying these standards are free to use different approaches, provided that they meet the requirements.

The examples deal with the use of structured surveys and are designed to illustrate different survey designs, sampling schemes, the calculation of sample size, and analysis of results.

Example 1 – One-stage structured survey (farm accreditation)

Context

A freshwater aquaculture industry raising fish in tanks has established a farm accreditation scheme. This involves demonstrating farm-level freedom from a particular (hypothetical) disease (disease X). The disease does not spread very quickly, and is most common during the winter months, with adult fish at the end of the production cycle being most severely affected. Farms consist of a number, from 2 to 20, of tanks each holding between 1000 and 5000 fish.

Objective

The objective is to implement surveillance that is capable of providing evidence that an individual farm is free from disease X. The issue of national or zone freedom, as opposed to farm freedom, is considered in the next example.

Approach

The accreditation scheme, which is based on OIE guidelines, establishes a set of standard operating procedures and requirements for recognition of freedom. These require farms to undertake a structured survey capable of producing 95% confidence that the disease would be detected if it were present. Once farms have been surveyed without detecting disease, they are recognised as free, as long as they maintain a set of minimum biosecurity standards. These standards are designed to prevent the introduction of disease X into the farm through the implementation of controls specific to the method of spread of that disease, and to ensure that the infection would be detected soon after it was introduced into the farm, based on evidence of adequate health record keeping and the prompt investigation of unusual disease events. Independent auditors evaluate the efficacy of these biosecurity measures with annual on-farm audits conducted.

Survey Standards

Sets of standards, which are based on the OIE guidelines, are established for the conduct of surveys to demonstrate freedom from infection with causative agent of disease X. These standards include:

- The level of confidence required of the survey is 95% that is, type I error = 5%.
- The power of the survey is arbitrarily set at 95% that is, type II error = 5%, which means that there is a 5% chance of concluding that a farm without infected animals is infected.
- The target population is all the fish on the farm. Due to the patterns of disease in this production system, in which only fish in the final stages of grow-out, and only in winter, are affected, the study population is defined as grow-out fish during the winter months.
- The issue of clustering is considered. As fish are grouped into tanks, this is the logical level at which to consider clustering. However, when a farm is infected, the disease often occurs in multiple tanks, so there is little evidence of strong clustering. Also, that it is difficult it is difficult on farms with few tanks to define design prevalence at the tank level, that is, the proportion of infected tanks that the survey should be able to detect on the farm. For these reasons, the entire grow-out population of each farm is treated as a single homogenous population.
- Stratification is also considered. In order to ensure full representation, the sample size

is stratified by tank, proportional to the population of each tank.

- The design prevalence at the animal level is based on the epidemiology of the disease. The infection does not spread quickly, but has been reported to infect at least 10% of fish in the defined target population if the population is infected. Adopting the most conservative approach, an arbitrarily low design prevalence of 2% is used. To have adopted a prevalence of 10% would result in a much smaller sample size, but the authorities would not accept that the population could still be infected at a level of say 5%, and infection still not be detected.

- The test used involves destructive sampling of the fish, and is based on an antigen detection enzyme-linked immunosorbent assay (ELISA). Disease X is present in some parts of the country, hence the need for a farm-level accreditation programme. This has provided the opportunity for the sensitivity and the specificity of the ELISA to be evaluated in similar populations to those on farms. A recent study using a combination of histology and culture as a 'gold standard' estimated the sensitivity of the ELISA to be 98% (95% confidence interval 96.7% – 99.2%), and the specificity to be 99.4% (95% confidence interval 99.2% – 99.6%). Due to these relatively narrow confidence intervals, it was decided to use the point estimates of the sensitivity and specificity rather than complicate calculations by taking the uncertainty in those estimates into account.

Sample Size

The sample size required to meet the objectives of the survey is calculated to take the population size, the test performance, the confidence required and the design prevalence into account. As the population of each farm is relatively large, differences in the total population of each farm have little effect on the calculated sample size. The other parameters for the calculation of sample size are fixed across all farms. Therefore, a standard sample size, based on the use of this particular ELISA, in this population, is calculated. The sample size calculations are performed using the FreeCalc software[5]. Based on the parameters listed above, the sample size required is calculated to be 410 fish per farm. In addition, the program calculates that, given the imperfect specificity, it is still possible for the test to produce up to 5 false positive reactors from an uninfected

population using this sample size. The authorities are not comfortable with dealing with false positive reactors, so it is decided to change the test system to include a confirmatory test for any positive reactors. Isolation and identification of the causal agent is selected as the most appropriate test, as it has a specificity that is considered to be 100%. However, its sensitivity is only 90% due to the difficulty of growing the organism.

As two tests are now being used, the performance of the test system must be calculated, and the sample size recalculated based on the test system performance.

Using this combination of tests, in which a sample is considered positive only if it tests positive to both tests, the specificity of the combined two tests can be calculated using the formula:

$$Sp_{Combined} = Sp_1 + Sp_2 - (Sp_1 \times Sp_2)$$

which produces a combined specificity of $1 + 0.994 - (1 \times 0.994) = 100\%$.

The sensitivity may be calculated by the formula:

$$Se_{Combined} = Se_1 \times Se$$

which produces a combined sensitivity of $0.9 \times 0.98 = 88.2\%$

These new values are used to calculate the survey sample size yielding a result of 169 fish. Attempts to improve the performance of a test, in this case increase specificity, generally result in a decrease in the performance of the other aspect of the test performance, sensitivity in this example. However, in this case, the loss of sensitivity is more than compensated for by the decreased sample size due to the improved specificity.

When using a test system with 100% specificity, the effective power of the survey will always be 100%, regardless of the figure used in the design. This is because it is not possible to make a type II error, and conclude that the farm is infected when it is not.

A check of the effect of population size on the calculated sample size is worthwhile. The calculated sample size is based on an infinitely large population. If the population size is smaller, the impact on sample size is shown in the following table:

Population Size	Sample Size
1000	157
2000	163
5000	166
10,000	169

Based on these calculations, it is clear that, for the population sizes under consideration, there is little effect on the sample size. For the sake of simplicity, a standard sample size of 169 is used, regardless of the number of grow-out fish on the farm.

Sampling

The selection of individual fish to include in the sample should be done in such a manner as to give the best chance of the sample being representative of the study population. A description of how this may be achieved under different circumstances is provided in Appendix 3. An example of a single farm will be used to illustrate some of the issues.

One farm has a total of 8 tanks, 4 of which are used for grow-out. At the time of the survey during winter, the 4 tanks have 1850, 4250, 4270 and 4880 fish, respectively, giving a total population of 15,250 fish.

Simple random sampling from this total population is likely to produce sample sizes from each tank roughly in proportion to the number of fish in each tank. However, proportional stratified sampling will guarantee that each tank is represented in proportion. This simply involves dividing the sample size between tanks in proportion to their population. The first tank has 1850 fish of a total of 15,250, representing 12.13%. Therefore 12.13% of the sample (21 fish) should be taken from the first tank. Using a similar approach the sample size for the other three tanks is 47, 47 and 54 fish, respectively.

Once the sample for each tank is determined, the problem remains as to how to select 21 fish from a tank of 1850 so that they are representative of the population. Several options are available.

- If the fish can be handled individually, random systematic sampling may be used. This is likely to be the case if, for example:
  - fish are harvested during winter and samples can be collected at harvest; or
  - routine management activities involving handling the fish, such as grading or vaccination, are conducted

If fish are handled, systematic sampling means selecting a fish at regular intervals.

For instance, to select 21 fish from 1850, the sampling interval should be 1850/21 = 88. This means that every 88th fish from the tank should be selected. To ensure randomness, it is good practice to use a random number, which in this case is between 1 and 88, to select the *first* fish, for example, using a random number table, and then select every 88th fish after that.

- If fish cannot be handled individually, which is by far the most common, and more difficult, circumstance, then the fish to be sampled must be captured from the tanks. Fish should be captured in the most efficient and practical way possible, but every effort should be made to try to ensure that the sample is representative. In this example, a dip net is the normal method used for capturing fish. Using a dip net, convenience sampling would involve capturing 21 fish by repeatedly dipping at one spot and selecting those easiest to capture, perhaps the smaller ones. This approach is strongly discouraged. One method of increasing the representativeness is to sample at different locations in the tank – some at one end, some at either side, some at the other end, some in the middle, some close to the edge. Additionally, if there are differences between the fish, an attempt should be made to capture fish in such a way as to give different groups of fish a chance of being caught, that is, not just trying to catch the small ones, but include big ones as well.

This method of collecting a sample is far from the ideal of random sampling, but due to practical difficulties of implementing random sampling of individual fish, this approach is acceptable, as long as the efforts made to select a representative sample are both genuine and fully documented.

Testing

Specimens are collected, processed and tested according to standardised procedures developed under the accreditation program. The testing protocol states that where a specimen that has a positive result to an ELISA is then tested for the causal agent with a positive result, that result indicate a true positive specimen, that is, that the farm is not free from disease. It is important to conform to this protocol. It is unacceptable to retest a true positive specimen, unless further testing is specified in the testing protocol, and the effect of such testing accounted for in the test system sensitivity and specificity estimates and therefore the sample size.

Analysis

If the calculated sample size of 169 is used, and no true positive specimens are found, then the survey will have a confidence of 95%. This can be confirmed by analysing the results using FreeCalc software[5], which reports a confidence level of 95.06%.

It may happen in some cases that the survey is not conducted exactly as planned, and the actual sample size is less than the target sample size. However, the size of the farm may also be smaller. In these cases, it is advisable to analyse the farm data on a farm-by-farm basis. For example, if only 165 specimens were collected from a farm with only 2520 fish, the resulting confidence would still be 95%. If only 160 fish were collected, the confidence is only 94.5%. If a rigid target of 95% confidence is used, then this survey would fail to meet that target and more evidence would be required.

Example 2 – Two-stage structured survey (national freedom)

Context

A country aims to declare freedom from disease Y of ornamental fish. The industry in this country is based on a relatively small number of importers and breeders, with a large number of retailers, and a very large number of domestic aquariums. The disease is reasonably highly contagious, and causes death of fish in the second half of their grow-out period. Affected animals show few characteristic signs and die within days. Almost all susceptible species in an infected tank will die once disease is introduced. It is more common in late summer, but can occur at any time of year.

Objective

The objective is to establish national freedom from disease Y. The surveillance system must be able to be practically implemented in this type of system, and with the existing constraints.

Approach

The aquaculture authorities decide to use a survey to gather evidence of freedom. The population is defined as all fish under the control of all commercial importers, breeders and retailers, as well as a list of amateur breeders and of those who participate in shows. Ornamental fish in domestic aquariums are excluded from the study population,

because there is little risk of transmission of disease from these populations. A two-stage sampling approach is used, sampling establishments at the first level and tanks at the second. Laboratory testing of specimens from a large number of establishments is not considered feasible, so a combined test system is developed to minimise the need for expensive laboratory tests.

The unit of observation and analysis is, in this case, the tank, rather than the individual animal. This means that the diagnosis is being made at the tank level, an infected tank or a non-infected tank, rather than at the animal level.

The survey is therefore a survey to demonstrate that no establishments are infected, using a random sample of establishments and making an establishment-level diagnosis. The test used to make an establishment-level diagnosis is another survey, this time to demonstrate that no tanks in the establishment are affected. A test is then performed at the tank level: owner observation followed, if necessary, by laboratory testing.

Survey Standards

- The confidence to be achieved by the survey is 95%. The power is set at 95%, but is likely to be nearly 100% if the test system used achieves nearly 100% specificity, as demonstrated in the previous example.
- The target population is all tanks with the susceptible species in the country during the study period. The study population is as described previously.
- Three tests are used. The first is owner observation, to determine whether fish are dying in a particular tank. If so, a second test is applied. The second test used is Polymerase Chain Reaction (PCR). Tanks that give a positive result to PCR are further tested using transmission experiments.
- Owner observation can be treated similarly to other tests. In this case, the observation of death is used a test for the presence of disease Y. As other diseases can kill, the test is not very specific. On the other hand, it is very unusual for disease Y to be present and not cause death, so the test is quite sensitive. A standard case definition is established for 'death', for instance, more than one fish being found dead. Based on this definition, owners are able to 'diagnose' each tank as having 'death'. Some owners may be over-sensitive, and decide that death is occurring when only a single fish is found dead, false positives, leading to a decrease in

specificity, while a small number of other owners fail to recognise death, decreasing sensitivity.

In order to quantify the sensitivity and specificity of owner observation of death, as a test for disease Y, a separate study is carried out. This involves both a retrospective study of the number of deaths in a population that is thought to be free from disease, as well as a study of owners presented with a series of prospects of death in given situations, to assess their ability to accurately identify a tank with dead fish. By combining these results, it is estimated that the sensitivity of owner-reported death as a test for disease Y is 87% while the specificity is 68%.

- When an owner detects a tank with dead fish, specimens are collected from moribund fish following a prescribed protocol.[6] Tissue samples from up to 20 affected fish are collected, and pooled for PCR testing. In the laboratory, the ability of pooled PCR to identify a single infected animal in a pool of 20 has been studied, and the sensitivity of the procedure is 98.6%. A similar study of negative specimens has shown that false positive results have occasionally occurred, probably due to occasional laboratory contamination, but maybe also because of the presence of non-viable genetic material from another source (fish-based feed stuffs are suspected). The specificity is therefore estimated at 99%.

- Published studies in other countries have shown that the sensitivity of transmission tests, which is the third type of test to be used, is 95%, partly due to variability in the concentration of the agent in inoculated material. The specificity is agreed to be 100%.

Based on these figures, the combined test system sensitivity and specificity is calculated using the formulae presented in example 1, first with the first two tests, and then with the combined effect of all three tests. The result is a sensitivity of 81.5% and a specificity of 100%.

- The design prevalence must be calculated at two levels. Firstly, the tank-level design prevalence, which is the proportion of tanks in an establishment that would be infected if infection were present, is determined. In neighbouring countries with this infection, experience has shown that tanks in the same establishment are quickly infected. It is unusual to see an infected establishment with fewer than 80% of tanks infected. Conservatively, a design prevalence of 50% is used. The second value for design prevalence applies at the establishment level, or the proportion of infected establishments that could be identified by the survey. As it is conceivable that the infection may persist in a local area without rapid spread to other parts of the country, a value of 1% is used. In this case, this is considered to be the lowest design prevalence value for which a survey can be practically designed.

- The population of eligible establishments in the country is 2302 according to available records. There are no records of the number of tanks in each of these establishments.

Sample Size
Sample size is calculated for the two levels of sampling, first the number of establishments to be sampled and then the number of tanks to be sampled. The number of establishments to be sampled depends on the sensitivity and the specificity of the test used to classify establishments as infected or not infected. As the 'test' used in each establishment is actually another survey, the sensitivity is equal to the confidence and the specificity is equal to the power of the establishment-level survey. It is possible to adjust both confidence and power by changing the sample size in the establishment survey, which is the number of ponds examined. This means that we can, within certain limits, determine what sensitivity and specificity we achieve.

This allows a flexible approach to sample size calculation. If a smaller first-stage sample size is desired, which is a small number of establishments, a high sensitivity and specificity are needed. This means that the number of tanks in each establishment that need to be examined is larger. Fewer tanks will result in lower sensitivity and specificity, requiring more establishments. The approach to determining the optimal, which is the least cost, combination of first- and second-stage sample sizes, is described in Appendix 3.

There is a further complication because each establishment has a different number of tanks, some with only a very small number. In order to achieve similar confidence and power, that is sensitivity and specificity, for each establishment, a different sample size may be required. The authorities choose to produce a table of sample sizes for the number of tanks

to sample in each establishment, based on the total tanks in each establishment.

An example of one possible approach to determining the sample size follows:

Using an extreme example, it is decided that the target sensitivity of the establishment level survey is 80% with specificity being 100%. This decision is made to minimise the number of tanks that must be examined. Using the FreeCalc software[5], with a design prevalence of 1%, which means that the survey is able to detect disease if 1% or more establishments are infected, the first-stage sample size is calculated as 350 establishments (when alpha = 0.05 and beta = 0.05). Within each establishment, the test used is the combined test system described above with a sensitivity of 81.5% and a specificity of 100%. Based on these figures the following table is developed, listing the number of ponds that need to be sampled in order to achieve 80% sensitivity.

Population	Diseased	Sample Size	Reactors	Type I	Sensitivity
< 3					
3	1	3	0	0.185	0.815
4	1	4	0	0.185	0.815
5	1	5	0	0.185	0.815
6	1	6	0	0.185	0.815
7	1	7	0	0.185	0.815
8	2	6	0	0.1333	0.8667
9	2	6	0	0.1901	0.8099
10	2	7	0	0.169	0.831
11	2	8	0	0.1527	0.8473
12	2	8	0	0.1951	0.8049
13	3	7	0	0.1537	0.8463
14	3	7	0	0.1853	0.8147
15	3	8	0	0.1608	0.8392
16	3	8	0	0.1883	0.8117
17	3	9	0	0.1662	0.8338
18	4	7	0	0.1914	0.8086
19	4	8	0	0.1628	0.8372
20	4	8	0	0.1833	0.8167
21	4	9	0	0.1586	0.8414
22	4	9	0	0.1769	0.8231
23	5	8	0	0.1634	0.8366
24	5	8	0	0.1798	0.8202
25	5	8	0	0.196	0.804
26	5	9	0	0.1682	0.8318
27	5	9	0	0.1827	0.8173
28	6	8	0	0.1772	0.8228
29	6	8	0	0.1907	0.8093
30	6	9	0	0.1621	0.8379
31	6	9	0	0.1743	0.8257
32	6	9	0	0.1863	0.8137
33	7	8	0	0.1869	0.8131
34	7	8	0	0.1984	0.8016
35	7	9	0	0.1681	0.8319
36	7	9	0	0.1785	0.8215
37	7	9	0	0.1888	0.8112
38	8	8	0	0.194	0.806
39	8	9	0	0.1634	0.8366
40	8	9	0	0.1725	0.8275
41	8	9	0	0.1816	0.8184

Population	Diseased	Sample Size	Reactors	Type I	Sensitivity
42	8	9	0	0.1906	0.8094
43	9	8	0	0.1995	0.8005
44	9	9	0	0.1678	0.8322
45	9	9	0	0.1759	0.8241
46	9	9	0	0.184	0.816
47	9	9	0	0.1919	0.8081
48	10	9	0	0.1641	0.8359
49	10	9	0	0.1714	0.8286
50	10	9	0	0.1786	0.8214
51	10	9	0	0.1858	0.8142
52	10	9	0	0.193	0.807
53	11	9	0	0.1676	0.8324
54	11	9	0	0.1742	0.8258
55	11	9	0	0.1808	0.8192
56	11	9	0	0.1873	0.8127
57	11	9	0	0.1938	0.8062
58	12	9	0	0.1705	0.8295
59	12	9	0	0.1766	0.8234
60	12	9	0	0.1826	0.8174

Population	Diseased	Sample Size	Reactors	Type I	Sensitivity

Sampling

First stage sampling, or the selection of establishments, is done using random numbers and a sampling frame based on the available data. The establishments are listed on a spreadsheet with each numbered from 1 to 2302 and a sample selected using a random number table (see Appendix 3) or software designed for the generation of random numbers such as EpiCalc.[7]

The second stage of sampling involves random selection of tanks within each establishment. This requires a sampling frame, or list of each tank in the establishment. The authorities recruit establishments by telephone. For each selected establishment, an officer calls to invite participation, and ask how many eligible tanks are in the establishment. A simple random sample of the appropriate number of tanks, between 3 and 9, from the table above, depending on the number of tanks in the establishment, is selected from this list. This process is described in Appendix 3. Identification of the actual tank is based on the owners' own numbering system for the tanks.

Testing

Once tanks have been identified, the actual survey consists of testing them. In practice, the owners observe the tanks during a defined period, say two months. Fisheries staff make regular phone calls to each establishment to check whether there are dead fish in any of the selected tanks. If any are seen, that is, the first test is positive, up to 20 moribund fish are collected for laboratory examination, first PCR, and then, if positive, transmission experiments.

Analysis

Analysis is performed in two stages. First, the results from each establishment are analysed to ensure that they meet the required level of confidence. If the target sample size is achieved and only negative results obtained, the confidence should be 80% or greater in each establishment. At the second stage, the results from each establishment are analysed to provide a country level of confidence. Again, if the target sample size, which is the number of establishments, is achieved, this should exceed 95%.

Comment

This example is somewhat contrived in its use of two-stage sampling, but is designed to illustrate the technique. In practice, one-stage sampling would be more likely to be used, with the establishment being the unit of observation, or cluster sampling with all tanks on the establishment being included in the survey. This is possible because of the low expense of the initial screening test of owner observation. If more expensive test procedures were applied to each tank, a two-stage approach may be warranted to limit the number of tanks in the survey.

Example 3 – Spatial sampling and the use of tests with imperfect specificity

Context

A hypothetical oyster culture industry, based primarily on rack culture of oysters, is distributed between 23 estuaries distributed along the coastline. In similar regions in other countries, disease Z causes deaths of oysters in late summer / early autumn. If an outbreak was to occur, experience from elsewhere has shown that a high proportion of oysters would be affected, however it is suspected that that the agent may be present at relatively low prevalence in the absence of disease outbreaks.

Objective

The national authorities want to determine whether the country is free from disease Z. If the disease should be detected, a secondary objective of the survey is to collect adequate evidence to support zoning at the estuary level.

Approach

The authorities conclude that clinical surveillance for disease outbreaks is inadequate because of the possibility of low-level subclinical infections. It is therefore decided to base surveillance on a structured two-stage survey, in which sampled oysters are subjected to laboratory testing. The first stage of the survey is the selection of estuaries. However, due to the objective of providing evidence for zoning, should infection be found in any of the estuaries, it is decided to use a census approach and sample every estuary. In essence this means that there will be 23 separate surveys, one for each estuary. A range of options for sampling oysters is considered, including sampling at harvest or marketing, or using farms (oyster leases) as a level of sampling or stratification. However, the peak time of activity of the agent does not correspond to the harvest period, and the use of farms would exclude the significant numbers of wild oysters present in the estuaries. It is therefore decided to attempt to simulate simple random sampling from the

entire oyster population in each estuary, using a spatial sampling approach.

Survey Standards

- The target population is all of the oysters in each of the estuaries. The study population is the oysters present during the peak period of risk of disease in late summer/early autumn. Wild and cultured oysters are both susceptible to infection, and may have associated with them different, but unknown, risks of infection. They are therefore both included in the study population. Sampling is based on mapping, which is described below. Therefore the study population can be described more accurately as that population falling within those mapped areas identified as oyster habitats.

- A design prevalence value is required only at the oyster level, as a census is being used at the estuary level. While the disease is often recognised with very high prevalence during outbreaks, a low value is used to account for the possibility of persistence of the agent in the absence of clinical signs. A value of 2% is selected.

- The test used is examination for microscopic lesions with immuno-staining techniques. This test is known to suffer from occasional false positives due to non-specific staining, but is very sensitive. Published studies indicate values of 99.1% for sensitivity and 98.2% for specificity. No other practical tests are available. This means that it is not possible to differentiate false positives from true positives, and that in a survey of any size, a few false positives are expected, that is, 1.8%.

- The confidence is set at 95% and the power at 80%. In the previous examples, due to the assumed 100% specificity achieved by use of multiple tests, the effective power was 100%. In this case, with imperfect specificity, there will be a risk of falsely concluding that a healthy estuary is infected, so the power is not 100%. The choice of a relatively low figure (80%) means that there is a 1 in 5 chance of falsely calling an estuary infected when it is not infected, but it also decreases the survey costs greatly, through a lower sample size.

Sample Size

Based on the assumption that the sampling procedure will mimic simple random sampling, the sample size, which is the number of oysters to sample per estuary, can be calculated with FreeCalc.[5] The population size, which is the number of oysters per estuary, is assumed to be very large. The calculated sample size, using the sensitivity, specificity and design prevalence figures given above, is 450. The use of FreeCalc also shows that, based on this sample size and the specificity of the test, it is possible to get 10 or fewer false positive test results, and still conclude that the population is free from disease. This is because, if the population were infected at 2% or greater, the anticipated number of positive reactors from a sample of 450 would be greater than 10. In fact, we would expect 9 true positives (450 * 2% * 99.1%) and 8 false positives (450 * 98% * 1.8%) or a total of 17 positives if the population were infected at a prevalence of 2%.

This illustrates how probability theory and adequate sample size can help differentiate between true and false positive results when there is no alternative but to use a test with imperfect specificity.

Sampling

The aim is to collect a sample of 450 oysters that represent an entire estuary. Simple random sampling depends on creating a sampling frame listing every oyster, which is not possible, and systematic sampling depends on being able, at least conceptually, to line up all the oysters. This again is not possible. The authorities decide to use spatial sampling to approximate simple random sampling. Spatial sampling involves selecting random points, defined by coordinates, and then selecting oysters near the selected points. In order to avoid selecting many points with no oysters nearby, the estuary is first mapped; the fisheries authorities already have digital maps defining oyster leases available. To these maps, areas with significant concentrations of wild oysters are also added, based on local expertise. Pairs of random numbers are generated such that the defined point falls within the defined oyster areas. Other schemes are considered, including using a rope marked at regular intervals, laid out on a lease to define a transect, and collecting an oyster adjacent to each mark on the rope, but the random coordinate approach is adopted.

Survey teams then visit each point by boat, using a GPS unit to pinpoint the location. A range of approaches is available for selecting which oyster to select from a densely populated area, but it should involve some effort at randomness. Survey staff opt for a simple approach: when the GPS receiver indicates that the site has been reached, a pebble is tossed in the air and the oyster

closest to the point where it lands is selected. Where oysters are arranged vertically, for example, wild oysters growing up a post, a systematic approach is used to determine the depth of the oyster to select. First, an oyster at the surface, next, an oyster halfway down, and thirdly, an oyster as deep as can be reached from the boat.

This approach runs the risk of bias towards lightly populated areas, so an estimate of the relative density of oysters at each sampling point is used to weight the results (see Appendix 3 for more details).

Testing

Specimens are collected, processed, and analysed following a standardised procedure. The results are classified as:

- positive, defined as showing strong staining in a highly characteristic pattern, possibly with associated signs of tissue damage,
- probable positive, defined on the balance of probabilities, but less characteristic staining, and
- negative.

Analysis

The interpretation of the results when using a test with imperfect specificity is based on the assumption that, in order to conclude that the population is free from infection, any positive result identified is really a false positive. With a sample size of 450, up to 10 false positives may be expected while still concluding that the population is free from infection. However, if there is one positive result, then the population cannot be considered free from infection. This is the reason for the classification of results that are not negative into positive and probable positive. If there are any positives at all, the population in that estuary must be considered infected. The probable positive results are consistent with false positive results, and therefore up to 10 may be accepted. Using FreeCalc, the actual confidence achieved based on the number of false positives detected can be calculated. For instance, if 8 probable positive results were detected, the confidence level for the survey would be 98.76%. On the other hand, if 15 probable positive results were detected, the confidence is only 61.9%, indicating that the population is likely to be infected.

Discussion

Normally, it may be assumed that a surveillance system aimed at demonstrating freedom from disease is 100% specific. This is because any suspected occurrence of infection is investigated until a decision can be made. If the decision is based on one or more positive results, then there is no basis for declaring freedom; the infection has been shown to be present.

This example presents a different situation where, due to lack of suitable tests, it is not possible for the surveillance system to be 100% specific. This may represent an unusual situation, but illustrates methods for dealing with this sort of problem. In practice, a conclusion that a country or estuary is free from infection, in the face of a small but statistically acceptable number of positive results, will usually be backed up by further evidence, such as the absence of clinical disease.

Appendix 3: Methods for sampling

This appendix provides guidance on specific sampling methods for use by people conducting surveys of aquatic animals in developing countries. Additional information is available.[8]

A) Use of Random Numbers

Random sampling is based on the concept of randomness, and the use of random numbers. Random numbers are best explained by an example. Dice have six numbered sides, 1 to 6, and when one is rolled, each side has the same chance of ending on top. However, at each roll, we do not know which number will come up. What we do know is that if we roll the dice again and again, on average, all numbers will appear equally often. Rolling dice is one way of generating random numbers, in this case, between 1 and 6.

Playing cards are another example of randomness. When we shuffle cards, we do not know what order they are in. But it is possible to predict what will happen over a large number of games of cards. This is because each card in the pack has the same probability of being on top of the deck.

When selecting a sample for a survey, we want to select members of the population in such a way that will ensure that each member has exactly the same chance of being selected. There are many different ways of doing this, and the aim of this chapter is to explain some methods that are useful to people conducting aquatic animal surveys in developing countries.

The examples of dice and cards are called physical randomisation techniques because we take physical objects and mix, shake, roll or shuffle them. This represents one of the simplest approaches to selecting a random sample. The problem with dice is that there are only 6 numbers, although there are decimal dice with 10 sides numbered 0 to 9. Blank cards are much more flexible, as shown by the following example.

A large shrimp farm is surveyed to determine whether it is infected with Taura syndrome. There are 30 ponds (the unit of interest) in the farm (the population), but only 8 are needed for testing. Each pond has a unique identification number. To select the sample of 8 ponds, 30 blank cards are used. The number of each pond is written on a card. The 30 cards are then shuffled well, and 8 cards selected. The ponds with the numbers selected are the ones to be included in the sample.

This is an effective method for random selection. However, when the group is too large, it can quickly become impractical. For instance, a national survey will examine fishing practices in 100 different villages. If there are 24,200 villages in the country, this would require writing the village name on 24,200 cards, shuffling and dealing out 100. Shuffling 24,200 cards could be difficult. For this reason, physical randomisation techniques are only used in small surveys. In larger surveys, the use of random numbers is more convenient.

Random numbers are numbers that have been generated randomly, or by chance alone. That means that for each digit, the chance of it being any number between 0 and 9 is the same. There are two sources of random numbers: random number tables, and computer generated random numbers.

Example: A computer was used to generate a random number, 39024. The number has 5 digits, 3, 9, 0, 2 and 4. When the computer selected the first digit, 3, the chance of it being any number between 0 and 9 was exactly the same. The number 3 was selected by chance. For the second digit, 9, again, the chance of any number between 0 and 9 being selected was the same. The fact that 3 had been selected for the first number made no difference to which would be selected for the second number and so on. It was simply a matter of chance. It is as if the computer is rolling special decimal dice (with 10 sides) and recording each digit as it is rolled.

Various computer programs include random number generators, as explained below. Computers and random number tables are equally useful to select a sample using the following steps:

Step 1: Make a list of all the members of the population. This list is called the sampling frame.
Step 2: Number each member on the list from 1 up to N, where N is the total number in the population.

Step 3: Using a computer or a random number table, select random numbers between 1 and N. Select one random number for each element to be selected in the sample.

Step 4: For each random number selected, find the corresponding element on the list. These are the ones to be included in the sample.

Random number tables are a convenient source of random numbers. An example of a random number table is shown below and full tables are available in various texts. There are sets of numbers grouped into 5. In the above example, 10 numbers were selected to pick 10 villages from a total of 75. One way to use a random number table to select random numbers is as follows:

Step 1:Choose a starting point and direction. You can start at the top of the table, or you can start anywhere in the middle. You can go across a row, or down a column. In this example, we will start at the top left number, and move across.

Step 2:Calculate the range for your random numbers. The numbers required in this example are between 1 and 75.

Step 3:Determine which digits to use from the numbers. The maximum number we want is 75, which has two digits. We therefore only need two of the five digits in each random number. To use the numbers efficiently, we can 'cut' them in half, and think of the first two digits (42) as the first number, and the third and fourth digits (53) as the second number. The last digit can be ignored.

Step 4:Search through the table for numbers in the required range. Any number between 1 and 75 is counted as one of our random numbers. Any number over 75 is ignored. Continue searching until enough numbers have been found (ten in this example).

Example: Using the table below, the first number is 42. This is between 1 and 75, so it is accepted. The second number is 53, and is also accepted. The next digit (9) is ignored. Moving to the right to the next group, the next number is 77. This is greater than 75 so is ignored. The next number is 68, which is accepted as our third random number. The last digit (6) is ignored. Continuing in this way we get a fourth (66), a fifth (52), a sixth (27), discard the next (79), a seventh (02, or 2), and eighth (47), a ninth (57) and a tenth (05, or 5).

42539	77686	66524	27792	02474	57058	61530	76108	49436
27030	88085	84744	32591	57804	54790	24545	73422	23337
50253	66592	66151	18506	04391	35824	35397	32031	67780
54127	25147	79021	54189	43708	08178	82187	72106	53795

It is helpful to distinguish selected numbers, by circling them, from discarded numbers, by crossing them off. Different sets of random numbers should be used for different surveys.

While using a random number table is quick and simple, the job can be done even more conveniently using a computer. Various programs are available to select random numbers, and one is included in EpiInfo.

B) Sampling Individual Aquatic Animals

There are many situations when, in order to understand what is going on within a system, it is necessary to sample individual animals. For instance, to determine whether a batch of shrimp post-larvae is infected with White Spot Syndrome Virus (WSSV) a sample of post-larvae may be selected and examined with a polymerase chain reaction test (PCR) for WSSV DNA.

The method for selecting a sample from a batch of 2000 post-larvae is similar to selecting farms from a district. For the inference that a sample will represent the population from which it was obtained to be valid, a method of random sampling must be used.

The same principles apply whether we are assessing the prevalence of Epizootic Ulcerative Syndrome in a pond, trying to estimate the population of fish in a reservoir, or calculating the incidence rate of QX disease in oysters in an estuary. In each situation, we need to examine individual animals, and these animals need to be randomly selected, in order to accurately represent the population.

There are three types of random sampling that may be applied to different situations. In summary, these are:
- Simple and two-stage random sampling, which is based on an existing or new sampling frame,
- Systematic random sampling, which is based on the ability to 'line up' the population into a sequence, and
- Spatial sampling, which is using location as a proxy for the sampling frame.

Whether considering farms, ponds, cages, villages or districts etc, it is generally possible to obtain an existing sampling frame or to create a new one. If not, systematic sampling may be used, and if this is still not possible, some form of spatial sampling is possible. The

characteristics of these populations that make these forms of sampling appropriate are:
- The size of the population is generally relatively small, or the population can be grouped into a relatively small number of sub-populations, for example, the number of villages in a province or the number of farmers in a village,
- The populations remain relatively constant over time,
- The populations are static (they do not change position in space),
- The individual elements in the population, for example, village or farmer names, pond or cage numbers, can be identified relatively easily, and
- The elements can be examined relatively easily, for example, looking at a pond, measuring a cage, interviewing a farmer.

In contrast, most aquatic animals share none of these characteristics:

- They often are grouped in relatively homogenous large populations, for example, thousands of fish in a pond,
- The populations change rapidly, for example, production cycles are relatively short, disease problems can cause very high mortality rates in a short time period,
- The populations are highly mobile, except for the mature stage of molluscs,
- It is not possible or practical to identify individuals in most populations, and
- It is difficult to examine individuals, as they are difficult to catch and examination often causes severe stress or death.

Another consideration is that aquatic animals are kept in a wide variety of different systems, ranging from the open ocean, reservoirs and rivers, to cages, ponds, raceways, nets, tanks and so on. The challenges and opportunities for random sampling animals in these different systems differ widely. There is no single method that is applicable in every situation. Instead, this section attempts to clarify the objectives of random sampling of individual aquatic animals and the principles of how it may be achieved, and presents some examples.

When planning a disease survey that includes sampling of individual animals, designing a sampling strategy that is practical and that obtains a representative sample is one of the most difficult challenges. In many cases, it is not possible to meet the requirements of random sampling. If random sampling cannot be used, there is the risk that any sample will

not be representative, and that the survey results will be biased. The essential task during survey design is to work out, within the practical constraints of the system, how to select a sample that minimises any possible bias.

Understanding the principles of random sampling will assist in ensuring that the survey is conducted in the correct way. Understanding the physical, cultural, economic and social constraints of conducting a survey clarifies what is possible. This will almost always fall short of what is known to be the ideal way to sample. The least compromise of the ideal and the possible will achieve the best design for the survey.

When the ideal method for sampling is not possible, it may be possible to determine what sort of biases are likely to be caused by the practical, imperfect strategy that has to be used. If the direction of any likely bias can be determined, this makes the results much more useful. For instance, if a sampling strategy was judged to be more likely to catch healthy fish than diseased fish, and a survey yields a disease prevalence of 18%, we could conclude that we don't know what the real prevalence is, because proper random sampling was not used, but whatever it is, it is certainly equal to or greater than 18%.

Strategies used to select a sample
An ideal sample is one that represents the population, with respect to the characteristic of interest. This can be achieved reliably only using random selection. Simple random sampling requires a sampling frame that contains every member of the population once and only once and that uniquely identifies them. It also requires that selected members of the population can be caught and examined. In aquatic animal populations, this is very rarely possible. Systematic sampling removes the need for a sampling frame, but requires that the population can be lined up, either physically, or conceptually, and that animals at a regular interval can be selected.
The third strategy that may be used is spatial sampling. This identifies random locations instead of elements of the population, and depends on the animals being relatively static in that location.
Other non-probability approaches to sampling may be used, when probability sampling is not possible. For instance, haphazard sampling is often able to select a representative sample, but is not as reliable as random sampling. Haphazard sampling is far better than

convenience or purposive sampling, which should not be used. Practical strategies employed in the field to get a representative sample generally attempt to overcome some of the constraints imposed by the characteristics of aquatic animal populations listed above, so that one of these four approaches can be applied.
Most of the techniques involve one or a combination of the following approaches:
- Lining the population up so that systematic sampling can be used;
- Decreasing the mobility of the animals;
- Using some form of spatial sampling.

A number of management activities provide an opportunity for systematic sampling. When such an opportunity exists, this is usually the best way to select a good representative sample. This will be discussed in the next section on management opportunities.
Decreasing the mobility of animals, and spatial selection strategies are used in situations where systematic sampling is not possible. These approaches will be discussed under the specific production system.

Sampling techniques and equipment
Farmers and fishermen have been catching fish and other aquatic animals for many thousands of years using a huge range of different tackle and equipment. These include nets, traps, lines and spears, with many variations on each. Poisons, electricity and explosives have been added to the list more recently. Several typical methods will be used as examples. However, there are two very important considerations when assessing equipment to use for sampling aquatic animals.

The first is that the survey team should know about, and fully understand, the common and traditional methods used for catching aquatic animals in the survey area. Before suggesting that a new approach be used, it is worth considering why the existing system is being used. In most cases, local tackle has evolved because, under the local circumstances, it is the most practical and efficient approach to catching aquatic animals. Understanding this, the survey team should have a very good reason for suggesting that something new be used instead.

The second consideration is to understand the differences between the aims of local producers and survey staff. Local farmers or fishermen may have a range of objectives. For instance, farmers may wish to collect all animals during a harvest. Alternatively, they

may use a series of partial harvests; in which case, they are only interested in catching fish of marketable size. Similarly for fishermen, they may be primarily interested in catching large fish, and deliberately avoid catching immature fish, to ensure the sustainability of their livelihood.

This preference for large rather than small fish may have other, unintended effects. For instance, large fish may tend to be healthy fish, whereas the diseased fish are smaller. The fish captured using traditional techniques may therefore not be representative of the population, but show a bias towards healthy fish. Any survey based on this type of sample would underestimate the level of disease. Conversely, some methods of catching aquatic animals may catch slower individuals more easily than faster ones. A sample collected in this way would be likely to catch more diseased animals, and the survey results would overestimate the level of disease in the population.

A decision on the appropriate tackle to use depends on an assessment of these two issues: what is the most practical and efficient way of capturing individual aquatic animals in the local environment, and what sorts of biases may this method introduce.

Opportunities for sampling
The variety of species and management conditions used in aquaculture and fisheries means that different sampling approaches need to be used in different situations. However, within a single management system, there are sometimes good and bad times to attempt random sampling. This is because, during the production cycle, there are a number of management activities that provide an opportunity for sampling, while during the rest of the cycle, sampling may be much more difficult. In general, these management opportunities for sampling provide situations where systematic sampling is possible because, in some sense or other, the population can be 'lined up'.

Taking advantage of these management opportunities enables the collection of a representative sample from a production system in which it would otherwise be impossible. However, there are some constraints.
The first constraint is the time of sampling. When sampling during specified management activities, the survey team has no control over when the sampling will occur. The time of

specific management activities are determined by the producers. If this fits in with the requirements of the survey, then there is no problem. However sometimes it may not. For instance, a disease may be most severe during the growing phase. During this phase, animals are left undisturbed. At harvest time, when sampling is easier, affected animals may have either died or have recovered, so there is little evidence of disease. Using harvesting as an opportunity to detect such a disease would not be appropriate.
The second constraint of sampling during specific management activities is the need for an excellent relationship with the producers. Harvesting, for example, is a busy time, involving a lot of work, and the time at which a farmer has a chance to earn income after months or years of investment. If survey activities make it difficult for the farmer to do their work, or if they damage the animals or cause the farmer to lose some of this income, then the involvement of the survey team at harvest time will not be welcome. Establishing a good relationship with the farmer, understanding their activities and needs, and ensuring that survey activities bring as much benefit and as little disadvantage as possible to the farmer are all necessary to take advantage of management opportunities for sampling.

Stocking
Stocking an aquaculture system or enhancing a wild-capture fishery with aquatic animals provides a good opportunity to evaluate the disease status of the juvenile animals. Because the fry, fingerlings, post-larvae or spats are small and worth less than adult animals, they are easier to handle and cheaper to collect. Testing animals at this early stage of life indicates whether specific pathogens have been introduced to a culture system, but doesn't help if we are interested in studying the progress or effect of disease.

Consider, for example, Silver Barb fry. Typically, fry for stocking are delivered from the hatchery or nursery in plastic bags, with a specified number of fry in each bag. Many farms may stock fish from a number of different hatcheries, either depending on availability or to spread the risk of receiving fish of poor quality. In a survey, the question of whether to treat one batch from one hatchery as a single population, or all the batches from different hatcheries as one population depends on the survey question being asked (see Chapter 10 of reference 6). However, usually each batch should be

considered as one population. A single batch may be made up of one or more bags.

Among the options for collecting a random sample, some will ensure that the sample is representative of the population, but require more effort, while others will be more practical, but less representative and so introduce bias.

One example of a simple approach is to use a two-stage sampling scheme. First select a number of bags from all the bags in the batch, and then select a number of fry from each selected bag. While this is fast and practical, there are a number of problems. If the sample size calculations are conducted using the methods outlined in Chapter 11 of reference 6, the number of bags will usually be high, meaning that most or all of the bags have to be sampled. The second problem is which fry to select from a bag.

A preferable approach is to treat the batch as a grouped population, and select a random sample using appropriate techniques. This will produce a list for inclusion in the sample that specifies we need fry numbers 10, 38, 43, 52, 57 and 69 from bag number 1, fry numbers 8, 23 and 48 from bag number 2, and so on (Chapter 11 of reference 6).

A third approach is to collect a systematic sample from all bags. For instance, if there were 5 bags, each with 300 fry, and a sample of 50 was required, then the population is 1500 and the sampling interval would be 1500/50 or every 30th fry. The first fry to select would be selected randomly from between one and 30.

Once the sampling approach and individuals to sample have been worked out, the next step is to select those individuals. There are various techniques to do this. The one most commonly used is disturb the fry and scoop some out. If 10 are needed from a bag, and there are more than 10 in the scoop, the rest are returned. If there are fewer, another scoop is taken. This approach is not random sampling but convenience sampling. It is very likely that the fry that are near each other are similar in some way. For instance, if the bag were stirred to try to mix the fry, the weaker ones would come towards the centre of the swirl. If the scoop was taken from there, it may be possible to collect all weak fry, and bias the results. Alternatively a scoop from the edge may collect all strong fry, with the opposite bias.

A similar and equally practical approach is to assess the number of fry on the basis of the volume of water. If a bag contains 300 fry in about 1 litre of water, and we want a sample of 100 fry from the bag, we might mix the bag, and pour out 330 mL of water into a second bag. This should contain about 100 fry. Again, this is a form of convenience sampling, and is unlikely to even get the right number.

A practical, but much more time consuming approach can be used to use either the systematic or grouped population random approach mentioned above. Often, when fry are delivered, the stated total number per bag is an approximation only. This is because the hatchery didn't count, but estimated the number in each bag, or else some of the fry died on the way to the farm and dead animals are unsuitable for diagnostic testing. To be sure of the number stocked and therefore the survival rate, many farmers may choose to count the fry before stocking. To do this they often use a small container, sieve, or perforated spoon. This container lets most of the water out, and holds a known number of fry of a certain size. For instance, one spoonful may contain, on average, 10 fry. By transferring the contents of the bag to another container, one spoon at a time, the farmer can count the total number of fry reasonably accurately and quite quickly. In the process, the fry can be thought of as being 'lined up', suitable for systematic sampling.

In our example of systematic sampling above, we wanted to sample every 30th fry. If the first random number chosen was 17, then 17 fry would be transferred, and then one selected. From then on, one fry every three spoonfuls would be taken for the sample, until the whole population was counted. This still leaves the question of which fry to take. This doesn't matter, as long as it is consistent. For instance, the one nearest the handle of the spoon may be taken.

This approach, while being laborious assures that a representative sample will be chosen, unlike the convenience approaches mentioned above.

The other random approach mentioned above, based on a grouped population and simple random sampling, can be achieved using a similar approach. All the fry are counted from one container to another. Instead of selecting every 30th, this time, fry corresponding to the randomly selected numbers are chosen. For instance, if the first fry to be collected is

number 27, two spoonfuls are transferred. On the third spoonful, 6 fry are counted out, and the 7th is taken for the sample. This approach is more time consuming, and is unlikely to produce a significantly better sample that the systematic approach.

The stated and the actual number of fry in a bag may differ. The only way to check this is to count the fry as described above. However, if the sampling design is based on the stated number in a bag, it may cause problems if the number is much larger or much smaller than claimed. For instance, if using systematic sampling and the number is much smaller than claimed, the total sample size will be smaller than expected. In this case it may be decided to test only those collected, or to collect more, until the desired sample size is reached. If collecting more fry, be sure to select fry distributed through the entire population so as not to introduce a bias. For instance, it is not appropriate to start again at the first bag, continuing to collect every 30th fry, as this means that the first bag is sampled twice, and the last bags only once.

If bags contain multiple species, whether all species are considered as part of one population, each is a separate population or only one species is studied, depends on the question to be answered.

Grading
Another good opportunity for sampling is during grading operations. Some species, such as macrobrachium shrimp or salmon, animals are graded during the growing phase. This is done to divide animals based on size or sex to obtain more even groups, ensuring fairer competition during feeding and more even growth rates.

Grading may be performed in a number of ways, such as manually, or using buckets or sieves with holes of a specified size to let smaller animals through. However it is conducted, grading involves handling every animal, and is therefore an excellent opportunity for systematic sampling. Sampling may be made even more efficient if it is stratified by size. This can be done by taking a systematic sample after grading rather than before.

Sampling at grading has an important advantage over most other management opportunities for sampling, as it occurs during the middle of the growth period, rather than at

the beginning or the end. If disease is likely to be active at this time, sampling during grading would be the most appropriate method of selecting a representative sample.

Farmers are busy during grading, so sampling then requires good cooperation between farmer and sampler. Handling risks damaging or stressing animals, so should be minimised.

Vaccination
In some more intensive systems, animals are vaccinated for particular diseases, such as vibriosis. Injectable vaccines are administered to individual animals, so systematic sampling is feasible during such vaccination.

Transfers
Some management systems require the transfer of fish from one cage to another for various reasons, such as cleaning or repairing cages. Transfer can be done in different ways, during which systematic sampling is feasible.

One approach is to use a fish-pump to pump the fish, in water, into the new cage. This effectively lines up the fish, although they may be passed through the pump unevenly, so that several come at once, and then none come for a period. Sampling may be possible by counting (even approximately) the fish emerging from the pump, and capturing fish at a regular interval. With large numbers of fish, it may be simpler to catch one fish at a regular time interval, rather than trying to count each fish as it emerges. Just as spatial sampling uses locations as a proxy for the sampling units, this approach uses time as a proxy. Catching one fish every 30 seconds is not an exact form of systematic sampling, but it is still likely to get a representative sample.

Another approach to transferring fish is to pass them through a race. The technique used may differ according to different physical layouts, but in some cases, using a scoop-net to take fish from the outlet of the race may be adequate, counting the fish as they emerge.

Transfers may be done manually, by lifting a net to crowd the animals, then catching them in containers, for example, buckets, and transporting them to a new site. This approach is very convenient, as each bucket can be assumed to hold an equal number of animals. Systematic sampling then involves the collection of one animal from every 6th (say) bucket, or alternatively, 2 animals from each bucket, depending on the sample size. Deciding which animal to choose from a

bucket should, of course, be done randomly – simply picking one off the top may introduce bias, for instance, if smaller animals tend to near the surface. This is discussed below, under Sampling Animals from Buckets or Crates.

Molluscs may be transferred from one place to another during their growth cycle, and can be sampled systematically at that time. During transfers, they will usually be gathered into small groups, such as trays or on sticks. Each group may be assumed to comprise an equal number. If time permits, simple random sampling may be used, based on the techniques for sampling from a grouped population. A faster approach may be to use systematic sampling, where one oyster is taken from a regular number of groups.

Harvesting

Harvesting provides one of the best opportunities for sampling aquatic animals, for the following reasons:
- Every animal is handled,
- Animals are fresh, but will soon die, so handling them, or keeping them out of water will not cause damage, or upset the farmer,
- They are often transported in groups, for example, in crates or buckets, which makes systematic sampling convenient, and
- In some systems, for example, shrimp, the product is graded at the time of harvest in order to determine the price. Grading is usually done by selecting a sample, for example, 10 kg of shrimp per 500 kg produced, and examining it in detail. The collection of the sample for grading is often done using haphazard sampling, but systematic sampling is possible without increasing the workload. A sub-sample of the graded shrimp may be collected using systematic sampling at the grading table.

Despite these advantages, there are a number of potential problems with sampling at the time of harvest:

- Only complete harvests give access to the entire population. In many systems, partial harvests may be used, and the harvested animals are very unlikely to be representative of the entire population. Progressive harvests are another form of partial harvest, where a small number of animals are taken out regularly. This is common practice in smallholder fish ponds.

- Harvesting is a busy time, and survey staff must work closely with the farmer to fit in and avoid disrupting the work.
- Harvesting may take an extended period of time, even several days. In order to collect a good sample, the survey staff need to be present throughout the harvest.
- The decision to harvest may be made at very short notice, leaving little time for preparation or travel for the survey staff. Many farmers will do an emergency harvest if they suspect there is a danger of a severe disease outbreak.

- Except in the case of some emergency harvests, harvesting usually occurs when animals are mature. This may not be the time of interest for studying the disease.

Capture Sampling

When a study requires that animals are sampled from a closed population in, for example, a pond, tank, cage, net, or reservoir, or an open population in a river or ocean, but none of the above management opportunities are available, it is necessary to capture the animals. In smaller closed populations, where every individual can be captured, for example, by draining a pond or tank, systematic sampling can be used.

Capture Techniques

This section deals with the more common situation where some but not all animals can be captured. Many techniques have been developed around the world, each designed to capture different species under different environmental and cultural conditions. Common examples include cast nets, dip nets, lines, traps, and trawl nets. Other techniques are poison, explosives and electro-fishing.

One of the important characteristics of fishing tackle and other capture systems is their selectivity, or their ability to capture different parts of the population, which introduces a bias into the sample captured. For example, when capturing shrimp with a net, the diameter of the opening, rate of movement of the net, and the mesh size all influence which shrimp will be captured. In addition, the way in which the tackle is used may also influence the bias. A net used at intermediate depths at the edge of a pond may capture mostly healthy shrimp, while one used in the middle of the pond, or at the bottom may capture mostly weaker or diseased shrimp.

The result is that all capture techniques are biased, and none is able to produce a reliably representative sample. This is because they are not able to meet the basic requirement of random sampling: that all members of the population have an equal chance of being selected in the sample. This poses a major problem for studying aquatic animal populations. The simple solution is to avoid sampling based on capture techniques, and use techniques described earlier in this section. However, when this is not possible, methods to minimise bias and ensure that a more representative sample is captured, are needed.

Capturing a more representative sample

Minimising bias
One of the best approaches is to minimise any bias that is caused by the choice of capture method. This requires a good understanding of the tackle being used and the factors influencing the catchability of different parts of the population. For example, it may be possible to use a series of nets of different mesh sizes, each with an ability to capture different segments of the population. Nets with a very fine mesh size may capture the smallest fish, whilst larger mesh sizes miss the smaller fish, but capture the larger ones.

Selecting the best capture system or combination of systems often requires the assistance of experts, with knowledge both of fishing tackle and the biology and behaviour of the species being targeted. This should be combined with local knowledge of the area being sampled, to identify all potential habitats.

Regardless of the care taken, there is still a large opportunity for the resulting catch to be a biased sample of the source population.

Some capture techniques that are commonly used to sample aquatic animals cause more bias than others. Shrimp are often sampled using a feed tray, but this will almost guarantee a biased sample – only actively feeding shrimp will be found in the sample, while sick shrimp will never be detected.

Spatial and temporal sampling
Aquatic animal populations regularly move from place to place, especially in open waters such as rivers, oceans and estuaries. These movements may be over relatively small distances, even just changes in the preferred depth, or can involve migrations over thousands of kilometres. Even in closed populations, such as a shrimp pond, different animals can be found in different locations at different times. Because of these movements of animals, sampling in a single location will usually not be able to collect a representative sample.

Knowledge of the movement patterns of the species being studied and of the factors that control these patterns can help control bias in a sample, and can also increase the efficiency of a sample markedly. For instance, understanding the migration patterns of fish along a river may mean that the entire population can be sampled at a single location at a particular time of the year, during the peak of their migration.

To capture a representative sample using spatial sampling (sampling at different locations) and temporal sampling (sampling at different times) are two main issues that must be considered. The first is the pattern and scale of movements. Some species may be found at different depths at different times, and may move distances of several kilometres, while others move over vast distances. The second is the factors affecting these movements. Long-distance migrations are usually seasonal, and have an annual cycle. Changes in depth or location may be related to day and night, or the phase of the moon, or other factors, such as mating behaviour. An understanding of all these factors will help identify the best location and time to sample animals.

While the use of a particular capture technique, such as a cast net, may not be a good approximation of random sampling, it is still possible to improve the representativeness of the overall sampling by selecting random locations and random times at which to use the capture technique.

Example: A survey is being conducted of one species in an estuary. Nets are used to capture the fish, and it is estimated that the net will have to be used 50 times to catch the required sample. In order to get the best sample, points are selected at random from all over the estuary, as described in chapter 5 of reference 6. One sample will be taken at each point. However, to take into account the possibility of different parts of the population being located in different places at different times, each point is assigned a random time, for example, a random hour between 1 and 24, at which that point will be sampled.

This approach increases the chance of a more representative sample. It can be thought of as sampling three aspects: Sampling the fish (using a non-random capture technique), sampling space at random, and sampling time at random.

With each of these three aspects, a 'population' and sample from it can be defined. For the fish, the population is all the fish of the relevant species in the estuary. For space, the population is all the points or locations in the estuary. For time, the population is all the hours in a day. For

seasonal movement patterns, the population is all the months in the year.

If we have a good understanding of the biology and distribution of the fish, we can further refine the definition of the population to make the sampling more efficient. For instance, if we know that some areas of the estuary are suitable habitats for the fish, but others are unsuitable, we might narrow the 'population of points' from which we sample, by sampling at random only from suitable habitats. Similarly, if we know the species is only present during one season, we may sample random weeks during that season.

Crowding
One of the reasons why capture sampling is almost always biased is because aquatic animals are able to move. When a cast net is thrown, some animals will be able to move out from underneath it, thereby avoiding capture. The animals that are not able to move quickly are often the weakest or smallest, which means that the sample is likely to be biased. One way of attempting to overcome this problem is to use crowding to decrease the mobility of the population.

Crowding can be used in closed populations, mainly for animals in cages and nets. It may also sometimes be used in tanks or ponds, when a large net is used to collect all the animals into a small part of the tank or pond. Once the animals are crowded together, they are much less able to move. It is then possible to use a spatial random sampling technique, by picking random locations, and perhaps random depths, from the area of crowded animals.

Example: A net is used to crowd fish in a pond into one side of the pond. Three decimal dice are used to pick random numbers. The first is the number of metres from the left end of the net. The second is the number of metres from the bank into the net. The third is the number of centimetres deep in the water. By rolling all three dice, a specific random position and depth is chosen. The fish at this location is selected for the sample.

The advantage of crowding is that the fish at the chosen location is not able to move away quickly, and is therefore somewhat easier to identify and catch. Large fish and small fish, health fish and sick fish will all be kept pressed together, and so have a roughly equal chance of being selected.

The main disadvantage of using crowding is that it risks stressing and damaging the animals, and may therefore be unacceptable to the farmers. If it is used, it should be done for the shortest possible time to minimise any damage. This means good planning and preparation, and selecting random locations before starting. In some cases, haphazard sampling is used rather than formal random sampling, as it is faster. However it must be remembered that this always runs the risk of introducing biases in the results.

Sub-sampling
Sub-sampling is the process of drawing a smaller sample from a larger sample. Sub-sampling is particularly useful when sampling open populations, using commercial capture techniques. Commercial techniques, such as trawling, are often able to capture very large numbers of animals quite quickly. If all the animals captured are included in the sample, either the required sample size will be captured from a single location in a short time, or, if many locations are sampled, the total catch will result in an enormous sample size, much large than is actually required. This is likely to increase the cost of the survey, particularly if some sort of laboratory examination is involved in the analysis. The solution is to take a sub-sample from the total catch and use that as the sample.

Sub-sampling may involve some work, but, as the total population (catch) to be sampled can be handled, either simple random sampling or systematic sampling ensures a representative and unbiased sub-sample. For example, when a catch of fish is dumped onto the deck of a trawler for sorting, every 20th fish could be selected for the sample, while the remainder are discarded, or retained for commercial purposes.

Sub-sampling can also be applied to commercial catches after they have been brought to shore. Systematic sampling of a catch during unloading at the dock, or random sampling of fish at markets both represent forms of sub-sampling the commercial catch.

While formal random techniques can be applied when sub-sampling, it is important to keep in mind how the initial sample (the commercial catch) was taken. Random sub-sampling ensures that the sub-sample is truly representative of the popluation from which it is drawn (in this case, the commercial catch). The sub-sample is not representative of the wild population from which the initial sample

was taken. If the capture technique used for commercial fishing is biased, as it usually is, then the catch is not a representative sample of the wild population, and nor is the sub-sample from the catch.

Understand bias

When a study is being conducted, and the only available sampling method is to capture the animals, it is often impossible to ensure that the sample is not biased. When surveying aquatic animals in the open waters of rivers, estuaries or the ocean, there are no management opportunities to allow systematic sampling and there is no ability to crowd the animals. The use of carefully selected and applied capture techniques, and of random spatial and temporal sampling to get good coverage of the population will both help decrease an bias, but it is often not possible to eliminate it. Even in closed environments, such as a shrimp pond, if a capture technique such as the use of a cast net is required for practical reasons, it is likely that the results will be biased. One way of dealing with this problem is to conduct studies allowing us to understand the bias, and take it into account when interpreting the results of the study.

Example: A study of shrimp is planned, and animals are to be sampled using a cast net. The research workers are aware that cast nets are more likely to capture sick and slow shrimp than healthy, fast shrimp, and that the results are of the study are likely to be biased. In order to understand what the bias might be, they conduct a preliminary study immediately before harvest. A cast net is used to sample several ponds, and an estimate the proportion of sick shrimp is obtained. The pond is then harvested, and a systematic sample used to estimate the proportion of sick shrimp. The two estimates from the two sampling techniques are compared. It is found that, by the (biased) cast net sample, about 9% of shrimp are found to be sick. Using the (unbiased) systematic harvest sample, only 6% of shrimp are found to be sick. The research workers conclude that the bias caused by the cast net sampling technique results in an overestimate of the prevalence of disease, by about 50%. They go ahead with their cast net study, but when reporting the results, they adjust their estimates to account for this bias.

This example demonstrates that bias is a problem only if its size and direction are unknown. If the bias caused by a sampling technique can be measured, then it is possible to adjust for it.

Re-define the population

To reiterate, it is almost impossible to draw a truly representative sample from an open population. If the sample is biased, then it is invalid to use estimates made from the sample to make inferences about the total population.

A practical solution to this problem is to re-define the population of interest. Normally, when studying a species, we are interested in all members of that species – big and small, young and old, sick and healthy – in order to understand the true effect of disease. However, if it is not possible to represent all groups of animals in our sample, we may ask: what does our sample actually represent? Often, the answer is "catchable fish", or those fish in the population that can be caught by the capture technique used. Our sample contains only a few fish, but, if we are careful with where and when we sample, it is likely to be representative of those fish that can be caught. Clearly, our sample does not represent any of those fish that could not be caught by our capture technique.

By re-defining the population, we can overcome the problem of bias and inference. If we find that 24% of fish in the sample have a disease, we may confidently presume that about 24% of the catchable fish in the population also have that disease. Unfortunately we are still left with an important question: – what about the others? How many uncatchable fish have the disease? It is rarely possible to provide a valid answer to this question.

Sampling less mobile animals

One last special case of sampling is worth considering. Most of the problems with sampling aquatic animals stem from their ability to move quickly in three dimensions. However mature molluscs are immobile, which offers opportunities to sample from farmed and wild populations.

Management opportunities

In many culture systems, molluscs such as oysters are moved from one location to another during different stages of growth and systematic sampling is convenient and easy. Often oysters are grown in trays or on sticks with a similar number in each group. To obtain a systematic sample, every 100th oyster, for example, may be sampled. If a tray contains on average 30 oysters, 3 trays are skipped, and in the 4th tray, the 10th oyster (counting systematically across, then down) is sampled.

The following 20 oysters on that tray and all oysters on the two next trays are skipped, and the 20th oyster is taken from the next tray – and so on.

Spatial sampling
If wild molluscs or cultured molluscs need to be sampled while they are in the water, spatial sampling techniques can be used, as follows:
1. Define the area to be studied, for example, an estuary or a single bay.
2. Identify and map areas suitable for the mollusc species being studied.
3. Generate and number random points in the study area, retaining only those that fall within areas suitable for molluscs.
4. Using a hand-held global positioning system (GPS) unit or a very accurate map, visit each random point.
5. Establish a minimum search distance around the point (for example, 5 m). If molluscs are found in this radius, select one at random. If no molluscs are present, move to the next point.

Using this approach will obtain a reasonably good representative sample. The estimate from this sample may be improved by weighting the data according to the density of molluscs around each selected point.

C) Weighting of spatial samples

Consider the example of selecting ponds at random using random points. If a point falls inside a pond, then that pond is selected. However, if a point falls outside a pond, which most points normally will, what should we do? One option is to find the nearest pond, and select it. This is known as nearest neighbour sampling. While this sounds sensible, it may again lead to serious bias. Imagine two shrimp farming areas, one very intensive with many ponds, and another with one pond and no other farms nearby. In the first, intensive area, a random point needs to be very close to a particular pond, in order to select that pond. If it is far away, then there will be another pond that is closer. This means that only a small number of random points will identify any particular pond as the closest. In the second situation with the one pond, random points that are located a very long way away will still select that one pond, because it is still the closest. When we use this type of sampling in areas where there are some units close to each other, and some that are sparsely spread, the result is that the sparse units are much more likely to be selected in the sample than the

densely spread units, because there are more random points that are closest to the sparsely spread units.

The way to overcoming this problem of bias towards sparsely located points in nearest neighbour sampling is to select only those units that are within a certain distance of a selected point. For instance, we could select ponds that are within 50 m of the selected point; if there are no ponds within that distance, then a new point is chosen. This is known as fixed distance weighted sampling. This approach solves the problem of bias towards very sparsely located ponds, but introduces yet another problem. What if there are two or more ponds within 50 m of the select point? The short answer is to choose one of them randomly. When there is more than one pond, we can think of the one we chose as representing all the ponds within 50 m of the point. Because the one pond chosen is representing other ponds, we need to give more emphasis to the results from this pond. This is done by weighting the results from that pond during analysis.

Ponds that fall within a fixed distance around each randomly selected point are eligible for inclusion in the sample. If there are no ponds within the distance, a new point is chosen. If there is one pond within the set distance, that one pond is selected. If there is more than one pond within the distance, for example, 6, one of the ponds is selected at random, for example, by using dice, and it is included in the sample. During analysis, this one pond from the dense area is given six times as much importance as the ponds from the other areas, by using weighting.

D) Least-cost two-stage sampling for freedom from disease

When using two-stage sampling, a survey can produce a result of the same accuracy by using a variety of combinations of first- and second-stage sample sizes. For instance, if a small number of cages are selected and many fish from each cage are tested, it is possible to get the same accuracy as if many cages are tested, and only a small number of fish are tested from each cage. By changing the type I and II error levels used for selecting the sample size for the second stage (testing fish within a cage), we are also changing the sensitivity and specificity of the cage test (used when selecting cages at the first stage). This enables

us to produce a variety of different sample size combinations, all of which will provide the same level of evidence for freedom from disease.

This flexibility is one of the advantages of two-stage sampling, because not all the combinations will cost the same. The overall cost depends on the costs of testing a single fish and a single cage. For surveys to demonstrate freedom from disease, the complexity of the calculations means that it is not possible to use a formula to work out the best combination.

Instead, it can be done using trial and error with the FreeCalc program. The following procedure may be used to calculate the best combination of first- and second-stage sample sizes:

Step 1:Determine the basic measures that cannot be changed. This includes the sensitivity and specificity of the laboratory test, the population of cages (first-stage population size), an estimate of the average fish population of the cages, the maximum acceptable prevalence of the disease among cages (first stage) and fish (second stage), and the overall type I and type II error levels for the survey. These are used when calculating first-stage sample size. The cost of testing a single fish, and the costs associated with sampling a single cage must be known.

Step 2:Pick starting values for the farm test sensitivity and specificity. The higher these values are, the fewer cages need to be tested, and the more fish need to be tested in each cage. If they are very high, there may not be enough fish in some cages to achieve this level. In general, try to make the specificity as high as possible.

Step 3:Calculate the number of cages needed using the selected cage test sensitivity and specificity.

Step 4:Now use the same figures to calculate the second stage sample size. Set the type I error to 1 - Sensitivity, and the type II error to 1 - Specificity. Change the sensitivity and specificity to those of the laboratory test, the population to the average cage size, and the prevalence to the maximum acceptable or minimum expected prevalence within the cage.

Step 5:Calculate the number of fish that need to be tested.

Step 6:Using the number of cages, and the number of fish, calculate by hand the total cost of the survey, based on the cost estimates, and record the result.

Step 7:Now go back to calculating the first-stage sample size, but change either the sensitivity or specificity or both. Repeat the calculations in steps 3 to 6, and record the sample sizes and total cost of this alternative combination.

Step 8:Continue testing new values until the one that gives the cheapest cost is found.

References

1. Martin SW, Meek AH, Willeberg P. *Veterinary epidemiology: Principles and Methods*. Iowa State University Press, Ames, 1987.

2. Dohoo I, Martin W, Stryhn H. *Veterinary epidemiologic research*. AVC Inc. Charlottetown, Prince Edward Island. 2003.

3. Office International des Épizooties. *Diagnostic manual for aquatic animal diseases*. 4th edn. Office International des Épizooties, Paris, 2003.

4. Office International des Épizooties. *Aquatic animal health code*. 6th edn. Office International des Épizooties, Paris, 2003.

5. Cameron AR. FreeCalc. Software for the calculation of sample size and analysis of surveys to demonstrate freedom from disease. Available for free download from http://www.ausvet.com.au.

6. Handlinger J. Collection and Submission of Samples for Investigation of Diseases of Fin Fish. Available for free download from http://www.affa.gov.au/aquaplan

7. Epicalc Software available for free download from http://www.myatt.demon.co.uk/epicalc.htm

8. Cameron AR. *Survey toolbox for aquatic animal diseases – a practical manual and software package*. Australian Centre for International Agricultural Research, (ACIAR) Monograph No 94. ACIAR, Canberra. 2002.