

Animal Health in Indonesia
Advanced Field Epidemiology
Manual



Table of Contents

Introduction to advanced topics in field epidemiology	4
1 Disease Investigation	5
1.1 Measures of disease frequency	5
1.1.1 Prevalence	5
1.1.2 Incidence	6
1.1.3 Attack rate	10
1.2 Data analyses to describe patterns of disease	11
1.2.1 Confirm the outbreak.....	11
1.2.2 Temporal patterns.....	11
1.2.3 Spatial patterns	14
1.2.4 Animal patterns	16
1.2.5 Measuring association between disease and risk factors.....	18
1.2.6 Advanced statistical analyses	26
1.3 Developing hypotheses and control measures	27
1.4 Role of tracing in disease investigation	31
1.5 Further field epidemiology studies	31
1.5.1 Types of epidemiologic studies.....	32
1.5.2 Choice of field study	39
1.5.3 Defining groups based on outcome and exposure.....	39
1.5.4 Selection of animals	40
1.5.5 Bias, confounding and interaction.....	42
1.5.6 Sample size.....	47
1.5.7 Planning a field epidemiological study	50
1.5.8 Interpreting field data and information	59
1.5.9 Preparing the report.....	60
1.6 References and other resources	61
2 Diagnostic tests	64
2.1 Measures of diagnostic test performance	64
2.1.1 Accuracy and precision	64
2.1.2 Ability of a test to detect disease states	66
2.1.3 Sensitivity.....	68
2.1.4 Specificity.....	68
2.1.5 Predictive Values.....	70
2.2 Multiple testing	71
2.2.1 Sensitivity and specificity for multiple tests	71
2.2.2 Conditional independence of tests.....	74
2.2.3 Application of series and parallel testing	77
2.3 Measuring agreement between tests	79
2.3.1 Proportional agreement of positive and negative results.....	81
2.4 Estimation of true prevalence from apparent prevalence	81
2.5 Group (aggregate) diagnostic tests	83
2.5.1 Calculating herd sensitivity and herd specificity	84
2.5.2 Risk of infection in test-negative animals	86
2.5.3 Demonstrate freedom or detecting disease?.....	89
2.5.4 Important factors to consider in group testing.....	89
2.6 Estimating test sensitivity and specificity	91

2.6.1	Gold-standard methods.....	91
2.6.2	Non-gold-standard methods	93
2.7	References – diagnostic testing	97
3	Disease control and eradication programs	99
3.1	What is meant by control or eradication of disease	99
3.2	Why have a regional control or eradication program?	100
3.3	Types of programs	102
3.3.1	Eradication programs	102
3.3.2	Control programs.....	102
3.4	Strategies that may be used for disease control	104
3.4.1	Detecting the disease agent	105
3.4.2	Reducing the number of infected hosts	107
3.4.3	Increasing resistance of susceptible hosts	108
3.4.4	Reducing contact between infectious and susceptible hosts	109
3.4.5	Supporting activities.....	112
3.5	Pre-requisites for a successful program	114
3.6	Application of control measures based on infection status	117
3.6.1	Control measures applied to an infected premise.....	117
3.6.2	Control measures that may be applied to a disease free premise	119
3.7	Example: Rabies in Bali	119
3.8	Designing an appropriate animal health program	122
3.9	Monitoring program performance	123
3.10	Economics of animal disease control	124
3.10.1	Data requirements for economic evaluation.....	124
3.10.2	Methods for economic evaluation	126
3.10.3	Macroeconomics vs microeconomics	130
3.11	References – economics of animal disease control	131

Introduction to advanced topics in field epidemiology

This manual is intended to complement the *Basic Field Epidemiology Manual*. The *Basic Field Epidemiology Manual* has been produced to as a resource for para-veterinary staff in Indonesia and to accompany training for para-veterinarians in routine tasks they undertake while providing diagnostic, treatment and disease prevention services for the benefit of Indonesian livestock and their owners.

This manual on *Advanced Topics in Field Epidemiology* is produced to provide a resource for veterinarians to aid in investigation of priority or emerging infectious disease, and for assessment and utilisation of iSIKHNAS disease data to conduct disease control at district, provincial, and central government levels.

This manual aims to assist veterinarians to be able to:

- Apply epidemiology skills in planning, implementing, analysing and reporting on disease investigation activities. This includes conducting additional field investigation studies to provide improved understanding of causes of disease and help inform on the optimal control strategies;
- Understand the interpretation of diagnostic test data and information. Apply interpreted diagnostic test results to disease investigation and control activities including the development and interpretation of new diagnostic tests;
- Describe, assess and compare options for control or eradication of animal diseases.

The information presented in this advanced manual is intended to be used in conjunction with the *Basic Field Epidemiology Manual*.

1 Disease Investigation

The *Basic Field Epidemiology Manual* provides information on the clinical examination, making a diagnosis, and the initial epidemiologic approach to disease investigation.

This section provides more advanced material to assist in the planning and implementation of epidemiologic investigations of disease and in the analysis of data collected from investigations. It is assumed that people reading this material will have read through the material above presented in the *Basic Field Epidemiology Manual*.

1.1 Measures of disease frequency

Counts of cases, non-cases, and population at risk can be used to estimate epidemiological measures of disease frequency. This allows you to describe disease events and compare disease events over time. The most used measures are prevalence and incidence.

1.1.1 Prevalence

The *prevalence* of a disease (or condition) is the proportion of cases in a population at a given point in time.

Budi looks after 15 cows; he has contacted you because 5 animals are sick with the same signs of disease. Soleh looks after 30 cows in the same area as Budi's cows and none of Soleh's cows are sick. All the cows all graze together in one area.

The prevalence of disease is

$$P = \frac{5}{15+30} = 0.11 = 11\%$$

$\textit{Prevalence} = \frac{\textit{number of cases}}{\textit{population at risk}}$
--

The simplest estimate of the population at risk is the sum of cases and non-cases (all animals) on the farm or farms where the investigation is taking place.

Prevalence may not always provide a good measure of the occurrence of new cases of a disease. An example of this is when cases are based on results of a serology tests measuring antibodies, a marker of prior exposure. In this situation a positive test may indicate prior infection from months or years previously and may not have anything to do with the current disease.

In order to understand whether prevalence may reflect recent disease cases or a mixture of old and recent cases, you will need to understand the nature of the disease being investigated, how animals respond to being infected and the type of diagnostic test that has been used to classify animals as cases or non-cases.

1.1.2 Incidence

The *incidence* is the number of **new** cases that arise in a population over a specified period of time. Unlike prevalence, incidence reflects *risk*, or the likelihood of an individual animal contracting the disease in a given period of time.

Incidence can be calculated in different ways:

- cumulative incidence (CI) or incidence risk
- incidence rate (IR) (or incidence density)

Cumulative incidence is the number of animals that develop the disease in a defined period of time divided by the number of healthy animals at risk at the beginning or start of the period of time. All incidence measures should be based only on new cases of disease that occur in the time period of interest.

Budi watches the 45 cows over the next 7 days. He calls and says that 8 more cows have become sick. The animals at risk are those that are disease free at the start of a period. At the start of the 7 day period there are 45 cows but 5 have already been diagnosed with the disease so this leaves 40 that were disease free at the start of the 7 day window.

$$CI = \frac{8}{40 \text{ (7 day period)}} = 0.20 = 20\%$$

Estimation of incidence is easiest when the population of animals does not change over time (no new animals arrive and no animals move away). This is called a *closed population*. Often the population may change over time because new animals arrive or some animals may be sold or moved away. This is called an *open population* or *dynamic population*.

If animals are lost to follow-up during the period of interest, then the denominator should be adjusted to take this into account. The same principle applies if any new animals enter the population at risk during the time period. There are several commonly used methods for

adjusting the count of population at risk in open populations. The two approaches most commonly used for cumulative incidence estimation are described below:

1. Number of disease free animals at the start of the follow-up period for closed populations (simplest approach but may be biased if there is a lot of animal movement).
2. Population size at mid-point of the follow-up period for open populations which may be estimated as

- $Average\ number\ at\ risk = N_{Start} + \frac{1}{2}N_{New} - \frac{1}{2}N_{Lost}$

- $Average\ number\ at\ risk = N_{Start} + \frac{1}{2}N_{New} - \frac{1}{2}(N_{Lost} + N_{Cases})$

$Cumulative\ Incidence = \frac{Number\ of\ new\ diseased\ animal\ in\ the\ time\ period}{Average\ number\ of\ animals\ at\ risk\ at\ the\ mid\ point\ of\ the\ period}$

At the start of the year there were 1000 cows in the population of interest, all were disease free. A total of 400 cows were sold half way through the year. During the year 20 cows died from anthrax.

$$N_{Start} = 1000$$

$$N_{New} = 0$$

$$N_{Lost} = 400$$

$$N_{Cases} = 20$$

$$Number\ at\ risk = 1000 - 0.5*400 - 0.5*20 = 790$$

$$Cumulative\ incidence = 20/790 = 0.025 = 2.5\%$$

Incidence rate is the number of new cases of disease that occur per unit of animal-time at risk during the defined follow-up period. Incidence rate requires more detailed estimation of animal-time at risk compared to cumulative incidence.

For closed populations (no movement of animals in or out), the denominator is the number of disease free animals at the start multiplied by the length of the follow up period (in days, weeks, months or even years).

For open populations (animal movement in or out) we need to adjust the denominator to take movements into account. Again the simplest way to do this is often to add half the number of animals that were added to the population at risk during the period and subtract half the number that left the population at risk, using the same approach as method 2 described above for cumulative incidence. The adjusted number is then multiplied by the length of the follow-up period.

If there is detailed follow-up data on individual animals, we can generate the denominator animal-time at risk by calculating the exact time at risk for each individual animal and summing these.

Example of incidence rate using approximate count for PAR

At the start of the year there were 20 cows in the population of interest that were all disease free. A total of 3 cows were sold half way through the 12 month period. During the year 2 cows died from haemorrhagic septicaemia.

$$N_{\text{Start}} = 20$$

$$N_{\text{New}} = 0$$

$$N_{\text{Lost}} = 3$$

$$N_{\text{Cases}} = 2$$

$$\text{Number at risk} = 20 - 0.5 \times 3 - 0.5 \times 2 = 17.5$$

$$\text{Time period} = 12 \text{ months}$$

$$\text{Animal time at risk} = 17.5 \times 12 = 210 \text{ animal months}$$

$$\text{Incidence rate} = 2/210 = 0.0095 \text{ cases per animal month}$$

We can change the units of the animal-time at risk to 100 animal-months

$$0.0095 \text{ cases per animal-month} = 0.95 \text{ cases per 100 animal-months}$$

Example of incidence rate using exact estimate of PAR

Assume 4 healthy cows were present on the farm at the start of the period and they were followed for 30 days.

1 cow was not sick at all throughout the 30 days = 1 animal month at risk

1 animal got sick on day 10 = 0.33 animal months at risk

1 animal got sick on day 20 = 0.67 animal months at risk

1 animal was sold on day 15 = 0.5 animal months at risk

Total animal time at risk = 2.5 animal months

Total new cases = 2

Incidence rate = $2/2.5 = 0.8$ cases per animal month

Some Important issues to remember:

- Incidence is a dynamic measure of disease whereas prevalence is only a static measure of disease
- Incidence and prevalence are related. The prevalence of disease in a population-at-risk reflects both the incidence of new cases of disease and the duration of disease in individual cases: **Prevalence = Incidence x Duration** under certain conditions.
- Changes in the incidence or the duration of a disease will change the prevalence. The incidence rate is usually greater than prevalence if the disease is short in duration and/or fatal. Prevalence is usually greater than the incidence if the disease is chronic in nature.
- Cumulative incidence (CI) rate provides a direct estimate of the likelihood of an animal experiencing the event of interest during the time period. CI has a meaning on an individual basis as well as on a population basis.
- Counting the denominator (animal time at risk) for incidence estimates can be problematic particularly if animals enter or leave the population during the time of interest. There are a number of ways to deal with this problem.

Table 1.1: Comparison of main features of prevalence incidence rate and cumulative incidence

	<i>Point prevalence</i>	<i>Period prevalence</i>	<i>Incidence rate</i>	<i>Cumulative incidence</i>
Numerator	All cases counted at a single occasion in time	Cases present at period start + any new cases during period	New cases during follow-up period	New cases during follow-up period
Denominator	All individuals examined	All individuals examined	Sum of time at risk for susceptible animals present at start of period	All susceptible individuals present at start of follow-up period
Time	Single point in time	Defined follow up period	Measured for each individual from start to end of period, until disease occurs or until animal exits the population	Defined follow up period
Study type	Cross-sectional	Cohort	Cohort	Cohort
Interpretation	Probability of disease at a point in time	Probability of having disease over a defined period	How quickly new cases develop over a defined follow up period	Probability of developing disease over a defined period

1.1.3 Attack rate

Attack rate (or called attack risk) is a specific type of incidence estimate (either a cumulative incidence or an incidence rate) which applies to outbreaks or situations where the period of observation is relatively short and where the population at risk is tightly defined eg the number of animals on the farm under investigation.

An attack rate is the number of cases of the disease divided by the number of animals at risk at the beginning of the outbreak (the outbreak covers a defined time interval).

$$\mathbf{Attack\ rate} = \frac{\mathit{number\ of\ animals\ affected}}{\mathit{number\ of\ animals\ exposed}}$$

For example the attack rate can be used to measure mortality due to yellow head virus infection in prawns. If over a 4 day period 3500 of the 5000 prawns in a pond die the attack rate is $3500/5000 = 0.7$ or 70%.

1.2 Data analyses to describe patterns of disease

It is assumed that the initial clinical examination and case definition have been completed and that data have been collected on cases and non-cases on those farms where diseased animals are located.

1.2.1 Confirm the outbreak

Often the major reason that an epidemiologic investigation is begun into a particular disease is because there is concern that there are more cases of disease than expected. Possible reasons may be an outbreak of a new disease, an exotic disease or some change in an endemic disease that has made it more infectious or capable of causing different disease signs.

As soon as a case definition is produced and counts are made of cases and non-cases it is important to produce some simple estimates of frequency (prevalence or incidence) to confirm that there is an increase in disease cases above what is expected and that further investigation is warranted.

1.2.2 Temporal patterns

Variation in the frequency of occurrence of cases of a disease over time is called its *temporal pattern*. There are three basic time spans that may be used to describe temporal patterns:

- an epidemic period, which is the time the start of a disease outbreak to the end of the outbreak (may vary from days to weeks or months or longer);
- a 12 month period to describe seasonal patterns; and
- a long period of many years to identify long-term trends.

The simplest temporal pattern for disease cases is an epidemic curve. The epidemic curve is a graph plotting the number of cases of the disease on the vertical axis against the time of onset of each case, either as a bar graph or frequency polygon. The first case identified for a particular outbreak is referred to as the index case. For infectious diseases, identifying the index case is important as information about the index case can be valuable in ascertaining the source of the outbreak and the incubation period.

*An **epidemic curve** is a graph of the number of cases of disease against the time of onset of each case*

In general, an epidemic curve has four distinct components and in some cases there may be a secondary occurrence of additional cases (a 5th component). These are displayed in the following figure.

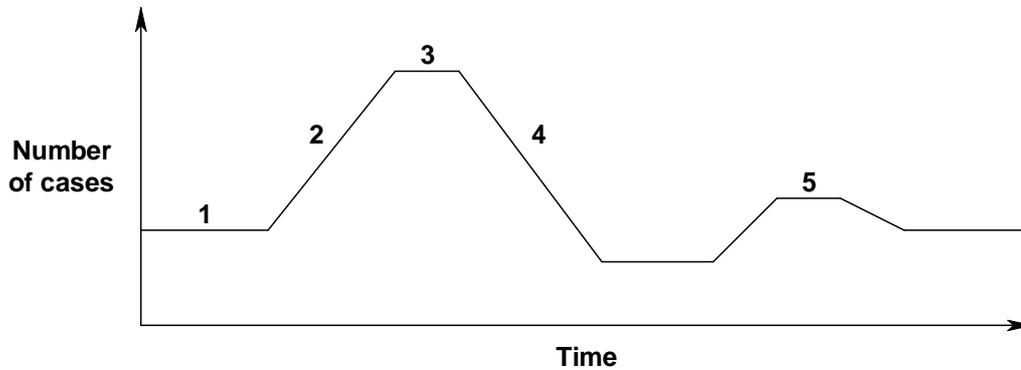


Figure 1.1: Diagrammatic representation of the components of an epidemic curve. 1= endemic level, 2=ascending branch of epidemic, 3=peak of epidemic, 4=descending branch of epidemic, 5=secondary peak.

The slope of the ascending branch can indicate the type of exposure (propagating or common source) or the mode of transmission and incubation period of the disease agent. If transmission is rapid and the incubation period short, then the ascending branch will be steeper than if transmission is slow or if the incubation period is long.

A *point-source epidemic* is one where all animals (units) are exposed to the source of disease (agent or toxin) over a very short period of time, resulting in a very steep ascending branch of the epidemic curve.

A *propagating epidemic* is one where transmission occurs among individuals in the population, so that the ascending branch ascends more gradually.

The length of the plateau and slope of the descending branch are related to the availability of susceptible animals which in turn depends on many factors such as stocking densities, introductions into the population, the changing importance of different mechanisms of transmission and the proportion of immunes in the population at risk.

Secondary peaks are usually due to the introduction of new susceptible animals in to the diseased area, spread of the disease into a new area containing susceptible animals, or a change in the mode of transmission.

The interval of time chosen for graphing the cases is important to the subsequent interpretation of the epidemic curve. The time interval should be selected on the basis of the

incubation period of the disease and the period over which the cases are occurring. Choice of a time interval may also depend on the frequency with which animals are being examined in order to determine when any one animal first shows signs of disease.

For many livestock diseases epidemic curves are often produced using one-day (daily) intervals on the horizontal axis, producing a plot displaying the number of new cases of disease each day. If there are multiple days in between occurrence of new cases then it may be sensible to aggregate the time to weekly or some other interval.

Menangle virus was first identified in 1997, following an investigation of an outbreak of mummified and stillborn foetuses in a commercial piggery at Menangle, New South Wales, Australia.

For the Menangle virus outbreak in an Australian piggery in 1997, temporal patterns were analysed on a weekly basis, because many piggery records are maintained as weekly averages and the epidemic extended over a >20-week period. In addition to the percentage of affected litters per week, average litter sizes and numbers of piglets that were live, mummified or stillborn were plotted, providing a comprehensive picture of the temporal pattern. All indices showed a very rapid rise from week 15 (of the calendar year), when the outbreak started, to week 21, when case numbers peaked. This pattern is strongly suggestive of a propagating epidemic with a rapidly spreading agent and relatively short incubation period (see Figure 2 from Love, et al., Australian Veterinary Journal 79(3):192-198, 2001 for a graphical representation of these patterns).

1.2.3 Spatial patterns

Spatial patterns refer to describing the outbreak in terms of where animals were located (place) when they first showed signs of disease (disease onset). Spatial patterns may assist with finding the source of the outbreak. It is often useful to consider place and time together. This can be done by drawing a plan of the spatial layout of the farm (or population), recording the location and dates when cases occurred. Such a diagram may also give a lead to whether the outbreak is a common source or propagating.

Where disease cases are occurring on a small geographic scale it may be easy and simple to draw simple maps showing disease cases and non-cases. This approach can be done by anyone and does not require special computer software or mapping data to allow computer mapping.

Where disease cases are occurring over a larger area or on a larger scale, it may be more effective to map cases using computer based mapping or Geographic Information Systems (GIS) software. This may require special expertise and additional background mapping files for that location.

Where disease cases are occurring over a more extended period (weeks or more) it is very useful to produce maps at daily, weekly or longer intervals to monitor progress of the epidemic and identify patterns of spread.

For example, the following Figure shows the layout of households in a Thai village, overlaid with the occurrence of cases of foot-and-mouth disease. From this map, it is apparent that this is a propagating epidemic, with the index case identified by a red circle, a small number of secondary cases identified in week 2 and additional cases in week 3. It also appears that infection has spread locally from the index case to a number of nearby households, as well as to some more remote households, where there has also been local spread. The initial spread was perhaps through utilisation of common grazing, allowing close contact between early cases and susceptible animals from elsewhere in the village. This was probably followed by local spread among clusters of households and perhaps from infected animals moving on laneways through the village.

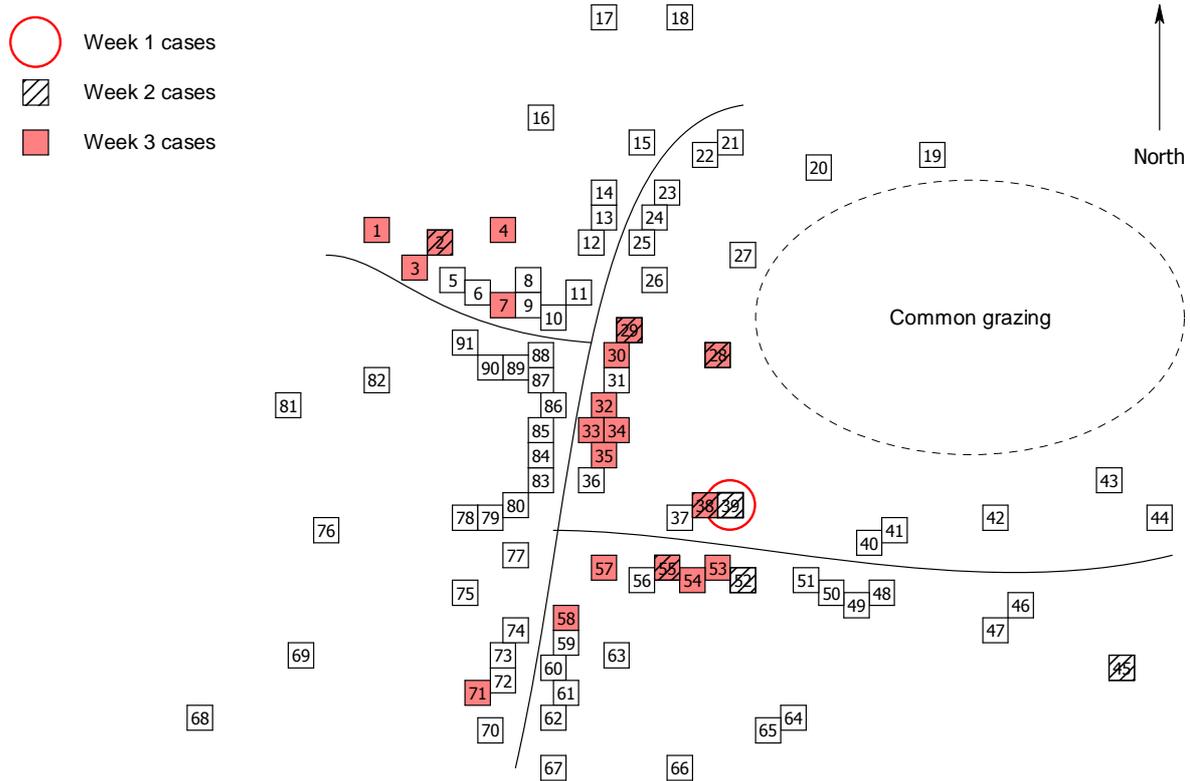


Figure 1.2: Spatial representation of the spread of foot-and-mouth disease over a 3-week period in a Thai village, adapted from Cleland, et al., (1991).

For the Menangle virus outbreak, the piggery comprised four separate management “Units”. Unit 1 was about 200m from Unit 2, while units 2 and 3 and 3 and 4 were each separated by about 50m (see Figure 1 in Kirkland, et al., 2001). Although all units were affected, 44% of litters were affected in Unit 2, compared to 28%, 26% and 37% for Units 1, 3 and 4 respectively. Analysis also showed that Unit 3 was affected first, in week 15. Other units were subsequently affected in weeks 23 (Unit 2), 24 (Unit 4) and 27 (Unit 1). It was also observed that a fruit bat colony (the hypothesised source of infection) was in close proximity to Units 3 and 4. Unit 1 was furthest from the hypothesised source and was the last unit to be affected, while Units 3 and 4 were closest to the hypothesised source.

1.2.4 Animal patterns

The term animal patterns is used to refer to some measure of disease frequency (prevalence or incidence for example) produced for different animal characteristics (species, breed, age, sex, weight class, vaccination status, stocking density) or for different levels of other management type things like location (paddock, pen, village), feed type or other variable.

We use some measure of disease frequency to describe animal patterns of disease. The most common approach in a disease outbreak is to estimate *attack rates* (AR) but other measures may be used. This approach allows investigation of possible risk factors for the disease

For example, in the foot-and-mouth disease example shown in the previous figure, 8 of 21 buffalo <1 year old were affected for an attack rate of 0.38 or 38% (Cleland et al., 1991). In contrast, 34 of 158 buffalo >1 year old were affected (attack rate = 0.215 or 21.5%). This suggests that young animals were almost twice as likely to be affected as older animals.

For example say there were deaths due to suspected *epizootic ulcerative syndrome* (EUS) in a pond and it appeared that small fish were at greater risk of having EUS than large fish. We might make the following calculations:

For small fish, AR1 =	No. of small fish with EUS
	Total small fish
For large fish, AR2 =	No. of large fish with EUS
	Total large fish

There were 1000 small fish in the pond and 300 had EUS and there were 1000 large fish of which 100 developed EUS during the outbreak. The attack rates here are 30% and 10% respectively, suggesting that small fish were 3 times more likely to develop EUS than large fish. This finding could lend support to a hypothesis that nutritionally stressed fish are more susceptible to infection.

Factor-specific attack rates for such factors as species, age, sex, feed, mob, management system etc can be explored. An example is shown below for EUS where size indicating nutritional stress is suspected.

Table 1.2: Table showing counts of fish arranged by fish size and EUS disease status along with attack rates for each class of fish size and relative risks comparing the risk of disease to that observed in a reference category (fish size = large).

Fish size (factor)	Count of fish		Epidemiologic measures	
	EUS cases	Total count	Attack rate (AR)	Relative risk (RR)
Small	30	100	30%	6
Medium	20	200	10%	2
Large	15	300	5%	reference

In the above table, attack rates are expressed as percentages. The last column is the *Relative Risk* or *Risk Ratio (RR)* which is the ratio of the attack rates comparing the AR in small and medium sized fish to the AR measured in large fish.

The higher the relative risk, the more impact the specific factor has in increasing the risk of disease. Small fish were 3 times more likely to have EUS than medium sized fish and 6 times more likely than large fish. Also, medium sized fish were at twice the risk of large fish.

Then we need to think about what differences in fish size mean. It may be that smaller fish are younger or under more nutritional stress and that these factors (age, feed availability, stress) may be the causes driving the apparent association between fish size and EUS disease risk.

1.2.5 Measuring association between disease and risk factors

Once the data are collected on temporal, spatial and animal-level patterns, it needs to be analysed to understand patterns and identify potential risk factors. The most commonly used measures for comparing disease-risk among groups are relative risk (or risk ratio) and the odds ratio. These are discussed briefly below.

1.2.5.1 2 x 2 tables

A 2x2 table is a simple way to present summary counts of diseased animals where the animals can be classified by disease status (case or non-case) and on some other possible risk factor with two levels (eg vaccination status classified as vaccinated or not-vaccinated). 2x2 tables are very commonly used in epidemiology to assess possible associations between disease occurrence and possible risk factors.

Often disease status (diseased or not-diseased) are arranged in columns and the risk factor (present=exposed, not present=unexposed) is assigned to the rows. A 2x2 table has the following format.

Table 1.3: Layout for a 2x2 table showing counts of animals arranged by disease status (columns) and risk factor status (rows).

	Disease status		
Risk factor	Diseased (case)	Not diseased (non-case)	Total
Exposed	a	b	a + b
Unexposed	c	d	c + d
Total	a + c	b + d	a + b + c + d

2x2 tables are commonly used to estimate measures of association such as relative risk or odds ratios.

1.2.5.2 Relative Risk

The *relative risk* (RR) is the ratio of a measure of incidence in the **exposed** group to the measure of incidence in the **unexposed** group. It can be based on the incidence rate or cumulative incidence and under some situations prevalence. Relative risks are the primary measure of disease association and should always be used when it is possible to estimate a population-at-risk.

Table 1.4: Table showing 2x2 layout and calculations required for estimation of a relative risk

	Diseased	Not diseased	Total
Exposed	a	b	a + b
Unexposed	c	d	c + d
Total	a + c	b + d	a + b + c + d

Incidence rate exposed = $IR_{exp} = a/(a + b)$ Incidence rate unexposed = $IR_{unexp} = c/(c + d)$

$$\text{Relative Risk} = RR = \frac{IR_{exp}}{IR_{unexp}} = \frac{\left[\frac{a}{a+b} \right]}{\left[\frac{c}{c+d} \right]}$$

There are a number of calculators or tools and apps that can analyse 2x2 tables and estimate Relative Risks or Odds Ratios. Typically these tools will also produce an estimate of the confidence interval and a p-value testing whether the RR is statistically different to 1.

- If the RR is greater than 1 and is associated with a significant p-value or if the 95% confidence interval for the RR does not include 1, then the risk factor is associated with an increased risk of disease.
- If the RR is less than 1 and is associated with a significant p-value or if the 95% confidence interval for the RR does not include 1, then the risk factor is associated with a reduced risk of disease. Often factors with a RR less than one are called *protective risk factors* because they are associated with a reduced risk of disease.

The numeric value of the RR is a measure of the strength of association between the risk factor and disease. If possible the 95% confidence interval for the RR should be examined. If

the confidence interval does not overlap one then we can be more confident that the association may be meaningful. A statistical test and p-value will provide similar information – telling you whether the RR is statistically different to one.

Note that the RR alone does not provide evidence of causality. It provides a measure of statistical association. Additional information is required before we can determine that a particular risk factor may be causal for a particular disease.

Factor-specific attack rates and corresponding relative risks for such factors as species, age, sex, feed, mob, management system, etc can be computed and arranged in an attack rate table as shown below. An attack rate table is simply a tabular presentation of attack rates for different risk groups, accompanied by relative and attributable risk values for comparison between groups.

Table 1.5: Attack rate table for risk factors for stillbirths in a group of Hereford heifers.

Factor	Levels	Stillborn	Live	Total	Attack rate	Relative risk	Attributable risk
Age	14 months at joining	14	25	39	35.9%	2.3	20.4%
	17 months at joining	16	87	103	15.5%		
Sire breed	Hereford	22	78	100	22.0%	1.0	0.8%
	Angus	7	26	33	21.2%		
Sex of calf	Female	18	48	66	27.3%	1.7	10.9%
	Male	11	56	67	16.4%		
Type of birth	Assisted	16	41	57	28.1%	1.7	11.6%
	Normal	14	71	85	16.5%		

In the above table, attack rates are expressed as percentages. The second last column is the Relative Risk or Risk Ratio (RR) which is the ratio of the attack rates and the last column is the difference in attack rates (the Attributable Risk).

In the example in Table 3.5, the highest relative risk is 2.3, indicating that younger heifers (14 months) were at 2.3 times the risk of having a stillborn calf compared to older (17 months) heifers. However, this has to be interpreted with caution, as the attack rate for older heifers was 15.5%, suggesting that other factors may also have been involved in causing this problem. Examination of the other relative risks list shows them all to be less than 2, suggesting that these factors are not very important. Therefore, from the data provided we can determine that younger heifers are at increased risk of stillbirth, but that there

are probably additional factor(s), for which we don't have data, that are contributing to this problem.

1.2.5.3 Odds Ratio

The *odds ratio* (OR) is another measure of association that is often used to approximate relative risk. The OR is usually estimated when the population-at-risk is unknown and therefore relative risks cannot be calculated.

OR can be estimated when you have population at risk data but they do not require population at risk data. Odds ratios only require data collected on cases and non-cases and often in these situations it is not possible to estimate a population at risk and therefore RR cannot be used.

Imagine a situation where you are investigating a rare disease. You visit farms every two weeks and identify any animals that develop the disease of interest. Each time you identify an animal with the disease, you select another animal from the same farm that does not have the disease (non-case). You then collect risk factor information on cases and non-cases.

This is a *case-control* study. Over time you build up a dataset with a number of cases and a number of non-cases. But, you do not ever have an estimate of population at risk. Using this dataset, you cannot estimate incidence, prevalence or relative risk. You can estimate OR as a measure of strength of association between disease and risk factors.

Odds ratios are calculated using the same 2x2 table structure as for relative risks but the formula is different. As an epidemiologist you will need to know when RR is able to be estimated and when it cannot be estimated.

Table 1.6: Table showing 2x2 layout and calculations required to estimate an odds ratio

	Diseased	Not diseased	Total
Exposed	a	b	a + b
Unexposed	c	d	c + d
Total	a + c	b + d	a + b + c + d

$$\text{Odds exposed} = \frac{a}{b}$$

$$\text{Odds unexposed} = \frac{c}{d}$$

$$\text{Odds Ratio} = \text{OR} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{ad}{bc}$$

The Odds Ratio is usually interpreted like a relative risk.

When a disease is rare, the numeric value of the odds ratio can be shown to be closer to the numeric value of the RR. Compare the relative risk and the odds ratio:

- a is small compared to b, and c is small compared to d
- therefore, (a+b) approximates b and (c+d) approximates d
- therefore, the RR approximates ad/bc
- so the OR estimates the RR as long as the disease is rare (say ~<10% but the approximation becomes better as the disease becomes rarer)
- When the disease gets more common the OR provides a poorer numeric approximation of the relative risk.

We can use confidence intervals and statistical tests to interpret the strength of association for OR in the same way that we can for RR.

1.2.5.4 Attributable risk (AR)

Attributable risk (also called *risk difference*) is a measure of association that is based only on the exposed population. It is the absolute difference between the two incidence rates from a 2 x 2 table. The AR tells the rate of disease in the exposed population that is **attributable** to being exposed. AR provides an estimate of the rate of disease that could be prevented if the exposure were removed completely from the population.

Table 1.7: Table showing 2x2 layout and calculations required for estimation of attributable risk

	Diseased	Not diseased
Exposed	a	b
Unexposed	c	d

$$IR_{exp} = \left[\frac{a}{a+b} \right]$$

$$IR_{unexp} = \left[\frac{c}{c+d} \right]$$

$$\text{Attributable Risk} = AR = IR_{exp} - IR_{unexp} = \left[\frac{a}{a+b} \right] - \left[\frac{c}{c+d} \right]$$

The AR has the same units as the IR and can theoretically vary from -1 to +1; the null value is zero. Remember that the RR has no units and has a null value of 1.0.

We can use confidence intervals and statistical tests to interpret the strength of association for AR in the same way that we can for RR and OR.

1.2.5.5 Attributable fraction (AF)

The *attributable fraction* (AF) is also only relevant for the exposed population and expresses the AR as a fraction of the incidence rate among the exposed - this measure indicates the proportion of disease in the exposed that could have been prevented had exposure not occurred.

$$AF = \frac{(IR_{exp} - IR_{unexp})}{IR_{exp}} = \frac{(RR - 1)}{RR}$$

$$AF \text{ for case - control} = \frac{(OR - 1)}{OR}$$

1.2.5.6 Biological meaningfulness and 95% Confidence intervals

Relative risk and odds ratio estimates provide measures of association between disease occurrence and a risk factor. They provide an estimate of how much exposure to the risk factor increases (or decreases) the rate or amount of disease in a population.

- RR or OR less than 1 (exposure is protective)
- RR or OR equal to 1 (no increase in risk or protectiveness from risk factor)
- RR or OR greater than 1 (exposure = increased risk)

Most statistical packages or online calculators that can estimate RR or OR will also provide a 95% confidence interval for the estimate and use statistical significance tests, such as the Chi-square test, to determine if the RR or OR is significantly different to 1.

An estimate that is associated with a significant p-value ($p < 0.05$) is considered potentially important but we need to consider both the p-value and the 95% confidence interval for the estimate. Statistical tests only tell us the probability that the observed result would have occurred due to chance alone – it tells us nothing about the biological importance of the risk factor. A risk factor may have a statistically significant effect in a particular study, but not be biologically important, and vice versa

The numeric estimate of the RR (or OR) and the 95% confidence interval often provide more useful information on the biological importance of the association.

When the total sample size is small, all estimates are likely to be not terribly useful.

Wildlife vaccines for rabies were dropped from planes and contained tetracycline markers so foxes eating the baits would be able to be identified as vaccinated based on the staining of their teeth. Over time foxes that were killed or found dead were assigned to a 2x2 table based on whether they were rabies positive or negative, and vaccinated or unvaccinated.

	Rabies +	Rabies -	Total
Unvaccinated	18	30	48
Vaccinated	12	46	58
Total	30	76	106

OR= 2.29 95% CI from 0.39 to 13.33

The OR suggests that the association is useful but the confidence interval ranges widely across one. This is probably because of the relatively small sample size.

The same wildlife study is continued for longer until additional samples are obtained and the analysis is repeated.

	Rabies +	Rabies -	Total
Unvaccinated	18	30	48
Vaccinated	12	46	58
Total	30	76	106

OR= 2.30 95% CI from 0.97 to 5.45

The OR is similar but the confidence interval is now almost all positive. It does extend just below 1 and the p-value is 0.06 (just not significant). However, the fact that the bulk of the OR is positive is reasonable evidence of a good association between the risk factor (vaccination status) and odds of rabies.

1.2.6 Advanced statistical analyses

Once potential risk factors have been identified and their importance assessed in unadjusted screening such as 2x2 tables it may be useful to undertake further statistical analyses.

Attack rates provide measures of disease occurrence – how much disease is occurring? Producing attack rate estimates for different levels of some other factors provide additional information on how much of the disease may be due to exposure to the particular risk factor.

Relative risk and odds ratio estimates then provide measures of the strength of association between disease and possible risk factors. We can use statistical testing to determine if the observed association is likely to be due to chance or not – we use this information to try and determine if the apparent relationship is important or not.

Potential risk factors can have a high relative risk but be statistically not-significant or vice versa, depending largely on sample size. Therefore, it is important to always consider high relative risk values (say >3) as being worth further investigation, even if they are not statistically significant. This is particularly true when the estimate is based on a small total count of animals (small sample size).

When there are multiple possible factors that may influence disease risk the first approach is to perform simple estimates of AR for each factor one at a time. These are called unadjusted or crude or screening associations. More advanced analyses may involve considering multiple factors at once (multivariable analyses) in order to adjust for interaction or confounding between different factors. These methods are beyond the scope of these notes and require advanced statistical expertise and special software.

1.3 Developing hypotheses and control measures

When you are investigating a disease outbreak and particularly in an emergency situation it is important to begin to develop hypotheses about the nature of the disease as soon as possible and to use this information to identify possible causes and interim recommendations that may help to control the disease and prevent further cases.

Generating hypotheses about the disease means using the description of the disease patterns and attack rates and other initial analyses to inform either a differential diagnosis list or if this is not possible then a guess at the sort of disease process that may be occurring (infectious or non-infectious, point-source or propagating outbreak).

Examples of hypotheses relevant for disease investigations include:

- the nature of the causal agent (eg toxin, infectious, viral, bacterial, etc)
- the source of the agent (eg environmental, species jump, introduced animals, endemic infection, etc)
- the method(s) of transmission (eg direct contact, food-borne, vector-borne, etc)
- why the incident has occurred (eg change in herd immunity levels, introduction of new disease, change in management practices, etc)
- risk factors for disease (eg exposure to specific feed components, or potential sources of infection)

The disease hypotheses are then used to inform interim recommendations for control measures. Control measures refer to any intervention aiming to reduce occurrence of disease or eradicate disease. Treatment of sick animals is one form of control measure and interventions to prevent spread or eradicate disease from an area are all types of control measures.

Disease hypotheses should be based on the facts gathered during the initial investigation. It may be possible to draw up a causal diagram for the disease showing how the various factors interact to cause the disease. This process helps to understand the disease process and can often lead to an improved understanding of the relationships between possible risk factors. Consideration of these relationships will often help identify points where intervention can be made to control and/or prevent the disease occurring.

In many situations, initial hypotheses can then be tested using further investigations conducted while the outbreak is still under investigation and while interim control measures are being implemented.

For example, if you suspected one or more specific disease agents you may be able to collect samples and send them off for testing to rule in or rule out those particular diseases. If the results confirm your initial hypotheses then your management of the situation may be clarified. If the results rule out the initial hypotheses then further epidemiologic investigation will be required.

Some measures may be implemented based on general precautions without any knowledge of what the disease might be. For example, isolation and quarantine of the affected properties and affected animals within a property, symptomatic treatment of affected animals and so on. In many situations the results of the initial investigation are used to inform interim control measures. As further information is collected these initial control measures can then be modified.

Actual measures implemented will depend on the individual circumstances, but could include one or more of the following:

- Specific treatments;
- Vaccination;
- Changes in nutrition, feed ingredients and/or management factors;
- Isolation or quarantine;
- Surveillance of the affected population and other at-risk populations for evidence of further spread and new cases;
- Changes in environment and/or housing;
- Safe destruction and disposal of contaminated waste or other infectious materials;
- Disinfection and decontamination; and
- Salvage sale or slaughter of animals.

Once hypotheses have been formulated, it is important to review and evaluate them. In particular:

- Do they explain the observations?
- Are they reasonable?

- Are there any facts that contradict the hypothesis and how can these be explained?
- Are there any unexplained aspects of the situation requiring further investigation and evaluation?
- What additional data do we need to test the hypotheses, or is there sufficient data already available?

In some cases, it will not be possible to stop an outbreak once it starts, but the detailed investigation of one or more outbreaks should provide valuable insight into possibly important "component" causes and support the development of strategies to prevent future outbreaks.

For Menangle virus, based on the observations during the outbreak, it was hypothesised that:

- The outbreak was a propagating epidemic of a previously unidentified virus causing infertility, stillbirths, mummified foetuses and congenital deformities
- The probable source of infection was from a fruit bat colony, either on fly-past or entry to sheds or laneways
- Spread within the piggery occurred via close contact and fomites during acute infection and at farrowing

An obvious conclusion from this was that the easiest way to prevent future outbreaks was to prevent any contact between pigs and fruit bats by enclosing and screening all sheds and laneways.

Sera and faeces were collected from fruit bats from the nearby colony, to test the hypothesis that the bats were the source of this virus. Forty of 80 (50%) serum samples were positive but virus was unable to be isolated from faeces from 55 bats.

Based on the findings from the various investigations it was not deemed possible to prevent continuing spread of the virus at the time of the outbreak. Instead, it was decided to undertake a staged eradication program once the main epidemic had burned out, including:

- Progressive eradication from the four production Units
- Segregation, depopulation and staged repopulation of Units
- Sheds and walkways "flying-fox-proofed" to prevent re-introduction
- Serological testing to monitor progress

Successful eradication was achieved and subsequently demonstrated by on-going monitoring of the population.

1.4 Role of tracing in disease investigation

Tracing of livestock movements is an important tool particularly for the detection of infected herds or flocks and particularly where there is interest in eradication of a disease. Tracing usually involves the identification of potentially infected farms through the tracing of movements of infected or exposed animals.

If there is no clear policy to control or eradicate the disease then there may be little justification to do tracing.

Further testing is usually undertaken on any other farms identified by tracing and considered to be at risk of having infected animals to establish their true infection status. If a farm's infection status cannot be determined immediately, quarantine measures may be imposed until the situation is resolved.

1.5 Further field epidemiology studies

The initial disease investigation aims to describe the disease cases, identify possible causes and implement control measures. In some cases, it may be possible to confidently diagnose a specific disease and implement effective treatment and control measures. In this situation there may be little need to conduct further investigation.

In many situations there will be a need to conduct further field investigations to provide additional information on one or more of the following:

- Gather more information on possible causes of the disease including identification of the infectious agent (if there is one) and identify other causes.
- Use tracing to identify the origin of the disease and risk of spread to other farms or locations.
- Monitor effects of control measures to measure efficacy of control measures and apply new measures if necessary, ensure that no new cases are occurring and that affected animals are recovering.
- Use field or experimental studies to test hypotheses arising from the initial investigations about possible causes and control measures.

Menangle virus was first identified in 1997, following an investigation of a serious outbreak of mummified and stillborn foetuses in a commercial piggery at Menangle, New South Wales, Australia. A high proportion of litters born to sows that were pregnant at the time of exposure to the virus were affected,

although clinical disease was not noticed at the time of infection. After an extensive investigation the infection was traced to a nearby colony of fruit bats (flying foxes), with a high proportion of bats sampled found to be seropositive for Menangle virus antibodies.

During the Menangle virus investigations, a wide range of additional investigations were undertaken, including:

- Detailed pathological, serological, microbiological and virological examination of affected and unaffected pigs to determine the likely cause and to rule out known infections and other diseases.
- Cross-sectional serological survey of all units/sheds to determine the extent and progress of infection.
- Surveys of pigs and piggeries in contact to determine whether infection had spread beyond the Menangle piggery.
- Sampling of unexposed piggeries to demonstrate freedom of the rest of the industry.
- Testing of archived sera (from this and other piggeries nation-wide) to demonstrate that it was a new infection not previously present in the Australian pig population.
- Interview and testing of piggery workers and others potentially exposed to evaluate public health risks.
- Serology on other species as potential sources.
- Serology and virus isolation on fruit bats to support the hypothesis that they were the likely source.

1.5.1 Types of epidemiologic studies

Epidemiologic studies can be broadly grouped into *observational study*, *intervention study* and *theoretical epidemiology*.

The characteristics of different study types are described briefly below and the advantages and disadvantages of each type are summarised. For more information on epidemiological study design, readers should consult standard epidemiology texts (Martin et al., 1987; Thrusfield, 1995; Rothman et al., 2008; Dohoo et al., 2010).

In disease investigations, almost all field studies will be observational studies.

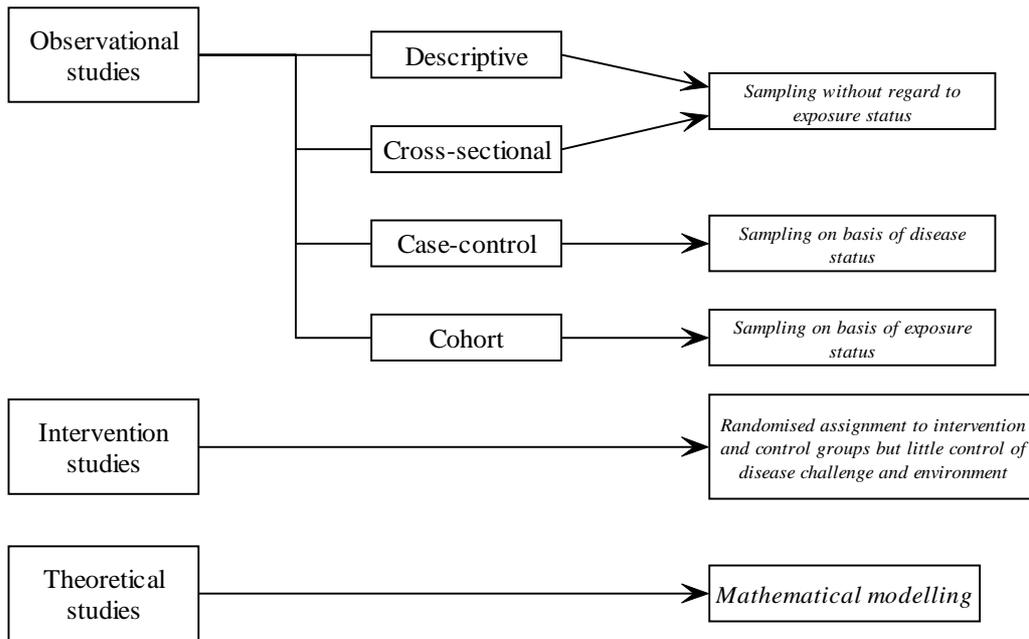


Figure 1.3: Figure showing classification of epidemiologic study types

1.5.1.1 Observational studies

In observational studies nature is allowed to take its course, and the study aims to collect data by observing what happens without intervention from the investigator. There are basically four types of observational study: *descriptive study*, *cross-sectional study*, *case-control study* and *cohort study*. Cross-sectional, case-control and cohort studies may also be termed analytic studies because they usually involve some form of statistical testing of various hypotheses.

Descriptive studies

A descriptive study can collect data to describe distribution and occurrence of a disease in a population but does not involve statistical hypothesis testing of possible risk factors. The first part of the investigation (describing disease in terms of time, place and animal) is largely descriptive but it can lead to hypotheses that may be tested and the data may be used in subsequent statistical testing of some sort.

Cross-sectional studies

Cross-sectional studies involve selection of animals at a point in time (or over a defined period) and then the prevalence of the disease in question is measured and data gathered on other factors to allow comparison between presence or absence of disease and presence or absence of various possible risk factors. Cross-sectional studies can be done quickly and cost-effectively but are less effective for testing hypotheses about causation of disease.

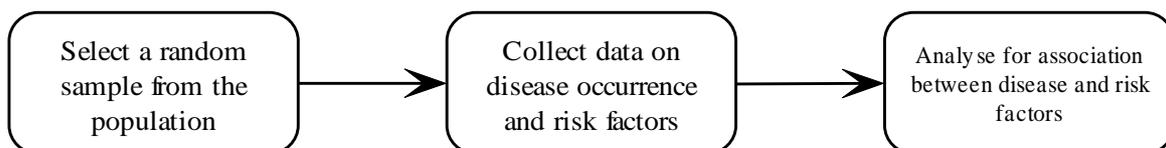


Figure 1.4: Cross-sectional studies

For example, you might undertake a randomised cross-sectional study of villages in a country for exposure to foot-and-mouth disease (FMD) virus. This would allow you to estimate the seroprevalence and to identify possible risk factors for exposure to support either follow-up studies and/or planning for future management of FMD.

Case-control studies

A “case” group is selected from animals with the disease of interest and a “control” group is selected from animals without the disease. The presence or absence of possible risk factors are then measured for the two groups and compared. Case-control studies are well suited to rare diseases and many suspected risk factors can be compared at the same time. They are relatively quick and inexpensive to perform but are susceptible to many biases and do not allow estimation of disease frequency (prevalence or incidence).

Case-control studies are very commonly done in disease investigation and particularly in the early stages because they are able to be implemented as soon as a case definition is completed and animals assigned to either cases and non-cases. This means that a case-control study can be begun while the very early stages of disease investigation are still going on.

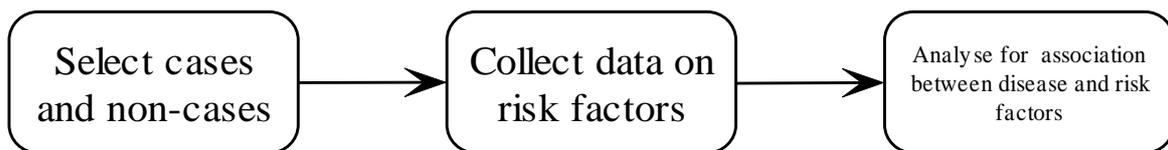


Figure 1.5: Case-control studies

For example, you might undertake a case-control study for foot-and-mouth disease occurrence in village livestock. Case villages would be selected from known affected villages while controls would be selected from unaffected villages in the same region. This would allow you to identify village-level risk factors for infection, to support planning for prevention and management of future outbreaks.

Cohort studies

The word *cohort* just means any group of animals that is followed over a period of time.

In a cohort study animals that are free of the disease of interest are selected based on presence or absence of one or more defined risk factors (presence of the risk factor = exposed group and absence of the risk factor = unexposed group). The selected animals are then monitored forward in time to measure the occurrence of the disease of interest in each group. In some cases where detailed retrospective records are available it may be possible to use retrospective data to perform a cohort study but the approach is still the same.

Cohort studies can provide incidence rates for the disease in the exposed and unexposed groups and they provide stronger evidence for causation of disease than either cross-sectional or case-control studies. They are also more expensive and take longer to plan and complete.

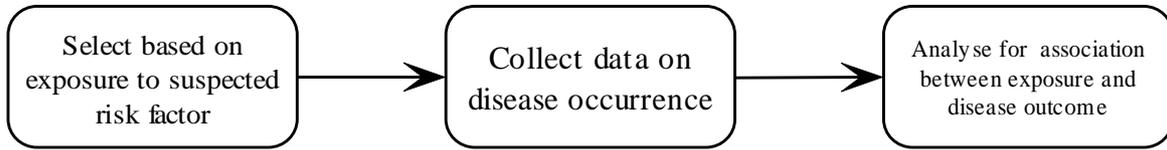


Figure 1.6: Cohort studies

The best known examples of cohort studies are numerous studies investigating health outcomes associated with cigarette smoking. Comparison of health outcomes between smokers and non-smokers has allowed researchers to quantify the increase in risk of lung cancer, cardio-vascular disease and other health problems associated with increased levels of smoking.

1.5.1.2 Intervention studies

A field intervention study is a type of clinical trial or experimental trial or study.

The key distinction from observational studies is that animals are randomly assigned to two or more treatment or intervention groups and then the effects of these different interventions is compared. Intervention studies may be used to test efficacy of various treatments for disease control (vaccination compared to no-vaccination, various different management or treatment strategies).

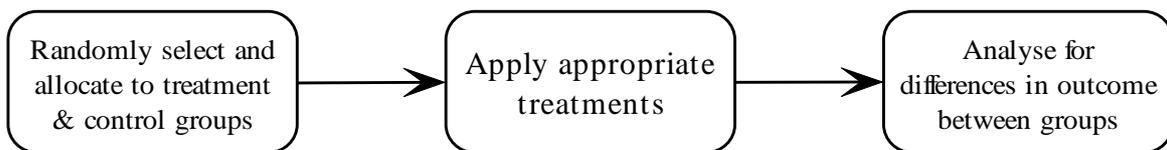


Figure 1.7: Intervention studies

For example, mineral deficiencies can often result in poor growth and even death of young sheep or cattle. Often you may suspect that a particular mineral is deficient but be unable to demonstrate this conclusively. One way of achieving this is to run a field trial, comparing growth rates in treated and untreated groups that are similar in all other ways.

1.5.1.3 Theoretical studies

Theoretical epidemiology studies are based on mathematical modelling using a computer and are designed to use simulation to answer "what-if" type questions. There are a wide variety of modelling methods used, but the primary aim is to reproduce a realistic simulation of disease behaviour in a population. The major benefits of models are that:

- The process of developing and interpreting the model often leads to valuable insights into disease epidemiology and behaviour that might not otherwise be apparent; and
- Models provide a structured and controlled environment in which hypothesised interventions can be tested and evaluated at significantly lower cost than undertaking field experiments or observations to achieve the same result (or for interventions that may not be practical to implement experimentally).

Models are particularly useful in examining the behaviour and impact of infectious diseases as well as the possible effects of a range of interventions. The results from such studies need to be confirmed with follow-up observational or intervention studies wherever possible.

Theoretical modelling generally involves advanced mathematical and statistical skills and custom software.

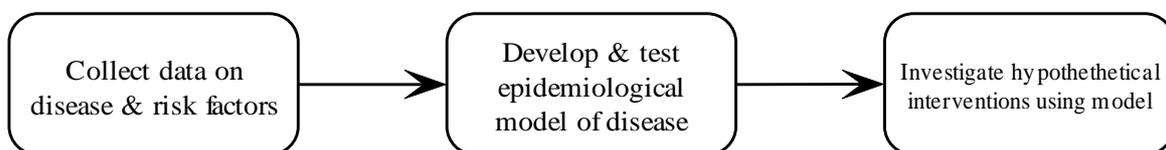


Figure 1.8: Theoretical studies

For example, simulation models of the spread of FMD have been used to help understand the behaviour of the 2001 outbreak in the UK and to predict the potential impact of alternative control strategies (Morris et al., 2001)

Table 1.8: Characteristics, strengths and weaknesses of main study types (adapted from Thrusfield, 1995)

Study type	Characteristics	Advantages	Disadvantages
Descriptive	<ul style="list-style-type: none"> • Observational • Describe patterns of disease in the population 	<ul style="list-style-type: none"> • Relatively quick and easy • Can generate hypotheses on possible risk factors for further investigation • Doesn't require random sampling or high degree of rigour 	<ul style="list-style-type: none"> • Doesn't support hypothesis testing or inference for possible risk factors • Can't estimate prevalence or incidence or exposure proportions • Subject to inherent biases and errors because of the nature of the data
Cross-sectional	<ul style="list-style-type: none"> • Observational • Observation at point in time • Outcome/exposure not considered in selection 	<ul style="list-style-type: none"> • Disease prevalence in exposed and unexposed populations can be estimated • Exposure proportions can be estimated • Relatively quick and cost-effective • Can study multiple factors at once 	<ul style="list-style-type: none"> • Unsuitable to investigating rare diseases • Less useful for acute diseases • May be difficult to control potential confounders • Incidence cannot be estimated • May be difficult to determine causality • May be problems with reliability of data/recall for historical data
Case-control	<ul style="list-style-type: none"> • Observational • Retrospective longitudinal • Selection based on outcome status 	<ul style="list-style-type: none"> • Good for rare diseases • Relatively rapid and cost-effective • Relatively small sample sizes • Often use existing data • Can study multiple factors at once 	<ul style="list-style-type: none"> • May be difficult to establish causality • Can't estimate prevalence or incidence or exposure proportions • Rely on access to historical data or recall • Difficult to validate data • May be affected by variables for which data is not collected • Selection of controls often difficult
Cohort	<ul style="list-style-type: none"> • Observational • Prospective longitudinal • Selection based on exposure status 	<ul style="list-style-type: none"> • Can calculate incidence in exposed and unexposed individuals • Can provide strong evidence for causality 	<ul style="list-style-type: none"> • Exposed/unexposed proportions cannot be estimated • Large sample sizes, particularly for rare diseases • Can only investigate small number of potential risk factors at any one time • Long duration of follow-up • Relatively expensive and time-consuming • Loss of individuals to follow-up • May be difficult to control potential confounders
Field/clinical trials	<ul style="list-style-type: none"> • Intervention • Longitudinal • Randomised selection 	<ul style="list-style-type: none"> • Relatively quick • Good for helping establish causation • Usually strong internal validity • Relatively small sample size and usually short duration • Can't estimate incidence/prevalence 	<ul style="list-style-type: none"> • May be problems with external validity, particularly to diverse target population • Can be expensive depending on the intervention and situation • Requires significant cooperation and rigorous management

1.5.2 Choice of field study

Field epidemiology for disease investigation will almost always involve observational studies and only very occasionally intervention studies.

Case-control and cross-sectional studies are very well suited to the early stages of a disease investigation because the investigation has to proceed through a case definition and then assignment of animals to either cases or non-cases followed by collection of data to describe disease patterns in time, place and animal.

Case-control studies do not have a population at risk. They are based on selection of cases and a separate selection of non-cases (controls). As a result case-control studies cannot be used to produce prevalence or incidence measures. Because of this, they cannot be used to produce relative risk measures. Data from case-control studies can produce odds ratios and these are the primary measure of strength of association for case-control studies. Case-control studies are particularly useful for studying rare diseases because cases can be detected and then enrolled in the study along with one or more controls.

Cohort studies are generally likely to be used in later stages of a disease investigation or in follow-up studies to test hypotheses about causal factors. Cohort studies start with a population that does not have the disease (population at risk) and detect new cases of disease. They can measure incidence rate and are used to produce relative risks as a measure of association between disease and risk factors. Cohort studies may also be more useful when there is interest in assessing specific exposures and particularly rare exposures since these can be defined and used to select the groups at the beginning of the study.

1.5.3 Defining groups based on outcome and exposure

The general purpose of field observational studies is usually to collect data on a number of animals with a particular focus on disease status (presence of disease = case, absence of disease = non-case), and on the presence or absence of one or more risk factors. We can then compare the diseased group to the non-diseased group to look for differences in risk factors.

The term *outcome* refers to whether or not an animal is recorded as a case (disease present) or non-case (disease absent). Information is provided in the *Basic Field Epidemiology Manual*

about development of the case definition and the use of the definition in conjunction with clinical examination and possibly laboratory tests or other information of animals to assign animals to confirmed case, suspect case and non-case categories.

When designing and analysing field studies many people use the term *exposure*. Exposure refers to the presence or absence of some risk factor or the level or category of a risk factor for each animal. Examples of possible risk factors that may be measured include species, breed, sex, body weight, age, vaccination status, recent treatments or other management procedures, place of origin, location of animals, recent climate (rainfall, temperature, humidity), feed, recent movements, pregnancy status etc.

Some risk factors do not change over time (breed, species and often sex) and may be described as fixed (unchanging). Other risk factors (age, body weight, feed, location, recent treatments) may change over time and may be measured at a defined time or used to develop categories (age less than 1 year or greater than 1 year) or averages.

Some risk factors are measured as categories (sex=male entire, male castrate, female), or scores (body condition score: 1=poor, 2=backward (thin), 3=moderate (no significant fat), 4=forward, 5=fat), or as a measurement on a continuous scale (body weight measured on a scale in kg).

1.5.4 Selection of animals

In a case-control study, the cases must have the disease being investigated, meaning that they must meet the case definition. The case definition must also be independent of the risk factors that are being studied. It is generally best to use the confirmed case definition and to try and restrict the selected cases to those animals that are likely to be recent cases as opposed to animals that may have been infected some time ago (chronic cases). If some of the cases die as a result of the disease then studying chronic cases may be more informative about factors that influence survival than factors that determine risk of occurrence of disease.

A comparable group of non-cases (controls) then has to be selected for comparison. Selection of controls for a case-control study is a complex subject and these notes provide brief coverage of simple points.

Controls should not have the disease (meet the definition for non-case) and should be identified independent of any possible risk factor of interest. The controls should generally be

representative of the population from which the cases arose. This means if the cases came from animals admitted to one or more veterinary clinics then the controls should be selected from those same clinics. The same applies to farms.

Controls should be eligible to get the disease and to be selected as cases if they had developed the disease. This means that if the disease is causing abortion in pregnant cattle, then male cattle would not be selected as controls. Any animal that may have had the disease earlier and recovered and considered immune would not be considered for selection as a control because they may not be expected to get the disease again.

Where the investigation involves a relatively small area (one farm or one village) it may be possible to assess all (or nearly all) animals at that location and assign them to either cases or non-cases. These two groups may then form the basis of a study. If you are able to collect data on all animals at the location and if you can confidently determine onset of disease then this will form a cohort study and not a case-control study and you will be able to estimate incidence or prevalence measures and use relative risk as the statistical measure of association between disease and risk factors.

Where the investigation involves multiple farms or a larger area or number of animals and it is not possible to assess and include every animal in a study, it will be necessary to select a subset or sample of cases and controls. If you have a list of non-case animals, then random sampling can be used to select a group of controls. More commonly you will not have a list of all animals to choose from. In this case it may be most practical to select one or more controls for each case from those animals that are close to the case (same paddock or pen) and that meet the definition for a non-case. In an open population (where animals are entering and leaving the local population), controls should be selected from those animals with a similar exposure time to the cases.

Most case-control studies select one control for every case. There may be some value in selecting 2-3 controls for every case because of the increased sample size but there is likely to be little value in selecting more than 3-4 controls per case.

Sometimes controls are selected in a process that involves matching on a defined attribute with each case. For example, a 2-year old cow with diarrhoea (case) may be matched with a 2-year old cow without diarrhoea (matched control). Matching is usually used to control for some form of confounding between the factor being used for matching (age) and the

association between some other possible causal factors and the disease of interest. When a factor is matched it cannot be analysed for an association with disease. Matching is a complex topic and should not be attempted without advice from a veterinary epidemiologist.

There are a range of issues relating to selection of animals for a cohort study as well. Animals eligible for inclusion in a cohort study should be free of the disease of interest at the start of the study and should be at risk of developing the disease. If animals may have previously had the disease and recovered and are now immune then these animals are no longer at risk of getting the disease and should be excluded from selection. Other diseases may recur once some time has elapsed for recovery. In these situations it may be possible to state that animals become eligible for inclusion in the study providing they have not had the disease within a defined period prior to the start of the study.

Cohort studies may be closed or open. A closed study is simpler and enrolls a defined selection of animals at the start and then only those animals are followed over time (no new animals are added) and every effort is made to follow all enrolled animals right through to the end of the follow-up period. An open study may allow animals to either enter or exit the study population at any time so that any one animal may be followed for a different time period.

1.5.5 Bias, confounding and interaction

Observational studies are subject to bias. Bias occurs in an epidemiological study when the observations do not reflect the true situation because of some systematic error. Knowledge about type of bias and strategies to minimise risk of bias is essential in designing and implementing epidemiologic studies.

***Bias** is any effect at any stage of an investigation tending to produce results that depart systematically from the true values i.e. a systematic error (lack of validity) rather than a random error (lack of precision).*

Although there are many different types of bias, they can be broadly classified into three general categories: *selection*, *measurement* and *confounding*. The differences between these categories are not always clear-cut and the strategies for preventing bias are not always exclusive to a single type. They may be viewed as a connected group of issues capable of interfering with inference.

1.5.5.1 Selection bias

Selection bias is a systematic error in the way that the samples of study units were drawn from their underlying populations, or in the way that study units were assigned to interventions

The potential for selection bias is high in many observational studies. For example, if in a cross sectional sample of prawns from a pond, the easy-to-catch prawns in the shallow water near the edge of the pond were caught the sample may not be representative of all prawns in the pond and a selection bias would result.

If, in an intervention study, an investigator can't describe a formal decision rule that he or she used to select subjects or assign treatments, then there is a risk of selection bias. Common types of selection bias in surveys include differences between volunteers vs randomly selected subjects, and responders vs non-responders. Another important source of selection bias arises from differences in access to extension activities and technical advice. For example, prawn farms that have regular input from trained specialists are very unlikely to be representative of all prawn farms.

There are many strategies for protecting against selection bias. These include having clear criteria for subjects to be eligible for inclusion in the study (see previous selection). *Randomised* selection uses chance to provide protection against selection bias.

1.5.5.2 Measurement bias

Measurement bias is a systematic error in the way that data were gathered or measured.

Measurement bias is often called misclassification bias. Animals may be misclassified with respect to disease (a case recorded as a control or vice versa) or with respect to a risk factor.

Misclassification may be differential (when one group is more or less likely to be misclassified than another) or nondifferential.

For example, if you are looking for evidence of previous exposure to a chemical that is suspected as a causal factor for a disease in fish, farms with and without the disease may be

investigated regarding chemical use. In such a case, all farms should be questioned with equal vigour, and with equal adherence to non-leading questions to avoid triggering "recall bias" as much as possible.

When investigating causal associations equal effort should be expended in searching for old records for the diseased and the non-diseased groups. If you're following farms or ponds which are exposed and not exposed to a suspect causal factor you must guard against checking the exposed groups twice as frequently or using more sensitive methods of disease detection in the exposed group. This is to avoid "diagnostic work-up bias".

There are lots of research design features that help to protect against measurement bias. Some of the more obvious include:

- Blind the measurer/data collector.
- Get better measuring equipment or tests.
- Standardise the protocol for data collection.
- Use prospective rather than retrospective data.
- Use objective rather than subjective measurement criteria.

1.5.5.3 Confounding

Confounding occurs when two risk factors are interrelated and it is incorrectly concluded that one of the factors is causally related to the disease in question. For example, it might be observed that shrimp in ponds with cloudy water do not grow as well as those in clearer water. We might conclude from this that light penetration of the water is important for normal growth. However, it may be that the cloudiness is due to the presence of particular algal species in the water which inhibit growth of the shrimp through toxin production. In this theoretical example, confounding has meant that we have incorrectly concluded that light penetration is associated with poor shrimp growth when the true cause was the presence of toxic algae. The relationships are represented below.

Confounding is a systematic error that results from unaccounted-for differential distributions of particular covariates. Confounding may be viewed as a form of bias.

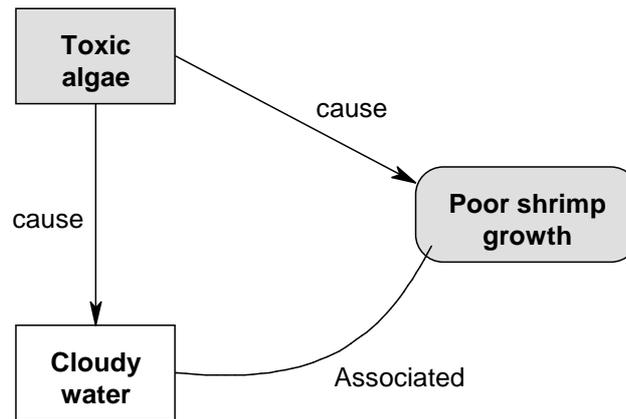


Figure 1.9: Example of relationships resulting in confounding leading to incorrect conclusions on the cause of poor shrimp growth

Confounding is one of the critical problems to watch for when undertaking an epidemiological study. It would probably be better to name the problem "confusing" as it occurs when the effects of two or more factors are mixed and it is difficult to determine which factors are truly "causal" in an epidemiological sense.

To be a confounder, an exposure factor must:

- be a risk factor for the disease in question;
- be associated with the exposure factor under study in the source population; and
- not be affected by the exposure factor or the disease. In particular, it cannot be an intermediate step in the causal path between the exposure and the disease.

Confounding is situation-specific, and you have to know something about the biology and logic of the situation to guess at things that should be explored as confounders. In general, in most studies you should at least think about the following kinds of variables: species, age, breed, season, sex and physiological status (eg spawning, nursing and growing), level of production.

One of the best protections against confounding bias is randomisation. Randomisation assures that, on average, most confounders will be distributed roughly evenly between treatment groups or sub-samples. The reason randomisation is useful is that randomisation is the only available method for controlling confounding due to unknown or unmeasured variables. All other

methods to control confounding assume that you know enough to have a measurement of the potential confounder. These other methods for controlling confounding include:

- Restriction of entry into the study
- Stratification (and its extreme: matching) in the design
- Standardisation of rates
- Stratification in the analysis
- Adjustment using multivariable statistical methods in the analysis

1.5.5.4 *Interaction (effect modification)*

Where two or more risk factors play a role in the causation of a disease, the possibility exists for interaction (also called effect modification) to occur between two or more of the factors. Interaction is different from confounding.

Interaction occurs when the incidence of disease in the presence of two or more risk factors differs from the incidence expected to result from their individual effects.

When interaction occurs, the effect can be greater than what we expect (positive interaction or synergism) or less than what we would expect (negative interaction or antagonism). The problem when evaluating effect modification is to ascertain what we would expect to result from the individual impacts of the different risk factors.

For example, say we find that the incidence of EUS is:

- 5% in fish in ponds with acidic water and a smooth lining;
- 2% in fish in ponds with non-acidic water and a rough lining; and
- 15% in fish in ponds with acidic water and a rough lining.

The 15% incidence is a lot higher than we would expect if the two factors of acidity and rough pond lining operated independently to increase the risk of EUS. We would therefore suspect synergy between these two risk factors and would need to investigate further.

In complex epidemiological studies, information is often collected on a wide variety of factors to identify the important risk factors for the disease of interest.

Assessment of interaction is commonly done during analysis and requires more advanced statistical skills.

1.5.6 Sample size

The number of measurements or animals included in a study (sample size) has the potential to influence a variety of measures including things like variance, confidence intervals and statistical significance. The smaller the sample size, the more likely it will be to generate results from analyses that may not be of much use in identifying causes for a disease. In some cases it may be possible that a study specifically assess one or more risk factors and fails to show any association with disease and yet if the same study had been performed with a larger sample size it might have identified the risk factors as causes of disease.

In practice the number of samples that may be able to be collected will be limited by available resources (labour, time, budget, sample storage and testing capacity).

It is possible to perform calculations to inform likely sample size estimates before a study is performed. The EpiTools website (<http://epitools.ausvet.com.au/>) has sample size calculators for a variety of study types including cohort and case-control studies.

The following information provides some simple rules of thumb for sample sizes for a field study aiming to identify a possible causal agent for a disease outbreak.

Where there is no single clear diagnosis for the disease or identification of a disease agent, there may be interest in conducting additional tests to confirm the specific disease and identify the infectious agent (if it is considered likely to be an infectious disease). This will typically involve laboratory testing of samples collected from animals that meet the case definition and from a comparison group of non-cases. The laboratory testing will generally look for detection of a candidate agent (virus or bacteria) and compare the prevalence of positive results in the two groups. If the results show a very low detection of the agent in the *free* or non-case group and a very high detection of agent in the *exposed* or case group then this would support a hypothesis of the infectious agent being a cause of the disease.

At least 10 animals at each stage of disease should be examined, but, if resources permit, this number should be extended to as high as 30. Statistical methods can then be used to assist in identifying which pathogen is the most likely.

The below table shows the number of animals that need to be examined to provide data for statistical analysis of association between disease and possible causal factors. Such analysis can assist in identifying which cause, from list of possible causes, is the most likely.

The dark shaded boxes show that examining 25-30 animals per case and non-case group are needed to ensure a high probability of identifying a difference between groups.

The light shaded boxes show the differences in prevalence required for a sample size of approximately 10 animals per group. In particular, where the difference between cases and non-cases is large (top left corner), only small numbers of animals are required to provide a high level of confidence that the observed association is not due to chance. In contrast, very large numbers are required if the difference between groups is likely to be small (diagonal from bottom left to top right).

Table 1.9: Number of animals per group to examine to determine if a particular finding is more common in cases than non-cases (95% confidence, 80% power, equal sizes for case and non-case groups and assuming a two-tailed test.

		Percentage of cases with pathogen									
		100	90	80	70	60	50	40	30	20	10
Percentage of non-cases with pathogen	1	4	6	7	9	12	15	21	30	50	121
	10	6	8	10	13	17	25	38	72	219	-
	20	7	10	13	19	28	45	91	313	-	-
	30	9	13	19	29	49	103	376	-	-	-
	40	11	17	28	49	107	408	-	-	-	-
	50	15	25	45	103	408	-	-	-	-	-
	60	20	38	91	376	-	-	-	-	-	-
	70	28	72	294	-	-	-	-	-	-	-
	80	44	219	-	-	-	-	-	-	-	-
	90	93	-	-	-	-	-	-	-	-	-
	100	-	-	-	-	-	-	-	-	-	-

Say we had taken specimens for detailed laboratory examination from 30 cases of a particular syndrome and 30 non-cases in a manner as previously described with the following microbiological results:

	Number (%) of infected		Number to detect observed difference
	Cases	Non-cases	
Organism 1	19 (63)	14 (47)	408
Organism 2	26 (87)	14 (47)	25
Organism 3	27 (90)	25 (83)	219

From the above results, and with reference the above table, Organism 2 is the only one which is statistically associated with an animal being a case, despite Organism 3 being isolated more frequently from cases. The reason for this conclusion is that the sample size of 30 is insufficient to detect a statistical difference in the isolation rates from cases and non-cases for Organisms 1 and 3, but is sufficient for Organism 2. This does not “prove” that Organism 2 is the primary pathogen (as it could be an opportunistic, secondary invader), but by examining a reasonable number (in this case, 30) of cases and non-cases we are much better able understand the relative importance of the three organisms.

1.5.7 Planning a field epidemiological study

It is important to use a structured and systematic approach.

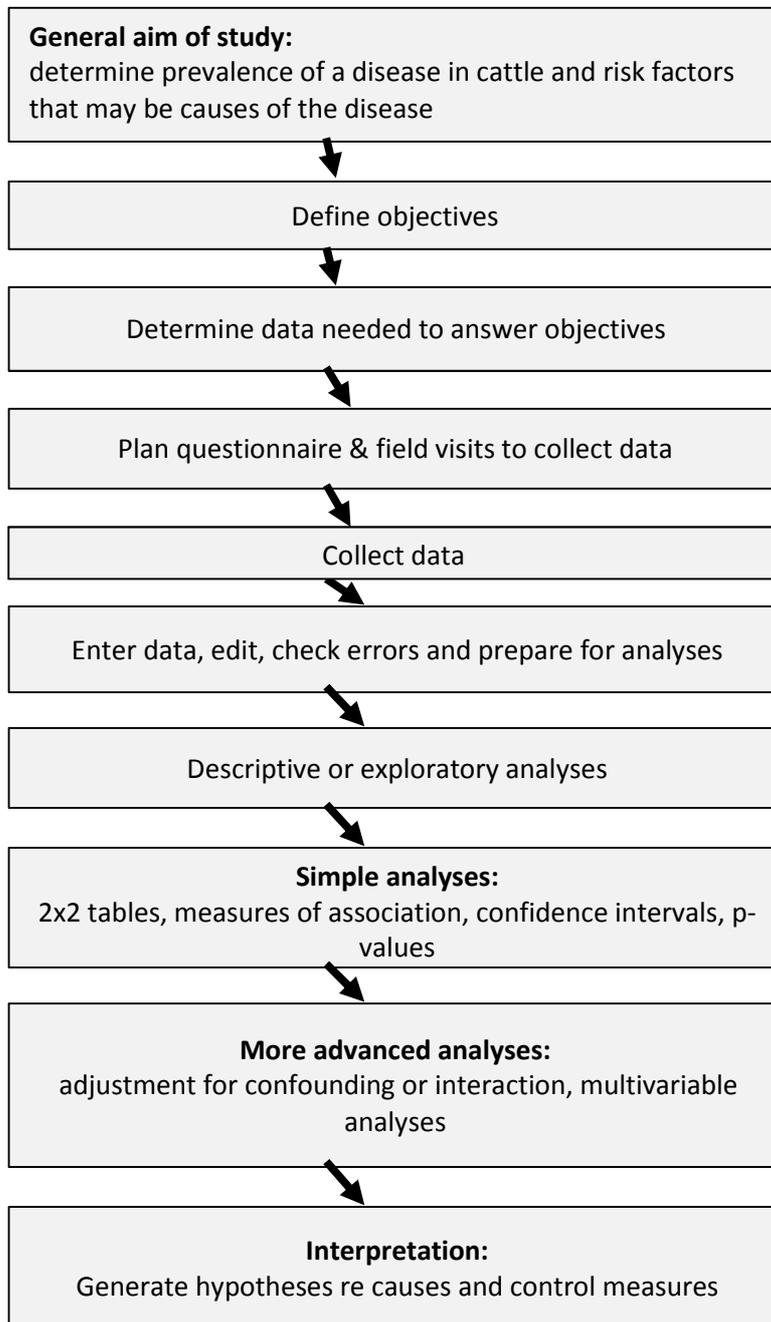


Figure 1.10: Steps in design and analysing field epidemiologic studies. From Gregg 2002.

1.5.7.1 *Identify the scope and responsibilities for any investigation*

The first step in any epidemiological analysis is to clearly define the problem and the scope, context and expected outcomes of the investigation. This might include determining if there is a disease problem and, if there is, to:

- determine the extent and impact of the problem
- identify possible and probable cause(s) and source(s) of the problem
- identify likely risk factors for the disease
- make recommendations for control and/or treatment and for future prevention

Where the analysis is undertaken at the request of a third party (for example government policy makers), it is important that any request is documented and that the terms of reference are clear and unambiguous.

During planning, it is also important to have clearly defined responsibilities (who is doing what and by when), deliverables (reports, software, information management system), and a detailed budget.

1.5.7.2 *SMART objectives*

Project objectives define the specific questions that the project will be expected to answer.

If the objective of an investigation is to estimate the prevalence of white spot disease virus in shrimp breeding stock, the study design should be directed at this objective, not at identifying risk-factors or looking for other viruses.

Before proceeding with the study you should (in consultation with others involved) define the objectives and expected outcomes of the study. SMART objectives are:

- **Specific:** meaning they are clear and well-defined. On completion of the investigation, it should be a straightforward process to determine whether or not the objectives have been achieved.
- **Measurable:** meaning each objective is associated with an outcome that can be measured to allow you to monitor and quantify progress toward achieving each objective and so you and others can determine when the objective has been achieved.

- **Achievable**: meaning that the objectives are practical and feasible and likely to be achieved with the skills and resources defined in the project.
- **Relevant**: Each objective should be relevant to the overall project goal or aim. Objectives that are not relevant risk wasting effort on producing a result that is subsequently ignored.
- **Time-bound**: meaning that each objective should include a timeline and milestones to be achieved within a given timeframe. Failure to specify a timeframe risks a project being continually delayed while projects that are perceived to be more urgent (those with specific deadlines), are progressed.

As the objectives are being developed it is important to consider and plan for how data and information may be collected and in what form. This information will in turn drive the types of analyses that will need to be done.

1.5.7.3 *Searching the literature and other sources*

A review of scientific literature may be performed as part of the initial investigation or to gather information when preparing or designing additional studies. A literature search might be useful to:

- identify previous studies that are relevant to the current task;
- gather additional data that might be of use in supplementing existing data for the study;
- develop a differential diagnosis list in a disease outbreak of unknown cause;
- see how others have approached similar tasks; and
- gather additional information to support your conclusions.

With widespread access to the internet and library services, searching for information is now relatively easy. Most search engines search using keywords that you enter and searches may be conducted against the title or abstract of a paper, authors or the entire content.

Searches can be refined by adding more terms and constructing logical search statements. Different search engines handle multiple terms differently, often using an ‘advanced search’ page to set detailed search parameters. In PubMed and Medline, terms can be combined in a search statement using AND and OR logical operators. For example: ‘dogs and hepatitis’; ‘“johne’s disease” or paratuberculosis’. If AND and OR operators are combined in one statement, the AND part will be processed first, then the OR, unless the OR is contained in parentheses.

Example: [cattle and johne’s disease or paratuberculosis] is different to [cattle and (johne’s disease or paratuberculosis)]. The first statement will retrieve all resources for Johne’s disease in cattle or paratuberculosis in any species, while the second returns only resources relating to Johne’s disease in cattle or paratuberculosis in cattle.

The ready availability of information via the internet means that often the bigger problem is not just finding information but finding those sources that are most relevant to your needs. It is important to compose and refine searches carefully, to make them highly specific for the desired topic. If this is not done, a large number of non-relevant articles are likely to be listed, making it very difficult to identify the important ones for closer scrutiny.

For example, a search on PubMed for “Johne’s disease” returns more than 800 matches. By refining the search to find references about vaccines in cattle (“Johne’s disease” and cattle and vaccine), this list can be reduced to less than 50. Additional terms can be added to further refine the search as necessary.

At the same time it is important not to get too specific, in case important papers have not been indexed on all the terms you have used.

Once a list of potential sources has been identified, selected items can usually be saved to a text file, or often to a reference manager. Abstracts of papers listed on PubMed and Medline are often available on-line free of charge, but copies of the full papers will usually need to be either purchased on-line, or obtained as downloads or photocopies through a library service (usually government agencies or universities).

A useful feature of Medline through Current Contents (for those with access to this service) is that it is possible to save regularly used searches for re-use or to be run on a weekly basis by the system, with new results each week forwarded as a text file to your email address. This feature is particularly useful if there are subject areas where you wish to stay abreast of the latest developments on an on-going basis.

1.5.7.4 Data collection

Once it is clear what is required, the next step is to plan for collection of information and data.

It is expected that most epidemiologic studies will focus on collecting raw data by making observations or measurements on animals or collecting samples for laboratory testing.

Data and information may also be obtained from talking to producers or other people, from the literature review, from government statistics on animal production and laboratory testing (including iSIKHNAS data).

Information gathered in looking at salmonellosis in sheep feedlots may come from a literature search, from which it is concluded that *Salmonella* is orally acquired and exposure dose is important – this may lead to identification of simple measures such as feeding in raised troughs which prevent faecal contamination of feed and ensuring good drainage to prevent slurry build-up. Alternatively, data might be available from feedlot and veterinary records, providing facts about cases (and non-cases) of salmonellosis that have occurred. This data would then need to be collated, summarised and interpreted to generate information from which to draw conclusions.

For an outbreak investigation, relevant data could include quantitative data on individual cases of disease, case histories on individual animals (both cases and non-cases) veterinarians' (or others) observations and impressions on cases, laboratory reports on testing undertaken on affected and unaffected animals, as well as potential sources of disease (such as samples of feed, water, soil and environment).

In other cases, the available data could comprise a series of paper files describing the issue of concern and providing relevant historical data. These files need to be read, collated and summarised to put the data into a form that can be easily understood and interpreted.

1.5.7.5 Entering, editing and analysing data

Once the relevant data are collected it is necessary to enter, collate and edit or check the data prior to any formal analyses.

1.5.7.6 Editing data

It is assumed that data are likely to be entered into either a spreadsheet or database for routine data management and preparation for subsequent analyses.

Where data are entered from paper records such as questionnaires or printed material, it may be useful to select a representative number of cells and check entered values against paper records as a form of quality assurance.

Table 1.10: Organisation of data into tabular format in preparation for analyses

FARMID	NUMPONDS	TYPE	FEED	ANN_PROD
001	3	Fish	Commercial	400
002	6	Shrimp	Commercial	6000
003	2	Shrimp	Self	800
004	1	Fish	Self	50

A variety of simple checks should be performed to try and detect errors and implausible or inconsistent values. Each column can be sorted and the top and bottom rows inspected to look for values that may be outliers or implausible (cow that weighs 45 kg).

Coding should be checked for all categorical variables. If sex is coded as M=male, MC=male castrate and F=female, then check to see if there are any cells with values that are not on this coding list. There may be cells entered as MF, male instead of M, heifer instead of female etc.

Often it is useful to examine two variables in combination for logic checks to detect problems like a female animal that is recorded as having been castrated, a non-pregnant animal recorded as having calved etc.

1.5.7.7 Developing the analytical approach

The design (way the data were collected) will inform the analytical approach. If the data were collected using a case-control study design then you should be planning to do 2x2 tables with odds ratios and possibly logistic regression to analyse the data. If the data were collected using a cross-sectional or cohort study design then you may be able to use 2x2 tables with relative risk measures and possibly other types of advanced analyses.

You should identify the outcomes of interest in the dataset, exposures or risk factors of interest and other variables that may be useful as confounders.

Complete descriptive or exploratory analyses. This may include simple summaries of the number of records in the dataset, description of each variable with coding system and type of data (continuous, ordinal, nominal or categorical), numbers of missing values in each variable, summary statistics for each variable (mean/median, counts by category etc), start and end date of data collection.

1.5.7.8 Simple analyses

2x2 tables (also called contingency tables or cross-tabulations) are simple and easy to perform and form the mainstay of initial analyses of field epidemiology data. Larger tables of attack rates can be produced for various risk factors.

Where data can be structured into a binary coding for disease (disease present, disease absent) and where the risk factor under consideration can also be classified using a binary approach (absent or present), then counts of the number of animals in each of these combinations can be entered into a 2x2 table and analysed to produce either odds ratio or relative risk estimates and associated confidence intervals and p-values.

Statistical tests provide a p-value (probability) that is interpreted as the likelihood of obtaining the results by chance alone if there was no association between risk factor and disease. Where the p-value is greater than a defined threshold, ($\alpha=0.05$ or in some cases 0.1), we interpret the findings as *not-significant* and as indicating that the result could have occurred by chance alone and that the evidence does not support an association between the risk factor and disease. When the p-value is less than the threshold then we interpret the finding as *significant* and as indicating that there is an association between the risk factor and disease.

It is important to note that sometimes statistical tests return a non-significant finding even though the factor may be associated with disease. This is more likely to occur when the sample size is small. Small sample size alone should not interfere with the point estimate of an OR or RR though it may affect the confidence interval and the p-value.

From an epidemiological perspective, estimates and confidence intervals and unadjusted screening tests (relative risks or odds ratios) may be more useful than technically more advanced multivariable statistical analyses. This is particularly the case in the early stages of a complex disease investigation. Over time as more carefully planned studies are designed

and implemented then advanced analyses may be more appropriate but these will generally take time to plan and perform.

It is also important to note that statistical significance or meaningful RR or OR estimates do not necessarily provide proof of causation. They provide evidence of statistical association. Where there is care and attention in the design of the study you may have more confidence in the causal interpretation of the results of statistical tests.

1.5.7.9 More advanced statistical analyses

Statistical advice should be sought before proceeding to more advanced analyses.

It is possible to perform stratified analysis of 2x2 tables using the *Maentel-Haenszel procedure* to adjust analysis for the confounding or modifying (interaction) effect of a different factor. The result may be an adjusted overall measure of association or separate measures of association for each level of the other factor.

Finally more advanced statistical models may be used to analyse larger or more complex datasets to produce adjusted measures of association between multiple factors in one model. There are a number of benefits of multivariable modelling in understanding associations between many factors and disease. The effects of any one factor in the model are adjusted for all other effects in the model, effects of interactions and confounding can be incorporated in the model and models can be expanded to incorporate dependencies amongst observations (clustering of units).

Logistic regression is very commonly used to analyse epidemiologic data from disease investigations where the disease outcome can be represented as a binary variable (0=no disease, 1=disease) and where multiple risk factors are being considered. Other types of analyses that may be considered include Poisson or negative binomial regression for count data (counts of the number of cases of disease), survival analysis for time to event (time to occurrence of disease) and possibly linear models when an outcome of interest is continuous (effect of disease status on body weight or growth).

1.5.7.10 Use of other information

In many cases, you could be asked to synthesise available information and make conclusions and recommendations with very little (or without any) quantitative data to analyse. In such

situations the 'data' is likely to consist of paper files, case reports, subjective observations or other 'soft' data.

Qualitative data is not amenable to the numerical methods used to summarise and make inference from quantitative data. Instead, a qualitative analysis is required, following a series of systematic steps, such as:

- thorough review and summarisation of the available material
- identification of consistent patterns or anomalies in the data
- identifications of strengths and limitations in the data
- identification of likely and logical explanations for the observed patterns

It is usually not possible to make definitive statements about cause-and-effect or other specific relationships. However, particularly where the data is of a reasonable quality and consistency it is often possible to arrive at a conclusion as to the most likely explanation (or a group of likely/possible explanations) for the observed patterns.

1.5.8 Interpreting field data and information

Elevated relative risk or odds ratio estimates may provide suspicion about possible causes of disease but should be interpreted with caution. Chance, bias, confounding and other sources of error (data entry error, incorrect analyses etc) should all be considered as alternative explanations for elevated or significant measures of association.

Where the data were derived from a carefully planned study with design attributes intended to prevent bias and other problems and where the data management and analyses have been conducted appropriately and the results have produced significant measures of association with meaningful relative risk (or odds ratio) values, and where the findings are biologically plausible and consistent with other studies or findings, then you may have increased confidence in the findings.

In many cases you will be expected to draw conclusions and make recommendations based on less than perfect data/information. When this happens it is essential not only to recognise the limitations of the available data and information, but also to continue with those analyses that the data will support and draw what conclusions you can. In many cases, your

recommendations are likely to include collection of additional data to provide further support (or otherwise) for your preliminary conclusions.

In 1994, an incident occurred in Queensland where a previously unidentified virus (since characterised as Hendra virus) was responsible for the death of 14 horses and one human (with a second affected human subsequently recovering), associated with a single racehorse stable (Baldock et al., 1996). During the investigation it became rapidly apparent that this was a previously unidentified disease, and that the aetiology was unknown. However, even before the causal virus was identified, it was possible to determine that it was probably infectious in nature; was most likely to be directly transmitted; was not highly contagious (either among horses or humans); and that it probably originated from an, as then, unidentified wildlife reservoir (Baldock et al., 1995). Just on one year after the Hendra outbreak, flying foxes (fruit bats) were identified as the presumptive natural host of the virus, with about 14% of flying foxes sampled being seropositive (Baldock et al., 1996). The virus was subsequently isolated from uterine fluids of a flying fox (Halpin et al., 1996). Flying foxes were known to feed in trees in a spelling paddock associated with the stable and in which the index case was grazing prior to becoming sick. The specific mechanism of transmission among bats and from bats to horses is still not known.

In the Hendra virus example, there was virtually no quantitative data available for analysis, and yet a remarkably accurate picture of what happened and the cause and source of the outbreak were generated by critical review and interpretation of the findings of medical and veterinary investigations of affected animals and humans.

1.5.9 Preparing the report

Effective communication of the findings of your investigation to the appropriate decision makers is critical. If the findings are not communicated in a manner that allows key stakeholders to understand the results and use the information to make good decisions, then the effort will have had little benefit.

A final report from any epidemiologic study should be prepared in a systematic manner with sections following scientific convention: introduction, outline of objectives, materials and methods, results, discussion and bibliography.

It is important to use a structured and systematic approach and always ensuring that the findings are consistent with the interpreted information and data available at the time. Describe and record your methods and findings so that any conclusions and recommendations are easily understood and the process of arriving at these conclusions is transparent and

apparent to others. This is essential so that the basis and limitations of the conclusions are understood by those responsible for implementing any response to your recommendations.

1.6 References and other resources

Baldock, F. C., Douglas, I. C., Halpin, K., Field, H., Young, P. L. & Black, P. F. 1996. Epidemiological investigations into the 1994 equine morbillivirus outbreaks in Queensland, Australia. *Singapore Veterinary Journal*, 29:57-61.

Baldock, F. C., Pearse, B. H. G., Roberts, J., Black, P., Pitt, D. & Auer, D. 1995. Acute equine respiratory syndrome (AERS): The role of epidemiologists in the 1994 Brisbane outbreak. In: *Proceedings of the Australian Association of Cattle Veterinarians*. Melbourne, Australia. Brisbane, Australia: Australian Association of Cattle Veterinarians, 174-177.

Bradford-Hill, A. 1965. The environment and disease: Association or causation? *Proc R Soc Med*, 58:295-300.

Cleland, P. C., Chamnanpood, P. & Baldock, F. C. 1991. Investigating the epidemiology of foot and mouth disease in northern Thailand. In: DJ, K., ed. *Epidemiology Workshop: Supplement to Epidemiology at work*. Proceedings 176, September 1991. Tanunda SA. Sydney Australia: Postgraduate Committee in Veterinary Science University of Sydney, 17-32.

Cutler, S. J., Fooks, A. R. & van der Poel, W. H. M. 2010. Public health threat of new, re-emerging, and neglected zoonoses in the industrialised world. *Emerging Infectious Diseases*, 16, 1-7.

Dohoo, I., Martin, W. & Stryhn, H. 2010. *Veterinary Epidemiologic Research*, Charlottetown, Prince Edward Island, Canada, VER Inc.

Frerichs, R. R. 2001. History, maps and the internet: UCLA's John Snow site. *Society of Cartographers Bulletin*, 34:3-7.

Gregg M.B. 2002. *Field Epidemiology*, New York, New York, Oxford University Press.

Halpin, K., Young, P. & Field, H. 1996. Identification of likely natural hosts for equine morbillivirus. *Communicable Diseases Intelligence*, 20, 476.

Hueston, W. D. 2003. Science, politics and animal health policy: *Epidemiology in action*. *Preventive Veterinary Medicine*, 60: 3-13.

Kirkland, P. D., Love, R. J., Philbey, A. W., Ross, A. D., Davis, R. J. & Hart, K. G. 2001. Epidemiology and control of Manangle virus in pigs. *Australian Veterinary Journal*, 79:199-206.

Lessard, P. 1988. The characterization of disease outbreaks. *The Veterinary Clinics of North America: Food Animal Practice*, 4:17-32.

Love, R. J., Philbey, A. W., Kirkland, P. D., Davis, R. J., Morrissey, C. & Daniels, P. W. 2001. Reproductive disease and congenital malformations caused by Menangle virus in pigs. *Australian Veterinary Journal*, 79:192-198.

Martin, S. W., Meek, A. H. & Willeberg, P. 1987. *Veterinary Epidemiology*, Ames, Iowa, Iowa State University Press.

Mayer, D. (ed.) 2004. Essential Evidence Based Medicine, 2nd edition. Cambridge UK: Cambridge University Press.

Morris, R. S., Wilesmith, J. W., Stern, M. W., Sanson, R. L. & Stevenson, M. A. 2001. Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain. *Veterinary Record*, 149:137-144.

Rothman, K. J., Greenland, S. & Lash, T. 2008. *Modern Epidemiology*, Philadelphia, PA, Lippincott, Williams, & Wilkins.

Snow, J. (ed.) 1855. *On the mode of communication of cholera*, London: John Churchill.

Stephen, C. & Ribble, C. S. 1996. Marine anemia in farmed chinook salmon (*Onchorhynchis tshawytscha*): Development of a working case definition. *Preventive Veterinary Medicine*, 25:259-269.

Thrusfield, M. 2005. *Veterinary Epidemiology* 3rd edition, Oxford, Blackwell Science.

Software resources

EpiCalc, 2000, v 1.2. A multi-function statistical calculator that works with pre-tabulated data. Available from: <http://www.brixtonhealth.com/index.html>.

EpiInfo, v 3.3. Database and statistics software for public health professionals. Available from: <http://www.cdc.gov/epiinfo/>.

EpiTools, 2004. AusVet's on-line epidemiological calculators and utilities. Available at: <http://epitools.ausvet.com.au>.

WinEpiScope, v 2.0. Software for quantitative veterinary epidemiology. Available at: <http://www.clive.ed.ac.uk/winepiscope/>.

Internet search engines and databases

Some of the commonly used, web-based, scientific databases include:

- Medline/PubMed – This indexes all major medical, veterinary, epidemiological and associated journals, and is freely available for all users through PubMed (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>).
- Medline is also available through Current Contents and other service providers through institutional library subscriptions.
- ScienceDirect (<http://www.sciencedirect.com/>) provides indexing and search facilities for a wide variety of scientific journals in the physical, life, health and social sciences.
- Biosis previews/Web of knowledge (http://thomsonreuters.com/products_services/science/science_products/a-z/biosis/) indexes a wide variety of journals, conference proceedings, books, review articles etc in the broad life sciences area. Available through institutional subscription.
- Sciverse Scopus (<http://www.scopus.com/home.url>) claims to be the world's largest abstract and citation database of peer-reviewed literature and quality web sources,

covering a multitude of topics. It is available through institutional subscription and some publishers provide temporary access to reviewers of journal papers.

- Agricola (<http://agricola.nal.usda.gov/>) is the catalogue of the National Agricultural Library of the USA and provides citations and abstracts for an extensive collection of agricultural literature.
- CAB Abstracts (CAB Abstracts) includes over 6.3 million records from 1973 onwards, with over 300,000 abstracts added each year, covering agriculture, environment, veterinary sciences, applied economics, food science and nutrition. Access is via institutional subscription or by time-based payment.
- JSTOR (<http://www.jstor.org/>) indexes more than 1,000 refereed journals from a wide variety of disciplines, including aquatic, biological and health sciences and statistics. Available through institutional or individual subscription.
- SIGLE (<http://www.opengrey.eu/>), or *System for Information on Grey Literature in Europe*, indexes more than 700,000 bibliographical references from the grey literature (research reports, doctoral dissertations, conference papers, official publications, and other types of non-refereed publications), produced in Europe. Open access to all users.

More general search engines include:

- Scirus (<http://www.scirus.com/srsapp/>) – This is a broader search engine covering a wide range of scientific information across disciplines and publication types. Scirus covers not only scientific journals, but also web publications and a range of other non-refereed sources.
- Google Scholar (<http://scholar.google.com.au/schhp?hl=en>) also supports broad searches of the academic and scientific literature. It allows for searching across many disciplines and sources and ranks documents according to relevance and quality or frequency of citation.
- Google (<http://www.google.com/>) and other internet search engines can be used, but the content returned is not limited in any way other than by your search. These engines will return news items, personal web pages and any internet content that is relevant to the search criteria (and some that is not!).

2 Diagnostic tests

In field epidemiology the term *diagnosis* generally means identification of a disease or condition that is affecting an animal. A diagnostic test is any procedure or process that may contribute to the development of a diagnosis. The term diagnostic test may be used to refer to a clinical examination of an individual animal or to a laboratory test performed on a sample collected from an animal (blood, faeces etc). The results of these tests are interpreted and used to determine if the animal has a disease or not.

Diagnostic tests may be applied to an individual animal (as above) or to a group or aggregate of animals such as a mob, herd or farm. The presence of one or more disease-positive animals in a mob may mean that the mob is declared infected (mob-level diagnosis).

2.1 Measures of diagnostic test performance

2.1.1 Accuracy and precision

Accuracy relates to the ability of the test to provide a result that is close to the truth (the true value). Accuracy is generally assessed in the long run, meaning that it can be thought of as the average of multiple test results. A test is therefore considered to be accurate when the average of repeated tests is close to the true value. Any one test result may itself not be as accurate as the average of repeated tests performed on the same sample.

Precision refers to how repeatable the test is. If the test is repeated and the result from different runs is always the same then the test is precise (regardless of whether the result is accurate or not).

*A **precise** test has a low level of random error i.e. a high level of **repeatability**.*

*An **accurate (valid)** test has a low level of systematic error (**bias**).*

A test can be precise without being accurate and vice versa. A good test is both precise and accurate. The concepts of precision and accuracy are most easily understood by thinking of shooting at a target as shown below. The term validity is often used to represent accuracy.

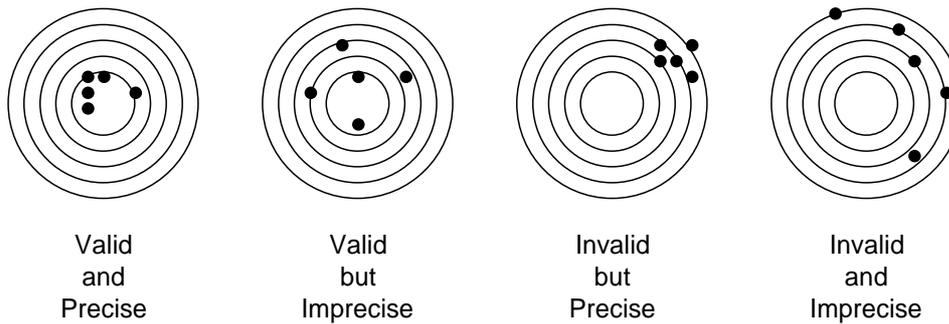


Figure 2.1: Diagram showing test accuracy (validity) and precision

There are lots of terms used to describe attributes of diagnostic tests.

- *Repeatability* generally refers to results from repeated testing of the same sample performed in the same lab
- *Reproducibility* refers to results from testing the same sample in different labs (splitting the original sample in to multiple sub-samples and sending them to different labs)
- *Agreement* refers to how well two different tests agree

Tests performed on identical material under apparently similar conditions do not, in general, yield identical results. This variation is attributed to random error inherent in every test procedure because factors that may influence the result of a test cannot all be completely controlled. When interpreting test results, this variability must be taken into account. There are many different factors which contribute to the variability of a test procedure, including:

- uniformity of test material
- transport and storage of test material
- reagents
- equipment and its calibration
- operator
- environmental conditions - temperature, humidity, light, air pollution

2.1.2 Ability of a test to detect disease states

All tests may produce errors in their results. To qualify as a test, the procedure should classify animals (diseased, not diseased) at least more accurately than a purely random procedure - such as tossing a coin.

The two types of **errors** that a test can make are:

- **false positive** — the test identifies an animal to have a disease when it does not
- **false negative** — the test identifies an animal not to have a disease when it does

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

- **Sensitivity** (Se) is the probability that a positive animal will be identified as positive by the test (1 – false negative rate) — this describes the test ability to detect a disease animal
- **Specificity** (Sp) is the probability that a negative animal will be correctly identified as negative by the test (1 – false positive rate) — this describes the tests ability to determine an animal is not diseased

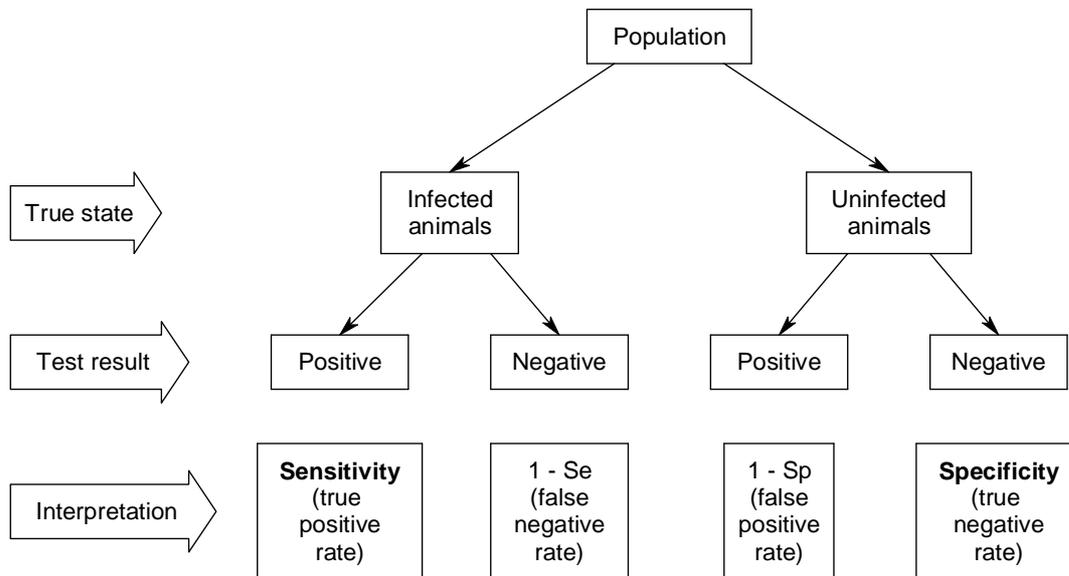


Figure 2.2: Diagram showing interpretation of sensitivity and specificity in relation to true disease status and diagnostic test outcome.

Assessing Se and Sp often includes something called a gold standard test. A gold standard is a test that is absolutely accurate – it correctly detects all diseased animals and it correctly classifies all non-diseased animals as being disease free.

Traditionally test performance was based on comparison to a gold standard. More recently methods have been developed that allow assessment of test performance in the absence of a gold standard but these methods are usually more complex.

Diagnostic test performance is often described using a 2x2 table approach.

Table 2.1: Table showing 2x2 layout used for assessment of diagnostic test performance. Columns display true disease status based on a gold standard test (+= positive, -=negative) and rows display the results from a diagnostic test.

	Gold Std +	Gold Std -	Total
Test +	a	b	a + b
Test -	c	d	c + d
Total	a + c	b + d	a + b + c + d

$$Se = \frac{a}{(a+c)} \quad Sp = \frac{d}{(b+d)}$$

If a new test were applied to 100 animals, made up of 60 healthy animals and 40 infected animals, the results in Table 4.2 might be obtained.

	Infected	Healthy	Total
Test +	36	10	46
Test -	4	50	54
Total	40	60	100

- the sensitivity of the test is $36/40 = 90\%$
- the specificity of the test is $50/60 = 83.3\%$

2.1.3 Sensitivity

Sensitivity is the proportion of animals with the disease (or infection) of interest which test positive (i.e. proportion of true positives).

Sensitivity (True Positive Fraction): The proportion of animals with the disease of interest that test positive. Sensitivity is also defined as the conditional probability that a test will correctly identify those animals that are infected ($Pr T+|D+$).

The false negative fraction is $1 - Se$.

2.1.4 Specificity

Specificity is the proportion of animals without the disease of interest which test negative (i.e. proportion of true negatives).

Specificity (True Negative Fraction): The proportion of animals without the disease of interest that test negative. Specificity is also defined as the conditional probability that a test will correctly identify those animals that are not infected ($Pr T-|D-$).

The false positive fraction is $1 - Sp$.

One way to remember the difference between Se and Sp is to think:

- $Se = "e"$ = false negatives and true positives
- $Sp = "p"$ = false positives and true negatives

There is an inverse relationship between Se and Sp for most tests and particularly those that are based on a continuous measure, such as an ELISA. This means that tests that have a very high Se will often have a lower Sp and vice versa.

Tests that produce a continuous measure (measuring antibody or enzyme concentration in blood for example) can have the cut-point altered to move test performance towards higher Se or higher Sp .

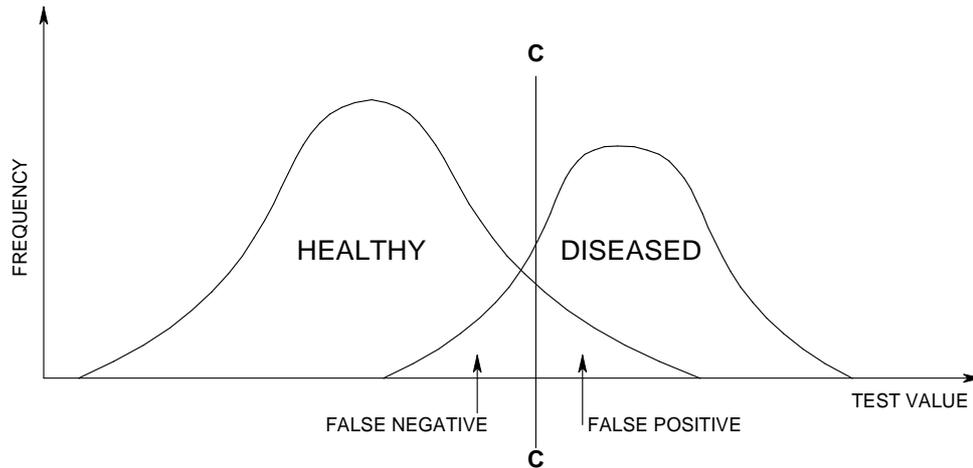


Figure 2.3: Plot showing a frequency measure of test results from application of a diagnostic test applied to healthy and diseased animals when the test output is measured on a continuous scale. The vertical line at C-C represents a cut-point to distinguish healthy animals (to the left of C-C) from diseased animals (to the right of C-C).

Results for animals that are disease free (D-) usually overlap with the results in the diseased population (D+). Animals to the right of the cut-point (C-C) are classified as reactors (diseased or infected) and animals to the left are classified as negative (non-infected). If fewer false positives are required, C-C is moved to the right; specificity increases and sensitivity decreases. However, if fewer false negatives are required, C-C is moved to the left: sensitivity increases and specificity decreases.

Selection of the appropriate cut-off value will depend on a number of issues including the relative cost of false positives and false negatives, the stage of an eradication program, if any, and the availability of other tests. An important consequence of imperfect specificity (i.e. $< 100\%$) is that if a large number of animals are tested from a population free of the disease in question, there is a significant chance of abnormal results. For example, if 10 independent samples were tested using a test with 90% specificity the probability of at least 1 positive test result occurring is 65%.

2.1.5 Predictive Values

Se and Sp are characteristics of the test when the test is applied to animals of known disease status. Se and Sp do not tell us how useful the test might be when applied to animals of unknown disease status.

For most people in the field, the practical situation is that they have performed the test on an animal with unknown disease status and they wish to use the test results to better classify the animal as disease + or disease -. Se and Sp are not helpful in this situation.

Predictive values are useful in this situation. Predictive values allow us to answer the two related questions:

- What proportion of the test positive animals are truly infected?
- What proportion of the test negative animals are truly not infected?

Predictive values are functions of *prevalence* and the test characteristics of *sensitivity* and *specificity*. As prevalence declines so does the positive predictive value. The converse is true for negative predictive value.

Positive predictive value = $a/(a+b)$ =	$\frac{\text{Prev} \times \text{Se}}{\text{Prev} \times \text{Se} + (1-\text{Prev}) \times (1-\text{Sp})}$
Negative predictive value = $d/(c+d)$ =	$\frac{(1-\text{Prev}) \times \text{Sp}}{(1-\text{Prev}) \times \text{Sp} + \text{Prev} \times (1-\text{Se})}$

With an understanding of the principles of predictive values, the following rules of thumb for using tests in the diagnostic process at the individual animal level can be recommended:

- If the objective is to *confirm a likely diagnosis* (the "rule-in" situation), then choose a test which has *high specificity* (~>95%) and at least moderate sensitivity (~>75%). If a positive result is returned, then it is highly likely the individual has the disease in question (PPV's are high for tests with high specificity). If a negative result is returned, then further diagnostic work up is required.
- If the objective is to *confirm that an individual is free from a particular disease* (the "rule-out" situation), then choose a test with high sensitivity (~>95%) and at least

moderate specificity ($\sim >75\%$). If a negative result is returned, then it is highly likely the individual is free from the disease in question. If a positive result is returned, then further testing is required with more specific tests to ascertain whether or not the result was a false positive result or not.

2.2 Multiple testing

Two or more tests can be used either sequentially or simultaneously and results interpreted in series or parallel. In parallel interpretation, an animal is considered positive if it reacts positively to either or both tests - this increases sensitivity but tends to decrease the specificity of the combined tests. In series interpretation, an animal must be positive on both tests to be considered positive - this increases specificity at the expense of sensitivity.

Parallel interpretation of tests means that both tests must give a positive result for the animal to be considered positive. Parallel testing increases sensitivity but tends to decrease the specificity, compared to using either test in isolation.

Series interpretation of tests means that the animal is considered positive if either of the tests gives a positive result. Series testing increases specificity at the expense of sensitivity, compared to using either test in isolation.

In general, the greater the number of tests involved, the greater the increase in sensitivity or specificity, depending on the method of interpretation that is used.

2.2.1 Sensitivity and specificity for multiple tests

Overall values for sensitivity for interpretation of tests in series or parallel, assuming conditional independence of the tests, can be calculated using the following example.

For this example the two tests are assumed to be independent and have the following characteristics:

Test 1 – Se = 50%; Sp = 98.7% Test 2 – Se = 60%; Sp = 98.6%

What are the theoretical sensitivities and specificities of the two tests used in parallel or series?

For sensitivity, we assume an animal is infected and that it is tested with both Test 1 and Test 2. For Test 1, the probability of a positive test result (given that the animal is infected) is $Se_1 = 0.5$ and the corresponding probability that it will give a negative result is $1 - Se_1$, also = 0.5 for this example. For Test 2, the probability of a positive test result (given that the animal is infected) is $Se_2 = 0.6$ and the corresponding probability that it will give a negative result is $1 - Se_2 = 0.4$.

For series interpretation, both tests must be positive for it to be considered a positive result. From the scenario tree this is the result for the first limb on the left, which has probability $P(+ / +) = Se_1 \times Se_2 = 0.5 \times 0.6 = 0.3$. Thus, the formula for sensitivity for series interpretation is $Se_{series} = Se_1 \times Se_2$ and for this example is 0.3 or 30%.

For parallel interpretation, the result is considered positive if either of the individual test results is positive. Alternatively, for a result to be considered negative both test results must be negative. Again this can be determined from the scenario tree, where the limb on the right represents both tests having a negative result and the probability of both negative results is $P(- / -) = (1 - Se_1) \times (1 - Se_2)$. Therefore the probability of an overall positive result for parallel interpretation is $Se_{parallel} = 1 - (1 - Se_1) \times (1 - Se_2) = 0.8$ (80%) for this example.

Similar logic can be applied to the example of an uninfected animal to derive formulae for specificity for series and parallel interpretation as shown below:

$Sp_{parallel} = Sp_1 \times Sp_2 = 0.973$ or 97.3% for this example and

$Sp_{series} = 1 - (1 - Sp_1) \times (1 - Sp_2) = 0.999$ or 99.9% for our example

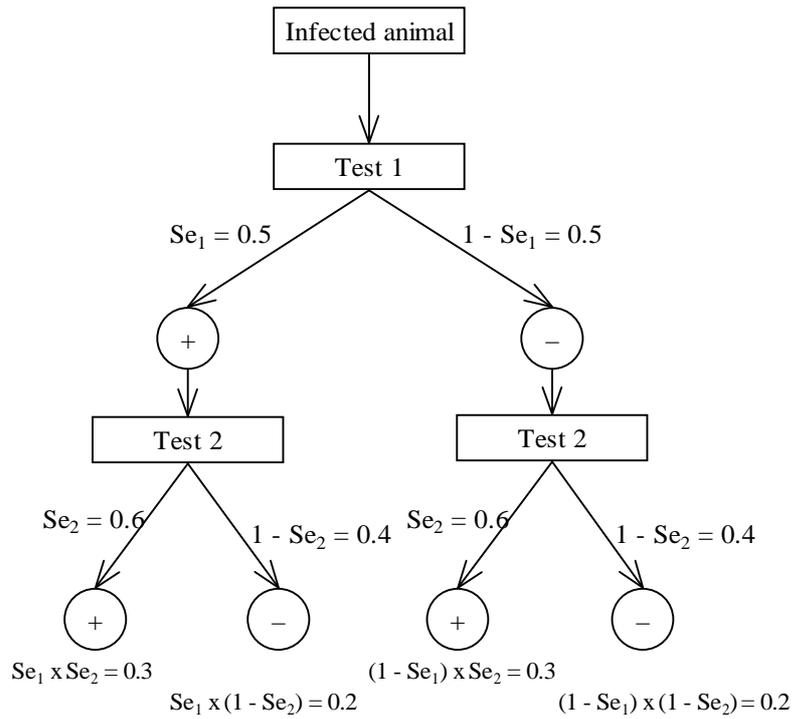


Figure 2.4: Scenario tree for calculating overall sensitivity for two tests interpreted in series or parallel

2.2.2 Conditional independence of tests

An important assumption of series and parallel interpretation of tests is that the tests being considered are conditionally independent. Conditional independence means that test sensitivity (specificity) remains the same regardless of the result of the comparison test, depending on the infection status of the individual.

If the assumption of conditional independence is violated then combined sensitivity (or specificity) will be biased. The “conditional” term relates to the fact that the independence (or lack of independence) is conditional on the disease status of the animal. Therefore sensitivities may be conditionally independent (or not) in diseased animals, while specificities may be conditionally independent (or not) in non-diseased animals.

Two tests are **conditionally independent** if test sensitivity or specificity (depending on disease status) of one test remains the same regardless of the result of the other (comparison) test

If tests are not independent (are correlated), the overall sensitivity or specificity improvements may not be as good as the theoretical estimates, because two tests will tend to give similar results on samples from the same animal.

For example, let us assume that the two tests described above were applied to 200 infected and 7,800 uninfected animals with the following results. What are the actual sensitivities and specificities for parallel and series interpretations and how do they compare to the theoretical values?

Test 1	Test 2	Infected	Uninfected
+	-	30	70
-	+	50	80
+	+	70	30
-	-	50	7620
Total		200	7800

Observed sensitivities and specificities of the two tests used in parallel or series are:

$$Se_{\text{series}} = 70/200 = 35\% \quad Se_{\text{parallel}} = 150/200 = 75\%$$

$$Sp_{\text{series}} = 7770/7800 = 99.6\% \quad Sp_{\text{parallel}} = 7620/7800 = 97.7\%$$

Sensitivity in series has dropped less than predicted (35% instead of 30% predicted), and sensitivity of parallel testing has increased less than predicted (75% compared to 80% predicted). The apparent difference between calculated and observed values for combined sensitivities suggests that these tests are in fact correlated.

This difference is due to correlation of the test sensitivities, so that infected animals that are positive to Test 1 are also more likely to be positive in Test 2, as shown by the substantial difference in sensitivity of Test 2 in animals positive to Test 1 (70/100 or 70%) compared to those negative to Test 1 (30/100 or 30%).

The differences in observed and predicted specificities are much smaller and in this case probably due to random variation.

Lack of conditional independence of tests is particularly likely if two tests are measuring the same (or similar outcome).

For example: ELISA and AGID are two serological tests for Johne's disease in sheep. Both tests measure antibody levels in serum. Therefore, in an infected animal, the ELISA is more likely to be positive in AGID-positive animals than in AGID-negative animals, so that the sensitivities of the two tests are correlated (not independent). This is illustrated in Table xx.5, where the sensitivities of both tests vary markedly, depending on the result of the other test. In contrast, serological tests such as ELISA and AGID are likely to be less correlated with agent-detection tests, such as faecal culture.

AGID	ELISA		Total
	+	-	
+	34	21	55
-	13	156	169
Total	47	177	224

All 224 sheep are infected, so we can calculate sensitivities of both ELISA and AGID as follows:

ELISA Se overall	$47/224 = 21.0\%$	AGID Se overall	$55/224 = 24.6\%$
ELISA Se in AGID +	$34/55 = 61.8\%$	AGID Se in ELISA +	$34/47 = 72.3\%$
ELISA Se in AGID -	$13/169 = 7.7\%$	AGID Se in ELISA -	$21/177 = 11.9\%$

2.2.3 Application of series and parallel testing

Series testing is commonly used to improve the specificity, and hence the positive predictive value, of a testing regimen (at the expense of reduced sensitivity).

For example, in large-scale screening programs, such as for disease control or eradication, a relatively cheap, high-throughput test with relatively high sensitivity and precision but only modest specificity may be used for initial screening. This sort of test can be applied to large numbers of animals (entire population) where the purpose is to be very confident that those animals that test negative are in fact disease free.

Any positives to the initial screening test are then tested using a highly specific (and usually more expensive) confirmatory test to minimise the overall number of false positives at the end of the testing process. For an animal to be considered positive it must be positive to both the initial screening test and the confirmatory follow-up test.

A good example of series testing is in eradication programs for bovine tuberculosis, where the initial screening test is often either a caudal fold or comparative cervical intradermal tuberculin test, which is followed up in any positives by a range of possible tests including additional skin tests, a gamma interferon immunological test or even euthanasia and lymph node culture, depending on circumstances.

In the above situation it is important to realise that even though the follow-up test is only applied to those that are positive on the first test, this is still an example of series interpretation. Because an animal must test positive to both tests for a positive overall result, the result of the second test in animals negative to the first test is irrelevant, so that the test doesn't actually need to be done. This is an important consideration in control or eradication programs, where testing costs are usually a major budget constraint and significant savings can be made by using a cheap, high-throughput screening test followed by a more expensive but highly specific follow-up test.

Parallel testing is less commonly used, but is primarily directed at improving overall sensitivity and hence negative predictive value of the testing regimen. Parallel testing is mainly applied where minimising false negatives is imperative, for example in public health programs or for zoonoses, where the consequences of failing to detect a case can be extremely serious. In contrast to series testing, every sample must be tested with both tests for parallel testing to be effective, so that testing costs can be quite high.

For example, in some countries testing for highly pathogenic avian influenza virus may rely on using a combination of virus isolation and PCR for detection of virus, with birds that are positive to either test being considered infected.

2.3 Measuring agreement between tests

There is often interest in comparing the diagnostic performance of two tests (new test compared to an existing test) to see if the new test produces similar results.

For the same specimens submitted to each of the two tests, the investigator records the appropriate frequency data into the 4 cells of a 2x2 table, a (both tests positive), b (test 1 positive and test 2 negative), c (test 1 negative and test 2 positive), and d (both tests negative). The value *kappa* (*k*), a measure of relative agreement beyond chance, can then be calculated using software such as EpiTools or using formulae in standard epidemiology texts.

Kappa has many similarities to a correlation coefficient and is interpreted along similar lines. It can have values between -1 and $+1$. Suggested criteria for evaluating agreement are (Everitt, 1989, cited by Thrusfield, 1995):

Table 2.2: Table showing interpretation of kappa values

<i>kappa</i>	Evaluation
$> 0.8 - 1$	Excellent agreement
$>0.6 - 0.8$	Substantial agreement
$>0.4 - 0.6$	Moderate agreement
$>0.2 - 0.4$	Fair agreement
$>0 - 0.2$	Slight agreement
0	Poor agreement
<0	Disagreement

Care must be taken in interpreting *kappa* – if two tests agree well, they could be equally good or equally bad! However, it may be possible to justify use of a newly developed test if it agrees well with a standard test and if it is cheaper to run in the laboratory.

Conversely, if two tests disagree, one test is likely to be better than the other although there may no way to tell which is better! The exception to this is where both tests have close to 100% specificity (i.e. no or few false positives). In this case the test with the larger number of

positive results is likely to be more sensitive. McNemar's Chi-squared test for paired data can also be used to test for significant differences between the discordant cells (b & c).

An example of kappa and agreement between tests

A comparison of two herd-tests for Johne's disease in sheep yields the following results (from Sergeant et al., (2002)):

	Test 2 results		
Test 1 results	+	-	Total
+	58	37	95
-	5	196	201
Total	63	233	296

How well do the two tests agree, and can you determine which test is better?

For these tests, kappa is 0.64, suggesting moderate-substantial agreement. However, McNemar's chi-squared is 22.88, with 1 degree of freedom and $P < 0.001$. This means that the discordant cells (37 and 5) are significantly different. From the data available it is not possible to say which test is better – the additional positives on Test 1 could be either true or false positives, depending on test specificity.

In this case, Test 1 was pooled faecal culture (specificity assumed to be 100%) and Test 2 was the agar gel-diffusion test with follow-up of positives by autopsy and histopathology (specificity also assumed to be 100%). How does this change the assessment of the two tests?

Considering that both tests have specificity equal (or very close) to 100%, there are likely to be very few false-positives. Therefore it appears that the sensitivity of Test 1 (pooled faecal culture) is considerably higher than that for Test 2 (serology), since Test 1 detected a greater number of positives overall.

2.3.1 Proportional agreement of positive and negative results

In some circumstances, particularly where the marginal totals of the 2-by-2 table are not balanced, *kappa* is not always a good measure of the true level of agreement between two tests (Feinstein and Cicchetti, 1990). For example, in the first example above, kappa was only 0.74, compared to an overall proportion of agreement of 0.94. In these situations, the proportions of positive and negative agreement have been proposed as useful alternatives to *kappa* (Cicchetti and Feinstein, 1990). For this example, the proportion of positive agreement was 0.78, compared to 0.96 for the proportion of negative agreement, suggesting that the main area of disagreement between the tests is in positive results and that agreement among negatives is very high.

2.4 Estimation of true prevalence from apparent prevalence

When we apply a test in a population, the proportion of positive results observed is the apparent prevalence. However, depending on test performance, apparent prevalence may not be a good indicator of the true level of disease in the population (the true prevalence). However, if we can estimate the sensitivity and specificity of the test, we can also estimate the true prevalence from the apparent (test-positive) prevalence (AP) using the formula (Rogan and Gladen, 1978):

$$\text{True prevalence} = \frac{AP + Sp - 1}{Se + Sp - 1}$$

which has a solution for situations other than when $Se + Sp = 1$. All values are expressed as proportions (between 0 and 1) rather than percentages for these calculations. Confidence limits can be calculated for the estimate using a variety of methods implemented in EpiTools. When true prevalence is 0, apparent prevalence = $1 - Sp$, the false positive test rate.

For example: Say we have conducted a survey with a test whose sensitivity is 90% (0.9) and specificity is 95% (0.95) and we find a reactor rate (apparent prevalence) of 15% (0.15). By using the formula, we can estimate the true prevalence to be 11.8% (0.118).

Another example

Suppose we have conducted a survey of white spot disease in a shrimp farm, using a test with sensitivity of 80% (0.8) and specificity of 100% (1.0). We have

tested 150 shrimp, and 6 shrimp tested positive. What is the estimated true prevalence?

The apparent prevalence is $6/150 = 0.04$ or 4% (Wilson 95% CI: 1.8% – 8.5%)

Therefore, true prevalence = $(0.04 + 1 - 1)/(0.8 + 1 - 1) = 0.04/0.8 = 0.05$ or 5% (95% CI: 1.1 – 8.9%)

What happens if we assume that sensitivity and specificity are both 90%?

If $Se = 0.9$ and $Sp = 0.9$:

Therefore, true prevalence = $(0.04 + 0.9 - 1)/(0.9 + 0.9 - 1)$
 $= -0.06/0.8 = -0.0625$.

The above example illustrates one potential problem with Rogan and Gladen formula, which is that in some circumstances negative estimates can be produced. However, a negative (<0) prevalence is clearly impossible, so for this scenario the assumptions about sensitivity and specificity must be incorrect. For example, if specificity was 90% (0.9), and you tested 150 animals, you would expect to have 0.1×150 or on average about 15 false positive results (even in an uninfected population). Therefore if only 4 positives were recorded, the specificity of the test must be much higher than 90% (a minimum estimate would be to assume all of the positives are false positives, so that specificity = $1 - \text{apparent prevalence} = 1 - 4\%$ or 96%).

Because prevalence estimates are proportions we should also calculate and present confidence intervals for the estimate.

2.5 Group (aggregate) diagnostic tests

The previous discussion describes the testing of individual animals. However, in epidemiological investigations, the study unit can often comprise a group of animals such as a herd of cattle, a flock of sheep, or a cage or pond of fish. For example, it is common practice to determine herd or flock status for some diseases based on the results of testing of a sample of animals, rather than testing the whole herd or flock.

In this situation, it is important to realise that testing for disease at the group or aggregate level incorporates a number of factors additional to those relevant to testing at the individual animal level. Thus, tests which may be highly sensitive and specific at the individual animal level can still result in misclassification of a high proportion of groups where only a small number of animals in each group are tested.

At the individual animal level, diagnostic test performance is determined by its sensitivity and specificity. The corresponding group-level measures are *herd sensitivity* and *herd specificity*. Herd sensitivity and herd specificity are affected by animal-level sensitivity and specificity, as well as the number of animals tested, the prevalence of disease in the group and the number of individual animal positive results (1, 2, 3 etc) used to classify the group as positive. Just as we do for individuals, we also want high sensitivity and high specificity in our group level interpretation.

Herd sensitivity (SeH) is the probability that an infected herd will give a positive result to a particular testing protocol, given that it is infected at a prevalence equal to or greater than the specified design prevalence.

Herd specificity (SpH) is the probability that an uninfected herd will give a negative result to a particular testing protocol (HSP)

2.5.1 Calculating herd sensitivity and herd specificity

The herd-level sensitivity (SeH) and specificity (SpH) with a cut-off of 1 reactor to declare a herd infected can be calculated as (Martin et al., 1992):

$$SeH = 1 - (1 - (Prev \times Se + (1 - Prev) \times (1 - Sp)))^m \text{ and}$$

$$SpH = Sp^m$$

Where Se and Sp are animal-level sensitivity and specificity respectively, $Prev$ is true disease prevalence and m is the number of animals tested. SeH is equivalent to the level of confidence of detecting infection in herds or flocks with the specified prevalence of infection. SeH and SpH can be easily calculated using EpiTools or other epidemiological calculators.

If test specificity is 100% (i.e. any reactors are followed up to confirm their status) calculation of SeH is simplified:

$$SeH = 1 - (1 - Prev \times Se)^m$$

An example

For example, assuming that we have tested 100 animals in a herd with a test that has $Se = 0.9$ and $Sp = 0.99$, what is the herd-sensitivity for an assumed prevalence of 5%?

$$SeH = 1 - (1 - (0.05 \times 0.9 + (1 - 0.05) \times (1 - 0.99)))^{100}$$

$$= 0.996 \text{ or } 99.6\%$$

This means that if disease is present at a prevalence of 5% or more, there is a 99.6% chance that one or more animals in the sample will test positively.

For this scenario, herd-specificity is:

$$SpH = 0.99^{100} = 0.37 \text{ or } 37\%$$

This means that there is a 37% chance that an uninfected herd will also have one or more animals test positively.

What happens if we assume that the prevalence of infection is 2% instead of 5%?

Herd-sensitivity:

$$\text{SeH} = 1 - (1 - (0.02 \cdot 0.9 + (1 - 0.02) \cdot (1 - 0.99)))^{100}$$

$$= 0.94 \text{ or } 94\%$$

SeH decreases as prevalence decreases.

Herd-specificity:

$$\text{SpH} = 0.99^{100}$$

$$= 0.37 \text{ or } 37\%$$

SpH is unaffected by prevalence because, by definition, SpH applies only to herds with zero prevalence (uninfected).

In the above example, increasing the cut-point number of reactors for a positive result from 1 to 2 (i.e. if there are 0 or 1 animals test positive the group is considered “uninfected” while if 2 or more test positive it is infected) results in an increase in SpH to 74% but a reduction in SeH to 77% (from EpiTools: <http://epitools.ausvet.com.au/content.php?page=HerdSens3>).

The formulae above assume that sample size is small relative to population size (or that the population is large). Similar formulae are also available for small populations or where the sample size is large relative to population size.

2.5.2 Risk of infection in test-negative animals

The only way to be 100% confident that no animals comprising a particular group are infected with a particular agent is to test every animal in the group with a diagnostic test which has perfect sensitivity and specificity. However, if only a low proportion of individual animals in the group are infected and only a small number are tested there can be quite a high chance that infected groups will be misclassified as uninfected. The following table shows the number of infected animals which may be present but undetected in a population of 100,000, despite a sample testing negative using a test with perfect sensitivity and specificity at the individual animal level.

Table 2.3: Number of diseased or infected animals which could remain in a group of 100,000 after a small number are tested and found to be negative using a test which has perfect sensitivity and specificity at the individual animal level for 95% and 99% confidence levels

No. of animals in sample tested from group of 100,000 and found negative	95%	99%
	100	2,950
500	596	915
1,000	298	458
10,000	29	44

The situation is further complicated where the test procedure being used has poor sensitivity, which is the case for many tests in regular use.

The probability of introducing infection in a group of tested-negative animals is the same as the probability that one or more animals in the group are infected but tests negative. This probability can be calculated as:

$$\begin{aligned} \text{Probability} &= 1 - \text{NPV}^m \\ &= 1 - [(1 - \text{Prev}) \times \text{Sp} / ((1 - \text{Prev}) \times \text{Sp} + \text{Prev} \times (1 - \text{Se}))]^m \end{aligned}$$

Where *NPV* is the negative predictive value of the test in the population of origin, *Se* and *Sp* are animal-level sensitivity and specificity respectively, *Prev* is true disease prevalence and *m* is the number of animals tested. As sample size increases the probability that the group will all test negative decreases, so that the overall risk associated with a group can be reduced by increasing the sample size. However, if all animals do test negatively the probability that one or more are actually infected increases (assuming that they are from an infected population), as shown in Figure xx.6.

For example: If 20 animals are selected from a herd or flock with a true prevalence of 0.05 (5%) and are tested using a test with $Se=0.9$ and $Sp=0.99$, and all 20 have a negative result, the probability that there are one or more infected animals in the group is about 0.1 (10%). In addition, the probability that all 20 animals will have a negative test result is about 0.33 (33%).

In simple language, there is a 1 in 3 chance that all animals test negative and also a 1 in 10 chance that there is one or more infected animals in the group, even if they have all tested negative.

Increasing sample size from 20 to 40 reduces the probability that all will test negatively from 33% to about 10%, but for those that are all negative, increases the probability that one or more are infected from 10% to 20% (1 in 5).

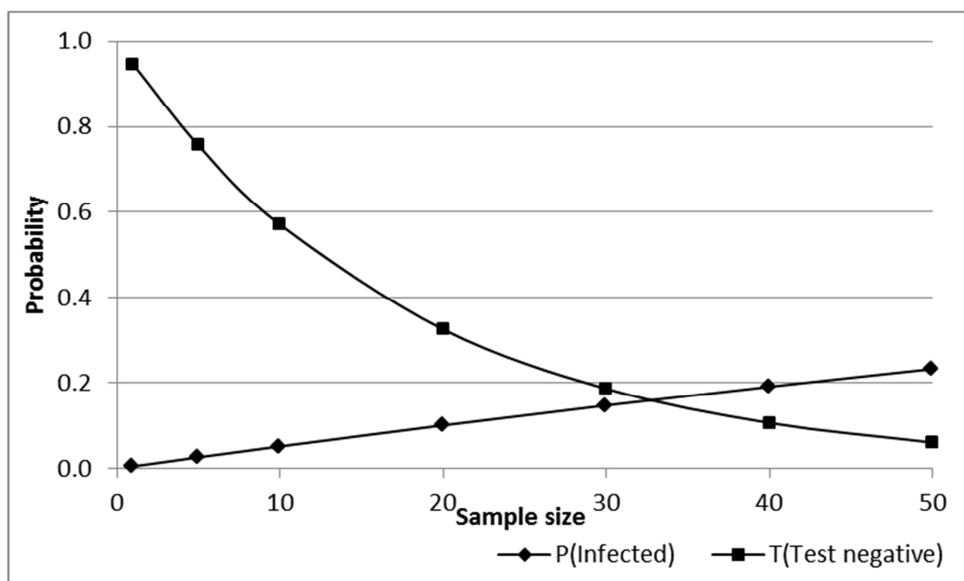


Figure 2.5: Effect of sample size on the probability that a group of test-negative animals will include one

or more infected (but test-negative) animals, and the probability that this will occur, for an assumed Se=0.9, Sp=0.99 and true prevalence=0.05 (5%) in the herd/flock of origin.

2.5.3 Demonstrate freedom or detecting disease?

It is impossible to prove that a population is free from a particular disease without testing every individual with a perfect test. However, demonstrating “freedom” from disease in a population is essentially the same as sampling to provide a high level of confidence of detecting disease at specified (design) prevalence. If we don’t detect disease, then we can state that we have the appropriate level of confidence that (if the disease is present) it is at prevalence lower than the *design prevalence*. Provided we have selected appropriate design prevalence, it can then be argued that if the disease were present it would more than likely be at a higher level than the design prevalence, and therefore we can be confident that the population is probably free of the disease.

The selection of appropriate design prevalence is obviously critical if it is too low sample sizes will be excessive, while if it is too high the argument that it is an appropriate threshold for detection of disease is weaker. For infectious diseases it is common to use a value equal to or lower than values observed in endemic or outbreak situations.

2.5.4 Important factors to consider in group testing

When testing a group of animals for the presence of disease, there are a number of important points to keep in mind:

- Individual and group level test characteristics (sensitivity and specificity) are not equivalent.
- The number of animals to be tested in the group (sample size) is relatively independent of group size except for small groups (<~1000) or where sample size is more than about 10% of the group size. Alternative methods are available for small populations or where sample size is large relative to group size.
- The number of animals required to be tested in the group depends much more on individual animal specificity than it does on sensitivity.
- The number of animals to be tested in the group is linearly and inversely related to the expected prevalence of infected animals in the group.
- As the required level of statistical confidence increases, so the required sample size increases. The usual level is 95%. If this is increased to 99%, there is an approximate increase of 50% in the required sample size. For a reduction from 95% to 90% confidence, there is a decrease in sample size by 25%.
- As the sample size increases, group level sensitivity increases.
- As the number of animals used to classify the group as positive is increased, there is a corresponding increase in specificity.

- As group level sensitivity increases, group level specificity decreases.
- When specificity = 100% at the individual animal level, all uninfected groups are correctly classified i.e. group level specificity also equals 100%.

2.6 Estimating test sensitivity and specificity

There are two broad approaches to estimating test sensitivities and specificities.

“Gold standard” methods rely on the classification of individuals using a reference test (or tests) with perfect sensitivity and/or specificity to identify groups of diseased and non-diseased individuals in which the test can be evaluated. In contrast, “Non-gold-standard” methods are used in situations where determination of the true infection status of each individual is not possible or economically feasible.

Regardless of the methods used for estimating sensitivity and specificity, a number of important principles must be considered when evaluating tests, as for any other epidemiological study (Greiner and Gardner, 2000):

- The study population from which the sample is drawn should be representative of the population in which the test is to be applied;
- The sample of individuals to which the test is applied must be selected in a manner to ensure that it is representative of the study population;
- The sample should include animals in all stages of the infection/disease process;
- The sample size must be sufficient to provide adequate precision (confidence limits) about the estimate; and
- Testing should be undertaken with blinding as to the true status of the individual and to other test results.

2.6.1 Gold-standard methods

Gold standard methods have the advantage of using a known disease status as the reference test. This allows for relatively simple calculations to estimate sensitivity and specificity of the test being evaluated, using a simple two-by-two cross-tabulation of the test against disease status. However, for many conditions a gold-standard test either does not exist or is prohibitively expensive to use (for example may require slaughter and detailed examination and testing of multiple tissues for a definitive result). In such cases the best available test is often used as if it were a gold standard, resulting in biased estimates of sensitivity and specificity. Alternatively, it may only be possible to use a small sample size due to financial limitations or the nature of the disease, resulting in imprecise estimates.

Gold standard test evaluation assumes comparison with the true disease status of an animal based on the results of a test (or tests) with perfect sensitivity and/or specificity

For example: The “gold-standard” test for bovine spongiform encephalopathy (BSE) is the demonstration of typical histological lesions in the brain of affected animals. However, false-negative results on histology will occur in animals in an early stage of infection. Therefore, if a screening test is evaluated by comparison with histology, specificity will be underestimated because some infected animals could react to the screening test but be histologically negative, resulting in mis-classification as false-positives. In addition, any infected but histologically-negative animals that are negative on the screening test will be mis-classified as true-negatives, resulting in over-estimation of the sensitivity.

If a disease is rare, or if the “gold standard” test is complex and expensive to perform, sample sizes for estimation of sensitivity are likely to be small, leading to imprecise estimates of sensitivity. If a disease does not occur in a country it is impossible to estimate sensitivity in a sample that is representative of the population in which it is to be applied. Conversely, if a disease does not occur in a country or region, it is relatively easy to estimate test specificity, based on a representative sample of animals from the population, because if the population is free of disease all animals in the population must also be disease-free.

Sometimes a new test may appear to be more sensitive (or specific) than the existing “gold standard” test (for example, new DNA-based tests compared to conventional culture). In this situation, the new test will find more (or fewer) positives than the reference test and careful analysis is required to determine whether this is because it is more sensitive or less specific. Even then, it is often not possible to reliably estimate sensitivity or specificity because there is no fixed reference point, so it may only be possible to say that the new test is more sensitive (or specific) than the old test, without specifying a value.

Gold-standard methods for estimating sensitivity and specificity of diagnostic tests and their limitations are discussed in more detail by Greiner and Gardner (2000).

2.6.1.1 *Estimating specificity in uninfected populations*

One special case of a gold standard comparison is for estimating test specificity in an uninfected population. In this case either historical information or other testing can be used to determine that a defined population is free of the disease of concern. This can be based on either a geographic region which is known to be free, or on intensive testing of a herd or herds over a period of time to provide a high level of confidence of freedom. If the population is assumed to be free, by definition all animals in the population are uninfected. Therefore, if a sample of animals from the population is tested with the new test, any positives are assumed to be false positives and the test specificity is estimated as the proportion of samples that test negatively.

For example, to evaluate the specificity of a new test for foot-and-mouth disease you could collect samples from an appropriate number of animals in a FMD-free country and use these as your reference panel.

Two drawbacks of this approach are: firstly that you cannot estimate sensitivity in this sample, since none of the animals are infected; and secondly that by using a defined (often geographically isolated) population there is a risk that specificity may be different in this population to what might be the case in the target population where the test is to be used.

2.6.2 Non-gold-standard methods

Non-gold-standard methods for test evaluation can often be used in situations where the traditional gold-standard approaches are not possible or feasible. These methods do not depend on determining the true infection status of each individual. Instead, they use statistical approaches to calculate the values of sensitivity and specificity that best fit the available data.

Non gold-standard test evaluation makes no explicit assumptions about the disease state of the animals tested and relies on statistical methods to determine the most likely values for test sensitivity and/or specificity

Although these methods don't rely on a gold standard for comparison, they do depend on a number of important assumptions. Violation of these assumptions could render the resulting estimates invalid. Non-gold-standard methods for estimating sensitivity and specificity of diagnostic tests have been described in more detail by Hui and Walter (1980), Staquet et al. (1981) and Enøe et al. (2000).

Available non-gold-standard methods include:

2.6.2.1 Maximum likelihood estimation

Maximum likelihood methods use standard statistical methods to estimate sensitivity and specificity of multiple tests from a comparison of the results of multiple tests applied to the same individuals in multiple populations with different prevalence levels (Hui and Walter, 1980; Enøe et al., 2000; Pouillot et al., 2002). Key assumptions for this approach are:

- The tests are independent, conditional on disease status (the sensitivity [specificity] of one test is the same, regardless of the result of the other test, as discussed in more detail in the section on series and parallel interpretation of tests);
- Test sensitivity and specificity are constant across populations;
- The tests are compared in two or more populations with different prevalence between populations; and
- There are at least as many populations as there are tests being evaluated.

2.6.2.2 Bayesian estimation

Bayesian methods have been developed that allow the estimation of sensitivity and specificity of one or two tests that are compared in single or multiple populations (Joseph et al., 1995; Enøe et al., 2000; Johnson et al., 2001; Branscum et al., 2005). These methods allow incorporation of any prior knowledge on the likely sensitivity and specificity of the test(s) and of disease prevalence as probability distributions, expressing any uncertainty about the assumed prior values. Methods are also available for evaluation of correlated tests, but these require inclusion of additional tests and/or populations to ensure that the Bayesian model works properly (Georgiadis et al., 2003).

Bayesian methods rely on the same assumptions as the maximum likelihood methods. In addition, Bayesian methods also assume that appropriate and reasonable distributions have been used for prior estimates for sensitivity and specificity of the tests being evaluated and prevalence in the population(s). For critical distributions where prior knowledge is lacking it may be appropriate to use an uninformative (uniform) prior distribution.

2.6.2.3 Comparison with a known reference test

Sensitivity and specificity can also be estimated by comparison with a reference test of known sensitivity and/or specificity (Staquet et al., 1981). These methods cover a variety of

circumstances, depending on whether sensitivity or specificity or both are known for the reference test. Key assumptions are conditional independence of tests, and that the sensitivity and/or specificity of the reference test is known.

In the special situation where the reference test is known to be close to 100% specific (for example culture or PCR-based tests), the sensitivity of the new test can be estimated in those animals that test positive to the reference test:

$Se(\text{new test}) = \text{Number positive to both tests} / \text{Total number positive to the comparison test}$

However, the specificity of the new test cannot be reliably estimated in this way, and will generally be under-estimated.

2.6.2.4 *Estimation from routine testing data*

Where a disease is rare, and truly infected animals can be eliminated from the data, it is possible to estimate test specificity from routine testing results, such as in a disease control program (Seiler, 1979). In this situation, test-positives are routinely subject to follow-up, so that truly infected animals are identified and removed from the population. It is also possible to identify and exclude tests from known infected herds or flocks. Specificity can then be estimated as:

$Sp = 1 - (\text{Number of reactors} / \text{Total number tested})$

In fact, this is an under-estimate of the true specificity, because there may be some unidentified but infected animals remaining in the data after exclusion of tests from known infected animals or herds/flocks.

For example: The flock-specificity of pooled faecal culture for the detection of ovine Johne's disease was estimated from laboratory testing records in New South Wales (Sergeant et al., 2002). In this analysis, there were nine test-positive flocks out of 227 flocks eligible for inclusion in the analysis. After exclusion of results for seven known infected flocks, there were 2/220 flocks positive, resulting in an estimated minimum flock-specificity of 99.1% (95% Binomial CI: 96.9% - 99.9%). In fact one or both of these flocks could have been infected, and the true flock-specificity could be higher than the estimate of 99.1%.

2.6.2.5 *Modelling approaches*

Several novel approaches using modelling have also been used to estimate test sensitivity and/or specificity without having to rely on a comparison with either a gold standard or an alternative, independent test.

2.6.2.6 *Mixture modelling*

One approach to estimating test sensitivity and specificity in the absence of a gold standard is that of mixture population modelling. This approach is based on the assumption that the observed distribution of test results (for a test with a continuous outcome reading such as an ELISA) is actually a mixture of two frequency distributions, one for infected individuals and one for uninfected individuals.

Using mixture population modelling methods, it is possible to determine the theoretical probability distributions for uninfected and infected sub-populations that best fit the observed data, and from these distributions to estimate sensitivity and specificity for any cut-point.

For example, this approach was used to estimate sensitivity and specificity for ELISA for *Toxoplasma gondii* infection in Dutch sheep (Opsteegh et al., 2010). ELISA results from 1,179 serum samples collected from sheep at slaughterhouses in the Netherlands were log transformed and normal distributions fitted to the infected and uninfected components. The resulting theoretical distributions allowed determination of a suitable cut-point with estimated sensitivity of 97.8% and specificity of 96.4%.

While this is a useful approach for estimating sensitivity and specificity in the absence of suitable comparative test data, it does depend on the assumptions that the test results follow the theoretical distributions calculated and that the sample tested is representative of the population at large. If the actual results deviate significantly from the theoretical distributions, or the sample is biased, estimates will also be biased.

2.6.2.7 *Simulation modelling of longitudinal testing results*

An alternative approach, using simulation modelling, has been used where no comparative test data was available, but results of repeated testing over time were available. In this example, the sensitivity of an ELISA for bovine Johne's disease was estimated from repeated herd-testing results over a 10-year period using a simulation model. Age-specific data from up to 7 annual tests in 542 dairy herds were used to estimate ELISA sensitivity at the first-

round test. The total number of infected animals present at the first test was estimated from the number of reactors detected at that test, plus the estimated number of animals that failed to react at that test, but reacted (or would have reacted if they had not died or been previously culled) at a subsequent test, based on reactor rates at subsequent tests. Reactor rates were adjusted for an assumed ELISA specificity of 99.8% to ensure estimates were not biased by imperfect ELISA specificity (Jubb et al., 2004). Age-specific estimates of ELISA sensitivity ranged from 1.2% in 2-year-old cattle to 30.8% in 10-year-old cattle, with an overall age-weighted average of 13.5%.

This approach depends on the assumption that most JD-infected animals become infected at a young age, and that all animals that subsequently reacted to the ELISA were in fact infected at the time of the first test. If adult infection occurred in these animals the estimated sensitivity could have substantially under-estimated the true value.

2.7 References – diagnostic testing

- Branscum, A. J., Gardner, I. A. & Johnson, W. O. 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modelling. *Preventive Veterinary Medicine*, 68:145-163.
- Cicchetti, D. V. & Feinstein, A. R. 1990. High agreement but low kappa: II. Resolving the paradoxes. *Journal-of-Clinical-Epidemiology*, 43:551-558.
- Enøe, C., Georgiadis, M. P. & Johnson, W. O. 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. *Preventive Veterinary Medicine*, 45:61-81.
- Everitt, R. S. 1989. *Statistical Methods for Medical Investigation*, New York, Oxford University Press.
- Feinstein, A. R. & Cicchetti, D. V. 1990. High agreement but low kappa: I. The problems of two paradoxes. [see comments.]. *Journal of Clinical Epidemiology*., 43:43-9.
- Georgiadis, M. P., Johnson, W. O., Singh, R. & Gardner, I. A. 2003. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. *Applied statistics*, 52:63-76.
- Greiner, M. & Gardner, I. A. 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. *Preventive Veterinary Medicine*, 45:3-22.
- Hui, S. L. & Walter, S. D. 1980. Estimating the error rates of diagnostic tests. *Biometrics*, 36:167-171.
- Johnson, W. O., Gastwirth, J. L. & Pearson, L. M. 2001. Screening without a "gold standard": the Hui-Walter paradigm revisited. *American Journal of Epidemiology*, 153:921-924.
- Joseph, L., Gyorkos, T. W. & Coupal, L. 1995. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *American-Journal-of-Epidemiology*, 141:263-272.

- Jubb, T. F., Sergeant, E. S. G., Callinan, A. P. L. & Galvin, J. W. 2004. Estimate of sensitivity of an ELISA used to detect Johne's disease in Victorian dairy herds. *Australian Veterinary Journal*, 82:569-573.
- Martin, S. W., Shoukri, M. & Thorburn, M. A. 1992. Evaluating the health status of herds based on tests applied to individuals. *Preventive-Veterinary-Medicine*, 14:33-43.
- Opsteegh, M., Teunis, P., Mensink, M., Zuchner, L., Titilincu, A., Langelaar, M. & van der Giessen, J. 2010. Evaluation of ELISA test characteristics and estimation of *Toxoplasma gondii* seroprevalence in Dutch sheep using mixture models. *Preventive Veterinary Medicine*.
- Pouillot, R., Gerbier, G. & Gardner, I. A. 2002. "TAGS", a program for the evaluation of test accuracy in the absence of a gold standard. *Preventive Veterinary Medicine*, 53:67-81.
- Rogan, W. J. & Gladen, B. 1978. Estimating prevalence from the results of a screening test. *American Journal of Epidemiology*, 107:71-76.
- Seiler, R. J. 1979. The non-diseased reactor: considerations on the interpretation of screening test results. *Veterinary-Record*, 105:226-228.
- Sergeant, E. S. G., Whittington, R. J. & More, S. J. 2002. Sensitivity and specificity of pooled faecal culture and serology as flock-screening tests for detection of ovine paratuberculosis in Australia. *Preventive-Veterinary-Medicine*, 52:199-211.
- Staquet, M., Rozenzweig, M., Lee, Y. J. & Muggia, F. M. 1981. Methodology for the assessment of new dichotomous diagnostic tests. *Journal-of-Chronic-Diseases*, 34:599-610.

3 Disease control and eradication programs

3.1 What is meant by control or eradication of disease

Control: is the reduction of the morbidity and mortality from a disease by

- Treating diseased animals to reduce the prevalence
- Preventing disease to reduce incidence and prevalence

Eradication: is the regional extinction of an infectious agent

Many animal diseases are *endemic* in a population and are either not sufficiently serious to warrant control, or are amenable to well recognised treatment, control or preventive measures implemented at the farm or individual level.

Most countries will have a list of animal diseases that must be reported to relevant authorities as soon as diagnosis is suspected or confirmed. In Australia these are called *notifiable animal diseases*. Diseases are generally included on a notifiable diseases list because of their potential to cause adverse impacts on animal health and production, international trade, biodiversity and human or ecosystem health.

There is little reason to invest effort in maintaining a list of notifiable diseases and the necessary infrastructure to detect and report diseases unless there is also a commitment towards controlling or eradicating these diseases.

Control or eradication programs may be applied to endemic diseases or as a planned response activity that is implemented only if specific exotic diseases occur in a particular country or part of a country that is normally free of that disease.

For example:

- Milk fever and grass tetany are affected by seasonal and management factors and are generally managed at farm and individual animal levels.
- Clostridial diseases of sheep and cattle are widespread and generally controlled by on-farm vaccination programs.
- Internal and external parasites in sheep and cattle are generally managed at the farm level, but can be very costly on an industry basis (10's – 100's of millions of dollars per year) and on-farm control may be supported by regional programs providing technical advice and support.

- Some diseases, such as ovine brucellosis or caprine arthritis-encephalitis virus, affect only some herds or flocks, and can be managed at a regional or industry level through voluntary quality assurance (QA) type programs.
- Control of Johne's disease in many countries is moving towards voluntary, industry-based programs.
- Zoonotic disease such as anthrax, rabies, bovine spongiform encephalopathy and highly pathogenic avian influenza are subject to strict regulatory programs in many countries.
- Brucellosis and TB in cattle have been eradicated from Australia and are subject to national eradication programs in some countries.
- Exotic disease outbreaks in some countries are usually subject to emergency eradication programs (for example foot-and-mouth disease) in countries where the disease doesn't usually occur.
- Global freedom from rinderpest was declared in 2011, following a lengthy eradication program.

3.2 Why have a regional control or eradication program?

Regional disease control or eradication programs have been an important facet of livestock production since at least the 18th century. Early programs were directed at eradication of outbreaks of severe diseases such as rinderpest and foot-and-mouth disease from Europe and the UK in the 18th and 19th centuries. Also in the 19th century, Australia eradicated sheep scab from its national flock, while in the mid-late 20th century contagious bovine pleuropneumonia, brucellosis and tuberculosis were also eradicated from the Australian cattle population. More recently, in 2011 international freedom from rinderpest was proclaimed after a protracted eradication campaign. This is only the second time global eradication of a disease has been achieved (following the eradication of smallpox in the mid 20th century) and the first time for an animal disease.

Regional or national programs may be implemented for any of the following reasons:

- To control or eradicate diseases with severe productivity and economic consequences including trade in animals and animal products (e.g. foot-and-mouth disease);
- To protect human health from zoonotic infections (e.g. bovine spongiform encephalopathy, bovine tuberculosis, anthrax);
- To maintain product quality (e.g. chemical residues);
- To protect unaffected producers or regions from disease that may be endemic in other regions (e.g. footrot, ovine brucellosis, Johne's disease, cattle tick);

- To reduce indirect effects of disease on unaffected producers who are not in a position to take action themselves to effectively prevent or control the impacts of the problem on their enterprise (e.g. chemical residues, Johne’s disease); and
- To reduce the impact of disease on affected herds and flocks (e.g. mastitis, internal parasites in sheep).

An important aspect of diseases requiring regional or group action to control or eradicate is that they are often diseases where producers can take individual action if they wish, but where the risk of re-infection or break-down of control because of external factors is sufficiently high to discourage individual action.

For example, many sheep producers in the Australian State of New South Wales were reluctant to attempt eradication of footrot in sheep until a regional program started and provided some reassurance that they were not likely to get reinfected.

Table 3.1 lists the characteristics of conditions that determine whether a disease is more suited to individual farm or regional control.

Table 3.1: Characteristics of conditions suited to farm-level or regional control (adapted from Hanson & Hanson, 1983)

Farm-level control	Regional control
Spread can be stopped by a physical barrier such as a fence	Physical barriers of limited effectiveness in preventing spread
Rate of transmission is slow enough to allow intervention before the entire herd is infected	Transmission is too fast for intervention before the entire (or majority of) herd is infected
Carriers are readily detectable on farm	Apparently healthy carriers can only be detected by laboratory tests
No public health, food safety or product quality implications	Condition is a public health, food safety or product quality risk
Low or no mortality rate	High morbidity and high mortality rates
Highly effective vaccine or treatment is available	Vaccine or treatment is only poorly to moderately effective

3.3 Types of programs

Regional animal health programs can vary substantially in their design, the tools used and the way they are implemented, depending on the rationale and objectives of the individual program. Programs can be broadly classified according to their objectives as either eradication or control programs.

3.3.1 Eradication programs

Eradication programs are generally directed at the elimination of a disease agent from a region. This is usually achieved by the implementation of measures directed at reducing prevalence on infected farms and interrupting spread from infected to uninfected farms.

Eradication programs often require a strong regulatory framework, with significant government input to the management and implementation of the program. Funding for eradication programs may be largely from governments, or shared by governments and affected industries, depending on the nature of the disease and the capacity and willingness of governments (and industry) to contribute.

Eradication programs are generally time-limited and aim to eliminate the disease within a relatively short or manageable period. Once disease is eliminated there is assumed to be no ongoing cost associated with eradication but there may be substantive ongoing costs associated with surveillance programs to prevent, detect and respond effectively to any future incursion of the disease into the area where it has been eradicated.

In cases of endemic diseases, eradication may be preceded by a period of control to reduce the prevalence of disease to a level where eradication becomes feasible and economic.

3.3.2 Control programs

As defined elsewhere, control implies any program directed at reducing the level of morbidity, mortality or production losses due to a disease. Control can be achieved by:

- treating diseased animals; and/or
- preventing infection occurring; and/or
- reducing the impact of disease in infected animals

Control programs are expected to have ongoing costs associated with disease detection and control, while-ever the disease or the reasons for its control persist.

3.3.2.1 Regulatory programs

Some control programs may be supported by government regulation to allow enforcement of compliance. Regulations may relate to movement controls, animal treatments, destruction of animals and compensation. Regulatory programs are more common for diseases that have a “public good” component, such as zoonotic diseases. Over time, if a control program is successful it can be extended and adapted into an eradication program. Examples of diseases where regulatory control programs are used include anthrax, rabies and bovine spongiform encephalopathy.

3.3.2.2 Voluntary (Industry-based) programs

Governments in many countries are moving to reduce regulation of the livestock industries, and this move is often accompanied by a move towards voluntary or industry-based control programs. These programs rely on farmers complying voluntarily with recommended practices to reduce disease risk to themselves and other producers, rather than using regulations to enforce compliance. Voluntary programs depend heavily on an effective communication and education program to change the behaviour and attitudes of farmer and their advisors and to get farmers to adopt the recommended practices.

Voluntary programs may have some regulatory support (for example legislative support for the use of vendor declarations or movement controls), but are being used increasingly as an alternative to regulatory programs, particularly where most of the benefits of the program flow to producers, rather than consumers or the general public. Examples of voluntary programs include the early stages of enzootic bovine leucosis eradication in dairy cattle in Australia and Johne’s disease control programs in many countries.

3.3.2.3 Assurance-based programs

Assurance-based programs rely on on-farm implementation of a quality assurance approach to management and production on some farms to provide a source of quality-assured stock for other producers. Quality assurance programs require participating farmers to implement a range of recommended practices to achieve a quality outcome and are supported by an audit process to ensure compliance and demonstrate program integrity. Stock from qualifying farms may be assured as low-risk for a particular disease or for chemical residues, depending on the program(s) in which they participate and the level they have achieved.

Although assurance-based programs may not significantly reduce the regional prevalence or impact of the disease or condition of concern, they can reduce further spread by providing sources of low-risk stock for producers who wish to avoid introducing unwanted diseases to their farm. They can also be used as part of a broader regulatory or voluntary control program. Examples of assurance-based programs include the various Johne's disease Market Assurance Programs in Australia and similar programs in other countries, as well as industry-based product quality programs.

3.4 Strategies that may be used for disease control

Maintenance of infectious disease in an animal population depends on presence of infectious individuals and herds, presence of susceptible individuals and herds, and contact between infectious and susceptible individuals and herds. Disease will persist in the population while ever these conditions remain.

The main strategies for control and eradication of animal diseases (Thrusfield 2005) include:

- **Quarantine:** Isolation of diseased animals (or animals suspected of being diseased) so the risk of spread to other susceptible animals is reduced. Often accompanied by other biosecurity measures relating to movement controls, hygiene and disinfection.
- **Slaughter of diseased animals:** May be accompanied by slaughter of high-risk contact animals in emergency disease control (eg Foot-and-Mouth disease outbreaks in some countries) and disposal of carcasses and other infectious material.
- **Vaccination:** May be to reduce spread of disease during an outbreak or as part of longer term eradication programs to reduce circulating infection.
- **Treatment:** Administration of drugs (antibiotics or anthelmintics) may be used as part of a control program or to reduce risk of diseases from occurring.
- **Control of animal movements:** Often part of quarantine measures to prevent disease spread. May also be used more routinely eg controlled grazing for management of internal parasites or movement of animals out of high risk areas at certain times of the year to avoid vector borne diseases or bringing animals indoors at night in Africa to minimise risk of exposure to African horse sickness virus carried by night flying midges.
- **Vector and reservoir control:** Infectious diseases may be transmitted by insect vectors or different reservoir hosts (Nipah virus). Control of the vectors or reservoir hosts will help in disease control.

- **Biosecurity measures:** Measures include hygiene, disinfection, and other management measures that may reduce disease spread. May be applied at animal, mob, farm or regional levels.
- **Genetic selection:** May be useful in control of some diseases by elimination of inherited diseases or selection of animals with increased resistance.

These strategies are generally applied through four separate activity pathways:

1. Detection of the infectious agent responsible for the disease;
2. Reduction in the number of infected hosts;
3. Increase in the resistance to infection of susceptible hosts; and
4. Reduction in contact between infectious and susceptible hosts.

3.4.1 Detecting the disease agent

3.4.1.1 Surveillance

The term *surveillance* describes an active process in which disease occurrence data is collected, analysed, evaluated, and reported to animal health agencies tasked with disease control. The term *monitoring* is usually used for a more passive process although in common usage both terms are often used interchangeably. Because of the substantial cost involved, programs often encompass several diseases at the one time.

An effective surveillance program will be able to answer a number of important questions relevant to disease control:

- Is the frequency of the disease remaining constant, increasing or decreasing?
- What is the relative frequency of one disease compared with another?
- Are there differences in the geographical pattern of the condition?
- Does the disease have any impact on productivity and/or profitability?
- Is the disease absent from a particular herd, region, or nation?
- Is a control or eradication program cost-effective?

The potential sources of data for surveillance programs include clinical evaluations, laboratory reports, slaughter inspection data, screening tests, owner reports, and on-farm screening programs.

Surveillance or monitoring programs may be developed at a number of different levels, depending on the level of need for the information. Some examples are listed below:

1. Individual farms - these usually include monitoring of economically significant production parameters, such as mortality rates, somatic cell counts in milk as an indicator of mastitis, growth rate, milk production, mortality rates, etc. Monitoring of temporal patterns of these variables is important for early detection of potential disease problems or failure of on-farm control programs.
2. Regional levels within a country (district, province, state etc) – including testing to detect infected animals or herds and to support disease freedom at a regional level.
3. National – National surveillance programs can be very costly. To help defray costs these programs may predominantly be based on passive surveillance (investigation initiated by the owner) or involve testing of only a sample of the national herd.

Surveillance to identify infected animals or infected herds/flocks is an essential component of any control or eradication program. For such programs, surveillance could be targeted at individual animals on-farm (for example, test-and-slaughter programs for brucellosis or bovine tuberculosis eradication), or could use aggregate samples, such as bulk-milk or pooled faeces, or could use off-farm sampling such as through milk factories or abattoirs (for example, milk-ring testing to identify brucellosis-infected herds).

Farmer notification of suspected cases also forms an important component of surveillance for case-detection. For surveillance to be effective, an economically justifiable test with known sensitivity and specificity should be used. Once an infected animal or farm has been identified, further action is likely using one or more of the other tools discussed below.

3.4.1.2 Tracing

Tracing of livestock movements is an important tool particularly for the detection of infected herds or flocks. For disease control purposes, tracing usually involves the identification of potentially infected farms through the tracing of movements of infected or exposed animals. Further testing is usually undertaken on the identified farms to establish their true infection status. If a farm's infection status cannot be determined immediately, quarantine measures may be imposed until the situation is resolved.

Tracing can involve any of the following activities:

- Identification of the property of origin of animals identified as infected or suspect through testing at abattoirs or saleyards (abattoir/saleyards traceback);

- Identification of the property of origin of animals suspected as a potential source of infection on an infected farm (trace-back);
- Identification of farms that have received possibly exposed animals from an infected farm (trace-forward)
- Identification of farms with animals potentially exposed during movement of infected animals, such as at saleyards or during transport;
- Identification of neighbouring farms or other farms potentially exposed to an infected farm by local movement of animals or infectious material; and
- Identification of vehicles used to transport potentially infected animals or vehicles, people or other fomites that have had possible contact with infected animals or environments.

Tracing activities are made much easier and more reliable by the consistent use of unique animal identification and national animal identification systems that are capable of tracking animal movements over time. An example of this sort of system is the National Livestock Identification System (NLIS) in Australia.

In the absence of a comprehensive database of animal movements, tracing relies on interviews with the owners of infected or exposed animals to identify potential animal or other movements that might have spread infection. Investigations may also include discussion and examination of records from livestock agents, stock selling centres, milk processors and abattoirs.

Effective tracing can also consume large volumes of resources for both the identification of movements to/from infected farms, and also the subsequent identification and investigation of the source or destination properties. However, examination of tracing records can often help understand the epidemiology and distribution of a disease during an outbreak.

3.4.2 Reducing the number of infected hosts

3.4.2.1 *Slaughter*

Slaughter of individual infected animals, in-contact animals or entire herds may be an option, depending on the nature of the disease and the program involved. Slaughter of infected animals and herds has an immediate effect of reducing the number of infected animals in the population and greatly reduces opportunities for further spread of the disease.

However, this comes at a significant cost in terms of surveillance to detect the infected animals and the costs of compensation and disposal if the animals are not salvaged through normal slaughtering.

Depending on the type and scale of the program, slaughtering of stock can be undertaken in a number of ways:

1. Immediate destruction of infected and in-contact animals generally in emergency situations such as response to an exotic disease outbreak (for example, foot-and-mouth eradication programs, bovine spongiform encephalopathy).
2. Test-and-slaughter programs have often been used in the past for eradication of specific diseases (bovine brucellosis and tuberculosis in Australia). In this case animals are tested and only those that are deemed to be disease positive are then slaughtered.
3. Herd depopulation may be used in extreme situations or for problem herds where eradication using other methods has failed (for example, foot-and-mouth disease, bovine spongiform encephalopathy in the UK; problem herds for bovine tuberculosis and brucellosis late in the Australian brucellosis and tuberculosis eradication programs)
4. Slaughter or early culling of individual animals may also be used in non-emergency situations as part of a voluntary or regulatory control program for some diseases (for example footrot or ovine brucellosis in sheep, chronic mastitis in dairy cows).

3.4.2.2 *Animal treatments*

Where available, treatments (either therapeutic or preventive) can be used to treat infected or exposed animals and reduce prevalence. For example, antibiotic preparations can be used to treat mastitis cases and teat disinfection preparations can be used to prevent new infections occurring.

3.4.3 Increasing resistance of susceptible hosts

3.4.3.1 *Vaccination*

Vaccination is an important tool for the control and eradication of many diseases, and can be used in two main ways:

- Routine animal management where commercially available vaccines may be used as part of routine on-farm disease control for diseases such as clostridial diseases, leptospirosis, vibriosis, Marek's disease, etc.
- Prevalence reduction where specific vaccines may be used either on individual farms or at a regional level to reduce the prevalence of disease as part of a regional control or eradication program, by increasing the level of *herd immunity*. This can be used solely for control purposes, or as a prelude to eradication, with eradication attempts

only proceeding subject to reducing prevalence of infection to an acceptable target level. For example, a key aspect of the brucellosis eradication program in Australia was the use of Strain 19 vaccine to reduce prevalence in high-prevalence regions before eradication commenced.

Progress of a disease in a population is affected by herd-immunity effects. Herd immunity effects appear when a meaningful proportion of the population is immune to a disease either from innate immunity (although this may not always have an immunological basis), natural infection or vaccination.

Herd immunity will slow the rate of transmission of a disease within a population, with the magnitude of the effect depending on the level of herd immunity. If herd immunity is high, infection may fail to establish or can be eliminated from the population. It is not necessary for all individuals in a group to be immune to eliminate infection. The level of herd immunity (proportion of immune animals in the population) must simply be sustained at a level which exceeds a critical threshold value at which the contact rate between infectious and susceptible individuals is insufficient to sustain the epidemic. This means that if a minimum critical proportion of animals can be kept immune to infection, a disease can be eliminated from the population. For many infectious diseases, effective vaccination rates of 70-80% provide sufficient herd immunity to prevent an epidemic being sustained.

3.4.3.2 Genetic manipulation

Many diseases have some level of genetic resistance or susceptibility. For these diseases it may be possible to breed for resistance to infection (for example internal parasites in sheep). However, any such breeding program is likely to be long-term, and must consider competing priorities for selection on production traits.

3.4.4 Reducing contact between infectious and susceptible hosts

3.4.4.1 Quarantine

Quarantine is the physical isolation of infected or potentially infected animals to prevent further spread of infection. Quarantine can be applied to farms that are known or suspected to be infected to prevent spread of infection to other farms. It can also be applied within farms to prevent spread between infected and uninfected groups of animals, or to isolate introduced animals until the farmer can be confident that they are disease free. Occasionally groups of

farms may also be quarantined, particularly if they are potentially exposed to a highly infectious disease.

3.4.4.2 Movement controls

In a similar way to quarantine of infected farms, regional or inter-property movement controls can be used to reduce the risk of spread of infection from areas of high prevalence to areas of lower prevalence. These movement controls can be supported by official disease “zones” and regulatory requirements for movements between zones, or by a less regulated approach and voluntary implementation of recommended movement controls to minimise disease spread by farmers.

If a regulatory program is implemented it is appropriate not only to have regulatory support for movement controls (including quarantine), but also the willingness and resources to enforce the regulations. Under such programs it may be necessary to have regulatory staff available to maintain movement check points, check movement documentation, carry out saleyard inspections and enforce other regulations, as appropriate. However, regulation does not necessarily mean that the program will be complied with. In fact, a voluntary program with effective education and ownership of the program by farmers may be more effective than an unpopular regulatory approach.

In a less-regulated or voluntary program, it is still important to know the level of farmer compliance with recommended control measures. Therefore, even in completely voluntary programs it is essential to monitor or audit compliance rates against targets on a regular basis.

If farmer compliance is poor, the program is unlikely to succeed and progress and future options should be urgently reviewed.

3.4.4.3 Vector control

For *vector*-borne diseases, control measures may be more easily directed at the vector than at the actual disease agent. For example, effective control of tick fever in cattle in many parts of Australia is achieved mainly by controlling its cattle-tick vector. Similarly, effective long-term control of liver fluke in sheep and cattle can be achieved by either eliminating the snail vector or restricting access of stock to the snail’s habitat area. Vector control also should include consideration of mechanical vectors such as syringes/needles, which can be important

vectors for some diseases such as enzootic bovine leucosis or caprine arthritis-encephalitis virus.

3.4.4.4 *Management measures*

Grazing or animal management

For some diseases, grazing management strategies can be used to reduce exposure of susceptible animals to contamination. For example, many internal parasite control programs are based on grazing susceptible young animals on pastures that have previously been grazed by low-risk older animals. Similar strategies have been tried for control of Johne's disease, although low-risk animals may be difficult to identify, and may be a different group to animals that are low-risk for parasites.

Many diseases are also affected by factors under the control of the farm manager, such as housing, nutrition, stocking rates, feeding practices, etc. For these diseases, effective control can often be achieved by changing management practices or housing to reduce the transmission or impact of the disease. For example, inadequate ventilation is an important contributor to respiratory disease in pigs, so that severe respiratory disease problems can often be overcome by improving shed ventilation. Similarly, bovine Johne's disease transmission relies on ingestion of contaminated faecal material by susceptible calves, so that the incidence of Johne's disease in dairy cattle can be reduced by changing management to minimise the exposure of young calves to adult faecal contamination.

Biosecurity

Biosecurity measures complement other control measures and generally involve two quite separate components, bioexclusion, aimed at keeping diseases out and biocontainment, aimed at preventing onward transmission from infected herds or flocks.

Bioexclusion is the implementation of measures to prevent the introduction of unwanted pathogens into a livestock (or other) population.

Biocontainment is the implementation of measures to prevent the onward transmission of unwanted pathogens from a (potentially) infected livestock (or other) population.

Bioexclusion measures are focussed on disease-free farms, and are made up of a range of measures designed to keep disease out. These can include isolation of introduced stock, only sourcing introductions from farms with a specified level of testing or assurance, disinfection

of equipment and clothes/boots coming onto the farm, management of boundary fences and contact with neighbouring stock, vaccination, testing of introductions and any other measures designed to keep disease out or for early detection and response to disease introduction.

Conversely, biocontainment measures are aimed at control of disease on infected farms to reduce prevalence and other measures to reduce the likelihood of onward transmission. Although quarantine is one important biocontainment measure, biocontainment is broader than just quarantine and includes a range of other measures, including many of the same activities as for bioexclusion. Specific additional measures include vaccination, culling or treatment of affected animals, selling animals for slaughter only, testing animals prior to sale, disinfection of people and equipment leaving the farm, maintenance of boundary fences, etc.

3.4.4.5 Disinfection

For highly infectious diseases such as foot-and-mouth disease, disinfection of premises and potential fomites (including veterinary equipment) is an essential component of any control or eradication program. Disinfection can also be an important part of on-farm biosecurity programs to keep farms free of disease.

3.4.5 Supporting activities

3.4.5.1 Communication, education and training

Support of producers and the general public for the program and compliance of producers with program requirements are essential requirements for a program's success. Without an effective communication and education program, high levels of producer support and particularly of producer compliance are unlikely to be achieved. Program messages must be simple and consistent, and in many cases a substantial effort will be required to change the attitudes of farmers and their advisors to disease control and also their actions in managing disease risk. Education and training are also critical elements, to inform and educate producers and advisors about technical aspects of the disease and the program.

This is increasingly important with the shift from regulatory to voluntary programs, so that farmers are being asked to voluntarily change their practices to reduce disease risk, possibly at a significant short-term cost to themselves.

3.4.5.2 Risk assessment

Traditional disease control programs have relied on regulatory management of quarantine and movement controls to limit the spread of disease, with the underlying assumption that the measures imposed would be effective. Movement controls were generally based on a perceived “no-risk” approach to prevent spread of infection.

With the move towards more voluntary programs, and the recognition that there is no such thing as a “no-risk” policy, risk assessment has become an important aspect of any control or eradication program. A risk assessment approach makes a thorough understanding of the epidemiology of the disease much more important, so that the true risk associated with various options can be properly evaluated and communicated.

It is also important to note that in risk analysis terminology, “risk” includes elements of both likelihood of occurrence of an event and the expected consequences, should it occur. This is in contrast to the epidemiological definition of risk, which relates to likelihood of occurrence only.

The increasing move to a risk-based approach and voluntary control programs has been developing has coincided with an environment of decreasing government expenditure on disease control, placing increased reliance on the livestock industries to fund and manage programs with fewer government inputs.

3.4.5.3 Economic analysis

Just as a cost-benefit analysis is essential in determining whether or not a program is worthwhile in the first place, it is also essential that any program is subject to ongoing economic analyses. Such analyses should be directed at determining if the achievement of the program objectives is still economic, as well as determining which are the most economic and cost-effective of a range of potential control options.

3.4.5.4 Animal Identification

Identification of individual animals to their property of origin (and even their property of birth) is an essential component of an effective surveillance program for the detection of infected herds and flocks.

For example, abattoir inspection of adult sheep is an important part of surveillance for ovine Johne’s disease in Australia. Australia’s flock

identification system allows rapid tracing of the origin of sheep that are inspected and found to be either positive or negative, so that an inspection history can be built up for each flock and region over time, providing better levels of assurance for low-risk flocks and areas and allowing estimation and monitoring of flock-prevalence on an area basis.

Many countries now have mandatory cattle identification and passport systems in place to support traceability of animals and product in the wake of the bovine spongiform encephalopathy outbreak.

Identification of animals to the property of origin is important both at the abattoir, and for sales between properties, to support rapid tracing of animal movements in cases of emergency disease outbreaks, such as for foot-and-mouth disease or bovine spongiform encephalopathy or for chemical residue incidents.

Permanent individual identification of animals on farms is also an important and useful tool in any program that depends on animal testing or examination. Unique animal identification allows animals requiring further action (such as culling or treatment) to be easily identified for such action as may be required.

3.5 Pre-requisites for a successful program

Before embarking on a potentially difficult, costly and often controversial disease control or eradication program, it is essential to evaluate the proposed program in terms of its technical feasibility and likelihood of success.

The critical elements required for a successful disease control or eradication program are summarised below (adapted from Yekutieli, 1981 and Thrusfield, 2005). Although it may be possible to successfully control or eradicate a disease without meeting all of the criteria listed, the likelihood of failure increases as more criteria remain unfulfilled.

1. Adequate knowledge about the cause of the disease and its epidemiology

Knowledge of the cause (at least in epidemiological terms) and the epidemiology of a disease is essential for the development of effective strategies for the prevention of transmission and spread of the disease and for the application of screening tests to detect cases.

2. Adequate veterinary infrastructure and resources, including administrative and operational personnel

Adequate infrastructure and veterinary staff are essential for the effective implementation of a program. Inadequate staffing of the program is likely to result in failures in the application of

the selected control measures and significant delays in meeting program objectives. Important components of the infrastructure required for a successful program include:

- field veterinary staff;
- lay staff to assist with field activities;
- administrative staff to manage the program; and maintain databases and reporting capability;
- regulatory staff to implement and enforce legislative support measures;
- diagnostic facilities and staff; and
- research facilities and staff.

3. Accurate, reliable and economic diagnostic tests

Reliable and cost-effective tests that have been adequately characterised for sensitivity and specificity are essential for the identification of infected animals and herds or flocks, for appropriate follow-up action. Reliable tests are also required for herd/flock classification and identification of low-risk replacement stock. A good understanding of test sensitivity and specificity and factors that may affect these characteristics is also required for the development of appropriate testing and surveillance strategies.

4. Epidemiological features which facilitate case detection and effective surveillance

Diseases that are mainly sub-clinical or for which diagnostic tests have a poor sensitivity are likely to be difficult and expensive to detect, making the reliable identification of cases and implementation of control measures difficult. Diseases which can be detected through screening of routinely available samples or by simple testing at the herd/flock level (for example abattoir screening, bulk milk samples) are more suited to an effective program than diseases which require on-farm testing of large numbers of individual animals for the identification of infected individuals and/or herds/flocks.

5. Control measures that are simple to apply, relatively inexpensive and highly effective at preventing transmission of infection

Any control or eradication program depends on the implementation of one or more control measures to interrupt transmission and reduce prevalence. While it is possible to control and even eradicate diseases with imperfect tools (for example brucellosis, TB), the more effective the measures are, the more likely a program is to succeed. The less that is known about disease transmission and on-farm control measures, or the harder it is to control on-farm, the more difficult it will be to control the disease on a regional or national level. Measures must

also be effective at preventing spread between farms, as well as at reducing or preventing transmission on infected farms.

Formal risk assessment should be completed of the risk (probability of disease event and consequence) of disease occurrence and spread under various scenarios (no control and under each control or eradication option).

6. A reliable source of sufficient numbers and quality of disease-free replacement stock for those destroyed or culled during the campaign

Any program requiring slaughter or compulsory culling of infected stock is heavily dependent on a source of disease-free replacements. If disease prevalence is high, this becomes more difficult. Also, if available tests have a poor sensitivity it may be difficult to reliably identify low-risk animals or populations as a source of replacements.

7. Support for the program amongst producers and the general public, and cooperation by producers with the requirements of the program

If there is not a high level of commitment to the program among producers it is likely to be affected by criticism, unrest and even active resistance, hampering implementation and potentially undermining the effectiveness of the program. This is even more important for voluntary programs, where farmer education, support and compliance are critical for program success.

8. Appropriate justification for eradication or control, supported by independent cost-benefit analysis

Without a clear and well-argued rationale for eradication or control, any program is likely to lack the support of producers, industry leaders and governments. The most common reasons for eradication or control have been discussed previously, but include public health effects or the cost of the disease to the industry or community. If eradication is proposed, there also must be a valid reason for recommending eradication rather than control.

For a program to be supported, a social cost-benefit analysis will generally be required, demonstrating that the program is economically justifiable and that the expected returns (in terms of savings in cost of disease or productivity losses) exceed the cost of the program over the longer term.

9. Supporting legislation to enable the program to proceed, including provision for compensation

Appropriate legislation is required to implement movement controls, compulsory slaughter, compensation and other measures included in regulatory type programs. However, even in

voluntary programs, some level of legislative backing may be required to provide a legal basis for area declarations and movement restrictions and for enforcement of program requirements.

10. The ecological consequences of the program must be assessed and addressed

There is increasing public concern over environmental and ecological issues, such that they must now be an important consideration in any animal health program. If the proposed program is likely to have adverse environmental or ecological effects it is unlikely to be supported by governments or the general public. However, programs that have a positive impact on the environment (for example by reducing the feral animal population) are likely to be well-supported.

11. Adequate funding committed to the program

Without adequate funding, any animal health program is doomed to failure. In the current economic climate, governments are reluctant to commit large amounts of public money unless there is a positive return on their investment and an obvious public benefit from the program. Where the livestock industries are the major beneficiaries of disease control, they are also expected to be the major funders in some countries. A requirement for industry contribution also raises the issue of how to collect money from producers at a State or regional level, usually through some form of levy at sale or slaughter.

3.6 Application of control measures based on infection status

One way to think about control and eradication measures is to consider measures that may be applied based on the infection status of an individual farm or village (Toma et al 1999).

3.6.1 Control measures applied to an infected premise

In Australia the most common unit at which these measures would be applied is the farm. A farm is generally one enterprise at a single location though it may cover a large area and have large numbers of animals of multiple species. In other areas of the world it may be more logical to think of a village or some other unit. A unit is likely to be a relatively small area where animals can co-mingle during feeding or management.

For contagious diseases the critical first step in control is to implement quarantine measures usually accompanied by restriction of movement of animals into and out of the premise. In many cases all movement of animals and animal products and other related material (animal

feeds, equipment, etc) may be stopped and even movement of people onto and off infected premises may be carefully controlled.

A second critical step for contagious diseases is generally to slaughter all susceptible animals on the infected premise and dispose of them in a way that eliminates infectious risk (burial or burning) along with any other infectious material such as bedding or other material. For diseases that are not highly contagious, it may be appropriate to test animals and only slaughter those that are known to be infected. Alternatively animals may be able to be sent for processing but not transferred to any other property.

These measures are accompanied by disinfection of the premise to minimise the probability of infectious agent surviving in the environment and application of biosecurity measures to reduce the risk of disease inadvertently being carried off the premise. This may include disinfection and changes of clothes and vehicles at entry and exit points to/from the property, control of movement of people and other things into and off the property.

Trace-forward and trace-back should be used to identify any high risk contacts of movements onto the property in the period before the disease was detected and off the property. Other properties identified through these procedures should be visited and examined to look for any evidence of infection. Knowing the incubation period of the disease, the likely date when the premise was infected and dates of movements onto or off the farm can all be used to identify windows of interest for tracing. If the date when the premise was likely to have been infected can be identified then all movements off the farm from that time up until quarantine is imposed should be checked.

Once infection has been eliminated from the premise, there is then usually a period when the premise may be left without any livestock. This is to help ensure that once animals are permitted back onto the farm (restocking) that they do not get infected from infectious agent that may have survived for some time in the environment. The period of time will depend on the longevity of the infectious agent in the environment.

If eradication is an aim of the response policy, then it will be necessary to demonstrate that eradication has been achieved before the area or country can declare itself free of the disease. Proof of freedom from disease is an important function of surveillance and will be discussed in more detail in a separate manual.

3.6.2 Control measures that may be applied to a disease free premise

A disease free farm can be considered like a remote and isolated castle with a moat around it. The principle is that measures have to be adopted that prevent disease from getting in.

Animals and any product capable of carrying infection should only be allowed to enter the premise if they can be confidently declared to be disease free. This may be difficult to achieve. It may be safest to not allow any animals to enter. If animals must enter then they should be subjected to sufficient testing and examination to be confident that they are not infectious.

Strict biosecurity measures should be imposed including restriction on visitors and disinfection and other measures (change of clothes, vehicle etc) at controlled entry and exit points.

If the disease is able to be spread through the air or via water (streams or overland flow), or wild animals, insects, birds, etc then each of these risks will need to be assessed and measures applied to mitigate risks. It may be useful to move animals away from the boundary of the farm and develop a destocked barrier region along the boundary. Additional control measures for rodents or other wild animals may be appropriate.

In many cases and particularly where the prevalence of infection around a premise may be high, even if all applicable control measures are implemented, the farm may become infected.

3.7 Example: Rabies in Bali

When the current rabies outbreak in Bali began in 2008, the island had no policies for rabies post-exposure prophylaxis (PEP), no dog bite surveillance, no rabies diagnostic facilities and no vaccination program for dogs. In subsequent years the Indonesian government provided PEP for humans and vaccines for dogs, diagnostic testing was established at the Disease Investigation Centre in Denpasar and surveillance was implemented of dogs that died or were killed either as part of culling programs or that were showing neurological signs. Culling of unconfined dogs was instituted in some areas but was found to be counter-productive because people reacted strongly by hiding or moving dogs to avoid culls and replacing dogs that had been culled.

While most dogs are owned, many are unconfined and this combined with the need for frequent booster vaccination using the locally produced vaccine meant that there were serious

problems in achieving effective vaccine coverage. Long-lasting vaccines were supplied from about 2009 and by mid 2011 about 250,000 dogs had been vaccinated (coverage >70%). A second island wide vaccination program was completed by late 2011.

Mass vaccination has been shown to be effective at reducing rabies incidence in dogs, reducing human exposures and reducing cases of rabies in people. Mass vaccination must be continued along with effective surveillance and control of inter-island movement of dogs if rabies control is to be maintained and spread of the disease to other islands is to be prevented.

A recent publication (Townsend et al 2013) applied the use of mathematical modelling to the question of whether mass vaccination might be able to successfully eradicate rabies from the island of Bali.

The authors concluded that a single mass vaccination program that achieved 60% or lower coverage of the dog population had no chance of successful eradication while two campaigns of 80% coverage or three campaigns of 60% coverage was predicted to achieve eradication in 90% of model runs.

A related publication (Agung Gde Putra et al 2013) indicated that there has now been two mass vaccination programs that have achieved ~70% coverage rates of dogs but that rabies continues to be present and dog bites of people (exposures) continue to be high (more than 4,000 per month). Townsend et al (2013) suggested that although the overall coverage in the first mass vaccination program was as high as 70% the long time taken to complete the program meant that the average island wide coverage was probably closer to 40% due to ongoing animal turnover and waning immunity.

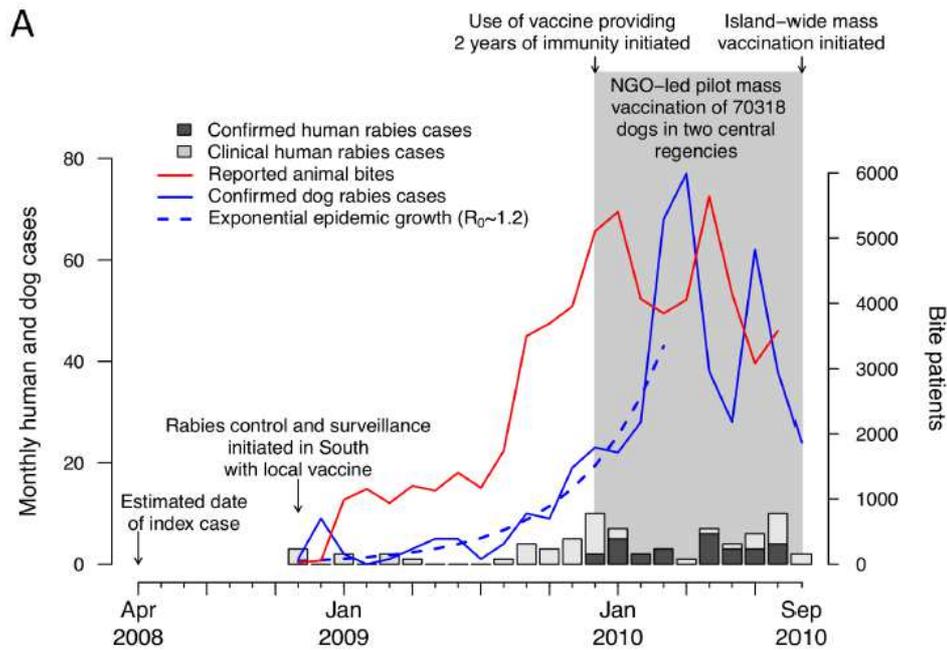


Figure 3.1: Rabies incidence on Bali prior to island wide mass vaccination. From Townsend et al (2013)

Townsend et al (2013) suggest that eradication is achievable with a third and possibly further mass vaccination campaigns, but that ongoing control and surveillance will be critical to prevent re-introduction of rabies in the future even if eradication is achieved. Eradication would save about 55 human deaths per year while requiring ongoing surveillance and control.

Continued control without eradication was predicted to save about 44 human deaths per year (some deaths would continue) and would require ongoing expenditure on mass vaccination programs and human PEP.

3.8 Designing an appropriate animal health program

The challenge for designing an animal health program is to bring together the most cost-effective mix of tools to achieve the desired goal.

Key issues in planning and designing an appropriate regional animal health program include:

- What is the current situation (how common is the disease, what inputs and tools are available, etc)?
- What is the desired situation?
- Is a regional program the right approach?
- Is a regional program feasible and likely to be successful?
- Is the proposed program likely to be a voluntary or regulatory type of program?
- What control tools are available for use in the program that are likely to be effective for the disease of concern?
- What level of resourcing is available for implementing the program?
- Is the proposed program feasible and likely to be successful?
- Who are the main beneficiaries of the program?
- How will the program be funded?
- How will the program be managed?

In most cases, any program will be made up of a one or more of the various strategies discussed above. Once the appropriate programs and strategies have been identified, and the ways in which they will be applied have been determined, detailed business and operational plans for the program should be developed.

The Program plan describes the overall management and operations of the program and should:

- Define the overarching goals or aims of the program;
- Identify specific objectives against which progress can be measured and reported;
- Provide a detailed description of how the program will be managed
- Define roles and responsibilities for participating organisations and key personnel;
- 3333Include a detailed budget and funding sources for the program;
- Identify supporting legislation and regulatory powers required or available to support the program

- Identify the resources required for implementation and where these resources will come from;
- Define timelines, targets and monitoring processes to evaluate progress of the program; and
- Provide decision points and criteria for key decisions as to whether to continue, modify or abandon the program.

In some cases a program plan may be split into a *business plan* that covers broad goals, management, responsibilities and funding; and a separate *operational plan* (often reviewed annually) which provides the specific details of targets, resourcing and day-to-day operational activities of the program.

3.9 Monitoring program performance

The success of animal health programs is highly variable, depending mainly on the factors outlined previously. However, if program performance is not monitored and regularly reviewed, stakeholders will not know whether it is succeeding or not. Therefore, ongoing monitoring of program performance and review of achievements against targets and objectives is essential for any animal health program.

It is also important that performance is monitored against both financial and animal health objectives. A program can be operating very efficiently on a financial basis, and remain well within budget, but fail to achieve any of its animal health objectives, and vice versa, either of which represents significant failure of the program.

As part of the planning process, milestones should be set, at which progress can be reviewed against targets. Failure to meet targets at a review point should trigger a response to identify why targets are not being met, and to implement measures to correct any deficiencies. In some cases the program business or operational plans and budgets may need review and refinement, or in severe cases a major overhaul of the program may be required.

3.10 Economics of animal disease control

Animal disease has a range of potentially adverse effects that can be presented in economic terms. Economics uses monetary units (dollars) to inform rational decision making about allocation of scarce resources between competing options. Economics is generally focused on the use of resources (inputs) that in turn produce goods (outputs). Outputs then produce some form of human benefit through a market (products used or purchased by consumers).

Economic analyses can be complex and difficult to understand, reflecting the complexity of animal production systems and the difficulty in describing and valuing the potential impacts at both farm and national levels.

3.10.1 Data requirements for economic evaluation

In order to perform economic analyses to compare options for disease control or eradication, it is often necessary to collect an extensive set of data and information including (Rushton et al 2012):

- The livestock production system or systems (if there are multiple separate systems) in sufficient detail to be able to describe production with and without disease. This will usually require some form of model to simulate the production system (number of females bred, proportion that get pregnant, number of calves born, annual loss rates, growth, turn off etc).
- The occurrence of disease and the effects of disease on the livestock production system (mortality, morbidity, production effects) and on factors outside the production system.
- Possible control measures including their effects on disease occurrence, livestock production and on market prices.
- Details of the costs associated with implementing different control options.

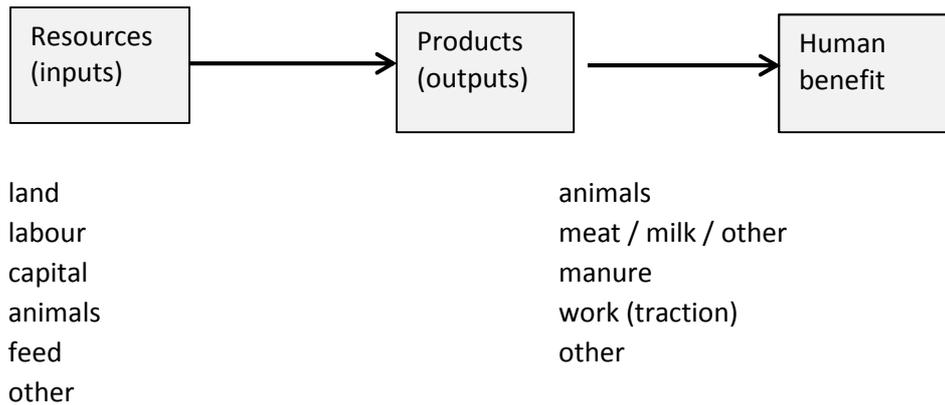


Figure 3.2: Livestock production system pathway showing effects of disease (from Otte and Chilonda, 2001)

Animal disease has the potential to produce adverse effects at each step along the livestock production pathway. Disease effects can be group as direct or indirect losses.

- Direct losses include:
 - Mortality of breeding or production animals
 - Reduced production efficiency eg reduction in feed conversion, fertility rate, growth rate etc. May be presented as a higher rate of inputs to sustain required outputs.
 - Reduction in product quantity (fewer offspring, less milk, less eggs, less meat or fleece etc), or reduction in product quality (poor hides because of tick damage, discarded milk because of mastitis etc)
 - Costs incurred in diagnosing and treating sick animals (veterinary fees and drugs costs).
- Indirect losses include:
 - Additional costs for disease control measures or eradication
 - Human health costs associated with health impacts from zoonotic diseases (BSE, HPAI, Salmonellosis) or from unintended consequences of control measures (chemical residues in products)
 - Negative animal welfare impacts of disease or control measures
 - Trade restrictions due to disease and control measures
 - Loss of consumer confidence in a market sector leading to reduced demand or altered consumer behaviour
 - Range of possible negative effects such as a move towards production systems that may be resistant to disease but that are relatively inefficient or have other

potentially negative effects (use of resistant genetics with reduced production efficiency).

The direct cost of disease has been defined as the sum of the production losses and expenditures that are incurred because of the disease. The components contributing to direct cost include $L + R + T + P$ (Bennett 2003):

- L = value of loss resulting from reduced output in the presence of disease when compared to no disease;
- R = increase in expenditure on non-veterinary resource resulting from the presence of disease (increased labour, feed, vehicle running costs, equipment etc);
- T = expenditure on diagnosis and treatment of disease in affected animals;
- P = expenditure on prophylactic measures to prevent infection and disease from occurring in healthy animals.

3.10.2 Methods for economic evaluation

Economic analyses of disease impacts may be performed at the micro-economic level (farm or household) or at the macro-economic level (industry sector or country). There is a bewildering array of terms and methods used in economic analyses of production systems and the effects of disease on these systems.

At the micro-economic level the most common approaches involve partial budgets and gross margins analysis. At the sector or national level it is more common to see benefit-cost analysis (BCA) of some form. These terms require some brief explanation.

The term *budget* simply means estimation of expected income and expenses.

A whole farm or *enterprise budget* estimates the income (outputs) and costs (inputs) for the enterprise or farm. Input costs include both fixed and variable costs.

Fixed costs for a farm or enterprise vary only in the long run and are still incurred even if output is zero. Fixed costs usually include permanent labour including paid staff and the owner's family, depreciation (infrastructure, vehicles, machinery, equipment), maintenance and repairs, fuel & oil costs (where they cannot easily be assigned to one enterprise), rent, interest.

Variable costs are those costs that are related directly to the amount of output produced and would decline to zero if output was set at zero. Variable costs are able to be allocated to specific enterprise activities (cattle production vs cropping for example). Variable costs include feed, veterinary inputs, seed, fertilizer, marketing costs and casual labour employed for specific jobs such as castration of calves. Vehicle running costs are generally not included in variable costs unless they can be clearly allocated to a specific enterprise. If the number of breeding cows doubles, then the variable costs associated with carrying the additional stock, such as feed costs and costs of medication (drench, vaccination) will also double.

Gross margins analysis is defined as the gross income from an enterprise less the variable costs incurred in achieving it and is generally calculated on a per-year basis. It does not include any fixed costs. The gross margin for an enterprise is the gross income minus the variable costs over a one-year period. Gross margins are generally produced in units such as \$ per animal or animal equivalent or per unit of land area (hectare). A gross margin is not a profit measure because it does not include fixed costs which have to be met regardless of enterprise size. Gross margins do allow comparison of similar enterprises and allow assessment of the impacts of changes in management practices.

A *partial budget* means summarising just those changes in expenses and in income when some minor change is made to management or some other input in the production system (using a new feed supplement or vaccinating/drenching animals). Partial budgets generally consider four components:

Table 3.2: Table showing components of a partial budget for assessing economic impacts of a minor change in management practice(s).

Costs	Benefits
new costs	costs saved
income lost	new income

Partial budgets are relatively simplistic and may not represent all the factors that might be relevant in a decision about investing in some change in management practices.

Table 3.3: Table showing approach to partial budget estimation

Change	Amount	Unit price	Gains	Costs
New feed supplement	S kg	\$s /kg		S*\$s
Additional feed required	F kg	\$f /kg		F*\$f
Additional labour	D days	\$d /day		D*\$d

Additional weight gain of cattle	W kg	\$w /kg	W*\$w
Additional manure sold	M kg	\$m /kg	M*\$m

Partial budget $[W*\$w + M*\$m] - [S*\$s + F*\$f + D*\$d]$

The example above shows a simple partial budget that attempts to assess the impact of a new feed supplement on weight gain and manure output in a cattle production enterprise. If the benefits are larger than the costs then the management change may be considered worthwhile on economic grounds.

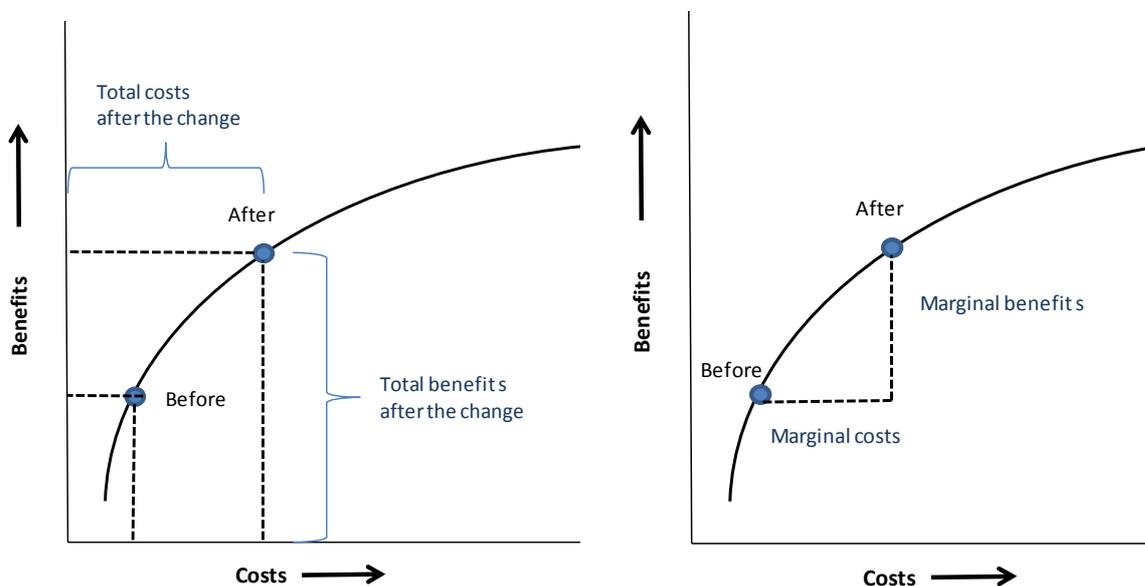


Figure 3.3: Production function showing budget impacts before and after a change in management practices. The left side shows measures or total benefits and costs and the right side shows marginal benefits and costs (Rushton 2009).

Often changes intended to control or eradicate livestock disease and benefit animal production measures may take years to fully implement. In addition costs may be higher in the beginning and then reduce over time and benefits may be lower in the beginning and slowly rise over time. It is very difficult to assess the impacts of these sorts of changes using partial budgets or gross margins analysis alone, mainly because of the time change in the value of money. One dollar earned (or spent) now is not the same as one dollar earned or spent in five years time, mainly because of the effects of things like inflation and discounting.

Often a method like gross margins analysis is extended using additional criteria that allow for the time-changing value of money. Converting future values (benefits or costs) generally involves application of a discount rate.

$$\text{Present Value} = PV = \frac{X_t}{(1 + r)^t}$$

where PV=present value

X_t= amount of money in year t

R is the discount rate expressed as a proportion (5%=0.05)

T= number of years from the present date

The discount rate is also described as the opportunity cost of money. There are many different approaches to setting the discount rate. A reasonable approach is to use the *real rate of interest* which can be estimated as the nominal interest rate (cost of borrowing money) minus the inflation rate. If the market interest rate was 7% and inflation was 1.5% then the real rate of interest would be 5.5%. An alternative approach is to use an estimate of the rate of return you could get if you invested the money in an alternative investment with a similar risk profile (ie investing in a bank or in a financial market).

If all current future benefits and costs are adjusted so that they are all measured in present value (PV), then it is possible to perform comparisons of different strategies that may have different patterns of benefits and costs over time. These comparisons are generally done using one of three criteria: net present value, internal rate of return or benefit-cost ratio.

The *net present value (NPV)* is the difference between the sum of the present value of all benefits and the sum of the present value of all costs. If the NPV is positive (present value of benefits is greater than present value of costs) then the investment is worth considering.

The *internal rate of return (IRR)* is defined as the discount rate that must be applied to make the NPV equal to zero. If the IRR is greater than the conventional discount rate than the project is worth considering because the findings are suggesting that the investment will provide a better return than if you had invested in an alternative investment.

A *benefit-cost ratio (BCR)* or *benefit-cost analysis (BCA)* or *cost-benefit analysis (CBA)* is calculated by dividing the present value of the benefits by the present value of the costs. If the ratio is greater than 1 then the benefits exceed the costs and the investment is worth

considering. Benefit-cost analysis is often of most value when performed at the sector or national levels.

Rushton et al (2012) also describe the use of *cost-effectiveness analysis* as an application of economic evaluation that can be applied in the early stages of a disease outbreak response and used to guide decisions relating to implementation of policy to achieve the most effective result per unit of investment. The principle is based on identifying possible interventions and associated costs and effects (outcomes), in order to achieve a pre-existing policy goal. The evaluation may be presented in terms of cost per positive animal detected or cost per animal saved and the application of this approach may guide decisions about which strategy to implement to achieve a given target in the most cost-effective manner.

It is also important to recognise that decisions at any level will only partially rely on rational economic measures. Farmers may choose one option over another because of risk perception or for other personal reasons rather than solely based on estimates of economic benefit. These issues may explain why individual farmers may choose options that are not necessarily associated with the highest NPV or best BCR based on economic analyses.

3.10.3 Macroeconomics vs microeconomics

Microeconomics refers to estimates of costs and benefits at the farm level.

Macroeconomics refers to economic analyses conducted at an aggregated level such as across an industry sector (livestock or agriculture) or across the national economy as a whole. Macroeconomics is a very complex area particularly when the approach involves trying to determine the interactions between animal disease (including effects of control or eradication programs) on different sectors in the economy such as domestic consumption, foreign trade, tourism, biodiversity and others.

Where animal disease or control programs targeting a disease have effects that are beyond the farm gate or even beyond the livestock production system, there is a strong case for government involvement in disease control programs to better manage investment in risk management for the benefit of the entire population of the country.

Countries may develop a shared responsibility for development and implementation of animal health and welfare policies including disease control programs. At one end of the scale where livestock producers are the primary beneficiaries of any improvement in

outcomes, then the producers may be expected to pay for most or all of the costs associated with the program(s). At the other end of the scale where the benefits of any outcomes may be considered to mainly involve people or areas other than the livestock producer (public health, animal welfare, environmental benefit), then there is a stronger case for having government bear some or most of the costs (Bennett 2012).

3.11 References – economics of animal disease control

Gde Putra, A.A., Hampson, K., Girardi, J., Hiby, E., Knobel, D., Wayan Mardianna, I., Townsend, S. & Scott-Orr, H. (2013). Response to a Rabies Epidemic, Bali, Indonesia, 2008-2011. *Emerging Infectious Diseases*, 19(4):648-651.

Bennett, R. (2003). The 'Direct Costs' of livestock disease: The development of a system of models for the analysis of 30 endemic livestock diseases in Great Britain. *Journal of Agricultural Economics*, 54(1):55-71.

Bennett, R. (2012). Economic rationale for interventions to control livestock disease. *EuroChoices*, 11(2):5-11.

Chilonda, P. and Van Huylenbroeck, G. (2001). A conceptual framework for the economic analysis of factors influencing decision making of small-scale farmers in animal health management. *Rev. sci. tech. Off. Int. Epiz.* 20(3):687-700.

Dohoo, I., Martin, W. & Stryhn, H. 2010. *Veterinary Epidemiologic Research*, Charlottetown, Prince Edward Island, Canada, VER Inc.

Hanson, R. P. & Hanson, M. G. 1983. *Animal Disease Control*, Ames, Iowa, Iowa State University Press.

Hueston, W. D. 2003. Science, politics and animal health policy: Epidemiology in action. *Preventive Veterinary Medicine.*, 60, 3-13.

Martin, S. W., Meek, A. H. & Willeberg, P. 1987. *Veterinary Epidemiology*, Ames, Iowa, Iowa State University Press.

Martin, S. W., Shoukri, M. & Thorburn, M. A. 1992. Evaluating the health status of herds based on tests applied to individuals. *Preventive-Veterinary-Medicine*, 14:33-43.

Otte, M.J. and Chilonda, P. (2001). *Animal health economics: an introduction*. Livestock Information Sector Analysis and Policy Branch, Animal Production and Health Division (AGA), FAO, Rome, Italy.

Rushton, J. *The Economics of Animal Health and Production* (2009). Wallingford, Oxfordshire, CAB International.

Rushton, J., Raboisson, D., Velthuis, A. & Bergevoet, R. (2012). Evaluating animal health investments. *EuroChoices*, 11(2):50-57.

Thrusfield, M. 2005. *Veterinary Epidemiology* 3rd edition, Oxford, UK, Blackwell Science.

Toma, B., Dufour, B., Sanaa, M., Benet, J.J., Ellis, P., Moutou, F. & Louza, A. (1999). *Applied Veterinary Epidemiology and the Control of Disease in Populations*. Paris, France, AEEMA.

Townsend, S.E., Sumantra I.P., Pudjiatmoko, B.G.N., Brum E., Cleaveland S., Crafter S., Dewi A.P., Dharma D.M., Dushoff J., Girardi J., Gunata I.K., Hiby E.F., Kalalo C., Knobel D.L., Mardiana I.W., Putra A.A., Schoonman L., Scott-Orr H., Shand M., Sukanadi I.W., Suseno P.P., Haydon D.T., Hampson K. (2013). Designing Programs for Eliminating Canine Rabies from Islands: Bali, Indonesia as a Case Study. *PLoS Negl Trop Dis* 7(8): e2372. doi:10.1371/journal.pntd.0002372.

Yekutieli, P. 1981. Lessons from the big eradication campaigns. *World Health Forum*, 2:465-490.