

FAO/ASTF Project: GCP/RAF/510/MUL:

Enhancing capacity/risk reduction of emerging Tilapia Lake Virus (TiLV) to African tilapia aquaculture: Intensive Training Course on TiLV

4-13 December 2018. Kisumu, Kenya

in cooperation with Kenya Marine Fisheries Research Institute (KMFRI) and Kenya Fisheries Service (KeFS)

Epidemiology Session Measuring Disease Frequency



Fernando O Mardones

DVM MPVM PhD (epidemiology)

Assistant Professor

Fac. Medicine

P. Universidad Católica, Chile

fomardones@gmail.com



Food and Agriculture
Organization of the
United Nations

Overview

- Measuring the amount of disease is important to assist management or understanding of disease.
 - How large is the disease problem
 - To compare the amount of disease between groups
 - To monitor the success of a disease control program
- Measuring disease can be done in many ways:
 - Counting disease events
 - Calculating the proportion of a population that is affected
- Also possible to compare the amount of disease between groups with ratios, and this can be useful to examine the effect of risk factors.

Counting in Epidemiological Studies

To understand and/or describe:

1. The mechanism(s) of spread of the disease and disease distribution or patterns by time, location, feeding habit, use of the animal, etc.
2. The impact of the disease on the study population and the risk of a given animal in the population having the disease or condition.
3. The implementation and effect of a control program to prevent or eradicate a disease from the population.

Also useful to evaluate the workload, cost, or the magnitude of resources required to provide adequate health care.

Counts of disease cases

- Often expressed as a fraction of the number of animals susceptible to the disease. The latter group of animals is called the population at risk (PAR).

e.g., PAR for retained placenta in cattle would be all cows and heifers that calve in the herd; for porcine parvovirus, the PAR would be serologically negative gilts and sows.

- What would be the population at risk for TiLV infection?

Population at risk

Before count, we must define

- The population at risk (see section 5, Patterns of Disease) is defined by where it is (geography), when (time period of interest), what it is like (description of the population by species, breed, age, sex, etc.) and susceptibility to the condition of interest.
- The unit(s) of study (see section 5, Patterns of Disease): are we observing individual animals, pens of animals, herds, villages, farms, etc.?
- A case definition (see section 3, Investigating Disease Outbreaks): this is a clearly defined description used to distinguish cases (animals, herds, farms, etc.) from non-cases. In some situations there may be multiple case definitions, depending on circumstances.

What do we count?

- **Clinical cases** – the number of cows with clinical signs of Johne's disease
- **Subclinical cases** – the number of apparently healthy chickens that are culture positive for salmonellosis;
- **Animals with certain characteristics** – the number of cows that conceived on the first breeding or the number of ewes that have two or more lambs; and
- **Combinations of the above** – the number of cows with clinical signs of Johne's disease and that are seropositive to paratuberculosis.

What do we count?

- Counting is also often applied at the group level, such as pens, farms, herds, etc.
- We also count the number of animals (or other units) without the disease in the study population (*i.e.* non-cases). Non-cases are any units that do not meet the definition for a case.
- Non-cases either do not show any clinical signs (possibly despite the presence of the organism), or show clinical signs but are negative on both histopathology and culture (or are uncultured).

Ratios, Proportions and Rates

- **Ratios:** expresses the relationship between **two independent numbers**.
The denominator usually does not include the numerator.
- Ratios can be expressed as fractions, but are often expressed simply as the ratio of the two numbers.
- Ratio = a/b or $a:b$, where a is not part of b .
e.g., the ratio of boars to sows in a pig herd is 1:20 or Feed to weight gain ratio is 2.5:1.

Proportions

- A fraction in which the numerator (frequency of disease or condition) is included in the denominator (population). This fraction can be multiplied by 100 in order to create percentages (%).
- **Proportion** = a/b , where a is part of b , or $a/(a+b)$ where a is not part of b .
e.g., the proportion of pregnancies ending in abortions on a dairy farm is $5/65$ or approx. 9%; the proportion of grower pigs (in a particular herd) with lameness is 2%.

Rates

- Expresses the relationship between a population at risk and the event under study over a specific time period.
- **Rate** = a/b where a is part of b per unit time = risk rate; or a is the number of cases and b is animal time at risk = true rate.

e.g., the rate of milk fever in a dairy herd was 10/420 calving cows per year; the incidence rate for foot abscess in baby pigs was 2.9 cases per 1000 pig days at risk.

Epidemiological Elements of Rates

1. The frequency of occurrence of the event

e.g., the number of cows infected with mastitis, the number of new cases of foot abscess occurring per week during the observation period.

2. The population at risk or non-cases at risk of having the disease

e.g., the number of lactating cows that do not have mastitis at the start of the observation period.

3. Time period of the event

i. The *external time component* is the whole time period of the study in relation to calendar time.

e.g., a study of lameness in dairy cows was undertaken during the months of August through November (inclusive).

ii. The *internal time component* is the time relative to a specific event.

e.g., the number of days or weeks post-calving.

Crude, Specific and Adjusted Rates

- **Morbidity** (illness) and **mortality** (death) rates may be classified as crude rates or specific rates (host-attribute-specific and/or cause-specific).
- Here, the term **rate** will be used as a general descriptive term to cover rates, ratios and proportions.
- Disease rates and proportions such as **prevalence** and **incidence** can be expressed as crude rates, specific rates or adjusted rates.

Crude rates

- A rate expressed for the entire PAR (*e.g.* crude mortality rate).
- The **advantage** of crude rates is that they are easy to calculate and to explain.
- They have the **disadvantage** that they ignore the potential influence of various host and management factors (*e.g.* 5 dystocias per 134 calvings).

Specific rates

- A rate for a specified subpopulation of the PAR, based on one or more characteristics such as age, breed or sex.
- For specific rates, both the numerator and denominator must have the specified characteristic.
 - e.g.* age-specific mortality rates: cases and non-cases counted separately for each age group, so that mortality rates can also be calculated for each age group.
- Specific rates allow **comparisons of subpopulations** but are more difficult to explain.
- They also make it harder to compare populations composed of multiple subgroups (like herds and flocks).
 - e.g.* 4 dystocias per 32 heifers calving; 1 dystocia per 102 cows calving.

Adjusted rates

- An adjusted rate (or standardized rate) is a rate calculated by adjusting the rate for each population to match a '*standard*' population structure for the characteristic of interest.
- Adjusted or standardized rates are used to compare disease rates between populations with different age, sex and/or breed structures.
- To calculate the adjusted or standardized rate for a population, specific rates are calculated for each level of the selected characteristic and then weighted by the proportion of the similar specific groups in the standard population.

Adjusted rates

- The standard population may be whatever structure you choose, but should approximate the structure for the overall population.

e.g. when comparing rates of different districts within a region the chosen standard population structure should be similar to the regional population structure.

Application of adjusted rates to two farms with apparently differing case percentages due to different underlying age structures

| Age (years) | Farm 1 | | Farm 2 | |
|--------------|--|-------|--------|-------|
| | PAR | Cases | PAR | Cases |
| Actual | different age structures between the two farms | | | |
| <2 | 30 | 3 | 20 | 2 |
| 2-5 | 50 | 5 | 30 | 3 |
| >5 | 20 | 8 | 50 | 20 |
| | 100 | 16 | 100 | 25 |
| Case % | | 16 | | 25 |
| Standardized | | | | |
| <2 | 25 | 2.5 | 25 | 2.5 |
| 2-5 | 50 | 5 | 50 | 5 |
| >5 | 25 | 10 | 25 | 10 |
| | 100 | 17.5 | 100 | 17.5 |
| Case % | | 17.5 | | 17.5 |

Measures of Morbidity (Illness) in a Population

- The key measures of the frequency of disease occurrence are **prevalence** and **incidence**.
- **Prevalence** = The proportion of existing cases of disease present in a population at a given point in time.
- **Prevalence** = number of cases/PAR
e.g., the prevalence of arthritis in adult pigs equals the number of cases of arthritis in adults divided by the total number of adults in the population.
- The prevalence of TiLV-infected Tilapia farms equals the number of infected farms divided by the total number of farms (with Tilapia).

Incidence

- The number of new cases that arise in a population over a specified period of time.
- **Incidence** = number of new cases in a given time period/total PAR
- 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 as a risk rate (or cumulative incidence) or a true rate (incidence density). In addition, attack rate is often used instead of incidence rate in outbreak investigations.

Cumulative incidence

- The number of animals that contract the disease in a defined period divided by the number of healthy animals at risk at the beginning of start of the time period.
- The length of the period has a large influence on the cumulative incidence: the longer the period, the higher the cumulative incidence.
- Therefore, it is essential that the relevant time period is quoted as part of the cumulative incidence, for example 1% per month, or 10% per year.

Incidence rate (IR; also called incidence density)

- Number of new cases of disease in a population during a certain period divided by the total number of animal-time-units at risk for all animals in the PAR.
- The time units may be animal-years, animal-weeks or any other suitable time unit.
- Only healthy animals contribute to the denominator because only healthy animals are at risk of contracting the disease under observation.
- However, a case can contribute to animal time at risk up until the point when it becomes a case.

Relation between CI and IR

$CI(t) = 1 - e(-IR \times t)$, where e is the base to the natural logarithm (2.718) and t is the time unit of concern. When the $CI < 0.10$ (10%), the formula approximate to:

$$CI(t) = IR \times t$$

The relationship between IR and prevalence is: $\frac{P}{1-P} = IR \times D$, where $\frac{P}{1-P}$ is the ratio of the proportion of diseased to healthy animals and where D is the average duration of disease. When P is small (< 0.05 or 5%) the above formula reduces to: $P = IR \times D$

Attack rate or attack risk

- A specific type of incidence rate that applies to outbreaks or situations where the period of observation is relatively short.
- 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).
- Attack rate = number of animals affected/number of animals exposed
e.g., the attack rate can be used to measure mortality due to ISAV virus infection in farmed Atlantic salmon. If, over a 10-day period, 35,000 of the 50,000 salmon in a cage die, the attack rate is 0.7 or 70%.

| | Incidence rate | Cumulative incidence | Prevalence |
|----------------|---|---|---|
| Numerator | New cases occurring during a period of time among a group initially free of the disease in question | New cases occurring during a period of time among a group initially free of the disease in question | Existing cases at a point in time |
| Denominator | Sum of time periods during which individuals could have developed disease | All at-risk individuals present at the beginning of the period | All at-risk individuals examined, including cases and non-cases |
| Time | From beginning of follow-up until disease occurs for each individual | Duration of period of observation | Single point in time |
| How measured | Prospective cohort study | Prospective cohort study | Cross-sectional study |
| Interpretation | Rapidity with which new cases develop over a defined time period | Risk of developing disease in defined time period | Risk of having disease at a particular point in time |

Do not forget!

1. Incidence is a dynamic measure of disease whereas prevalence is only a static measure of disease.
2. Incidence and prevalence are related. The prevalence of disease in a PAR reflects both the incidence of new cases of disease and the duration of disease in individual cases:

Prevalence = incidence x duration under certain conditions.

3. 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.

Do not forget!

4. True rate describes the average speed at which the event of interest occurs per unit of animal time at risk. It is often called incidence rate. True rate has no meaning on the individual level. However, it can be interpreted on a population basis.
5. Risk rate (cumulative incidence rate) provides a direct estimate of the likelihood of an animal experiencing the event of interest during the internal time period. Risk rate has a meaning on an individual basis as well as on a population basis.

Do not forget!

6. Counting the PAR (*i.e.* the denominator):
 - i. With prevalence, the total number of animals examined during the time you counted the frequency of disease is the denominator.
 - ii. With incidence rates, however, we are looking at a population over a period of time; therefore, the number of animals at risk can change. There are a number of ways to deal with this problem, but the two most common are:
 - Use an estimate of the population, either by counting the population at a time midway in the time interval, or by taking the average of the population at the beginning and end of the time interval.
 - Calculate the population on each day of the time interval and arrive at the number of animal-days-at-risk (incidence density rate).

Measures of Mortality (Death) in a Population

Measures of mortality in the general (healthy) population

1. Crude death rate:

$$\frac{\text{NUMERATOR}}{\text{DENOMINATOR}} = \frac{\text{deaths in a given time}}{\text{total population at risk}}$$

2. Cause-specific death rate (a measure of the risk of death from a specific cause):

$$\frac{\text{NUMERATOR}}{\text{DENOMINATOR}} = \frac{\text{deaths in a given time due to the disease of interest}}{\text{total PAR}}$$

3. Age/cause-specific death rate (limits numerator and denominator to specific age/cause of interest):

$$\frac{\text{NUMERATOR}}{\text{DENOMINATOR}} = \frac{\text{deaths in a given time in the group of interest}}{\text{total PAR for the group of interest}}$$

Measures of disease attributes among the ill or dead animals

1. Case recovery rate (actually a proportion rather than a true rate):

$$\frac{\text{NUMERATOR}}{\text{DENOMINATOR}} = \frac{\text{number of cases recovering}}{\text{total cases for which outcome known}}$$

2. Case fatality rate (a proportion rather than a true rate):

$$\frac{\text{NUMERATOR}}{\text{DENOMINATOR}} = \frac{\text{number of cases dying}}{\text{total cases for which outcome is known}}$$

3. Proportional mortality rate: The proportion of total deaths attributable to a specific cause:

$$\frac{\text{NUMERATOR}}{\text{DENOMINATOR}} = \frac{\text{deaths due to specific cause of interest}}{\text{total deaths in population}}$$

Comparing Disease Frequencies

- Since incidence rates reflect risk, then the incidence rates (or attack rates) of two different groups may be compared in a ratio called the risk ratio or relative risk (RR).
- The RR compares disease among individuals of the one group to another group.
- Relative risk and a number of other commonly used measures can be used to compare disease frequency between risk groups in the population.

Relative risk

- The relative risk (or risk ratio, relative incidence rate ratio, etc.) is the ratio of the incidence rate (IR) in the exposed group to the IR in the unexposed group.
- You can use cumulative incidence, incidence density or attack rate for the calculations, as long as you use the same type of measure in both parts of the ratio.
- Since RR is the ratio of incidences, RR cannot be calculated for case-control studies (because incidence cannot be calculated in case-control studies).

Calculation of relative risk from an attack rate table

| | Diseased | Not diseased | Total |
|-------------|----------|--------------|-----------|
| Exposed | a | b | $a+b$ |
| Not exposed | c | d | $c+d$ |
| Total | $a+c$ | $b+d$ | $a+b+c+d$ |

$$\text{Incidence (exposed)} = \frac{a}{a+c}$$
$$\text{Incidence (unexposed)} = \frac{c}{c+d}$$
$$\text{Relative risk} = \frac{a/(a+c)}{c/(c+d)}$$

Relative Risk can vary from zero to infinity

- An estimate of how much more likely disease is to occur in the exposed group compared to the unexposed group and has a null value (no association or no increase in risk) of 1, which is equivalent to equal incidence rates.
- If **RR is >1** the factor increases the risk of disease. If **RR is <1** the factor decreases the risk of disease.
- However, a **confidence interval** for the estimate should always be calculated and the value can only be considered to vary significantly from 1 if the confidence interval does not include 1.
- As a rule of thumb, **RR values greater than about 3** (or less than about 0.33) are considered potentially biologically important.

In class exercise

| Higher Tilapia density | Diseased | Not diseased | Total |
|------------------------|----------|--------------|-------|
| Yes | 43 | 11 | 54 |
| No | 3 | 18 | 21 |
| Total | 46 | 27 | 75 |

Please, report the incidence in exposed, non exposed and the RR...

Interpretation

- Tilapia farms that were at higher densities were almost six times more likely to become sick than were those that were at lower densities.
- In fact, 95% confidence limits for the RR estimate are 1.96 to 16, suggesting that this difference is likely to be statistically significant.
- We could also use a Chi-square statistical test to see if this relationship is significant.

Odds ratio (OR)

- Measure of the strength of association that is very useful in epidemiological studies of all types (cohort, case-control, cross-sectional).
- As the name implies, this is a ratio of the odds of exposure:non-exposure in disease-specific groups or the ratio of the odds of disease:no disease in exposure-specific groups.

Odds Ratio

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

| | Diseased | Not diseased | Total |
|-------------|----------|--------------|-----------|
| Exposed | a | b | $a+b$ |
| Not exposed | c | d | $c+d$ |
| Total | $a+c$ | $b+d$ | $a+b+c+d$ |

Exercise in class

- Please, estimate OR for density example

Interpretation OR

- Similar to the relative risk: values >1 indicate increased risk, while values <1 indicate a protective factor. Just like a RR, the null value of the OR is 1, and the OR has no units.
- The OR's significance can also be tested using a Chi-square statistical test or confidence intervals. In the above example the 95% confidence interval is from 5.8 to 94.
- The odds of developing disease were 23.5 times greater for those farms that were under high densities compared with those that did not were exposed to higher densities.

Comparing the RR and the OR

if the disease is rare a number of approximations start to hold true:

- a is small compared to b ;
- c is small compared to d ;
- $(a + b)$ approximates b ; and
- $(c + d)$ approximates d .

| | Diseased | Not diseased | Total |
|-------------|----------|--------------|-----------|
| Exposed | a | b | $a+b$ |
| Not exposed | c | d | $c+d$ |
| Total | $a+c$ | $b+d$ | $a+b+c+d$ |

- Therefore, if the disease is rare, the OR approximates the RR.
- Rule of thumb: CI <10% but the approximation becomes better as the disease becomes rarer.

Attributable risk

- Attributable risk also called risk difference is the absolute difference between the two incidence rates.
- $AR = IR_{exposed} - IR_{unexposed} = \left(\frac{a}{a+b}\right) - \left(\frac{c}{c+d}\right)$
- The AR tells us how much of the disease in the exposed group is attributable to being exposed.
- It implies the rate of disease that could be prevented if the exposure were removed completely from the population.

Attributable Risk

- If you get a negative AR, the AR is telling you the rate of disease that was prevented by the exposure.
- 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.
- Please, calculate the AR from our example.

Calculations from our example

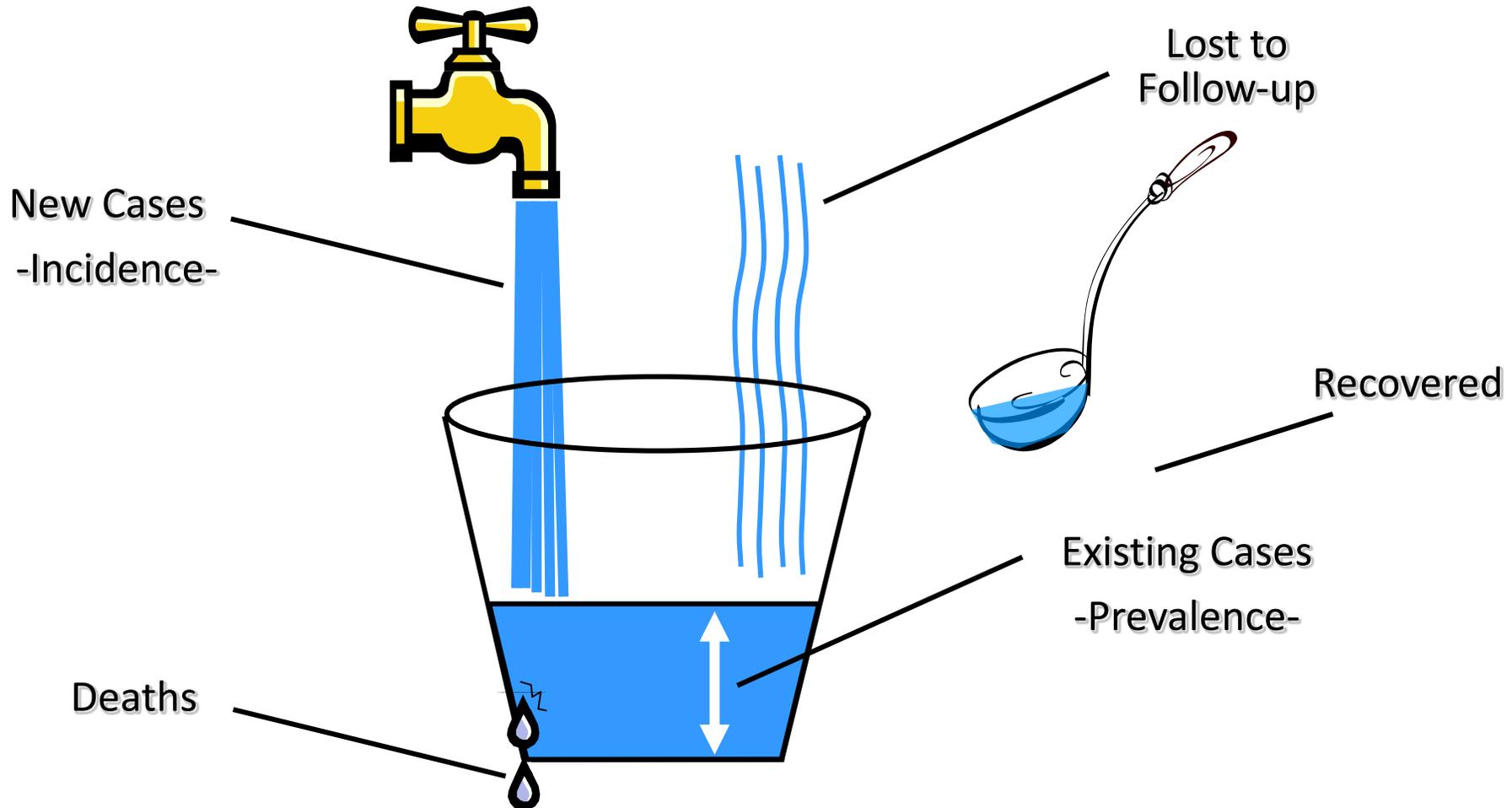
- $IR_{exposed} = \left(\frac{a}{a+b}\right) = 0.8$

- $IR_{unexposed} = \left(\frac{c}{c+d}\right) = 0.14$

- $AR = IR_{exposed} - IR_{unexposed} = 0.8 - 0.14 = 0.66$

- Our interpretation is that the amount of disease above the background rate that is associated with exposure to higher densities is 0.66.

Measuring Disease in a Population – Inputs and Outputs –



FAO/ASTF Project: GCP/RAF/510/MUL:

Enhancing capacity/risk reduction of emerging Tilapia Lake Virus (TiLV) to African tilapia aquaculture: Intensive Training Course on TiLV

4-13 December 2018. Kisumu, Kenya

in cooperation with Kenya Marine Fisheries Research Institute (KMFRI) and Kenya Fisheries Service (KeFS)

Epidemiology Session Measuring Disease Frequency



Fernando O Mardones

DVM MPVM PhD (epidemiology)

Assistant Professor

Fac. Medicine

P. Universidad Católica, Chile

fomardones@gmail.com



Food and Agriculture
Organization of the
United Nations