

World Organisation for Animal Health

Australian Aid

Training Workshop on Animal Disease surveillance

Sample Size Calculation Methods in Animal Health Surveillance

Shankar Yadav

12-13 November 2026
Manila, The Philippines

Learning Objectives

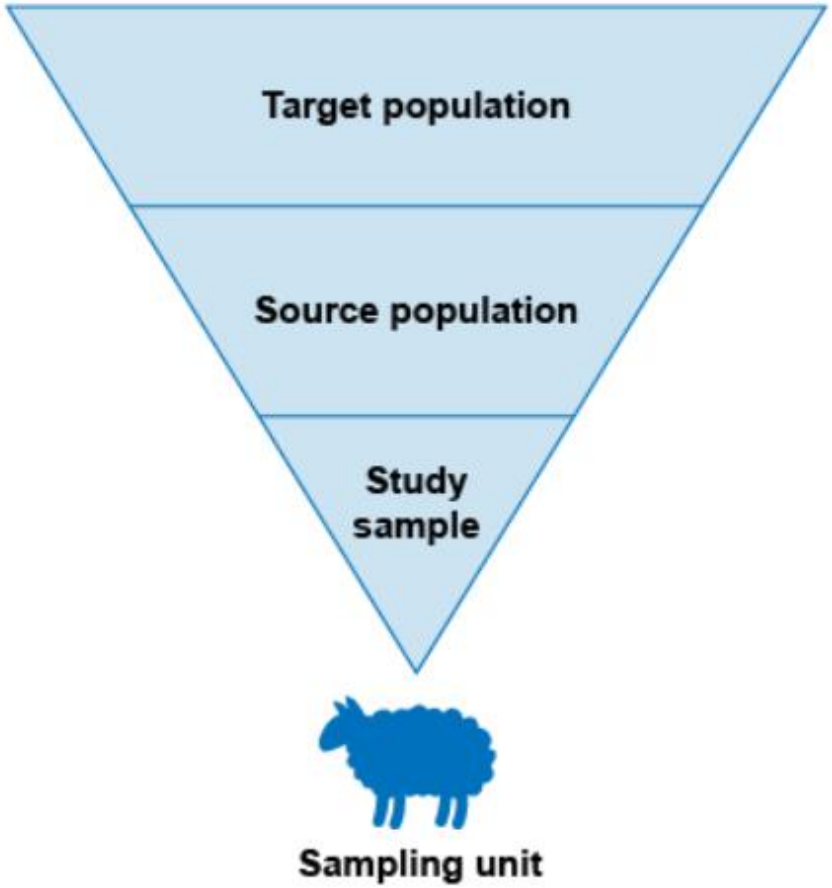
- By the end of this presentation, participants should be able to:
  - Explain the importance of sample size determination.
  - Understand factors influencing sample size estimation.
  - Identify different sample size calculation methods.
  - Calculate sample sizes for disease prevalence estimation.
  - Determine sample sizes for disease detection/freedom from disease.
  - Understand herd-level and animal-level surveillance.
  - Apply sample size calculations to practical animal health scenarios.

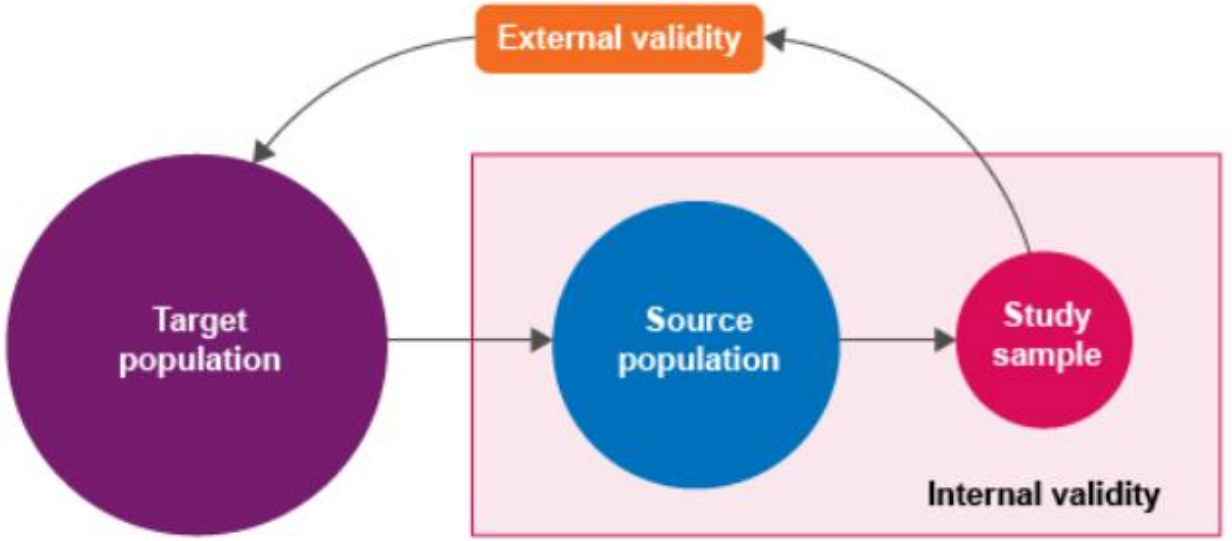
Why is Sample Size Important?

Sample size determines how many animals or herds must be sampled to achieve **reliable surveillance results.**

- **Importance:**
  - Ensures adequate disease detection.
  - Improves the accuracy of prevalence estimates.
  - **Reduces sampling costs.**
  - Supports evidence-based disease control.
  - Helps demonstrate disease freedom.
  - Improves confidence in surveillance findings.

Population and Sample Size





Selection Bias
External validity
Internal validity

Selection Bias

Selection bias occurs when the animals or farms included in a study are different from those that should be represented.

Example: Foot-and-mouth disease (FMD)

- A researcher wants to estimate FMD prevalence in a country but collects **samples only from large commercial cattle farms.**
  - Smallholder farms are excluded.
- **Why is this a problem?**
  - Smallholder farms may have different biosecurity practices.
  - Animal movement patterns may differ.
  - Disease exposure may be different.
- **The estimated prevalence may not represent the entire cattle population.**
- **Effect:** Can affect **internal validity and external validity**, depending on the study question and sampling process.

Internal Validity

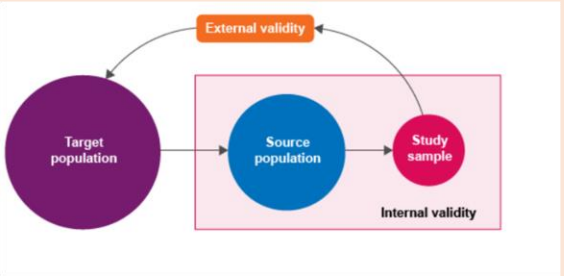
Internal validity refers to **how accurately a study measures** what it intends to measure in the animals or population included in the study.

In simple terms: *Are the results of this study correct and free from systematic errors?*

Example

- Suppose a researcher investigates the association between housing conditions and mastitis in dairy cows.

- The researcher:
  - Selects 20 dairy farms.
  - Collects information about housing conditions.
  - Tests cows for mastitis.
  - Compares mastitis occurrence between different housing systems.

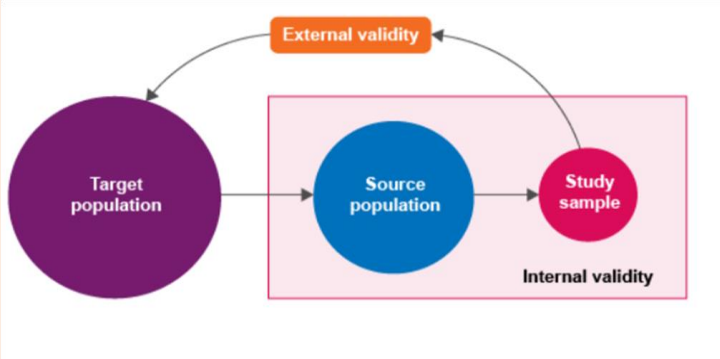


- **However, the researcher does not account for differences in milking hygiene between farms.**
- Milking hygiene may influence mastitis occurrence and may also differ between housing systems. This could introduce bias and produce misleading results.
- Therefore, **the study may have poor internal validity**.

External Validity

External validity refers to the extent to which the **findings of a study can be generalized or applied to other animal populations, geographical areas, production systems, or time periods**.

- Suppose a researcher investigates the prevalence of bovine tuberculosis in 10 intensive dairy farms in country X.
  - The study finds a prevalence of 5%.
- Can we conclude that bovine tuberculosis prevalence is also 5% in:
  - Beef cattle farms?
  - Smallholder dairy farms?
  - Extensive grazing systems?
  - Dairy farms in a different country Y?



- **Not necessarily.**
  - These populations differ in **management practices, animal movement, biosecurity, environmental conditions, and disease exposure**.
  - **Therefore, the study may have limited external validity.**

A Scenario on Internal versus External Validity

Imagine a country has 1,000 poultry farms, but its surveillance system collects samples from only 100 large commercial farms.

- The surveillance reports a low prevalence of avian influenza.

Internal validity

- If samples are properly collected, diagnostic tests are reliable, and cases are correctly identified, the results may accurately describe the 100 participating farms.
- Potentially good internal validity

External validity

- If backyard poultry and small farms are not included, the findings may not represent all 1,000 farms.
- Potentially poor external validity

- **This illustrates why a surveillance system can produce accurate results for the population it monitors but still fail to represent the wider animal population.**

Factors Affecting Sample Size

- **Population size:**
    - Total number of animals in the target population.
- **Power:**
    - How likely is a study to identify a statistically significant result if there is an association or difference? **Example,** assume a vaccine study with power of 80%.
    - If the vaccine truly has the effect specified in the study design, the study has an **80% probability of detecting that effect as statistically significant.**
    - There is a 20% probability of failing to detect that effect.
- **Expected prevalence:** What researchers believe or anticipate the disease prevalence will actually be.
- **Design prevalence:** The minimum disease prevalence that the surveillance system is designed to detect at a specified confidence level.
- **Confidence level:** Usually 95%. A 95% confidence level means that a surveillance or sampling procedure is designed to provide **a 95% probability of detecting at least one infected animal** when the disease is present at or above the specified design prevalence, under the stated assumptions.

Factors Affecting Sample Size

- **Precision:**
  - Acceptable margin of error.
  - Precision refers to how close repeated measurements or estimates are to one another.
  - **5% precision usually means an acceptable margin of error of ±5 percentage points around an estimated prevalence.**
- **Test sensitivity:** Ability of a diagnostic test to identify infected animals.
- **Test specificity:** Ability to identify non-infected animals correctly.
- **Sampling design:** Simple random, cluster, or multistage sampling.
- **Design effect:** Adjustment for similarities among animals within herds.

Sample Size for Estimating Disease Prevalence

- This method is used when the objective is to estimate the prevalence of a disease in an animal population.
- **Formula:** $n=Z^2P(1-P)/d^2$ Where:
  - n = required sample size
  - Z= Z-score for confidence level
  - P = expected prevalence
  - d = desired absolute precision

Confidence level	Z-value
90%	1.645
95%	1.96
99%	2.576

EXAMPLE

Suppose the expected prevalence of bovine brucellosis is 10%, with 95% confidence and 5% precision.

$n=1.96^2(0.10)(0.90)/0.05^2$

$n=138.3$

Therefore, at least 139 cattle should be sampled, assuming simple random sampling and a sufficiently large population.

Sample Size for Disease Detection- Infinite or Unknown Population

This method determines how many animals should be tested to have a specified probability of detecting at least one infected animal. Appropriate for large or unknown population.

Formula:

$n = \ln(1-C)/\ln 1-P$ Where:

- n = sample size
- C = desired confidence
- P = design prevalence

This formula assumes:

- Random sampling.
- Independent observations.
- A large population.
- A perfect diagnostic test (100% sensitivity and specificity).

Example:

If disease prevalence is expected to be 5% and the desired confidence of detection is 95%:

n=ln(0.05)/ln(0.95)

n=58.4

At least 59 animals should be tested to detect at least 1 infected animal.

Assumes perfect diagnostic sensitivity, independent random sampling and a large population.

Adjustment for Diagnostic Test Sensitivity

- If the diagnostic test has less than 100% sensitivity, the sample size must be increased.
- A commonly used approximation is:

$$n = \frac{\ln(1 - C)}{\ln(1 - P \times Se)}$$

Where:

- C = desired confidence.
- P = design prevalence.
- Se = diagnostic test sensitivity.

Example:
If disease prevalence is expected to be 5% and the desired confidence of detection is 95%: now the test sensitivity is 90 %

$n = \ln(0.05)/\ln(1 - 0.05 \times 0.90)$
$n \approx 66$

Lowering the sensitivity of the test from 100 % to 90 % , the required sample size increased from 59 to 66.

Sample Size for Disease Detection- Finite Population

- When the total animal population is known and **relatively small, a finite-population approach** is more appropriate.
- A formula presented in veterinary epidemiology for this purpose is:

$$n = \left[1 - (1 - p_1)^{1/d}\right]\left(N - \frac{d}{2}\right) + 1$$

Where:

- n = required sample size.
- p_1 = desired probability of detecting at least one case.
- N = total population size.
- d = minimum number of diseased animals expected in the population.

For example, if:

- Total cattle population = 500.
- Minimum expected prevalence = 5%.
- Desired detection probability = 95%.

Then:

$d=500\times0.05=25$

$n=[1-(1-0.95)^{1/25}](500-25/2)+1$

$n\approx58$

The finite-population calculation gives approximately 58 animals.

Sample size Comparing Two Proportions

Sample size estimation for comparing **two proportions in animal disease studies** is commonly used when comparing disease prevalence, incidence, mortality, or treatment success between two animal groups.

For example:

- Comparing disease prevalence between vaccinated and unvaccinated animals.
- Comparing infection rates between two farms.
- Comparing mortality between treated and untreated animals.
- Comparing diagnostic test positivity between two animal populations.

Sample size formula for comparing two independent proportions

$$n = \frac{\left[Z_{\alpha/2}\sqrt{2\bar{P}(1-\bar{P})} + Z_{\beta}\sqrt{P_1(1-P_1)+P_2(1-P_2)}\right]^2}{(P_1-P_2)^2}$$

Where:

- n = required sample size per group.
- P_1 = expected disease proportion in group 1.
- P_2 = expected disease proportion in group 2.
- $\bar{P} = (P_1 + P_2)/2$ = average expected proportion.
- $Z_{\alpha/2}$ = Z-value for the significance level (usually 1.96 for 5% significance).
- Z_{β} = Z-value corresponding to statistical power (0.84 for 80% power).

Study assumptions:

Sample Size Comparing Two Proportions

Suppose you want to compare the prevalence of a particular animal disease between:

- Group 1: Animals from non-commercial production system, expected prevalence = 30%.
- Group 2: Animals from commercial production system, expected prevalence = 15%.

Parameter	Value
Expected proportion in group 1	0.30
Expected proportion in group 2	0.15
Significance level (α)	0.05
Statistical power	80% ; *Zbeta;-0.84*
Allocation ratio	1:1

$$n = \frac{\left[Z_{\alpha/2}\sqrt{2\bar{P}(1-\bar{P})} + Z_{\beta}\sqrt{P_1(1-P_1)+P_2(1-P_2)}\right]^2}{(P_1-P_2)^2}$$

Where:

- n = required sample size per group.
- P_1 = expected disease proportion in group 1.
- P_2 = expected disease proportion in group 2.
- $\bar{P} = (P_1 + P_2)/2$ = average expected proportion.
- $Z_{\alpha/2}$ = Z-value for the significance level (usually 1.96 for 5% significance).
- Z_{β} = Z-value corresponding to statistical power (0.84 for 80% power).

$\bar{P}$ = (0.30+0.15)/2=0.225

Substituting into the formula:

$n \approx 121$

Estimated sample size

- Group 1: 121 animals
- Group 2: 121 animals

Total: 242 animals

Sample size for Cluster Sampling

- Epidemiological studies deals with clustered data such as cows are clustered within herd.
- So, the cows in the same cluster are more similar compare to the population.
- If our study interest is at animal level, then cluster or herd may not be an issue. For example, days from calving to conception is measured at the animal level, therefore, clustering does not present a problem when computing the necessary sample size
- If the factor of interest is something that occurs at the herd level such as impacts of barn types on health outcomes; mastitis as an example. In this context, the number of herds become more critical than the number of animals- even though the outcome is measured at the animal level, but the disease of interest is impacted by barn type.

Design effect formula:

DEFF = 1 + (m—1)ρ Where:

- DEFF = Design effect.
- m = Average number of animals sampled per cluster.
- ρ= Intracluster correlation coefficient (ICC).

Adjusted sample size: ncluster = n × DEFF

This adjustment helps account for the reduced statistical information obtained from sampling animals within the same cluster.

Sample Size for Cluster Sampling

- In animal health surveillance, animals are often grouped within farms, villages, or herds.
- Animals within the same herd may have similar disease risks. Therefore, cluster sampling must account for this similarity.
- **Design effect formula:**
- DEFF = 1+(m−1)ρ Where:

DEFF = Design effect.

m = Average number of animals sampled per cluster.

P = Intracluster correlation coefficient (ICC).

Adjusted sample size: ncluster = n×DEFF

This adjustment helps account for the reduced statistical information obtained from sampling animals within the same cluster.

Assume:

•Average cluster size = 10 animals per farm.

•ICC = 0.05.

DEFF= 1+(10−1)×0.05

DEFF= 1+0.45 =1.45

The original sample size **ON SLIDE 16** was n=121 animals per group.

Therefore:

ncluster=121×1.45

ncluster=175.45

Thus, approximately 176 animals are required per group.

Scenario: African Swine Fever Surveillance

- A veterinary authority wants to estimate ASF prevalence in pigs.
- Given:
  - Expected prevalence = 5%.
  - Confidence level = 95%.
  - Desired precision = 3%.
- Calculate the sample size.

Scenario: African Swine Fever Surveillance

- A veterinary authority wants to estimate ASF prevalence in pigs.
- Given:
  - Expected prevalence = 5%.
  - Confidence level = 95%.
  - Desired precision = 3%.
- Calculate the sample size.
- $n=1.96^2(0.05)(0.95)/0.03^2$
- n=203

Formula: $n=Z^2P(1-P)/d^2$

Where:

n = required sample size

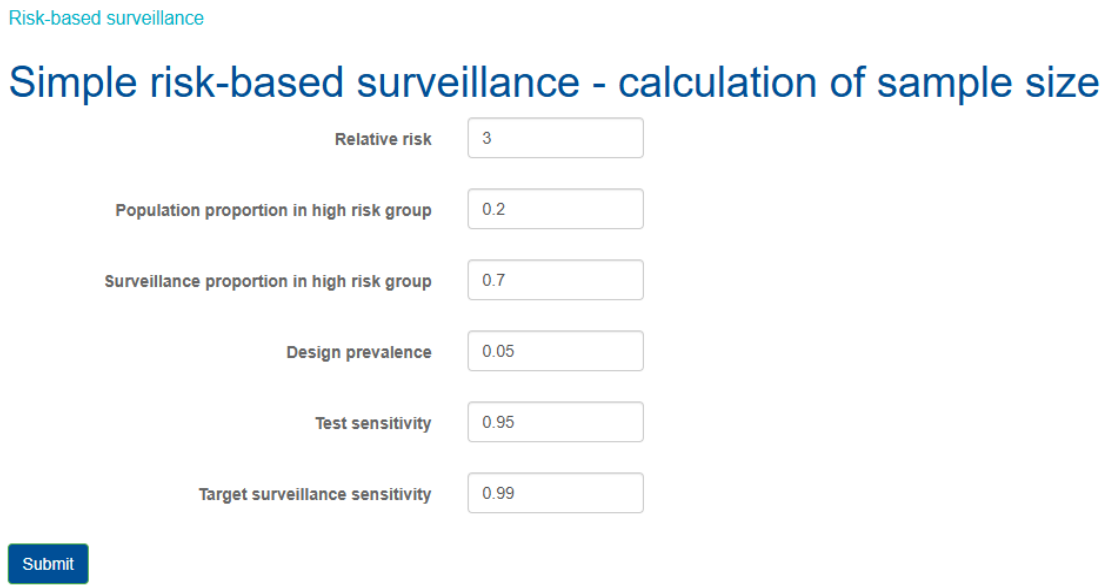
Z= Z-score for confidence level

P = expected prevalence

d = desired absolute precision

EPITOOLS for Sample Size Calculation-Risk-based Surveillance-1-Stage Survey with Single Risk Factor

- This method calculates the sample size for simple risk-based surveillance, for instance, **a survey in which a high-risk population is targeted.**
- This analysis assumes that there is **no clustering** of disease (for instance, we are working at the herd level), and that the effective specificity of the surveillance system is equal to one (all positives are followed up to ensure that they are not false positives):



	Sample size	High risk	Low risk
Risk-based	**55**	**39**	**16**
Representative	**95**		
Saving	**42.105%**		
EPI (High risk)	**10.7%**		

EPITOOLS for Sample Size Calculation-Risk-based Surveillance-2-Stage Survey with Single Risk Factor at Cluster Level

- This method calculates the sample size for simple risk-based surveillance with 2-stage sampling, for instance, **a survey in which high risk herds are preferentially targeted for testing, but with representative sampling of animals within selected herds.**
- The analysis **assumes 2-stage sampling** to account for clustering of disease (for example at the herd, flock or village level) and that the effective specificity of the surveillance system is equal to one (all positives are followed up to ensure that they are not false positives).

Simple 2-stage risk-based surveillance - calculation of sample size

Relative risk	3
Population proportion in high risk group	0.2
Surveillance proportion in high risk group	0.7
Herd-level design prevalence	0.05
Animal-level design prevalence	0.1
Test sensitivity	0.9
Target herd-sensitivity	0.95
Target surveillance sensitivity	0.99

	Herds to test	Animals to test per herd	Total animals to test
High-risk herds	39		
Low-risk herds	16		
Risk-based total	55	32	1760
Representative	95	32	3040
Saving	42.1%		
Herd-level EPI in high-risk group	10.7%		

EPITOOLS for Sample Size Calculation-Risk-based Surveillance-2-Stage Survey with Risk Factors at Both Cluster and Unit Levels

- This method calculates the sample size for **more complex risk-based surveillance**.
- This analysis allows for clustering at the herd level, two risk factors (at herd and animal level) and a factor influencing the performance of the test. It assumes that these factors are independent and that the effective specificity of the surveillance system is equal to one (all positives are followed up to ensure that they are not false positives):
- Two risk factors are considered, one at the herd level and one at the animal level.

Complex risk-based surveillance - calculation of surveillance sample size

Herd-level risk factor

Relative risk	3
Population proportion in high risk group	0.2
Surveillance proportion in high risk group	0.7
Design prevalence	0.1

Animal-level risk factor

Relative risk	2
Population proportion in high risk group	0.5
Surveillance proportion in high risk group	0.9
Design prevalence	0.1

Diagnostic test

Population proportion in high sensitivity group	0.4
Surveillance proportion in high sensitivity group	0.8
Sensitivity in high sensitivity group	0.95
Sensitivity in low sensitivity group	0.9
Target surveillance sensitivity	0.95

Submit

	Sample size
Risk-based	**145**
Representative	**324**
Saving	**55.182%**

Sample Size Increases

- **When:**
  - Expected disease prevalence is low.
  - Higher confidence is required.
  - Greater precision is required.
  - Diagnostic test sensitivity is low.
  - Animals are clustered within herds.
  - Non-response or missing samples are expected.

KEY MESSAGES

- Sample size depends on the surveillance objective.
- Prevalence estimation and disease detection require different calculations.
- Herd-level and animal-level sampling must be distinguished.
- Clustering should be considered in livestock populations.
- Diagnostic test performance affects surveillance sensitivity.
- Appropriate sample size improves the reliability and cost-effectiveness of animal health surveillance.
- **Design prevalence for disease detection looks like a highly sensitive parameter; therefore, the choice of this value would require due attention and consideration.**

Recommended references

- Thrusfield, M. Veterinary Epidemiology. 4th Edition.
- Dohoo, I., Martin, W. and Stryhn, H. Veterinary Epidemiologic Research.
- World Organisation for Animal Health (WOAH) – Terrestrial Animal Health Code.
- FAO – Animal disease surveillance guidelines.
- Cannon, R.M. and Roe, R.T. (1982). Livestock Disease Surveys: A Field Manual for Veterinarians.
- Stevenson MA. Sample Size Estimation in Veterinary Epidemiologic Research. Front Vet Sci. 2021 Feb 17;7:539573. doi: 10.3389/fvets.2020.539573.
- https://epitools.ausvet.com.au/samplesize

Thank you

WOAH Sub-Regional Representation
for South-East Asia

srr.seasia@woah.org

