Food and Agriculture Organization of the United Nations

SUSTAINABLE DEVELOPMENT GOALS

VIRTUAL COURSE

26 March to 15 April 2021

Design of an Active Surveillance for Tilapia Lake Virus (TILV) Disease and Its Implementation

TCP/INT/3707: Strengthening biosecurity (policy and farm level) governance to deal with Tilapia lake virus

Food and Agriculture Organization of the United Nations

SUSTAINABLE DEVELOPMENT GOALS

CHECKLIST 7

7 April 2021

Study design and sampling

Nihad Fejzic
nihad.fejzic@vfs.unsa.ba

Fernando Mardones
fomardones@gmail.com

TCP/INT/3707: Strengthening biosecurity (policy and farm level) governance to deal with Tilapia lake virus

Learning objectives

- To understand the requirements and criteria for Checklist 7
- To design a cross sectional study for TiLV active surveillance and its requirements

Presentation topics

- Observational studies
- Survey design
- Representativeness
- Non representative sampling
- Selection method, sampling frame, random sampling

Activities in each surveillance:

- Determine what questions is being asked and how best to answer that
- Identify target population
- Choose the **right survey design**
- Prepare questionnaires and data collection form
- Decide if the survey is to use stratification
- Calculate **the best sample size**
- Plan field activities
- Train survey team

CONTINUED IN NEXT SLIDE

Activities in each surveillance:

- Conduct a pilot survey
- **Select the sample**
- Carry out field work
- Collect information (from producers and/or animals)
- Process specimens ready for analysis
- Send the specimens to the laboratory
- Check the data for completeness and accuracy

CONTINUED IN NEXT SLIDE

Activities in each surveillance:

- Enter the survey data and laboratory results into computer
- Check the data for mistakes during data entry
- Collect information (from producers and/or animals)
- Analyze the data to calculate estimates
- Report the data

Surveillance approaches

- **Passive surveillance**

- **Active surveillance**
  - Prevalence surveys
  - Production surveys
  - Incidence rate surveys
  - Surveys to demonstrate freedom from disease

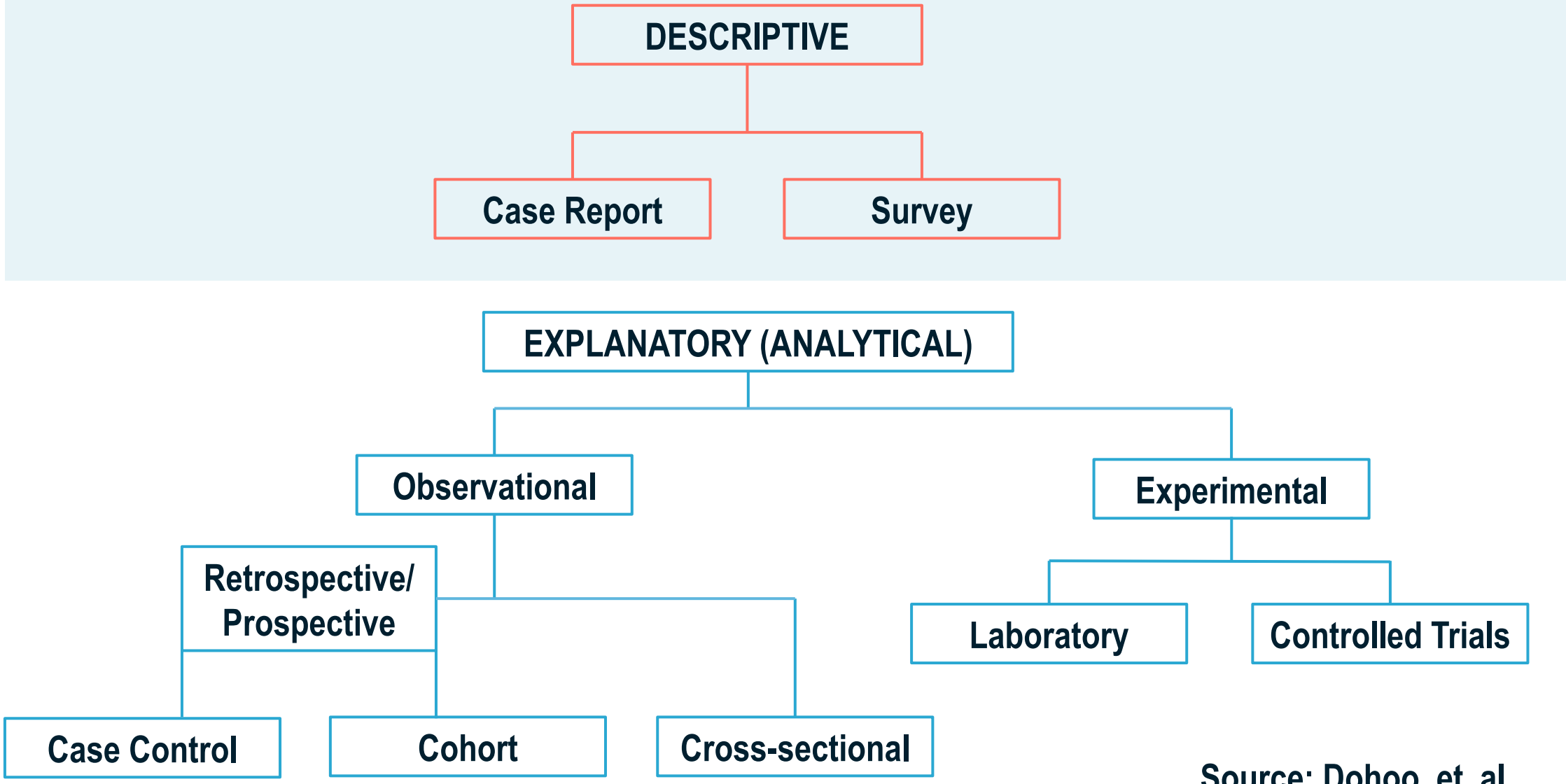
- **Risk based surveillance**

Surveillance and surveys

- **Surveillance:** systematic collection, analysis and dissemination of information on diseases

- **Surveys/studies**: tools for collection of information
  - Descriptive and explanatory
  - Experimental and observational
  - Retrospective and prospective
  - Cross section and longitudinal
  - Case – control and cohort

Survey (study) types



Source: Dohoo, et. al

Descriptive studies

- Describe the characteristics of the condition in time and/or space
- Does not always answer a research question
- No group comparison
- Forms the basis of hypothesis building
- Often limited resources needed compared to analytical studies

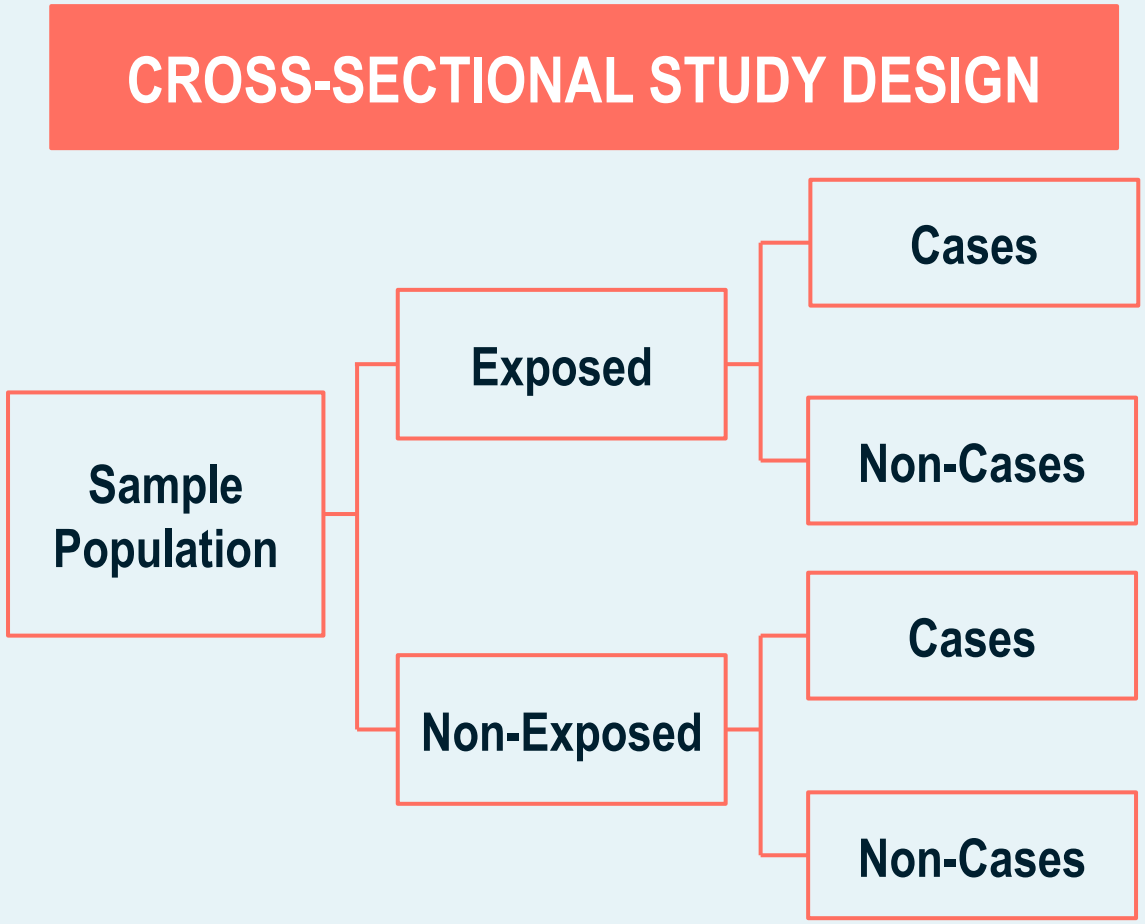
Analytical studies

- Involve analysis of the characteristics in question
- Hypothesis testing (test an association)
- Comparison of exposure/non-exposure (dose) for cases/non-cases
- Sometimes used to collect evidence for possible causation

Main characteristics of different types of survey

- **Cross sectional study**
  – prevalence study/outbreak investigation

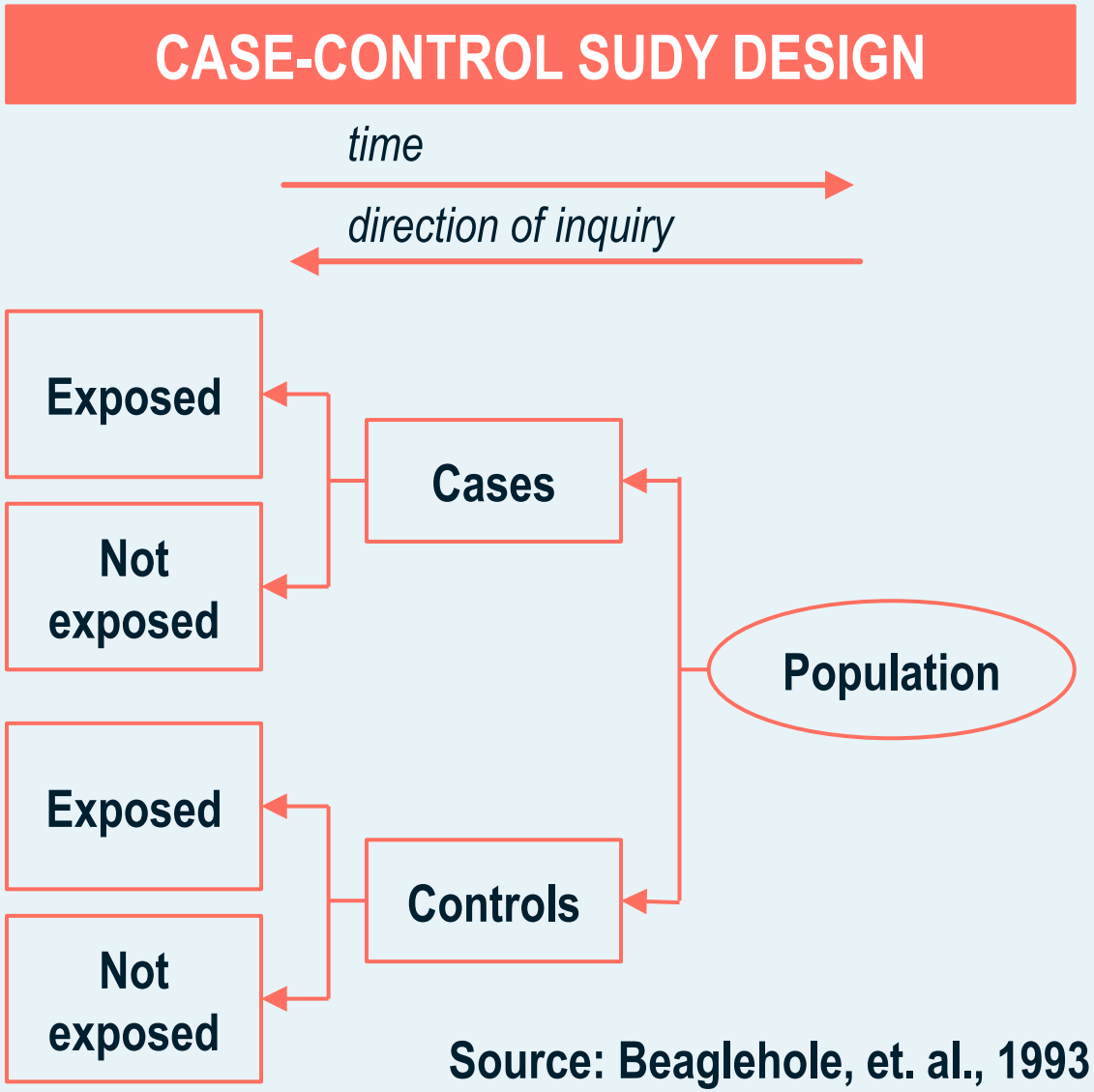
  - Useful for common conditions
  - Short time frame
  - Low cost
  - Good for multiple risk factors or outcomes
  - Prospective or retrospective
  - No measure of disease incidence



Main characteristics of different types of survey

Case – control studies

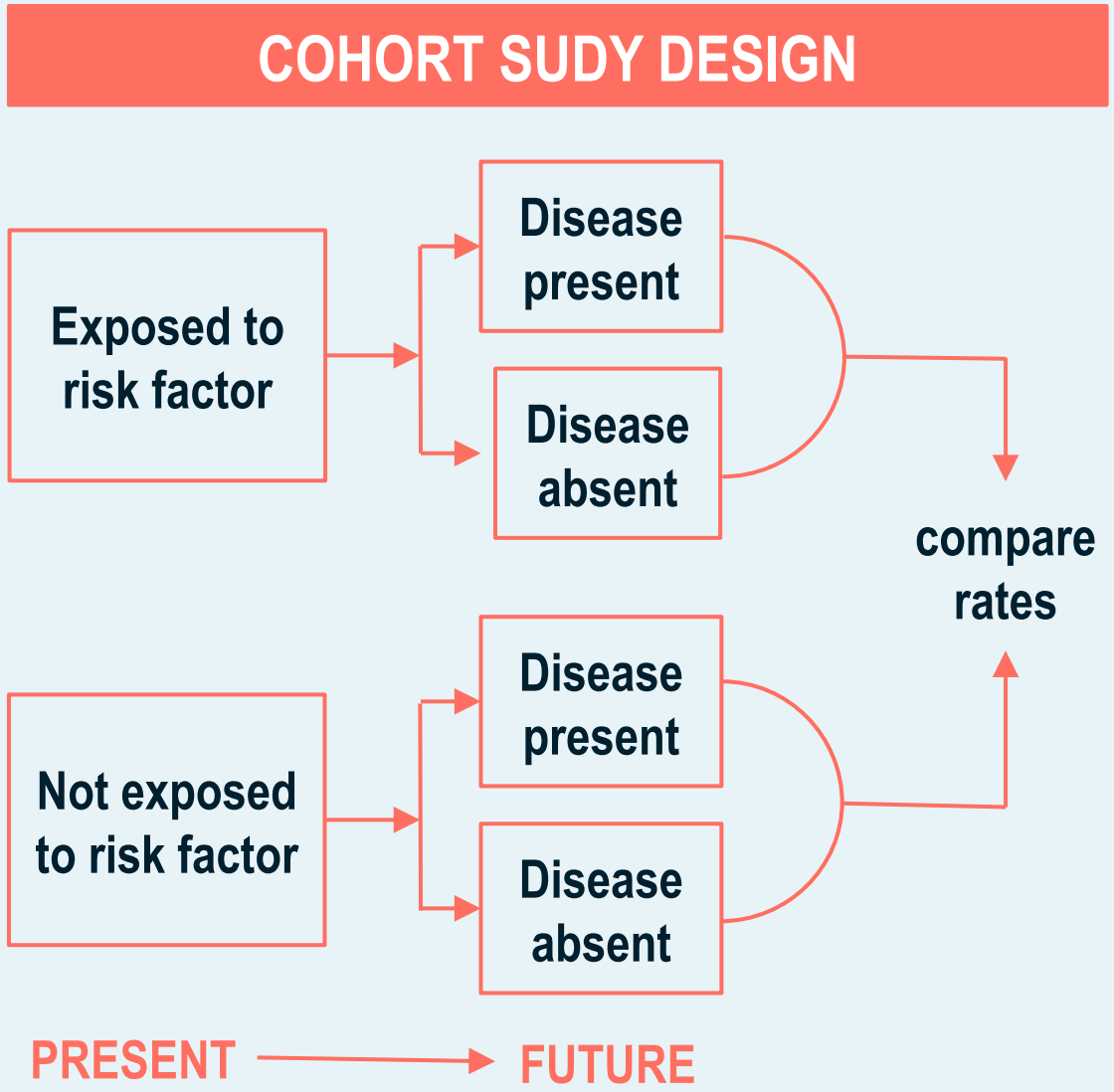
- Groups according to disease status
- Mostly retrospective
- Useful for rare diseases
- Short time – low cost
- Good for multiple risk factors, not for multiple outcomes
- Non representative, prone to bias
- Can not calculate RR



Main characteristics of different types of survey

Cohort study

- Groups according to exposure
- Mostly prospective
- Useful for rare exposures
- Long time – high cost
- Good for multiple outcomes, not for multiple diseases
- Direct estimate of effect (incidence, RR)
- Control of bias



Risk based surveillance

- More disease burden – less resources
- Combines epidemiology, public health, economy, trade consequences
- Looks for disease where is expected – risk assessment
- Higher benefits-cost ratio
- Estimation of effectiveness!?

Summary: study design

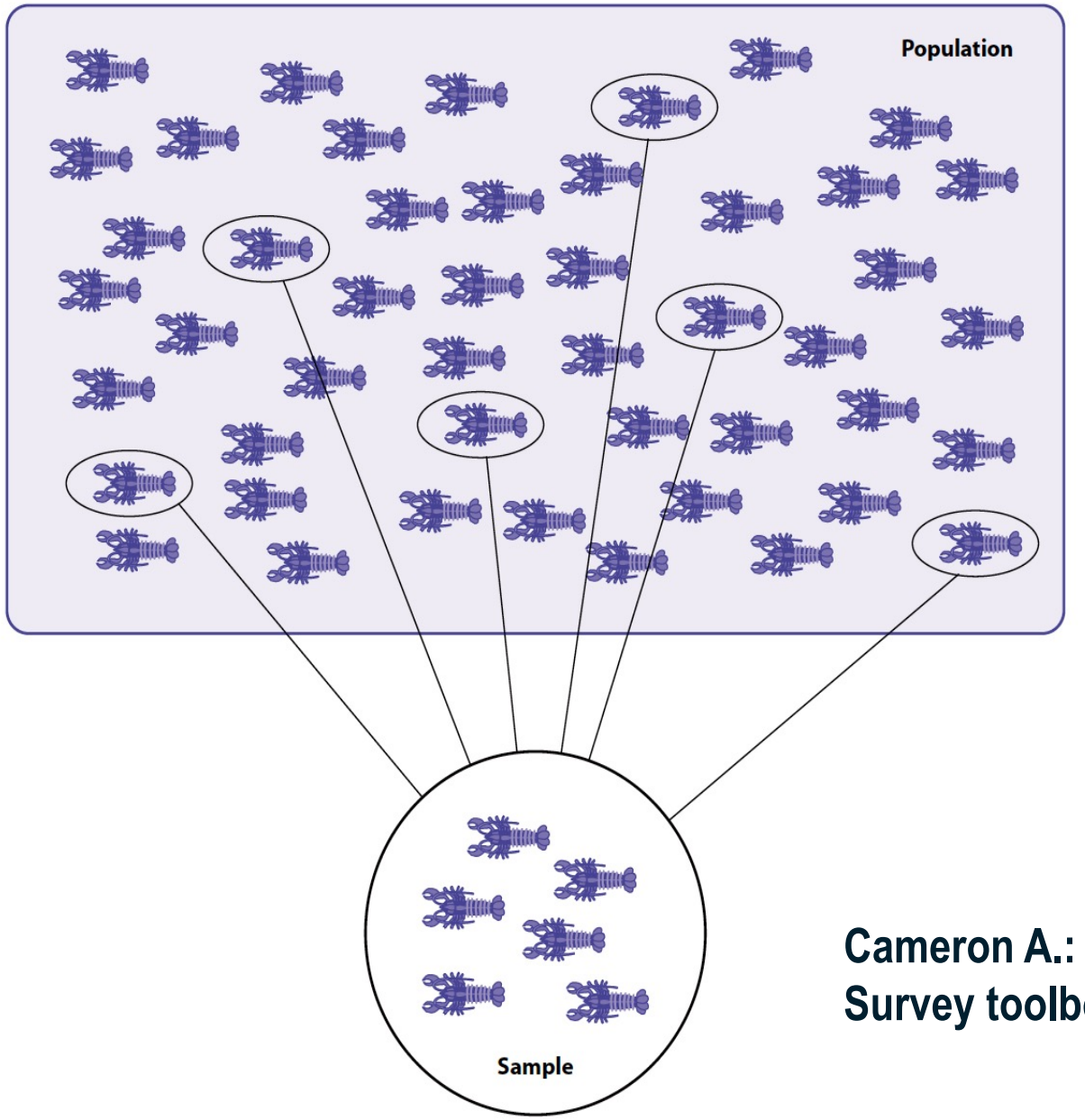
- Related with surveillance scenario (Checklist 1)
- Related with surveillance **objectives** (Checklist 2)
- Accounting for population data and disease clustering (Checklist 3 and 4)
- Accounting for **case definition** (Checklist 5)
- Most commonly used in aquatic disease surveillance: cross sectional study

Sampling

- For surveillance of aquatic diseases we commonly need to be able to select:
  - Fish farms - within country/zone
  - Fish within farms
  - Fish in water bodies

- Need to determine a sample size to (randomly) sample enough farms and fish to optimize your estimate
  - not too many –> waste of resources
  - not too few –> won't give you the results you want

- Sampling principles
- Sampling methods
- Sample size



Cameron A.:
Survey toolbox

Sampling principles

- Sampling is act of collecting samples to be analyzed via diagnostic testing (Level I, II III)
- The theory behind sampling is in most of cases when population is large it is not cost effective of feasible to sample whole population (census)
- When done properly sampling will give a representative samples for accurately assessing the health status of the population of interest

Sampling method

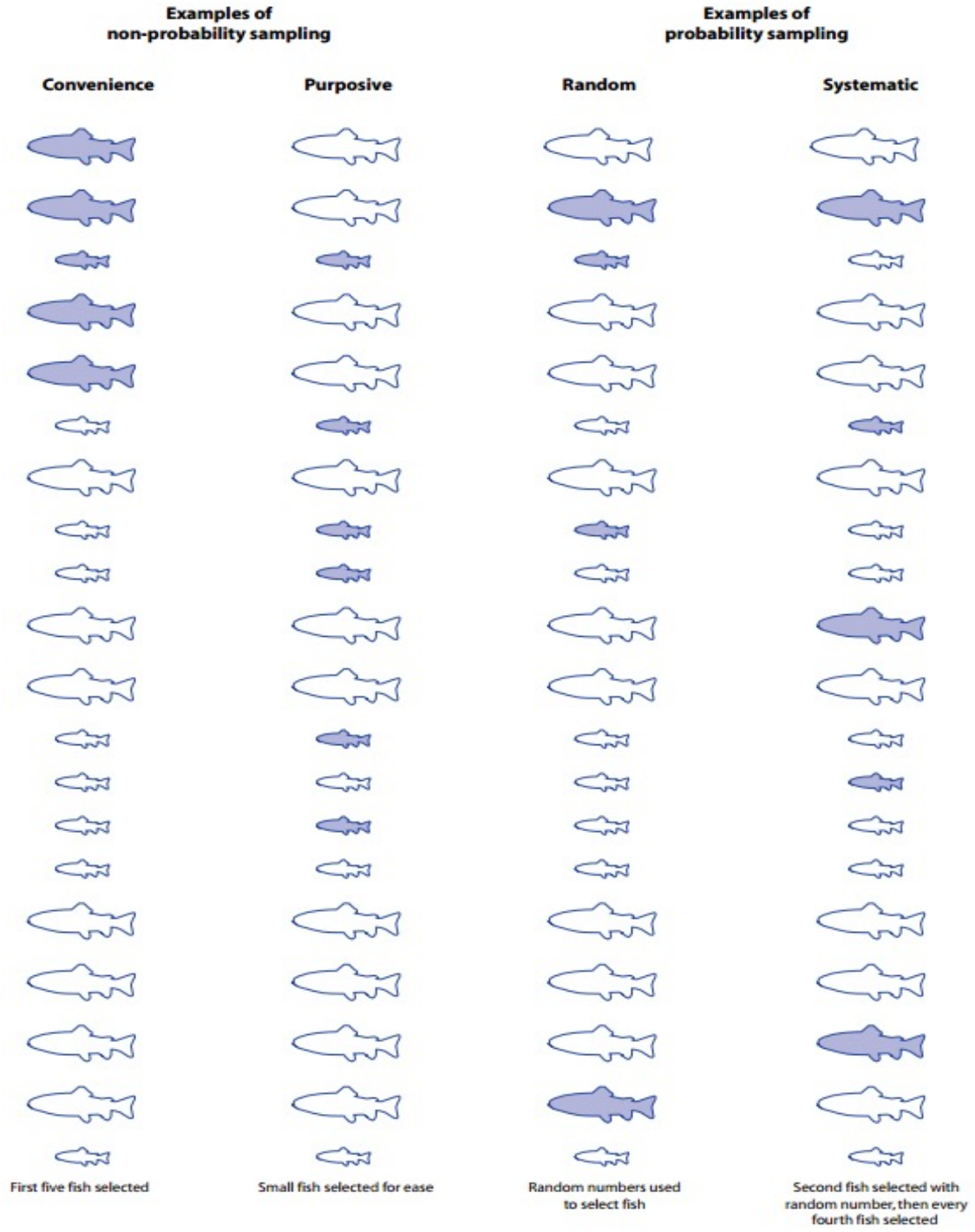
- Test entire population – census
- Test sample (provide the greatest likelihood that the sample will be representative of the population)
  - Representative from population
  - Non representative
- Large populations - sampling frame not available – multi-stage sampling

Representative sampling – each individual in population has same and equal probability being selected into sample

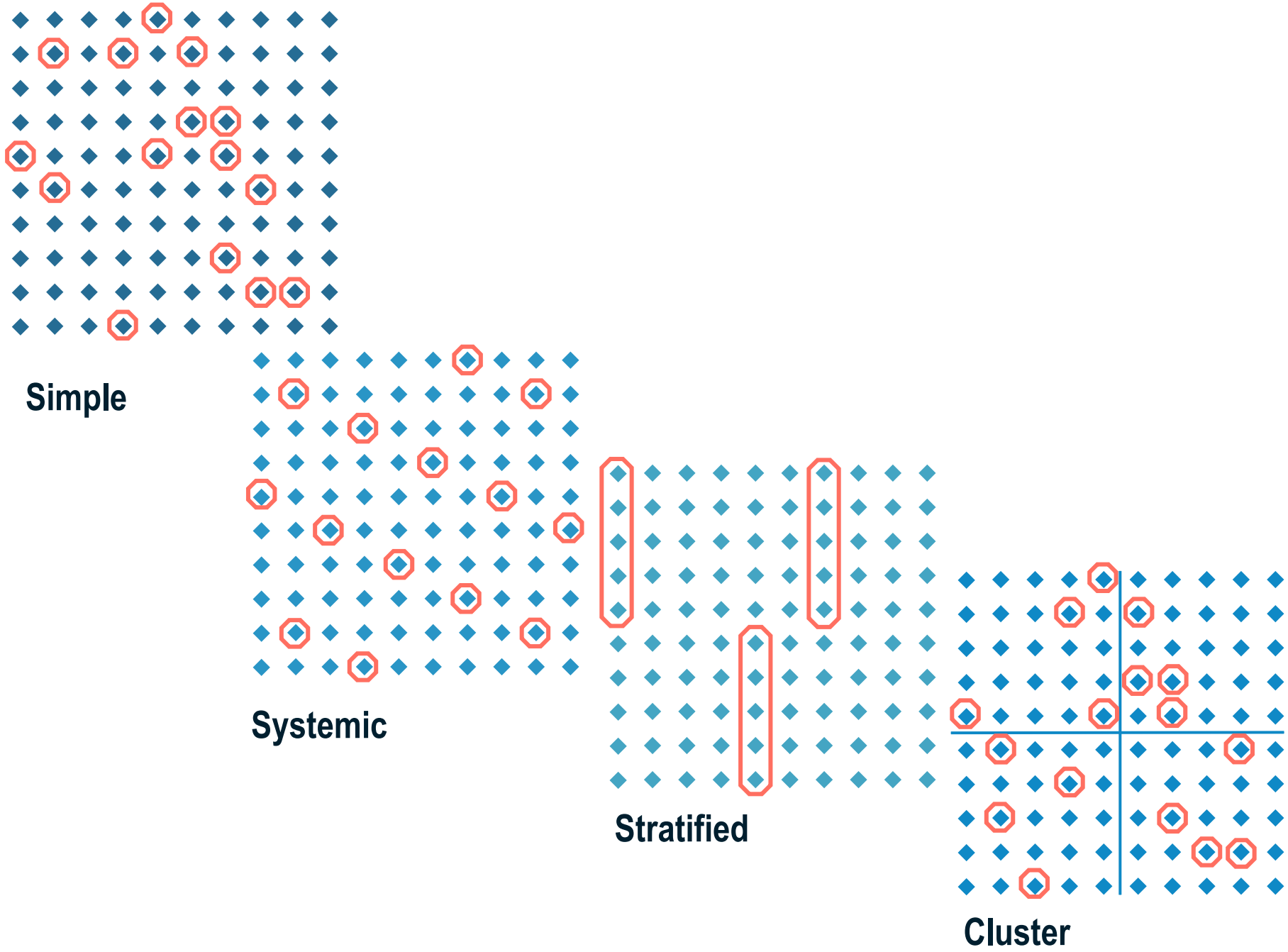
- Simple random sapling
- Systemic random sampling
- Stratified sampling
- Proportionally stratified sampling
- Spatial sampling

However:

- Random samples are rarely possible in fish farms
- Dead, sick, moribund, apparently healthy, healthy
- Different ease of sampling
- Fully representative samples are rarely achieved



Cameron A.: Survey toolbox


Simple
Systemic
Stratified
Cluster

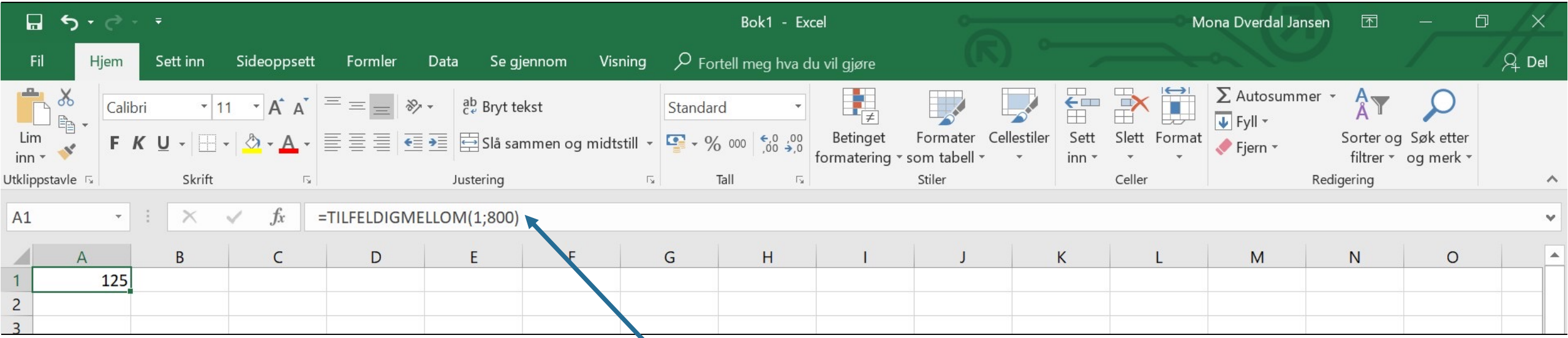
Random sampling

- Many farmed terrestrial animals are identified by an individual number
- Sampling frame in aquatic animals is different
- Random sampling can be applied using management practices (during grading or transfer of fish, during vaccination, during harvest)
- Most frequently used method in farm is capture sampling
- Likely to introduce some bias into the sample and it is important to be aware of the direction of bias.

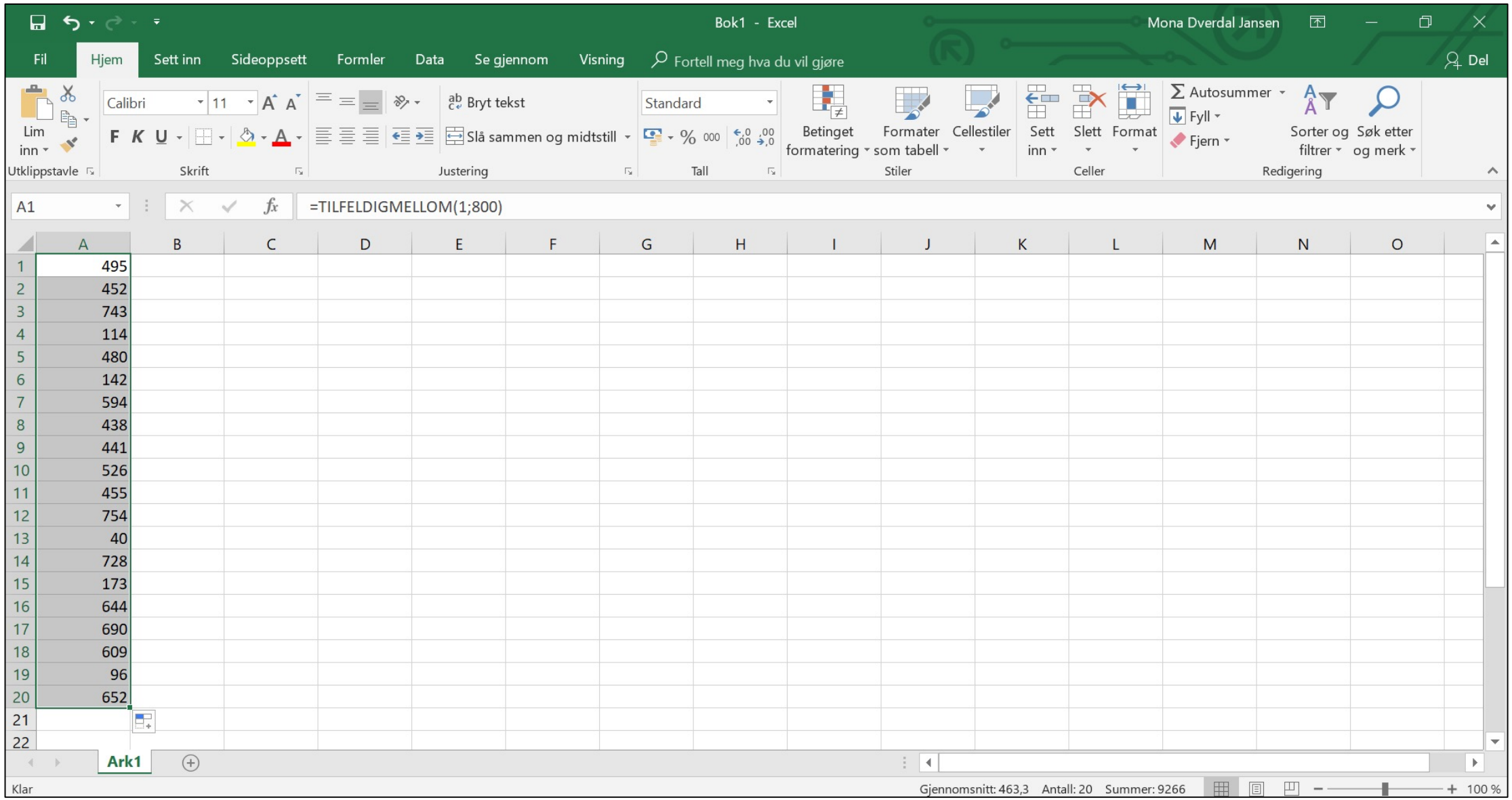
Simple random sampling

- Sampling from a population (e.g. farms) that is homogeneous in relation to disease distribution
- Requires a sampling frame
  - complete list of all sampling units available in the source population
  - requires individual identification for each sampling unit
- Each individual is selected using a random process so that each has an equal chance to be selected
- Collect the units with ID indicated by the random process
- NOTHING IS RANDOM IN RANDOM SAMPLING!!!
- Use any method to achieve random selection – documented and described
- Convenience sampling never acceptable

Random number generator



=RANDBETWEEN(1;800)


Bok1 - Excel
Mona Dverdal Jansen
Fil
Hjem
Sett inn
Sideoppsett
Formler
Data
Se gjennom
Visning
Fortell meg hva du vil gjøre
Del
Lim inn
Utklippstavle
Calibri
11
Skrift
Bryt tekst
Slå sammen og midtstill
Justering
Standard
Tall
Betinget formatering
Formater som tabell
Cellestiler
Stiler
Sett inn
Slett
Format
Celler
Autosummer
Fyll
Fjern
Sorter og filtrer
Søk etter og merk
Redigering
A1
=TILFELDIGMELLOM(1;800)
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
1 495
2 452
3 743
4 114
5 480
6 142
7 594
8 438
9 441
10 526
11 455
12 754
13 40
14 728
15 173
16 644
17 690
18 609
19 96
20 652
21
22
Ark1
Klar
Gjennomsnitt: 463,3
Antall: 20
Summer: 9266
100 %

Systematic random sampling

- Sampling from a population that is homogeneous in relation to disease distribution
- Units are sampled at a regular interval after a random start
  - No need for a sampling frame
- All units have to be sequentially accessible and each unit has an equal chance of being selected
  - Farms
  - Fish e.g. at slaughter, during vaccination
- Subject to bias in the random start, the interval or in the listing order

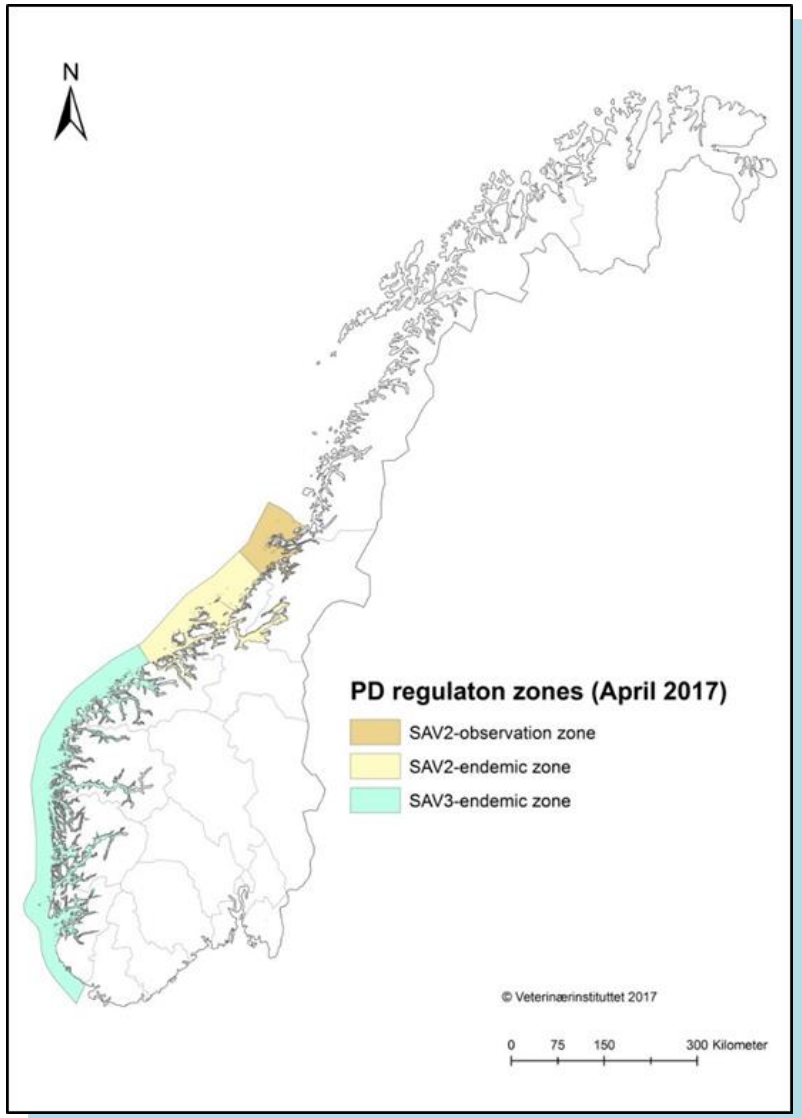
Systematic random sampling - example

1. **Calculate the Interval j = Total number of animals / Sample size**
   - We want 100 samples from 10 000 fish unit
   - j = 10 000/100 = 100
2. **Randomly pick a number from within j**
   - Starting with random in first 1-100 e.g. 27
3. **Sample every j^{th} animal**
   - E.g. 27, 127, 227,….., 9 927

Stratified random sampling

- Sampling from a population that is heterogeneous regarding the disease distribution

- The population can be categorized into groups where the units are more homogeneous than between groups



Spatial sampling

- May be used to generate a random sample when there is no sampling frame (for simple random sampling) and population is not lined up (for systematic sampling)
- Population has to be relatively stationary
- It is similar as simple random sampling, but instead of selecting individuals from a sampling frame, we select random location form area (randomly selected locations)

Non representative sampling

- Probability of a member of the population being selected in the sample is not known, and some groups are more likely to be selected than others

- The result from surveys using non representative (non probability) sampling are likely to be biased

- Sampling of available animals:
  - Convenience sampling
  - Purposive sampling
  - Haphazard sampling

Design of sampling

One stage sampling

- Sample frame available
- Random sample is collected (simple random sampling)

Two stage sampling

- **Design 1:**
  Probability proportional to size
- **Design 2:**
  Simple random sampling
- **Design 3:**
  Random geographic coordinate sampling

Now we know how we want to sample …

… how many should be included?

- How many farms?
- How many fish?

Sample size – general considerations

- Rarely possible to measure the entire population
- Parameters of interest calculated from subset of population
- Important to have enough samples to have sufficient certainty to aid decision-making
- Almost always restricted by the available budget
- Need to get sufficient sensitivity at farm level (or other unit-level)
  - Enough fish at each site
- Often result in:
  - Calculation of the cost per site
  - This then determines the maximum number of sites that can be sampled

Sample size calculation

- The number of units to be sampled from a population should be calculated using a statistically valid technique, considering key factors into account:
  - **Imperfection of diagnostic test/s (Se, Sp)**
  - **the design prevalence/s**
  - **the level of confidence**

- Other factors:
  - Population size (acceptable to assume infinitely large population)
  - The desired power of the survey

Sample size calculation

- Free software available at:
  - http://www.winepi.net/uk/index.htm
  - http://epitools.ausvet.com.au/content.php?page=home

- Published tables and equations in epidemiology text books, manuals

Design prevalence

- Design prevalence is not disease prevalence
- It form part of definition of the null hypothesis
- It is abstract statement of what may be present in nature
- Design prevalence: minimum expected prevalence, maximum acceptable prevalence, minimum detectable prevalence

Design prevalence (DP)

- OIE Manual specified DP for certain terrestrial disease, not for aquatic
- At the individual animal level, the DP should be based on the biology of the infection
- A suitable DP value at the animal level may be
- 1% - 5 % for infections that are transmitted slowly
- Over 5 % for more contagious infections
- At higher levels (cage, pond, farm, village, etc) the DP usually reflect the prevalence of infection that is practically and reasonably able to be detected by a surveillance system.
- A suitable DP prevalence value for the first level of clustering (e.g. Proportion of infected farms in a zone) may be up to 2%

- For example, if a DP of 50% is used when assessing prevalence of EUS on fish farms in country X, negative result will mean that we are 95% confident that the prevalence of infection is less than 50% !!!!!

- If a value is 1%, negative result means that it is still possible that infection exist, but at prevalence of no more than 1%

- Freedom of disease is zero prevalence (no one infected animals in country or zone)

Example of sampling plan in closed system aquaculture

Type of sampling	Frequency Example of sampling plan in closed system aquaculture	Class	Test type	Pathogens	Sample
Routine moribund	Quarterly, ongoing	Moribund or fresh mortality	Cell culture	Multiple exotic or emerging	70/year
Pathogen specific	Biannual, temporary	Random selection	RT-PCR	SAV, ISA, ...	165/year
Observational	Routine (daily, weekly, ongoing)	Moribund or fresh mortality	Veterinary investigation if above threshold	Pathogen clinical in this species	5 samples per event

EpiTools epidemiological calculators

This site is developed and maintained by Ausvet. The site is intended for use by epidemiologists and researchers involved in estimating disease prevalence or demonstrating freedom from disease through structured surveys, or in other epidemiological applications.

Surveillance utilities

- **1-Stage representative freedom surveys**
- **2-Stage representative freedom surveys**
- **Risk-based freedom surveys**
- **Random sampling from a population**
- **Estimating true prevalence**
- **Pooled prevalence calculator**
- **Survey Toolbox for livestock diseases and freedom in finite populations**
- **HerdPlus module for herd-sensitivity and freedom in finite populations**

Epidemiological studies

- **Sample size calculations**
- **Summarise categorical or continuous data**
- **Statistical significance testing**
- **Probability distributions**
- **Bioequivalence analysis**

Diagnostic tests

- **Application of diagnostic tests**

Case study data

- **GIS case study data from *Epidemiology for Field Veterinarians* text**
- ***Epidemiological Problem Solving* case studies and model answers**

Suggested citation: Sergeant, ESG, 2019. Epitools epidemiological calculators. Ausvet Pty Ltd. Available at: http://epitools.ausvet.com.au.

If you cite Epitools in your publications, please email the details or a copy of your paper to Evan Sergeant for inclusion in the reference list.

Sample size calculations

These utilities can be used to calculate required sample sizes to estimate a population mean or proportion, to detect significant differences between two means or two proportions or to estimate a true herd-level prevalence.

Epidemiological studies

To estimate a single proportion

To estimate a single mean

Two proportions

Two means with equal sample size and equal variances

Two means with unequal sample size and unequal variances

To estimate true prevalence (at animal or herd-level)

Sample size for a cohort study

Sample size for a case-control study

Sample size to demonstrate disease freedom

Sample size assuming perfect test specificity

Sample size for pooled sampling in a large population

Sample size to achieve target confidence of freedom

Design prevalence required to achieve target population sensitivity for given sample size

FreeCalc sample size calculation for imperfect tests

2-stage sampling, assuming **perfect** test specificity:

Least-cost sample sizes where cluster sizes are known (and select clusters for testing).

Least-cost sample sizes where cluster sizes are *NOT* known.

Sample size calculations

Sample size to estimate a proportion or apparent prevalence with specified precision

Estimated true proportion

0.3

Desired precision (+/-)

0.05

Confidence level

0.95

Population size (for finite populations)

Submit

This utility calculates the sample size required to estimate a proportion (or prevalence) with a specified level of confidence and precision.

Inputs are the assumed or estimated value for the proportion, the desired level of confidence, the desired precision of the estimate and the size of the population for limited population sizes. The desired precision of the estimate (also sometimes called the allowable or acceptable error in the estimate) is half the width of the desired confidence interval. For example if you would like the confidence interval width to be about 0.1 (10%) you would enter a precision of +/- 0.05 (5%).

The program outputs the sample sizes required to estimate the true value with the desired precision and confidence, for both an infinite population and for a population of the specified size. If population size is left blank or zero, only the sample size for an infinite population is calculated.

Note: Adjustment for finite population size may underestimate required sample size unless this is also taken into account when estimating variance and resulting confidence interval

Sample size to estimate a simple proportion (apparent prevalence)

Analysed: Tue Oct 08, 2019 @ 18:26 UTC

Inputs

Estimated Proportion	0.02
Desired precision of estimate	0.05
Confidence level	0.95
Population size	N/A

Results

Sample size required for specified inputs

Large population	31

Example of EUS (if probability sampling is achievable)

SAMPLING FRAME	EXIST			DON'T EXIST		
Surveillance scenario	Infected	Unknown	Considered free	Infected	Unknown	Considered free
Sampling	Probabilistic			Convenience/purposive/hap hazard		
Design Prevalence	>5 %	50 %	< 2%			
Dg test Sn/Pp			100 %			
Confidence	95 %	95%	95%			
Sample size*						
First stage (epi unit)	73	385	165			
Second stage (within farm)	165	165	165			
Methods	Simple random	Simple random	Simple random			

Sample size, based on defined parameters, calculated using http://epitools.ausvet.com.au/content.php?page=home

Example of TiLV
(if probability sampling is not achievable)

- Epidemiological Unit: A tilapia farm
- Unit of sampling: A mix of 30 moribund or sick tilapia from ponds at the farm.
- Total number of enrolled and participant farms: 40 to 60 tilapia farms, which should be visited twice (total field visits = 80 to 120, per country - at least 1+ve farm, at 2% Prev).
- Dynamic of sampling: 1st and 2nd semesters

Sampling techniques and equipment

- The survey team should know about, and fully understand, the common and traditional methods used for catching aquatic animals in survey area

- Understand the differences between the aims of local producers and survey staff.

- Capture techniques:
  - Cast net
  - Dip nets
  - Traps and nets

Summary: sampling requirements

- Study design decided
- Sampling unit defined
- Sampling method described
- Explained consideration regarding sampling size
- Sample selection process described

Food and Agriculture Organization of the United Nations

SUSTAINABLE DEVELOPMENT GOALS

Thank you for your attention!

Nihad Fejzic
nihad.fejzic@vfs.unsa.ba

Fernando Mardones
fomardones@gmail.com

TCP/INT/3707:
Strengthening biosecurity (policy and farm level) governance to deal with Tilapia lake virus

This was also made possible with the support of the Norwegian Agency for Development Cooperation under the project GCP/GLO/979/NOR Improving Biosecurity Governance and Legal Framework for Efficient and Sustainable Aquaculture Production.

Norad