



# Research, Seminar, and Workshop Club

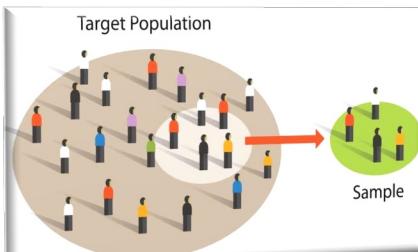
Department of Economics, Begum Rokeya University, Rangpur

## Sampling

Population, Sampling, Sample Size Determination & Ethics



Md. Siddikur Rahman, PhD  
Associate Professor  
Dept. of Statistics, BRUR



**SAMPLE SIZE DETERMINATION**

QuestionPro

One Research



## **LEARNING OBJECTIVES**

- **Learn the reasons for Sampling**
- **Understanding some key terms related to Sampling**
- **Develop an understanding of Sampling Methods**
- **Determination of Sample Size**
- **Ethics in Sampling/Research**

# INTRODUCTION

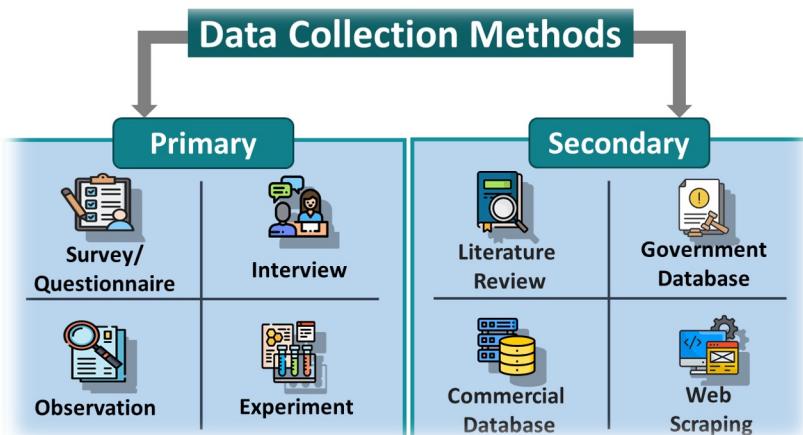
- The objective of statistical analysis is to draw conclusions about a population.
- Analysis is done based on data collected from the population under study.



# METHODS OF DATA COLLECTION

- **Census method or Complete enumeration survey**
- **Sampling method**

## SAMPLING OR CENSUS?



Types of Data Collection Methods		
Quantitative	Mixed	Qualitative
<ul style="list-style-type: none"><li>Number-based</li><li>Involves measuring and counting</li><li>More time is consumed for planning as compared to the analysis phase</li><li>Objective approach</li><li>Data is collected from:<ol style="list-style-type: none"><li>Surveys</li><li>Statistical experiments</li><li>Content analysis</li></ol></li><li>Example- The yield of the final product after extraction was observed to be a white powder. The yield of this product was 78%.</li></ul>	<ul style="list-style-type: none"><li>Combination of qualitative and quantitative data collection methods</li><li>The planning and analysis phase takes time</li><li>Data is collected from both the methods</li><li>Example- The final product after extraction was observed to be a white powder. The yield of this product was 78%.</li></ul>	<ul style="list-style-type: none"><li>Behavior-based</li><li>Involves interviews and observation</li><li>Less time is consumed for planning as compared to the analysis phase</li><li>Subjective approach</li><li>Data is collected from:<ol style="list-style-type: none"><li>Interviews</li><li>Case studies</li><li>Ethnography</li></ol></li><li>Example- The final product after extraction was a white powder.</li></ul>

# **KEY TERMS RELATED TO SAMPLING**

**POPULATION:** A population is the collection of all elements or items of interest in a particular study about which we wish to make an inference.

**SAMPLE:** A sample is a representative part of a population.

**SAMPLE SIZE**

Number of units contained in a sample.

**POPULATION SIZE**

Number of units which constitute the population.

# **KEY TERMS RELATED TO SAMPLING**

## **SAMPLING UNIT**

- **Well-defined, distinct and identifiable elements or group of elements on which observations can be made.**
- **For a household survey, the housing apartments or families may constitute the sampling units.**

## **SAMPLING FRAME**

**A complete list of all population sampling units.**

# **CONCEPT OF SAMPLING**

**SAMPLING:** Sampling is a technique used to select representative units from population units.

**Example:**

The Ministry of Health and Family Welfare might be interested in knowing the status of knowledge among the adult population in Rangpur Division on the danger of environmental pollution by interviewing a few selected adults in Rangpur.

# **CONCEPT OF SAMPLING**

## **Why Sample?**

- **Why not study everyone?**
- **Debate about Census vs. sampling**

## **Why Sample the Population?**

- **Less time**
- **Less cost**
- **More reliable result**
- **More detailed information**
- **Accuracy and Feasibility**

# SAMPLING OR CENSUS?

- Sampling is preferred to Census because the effort, money, and time required for carrying out a census will generally be extremely large.



## **STAGES OF SAMPLING PROCESS**

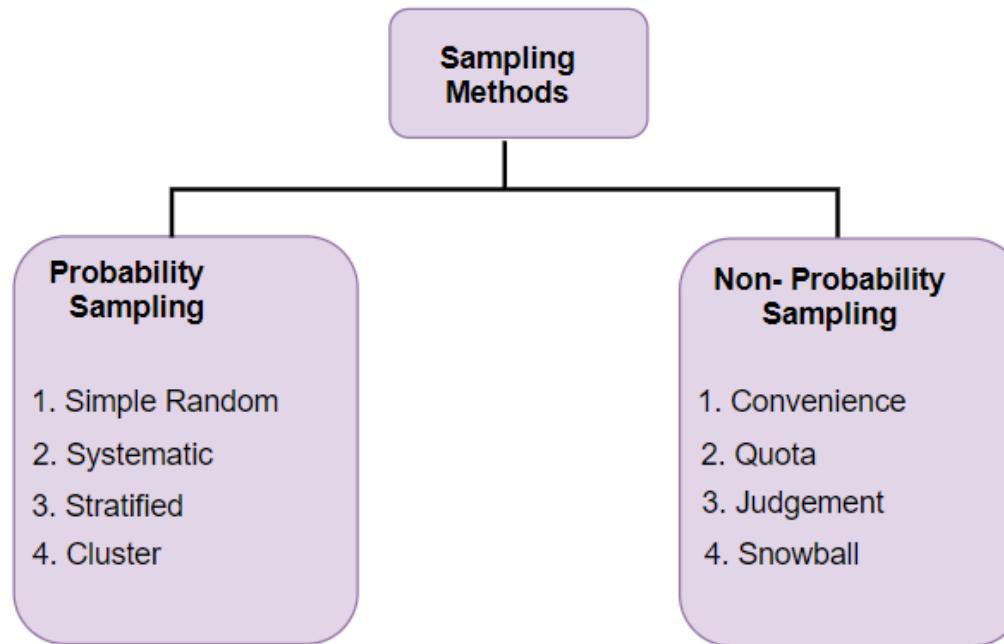
- **Defining the population of concern**
- **Specifying a sampling frame**
- **Specifying a sampling method for selecting items or events from the frame**
- **Determining the sample size**
- **Implementing the sampling plan**
- **Sampling and data collecting**
- **Reviewing the sampling process**

## **PROBLEMS IN SAMPLING**

- **What problems do you know about?**
- **What issues are you aware of?**
- **What questions do you have?**

- **Subject to sampling error**
- **A sample may seriously over-represent, under-represent, or even fail to represent the population.**
- **Estimates provided by such surveys are liable to a larger margin of errors.**

# SAMPLING METHODS



# **SIMPLE RANDOM SAMPLING**

- The probability of selection is the same for every unit in the population.
- Applicable when the population is small, homogeneous & sampling frame is readily available.

## **EXAMPLE OF A SRS?**

Suppose DESA is planning to identify the reasons why some customers are not regularly paying their electricity bills in the Rangpur division. A simple random sample of customers may be used to reach a conclusion.

## METHODS OF DRAWING SIMPLE RANDOM SAMPLING

Make a numbered list of the population



Decide the size of the sample



Select the samples using one of the following methods:

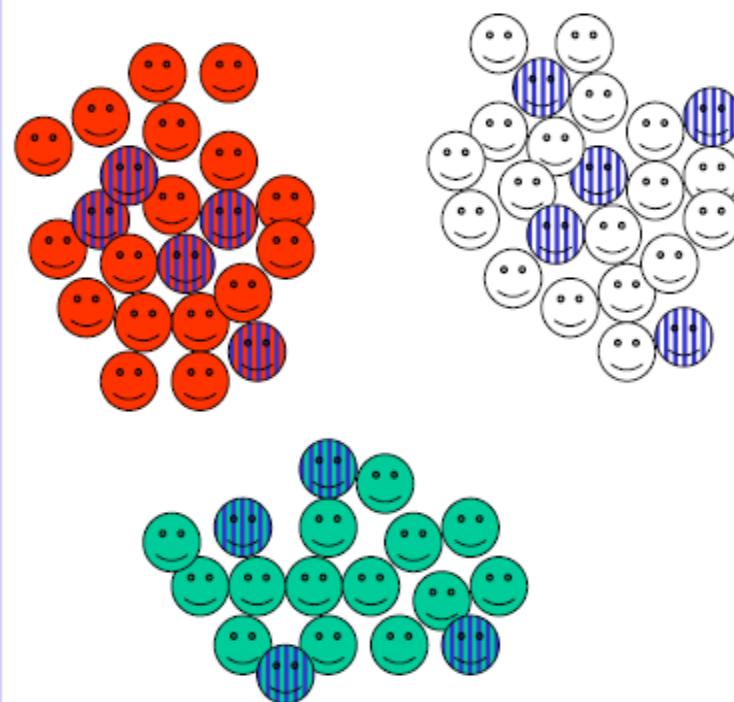
- Lottery method
- Use of Random Number Table
- Computer Random Number Generator
- Mechanical device

## STRATIFIED SAMPLING

- A population is divided into sub-groups called strata.
- A sample is selected from each stratum so that units within the strata are homogeneous but between strata are heterogeneous.
- **Example:** In studying the opinion on any issue among the students of BRUR, academic performance score (high, medium, low) or place of residence (urban, rural) may serve as stratification variables since each of these variables is believed to be related to the students' television viewing habits.

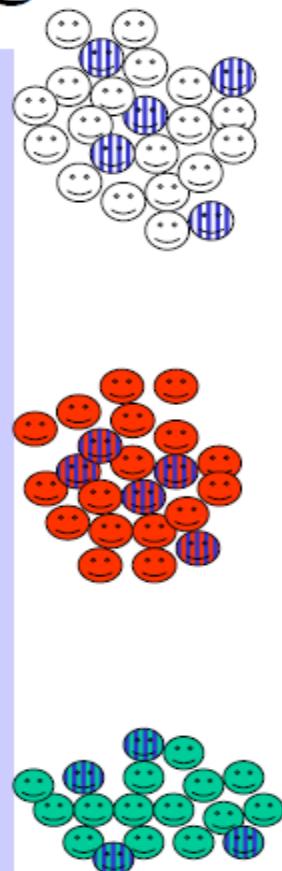
# Stratified Random Sampling-1

- Divide population into groups that differ in important ways
- Basis for grouping must be known before sampling
- Select random sample from within each group



## Stratified Random Sampling-2

- For a given sample size, reduces error compared to simple random sampling IF the groups are different from each other
- Tradeoff between the cost of doing the stratification and smaller sample size needed for same error
- Probabilities of selection may be different for different groups, as long as they are known
- Oversampling small groups improves inter-group comparisons



## SYSTEMATIC SAMPLING

- A systematic sampling method is applied when the population is very large, SRS is difficult, or lists are already grouped into sections or classes.



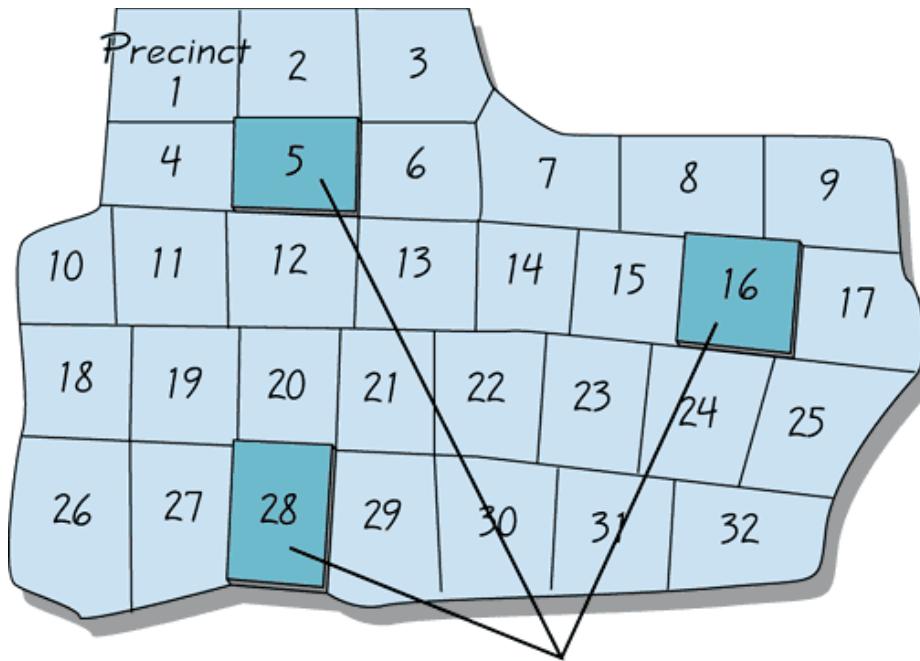
# **CLUSTER SAMPLING**

- The population is divided into clusters
- Clusters are randomly selected and a sample is collected by randomly selecting from each cluster.

## **Example:**

- Suppose we want to interview the industrialists about the environmental policies. Selecting the industrialists from the 64 district centres would be time-consuming and expensive.
- Instead, we can employ cluster sampling, using a simple random sampling of 64 districts.

# CLUSTER SAMPLING



## Cluster Sampling

- Divide the population into groups (called clusters), randomly select some of the groups, and then collect data from ALL members of the selected groups
- Used extensively by government and private research organizations

# MULTISTAGE SAMPLING

- The selection of sample is done in more than one stage

## Multi-stage sampling scheme

Stage	Sampling units	Referred to as
1	Thana	Primary sampling unit
2	Union	Secondary sampling unit
3	Village	Penultimate sampling unit
4	Household	Ultimate sampling unit

## **NONPROBABILITY SAMPLING**

**Non-probability sampling, unlike the probability sampling, is a non-random and subjective method of sampling where the selection of the sample depends on the personal judgment of the sampler.**

# **CONVENIENCE SAMPLING**

- **Restricted non-probability samples are known as convenience samples. Field workers who are restricted have the freedom to choose whomever they find, hence the name.**
- **Convenience sampling is the simplest and least reliable form of non-probability sampling.**

## **Example:**

**Only adult patients should be taken to study attitudes towards the family planning services provided by FWA because this is more convenient than taking a random sample of people in the villages.**

## JUDGEMENT SAMPLING

A non-probability sampling method that conforms to certain criteria is called judgment sampling.

**Example:** A cricket team will be formed from BRUR for any inter-university cricket tournaments. Then, we select, say, 15 students from various departments who actually play well and consider body fitness. Actually, for a specific purpose, we select samples.

## **QUOTA SAMPLING**

- Selection of a certain number of individuals from certain sub-groups or strata
- The formation of the strata is usually based on such characteristics such as sex, age, social status, region of residence etc.

## **SNOWBALL SAMPLING**

**The persons initially chosen for the sample are used as informants to locate other persons with the necessary characteristics to make them eligible for the sample through a referral network. Snowball sampling has been particularly used to study**

- **drug cultures,**
- **heroin addiction,**
- **teenage gang activities, etc.**

# SAMPLE SIZE DETERMINATION

- Having decided how to select our sample, we now have to determine our sample size.
- Researchers widely believe that the bigger the sample, the better the study.

**But it is wrong???**

The feasible sample size is determined by the availability of resources:

- Time
- Manpower
- Transport
- Money

# Determination of sample size for estimating mean

Sample size in case of infinite population

$$n = \frac{z^2 \sigma^2}{e^2}$$

Where n = sample size

Z = The value of the standard normal variate  
at a given confidence level

$\sigma$  = standard deviation of the population

e= acceptable error (level of statistical significance)

## Sample size in case of finite population

$$n = \frac{z^2 N \sigma^2}{e^2 (N - 1) + z^2 \sigma^2}$$

Where n = sample size

Z = The value of the standard normal variate  
at a given confidence level

$\sigma$  = standard deviation of the population

N = size of population

e = acceptable error (level of statistical significance)

# Determination of sample size for estimating proportions or percentage

**Sample size incase of infinite population (population is 10000 and above)**

$$n = \frac{z^2 pq}{e^2}$$

Where n = sample size

Z =The value of the standard normal variate  
at a given confidence level

P=sample proportion or percentage of incidence  
or prevalence q =1-p

e= acceptable error (level of statistical significance)

## Sample size incase of finite population

$$n = \frac{z^2 pqN}{e^2(N-1) + z^2 pq}$$

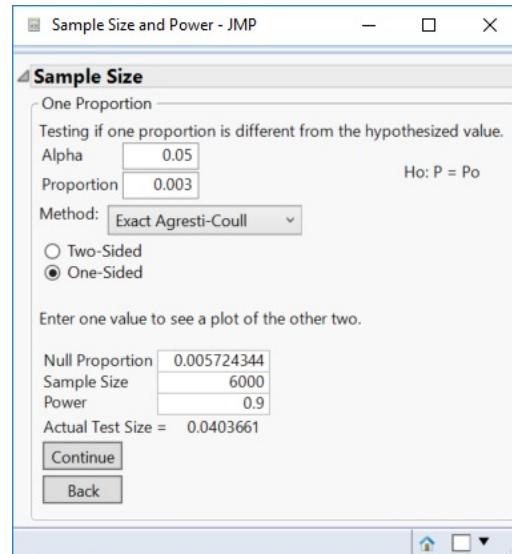
Where n = sample size

Z =The value of the standard normal variate  
at a given confidence level

P = sample proportion or percentage of incidence  
or prevalence q =1-p

e= acceptable error (level of statistical significance)

# ONLINE SAMPLE SIZE CALCULATOR



<https://www.calculator.net/sample-size-calculator.html>

<http://www.raosoft.com/samplesize.html>

# Declaration of Helsinki research Protocol?

- The Declaration of Helsinki is a set of ethical principles for research involving human subjects.
- Authors, editors and publishers all have ethical obligations with regard to the publication of the results of research.

# Manuscript/Thesis

1. Materials and methods
2. Study design and population
  - 2.1. Data collection
  - 2.2. Data management and statistical analysis



# Published Papers





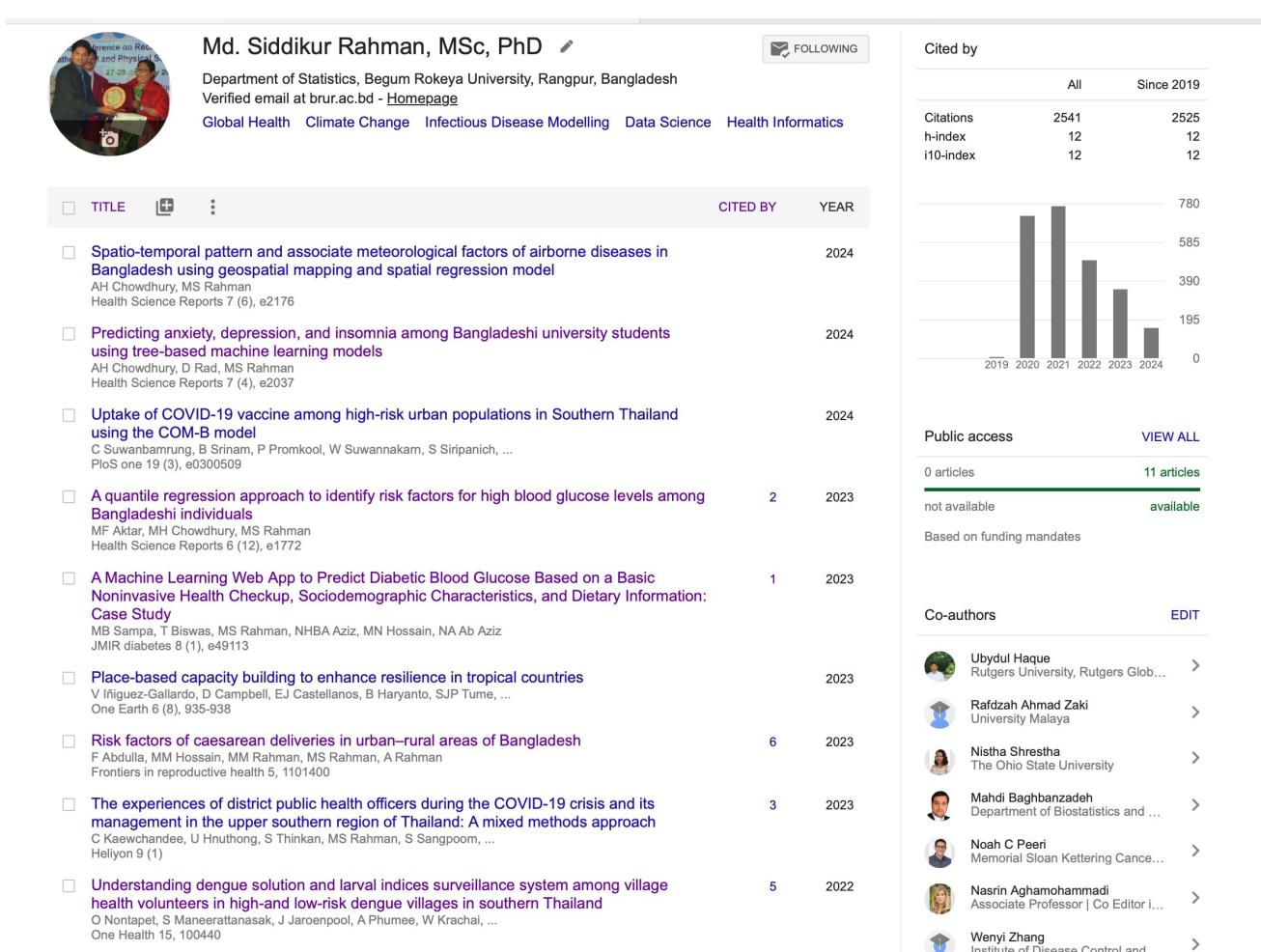
ELSEVIER

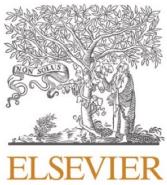
Springer

WILEY

OXFORD  
UNIVERSITY PRESS

OXFORD  
UNIVERSITY PRESS





Contents lists available at [ScienceDirect](#)

## Environmental Research

journal homepage: [www.elsevier.com/locate/envres](http://www.elsevier.com/locate/envres)



# Knowledge, attitudes, and practices on climate change and dengue in Lao People's Democratic Republic and Thailand

Md. Siddikur Rahman<sup>a,b</sup>, Hans J. Overgaard<sup>a,c,\*</sup>, Chamsai Pientong<sup>a,d</sup>, Mayfong Mayxay<sup>e,f,g</sup>, Tipaya Ekalaksananan<sup>a,d</sup>, Sirinart Aromseree<sup>a,d</sup>, Supranee Phanthanawiboon<sup>a,d</sup>, Sumaira Zafar<sup>h</sup>, Oleg Shipin<sup>h</sup>, Richard E. Paul<sup>i</sup>, Sysavanh Phommachanh<sup>e</sup>, Tiengkham Pongvongsa<sup>j</sup>, Nanthasane Vannavong<sup>k</sup>, Ubydul Haque<sup>l</sup>

<sup>a</sup> Department of Microbiology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

<sup>b</sup> Department of Statistics, Begum Rokeya University, Rangpur, Bangladesh

<sup>c</sup> Faculty of Science and Technology, Norwegian University of Life Sciences, P.O. Box 5003, Ås, Norway

<sup>d</sup> HPV & EBV and Carcinogenesis Research Group, Khon Kaen University, Khon Kaen, Thailand

<sup>e</sup> Institute of Research and Education Development (IRED), University of Health Sciences, Vientiane, Lao Democratic People's Republic

<sup>f</sup> Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Vientiane, Lao Democratic People's Republic

<sup>g</sup> Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

<sup>h</sup> Asian Institute of Technology, Bangkok, Thailand

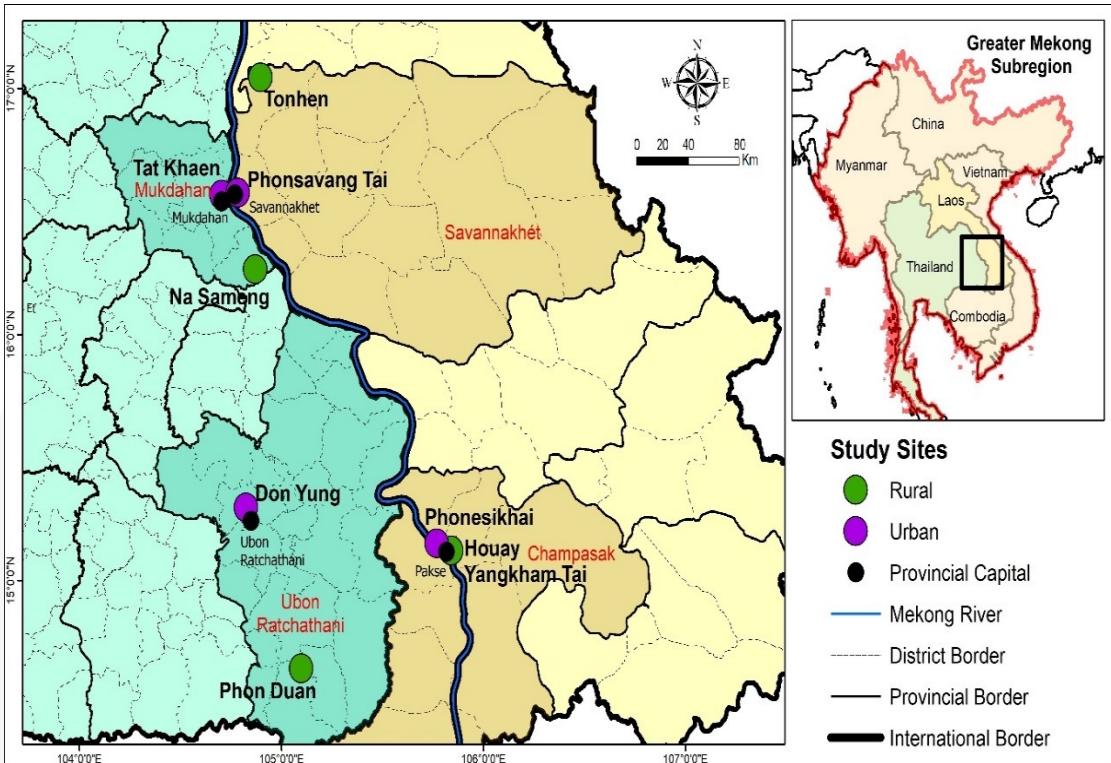
<sup>i</sup> Institut Pasteur, Functional Genetics of Infectious Diseases Unit, UMR 2000 CNRS, Paris, France

<sup>j</sup> Savannakhet Provincial Health Office, Savannakhet Province, Lao Democratic People's Republic

<sup>k</sup> Champasak Provincial Health Office, Pakse, Lao Democratic People's Republic

<sup>l</sup> Department of Biostatistics and Epidemiology, University of North Texas Health Science Center, North, TX, Fort Worth, TX, 76107, USA

# KAP study sites & data collection



**HH Survey**  
Thailand (10, Sep-15 Oct, 2018)  
Laos (30 November, 18-14 Feb, 19)  
Total: 720 households

**KAP Survey**  
Thailand (23, Feb-27 April, 2019)  
Laos (28 Feb -12 March, 2019)  
Total: 719 households

**Government Offices**  
Total: 20



One Health 13 (2021) 100358



ELSEVIER

Contents lists available at ScienceDirect

## One Health

journal homepage: [www.elsevier.com/locate/onehlt](http://www.elsevier.com/locate/onehlt)



# Mapping the spatial distribution of the dengue vector *Aedes aegypti* and predicting its abundance in northeastern Thailand using machine-learning approach

Md. Siddikur Rahman <sup>a,b, \*\*</sup>, Chamsai Pientong <sup>a,c</sup>, Sumaira Zafar <sup>d</sup>, Tipaya Ekalaksananan <sup>a,c</sup>, Richard E. Paul <sup>e</sup>, Ubydul Haque <sup>f</sup>, Joacim Rocklöv <sup>g</sup>, Hans J. Overgaard <sup>a,h,\*</sup>

<sup>a</sup> Department of Microbiology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

<sup>b</sup> Department of Statistics, Begum Rokeya University, Rangpur, Rangpur-5404, Bangladesh

<sup>c</sup> HPV & EBV and Carcinogenesis Research Group, Khon Kaen University, Khon Kaen, Thailand

<sup>d</sup> Environmental Engineering and Management Program, Asian Institute of Technology, Pathumthani, Thailand

<sup>e</sup> Unité de la Génétique Fonctionnelle des Maladies Infectieuses, Institut Pasteur, CNRS UMR 2000, 75015 Paris, France

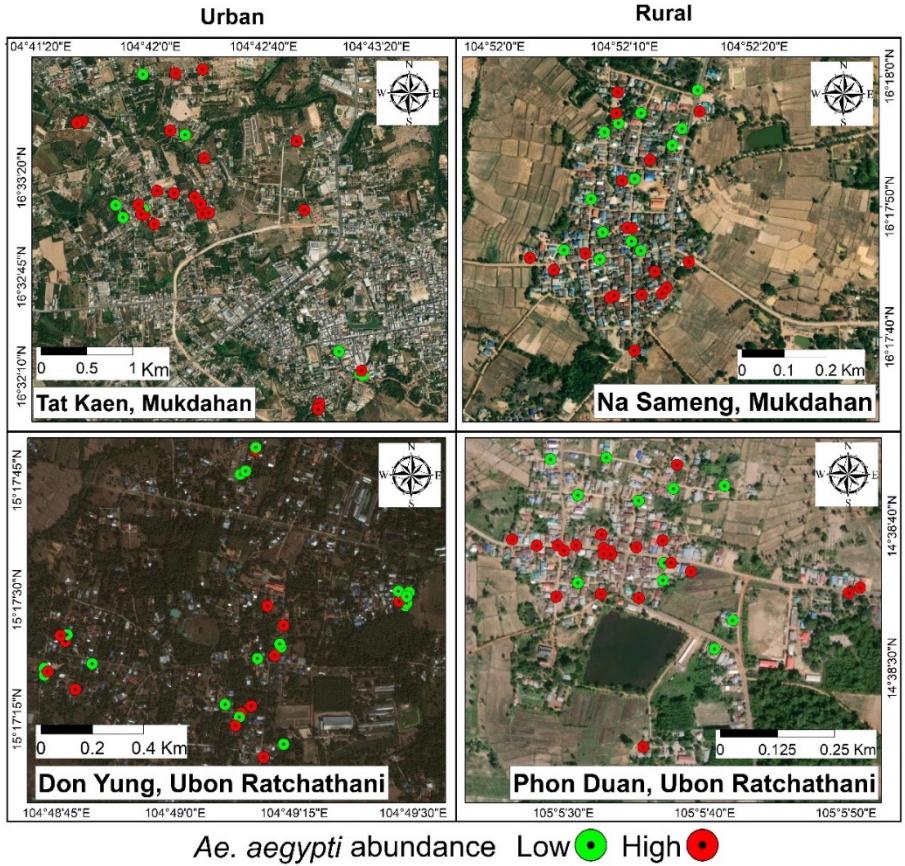
<sup>f</sup> Department of Biostatistics and Epidemiology, University of North Texas Health Science Center, Fort Worth, TX 76177, USA

<sup>g</sup> Department of Public Health and Clinical Medicine, Umeå University, 90187 Umeå, Sweden

<sup>h</sup> Faculty of Science and Technology, Norwegian University of Life Sciences, P.O. Box 5003, Ås, Norway



**Figure 1.** Location of study sites and Spatial distribution of the dengue vector, female adult *Ae. aegypti* abundance



- A total of 1,066 females adult *Ae. aegypti* mosquitoes were collected in both urban (551) and rural (515) sites.
- The numbers collected during the wet season (May–October) were 345 and 284, respectively.
- Out of 128 households, 75 (59%) of the households had a high abundance of females adult *Ae. aegypti* mosquitoes.



Thanks



Any  
Questions?

[siddikur@brur.ac.bd](mailto:siddikur@brur.ac.bd)

