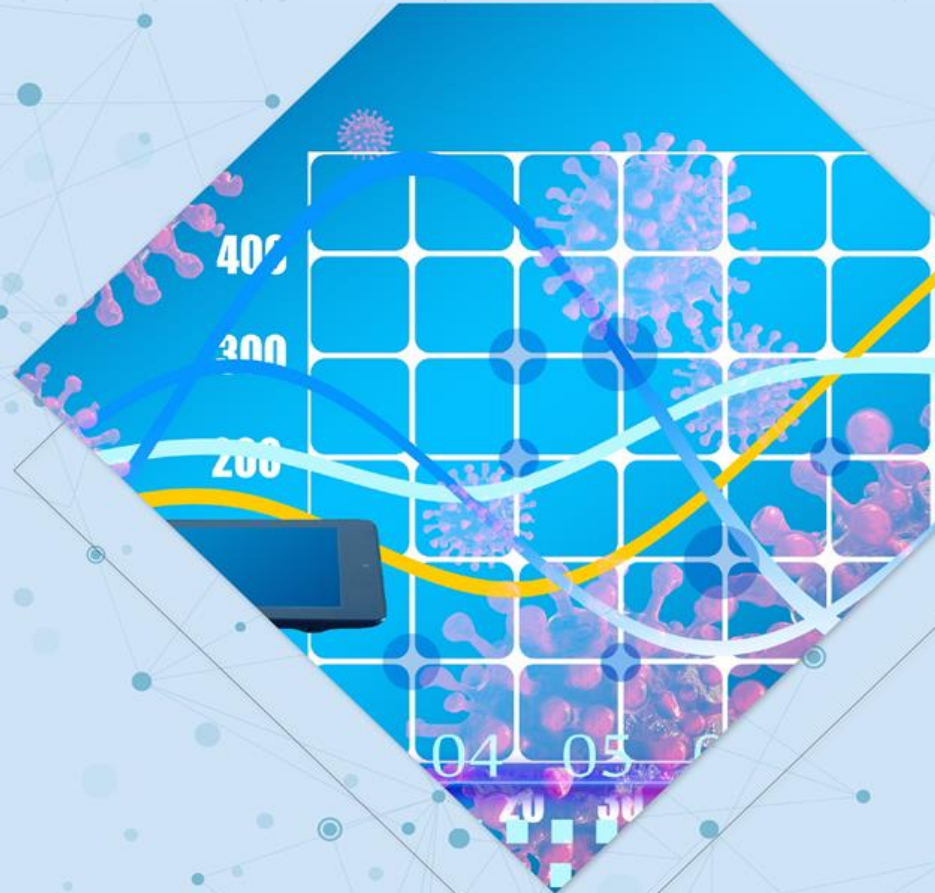




CHAPTER 3

Biostatistics

(ชีวสถิติ)



Episode 1



Wanpen Waelveerakup, Dr.P.H.
Email: wanpenw@webmail.npru.ac.th



THE CONTENTS OF THIS CHAPTER



1

Overview of biostatistics

2

The epidemiological measurements

2.1 Measures of disease frequency



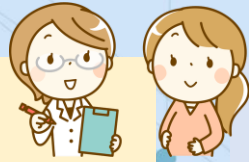
Overview of Biostatistics



What are statistics and biostatistics?



❖ **Statistics** is the practice or science of collecting and analyzing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample.



❖ **Biostatistics** is the application of statistical techniques to scientific research in **health-related fields**, including medicine, biology, and public health, and the development of new tools to study these areas.

Biostatistics



- **Biostatistics deals with the study and methods of collection, presentation, analysis, and interpretation of data for biological research.**
- Biostatistics is considered one of the foundational disciplines in public health.
- **Biostatisticians conduct quantitative and qualitative research to help identify population health trends and risk factors.**



Types of biostatistics



Descriptive statistics

พรรณนา

Descriptive statistics summarize and organize the characteristics of a data set.

A data set is a collection of responses or observations from a sample or entire population (ประชากรทั้งหมด).

Two types of biostatistics

Inferential Statistics

อ้างอิง

Inferential statistics have two main uses:

1. Making estimates about populations
 2. Testing hypotheses to draw conclusions about populations
- ✓ Hypothesis tests
 - ✓ Confidence intervals

Inferential Statistics

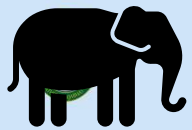
Inferential statistics help you to make predictions (“inferences”) from that data. With inferential statistics, you take data from samples and make generalizations about a population.



Symbols used for Corresponding Statistics and Parameters

| Statistical Focus | Statistics (in the Sample) | Parameter (in the Population) |
|----------------------------|----------------------------|-------------------------------|
| Mean | \bar{x} | μ |
| Variance | s^2 | σ^2 |
| Standard Deviation | s | σ |
| Proportion | p | P |
| Product-moment Correlation | r | ρ |

Source: <https://www.pinterest.com/pin/326299935485508303/>



Types of data

Qualitative or
Categorical data

Quantitative or
Numeric data

Normal scale

Ordinal or
Ranking scale

Interval scale

Ratio scale

- ✓ Gender
- ✓ Home town

- ✓ Socio-economic status ("low", "middle", "high")
- ✓ Education level ("high school", "BS", "MS", "Ph.D.")

✓ Temperature

✓ pH

✓ SAT score

✓ Pulse

✓ Weight

✓ Length



There are four steps of the statistical process:

1. data collection,
2. data analysis,
3. data presentation, and
4. interpretation.

data analysis

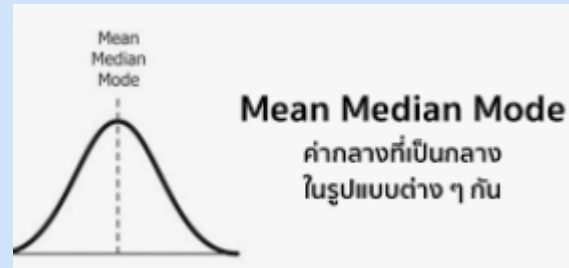


อ้างอิง

Inferential Statistics

Two types of biostatistics

- Mean (ค่าเฉลี่ย)
- Median (มัธยฐาน)
- Mode (ฐานนิยม)



Descriptive statistics

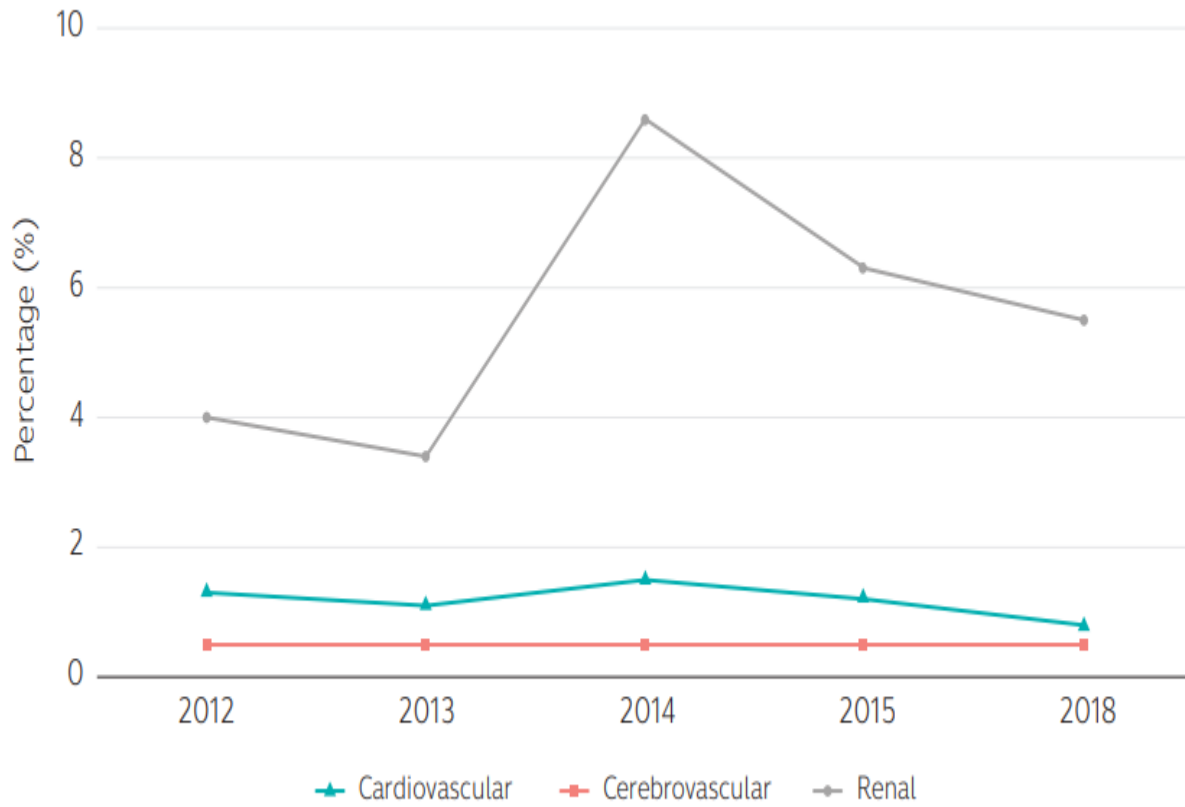
พรรณนา

- Standard deviation
- Range (พิสัย)
- Variance (ความแปรปรวน)

| Comparison test | Parametric? | What's being compared? | Samples |
|------------------------------------|-------------|------------------------|------------|
| t-test | Yes | Means | 2 samples |
| ANOVA | Yes | Means | 3+ samples |
| Mood's median | No | Medians | 2+ samples |
| Wilcoxon signed-rank | No | Distributions | 2 samples |
| Wilcoxon rank-sum (Mann-Whitney U) | No | Sums of rankings | 2 samples |
| Kruskal-Wallis H | No | Mean rankings | 3+ samples |

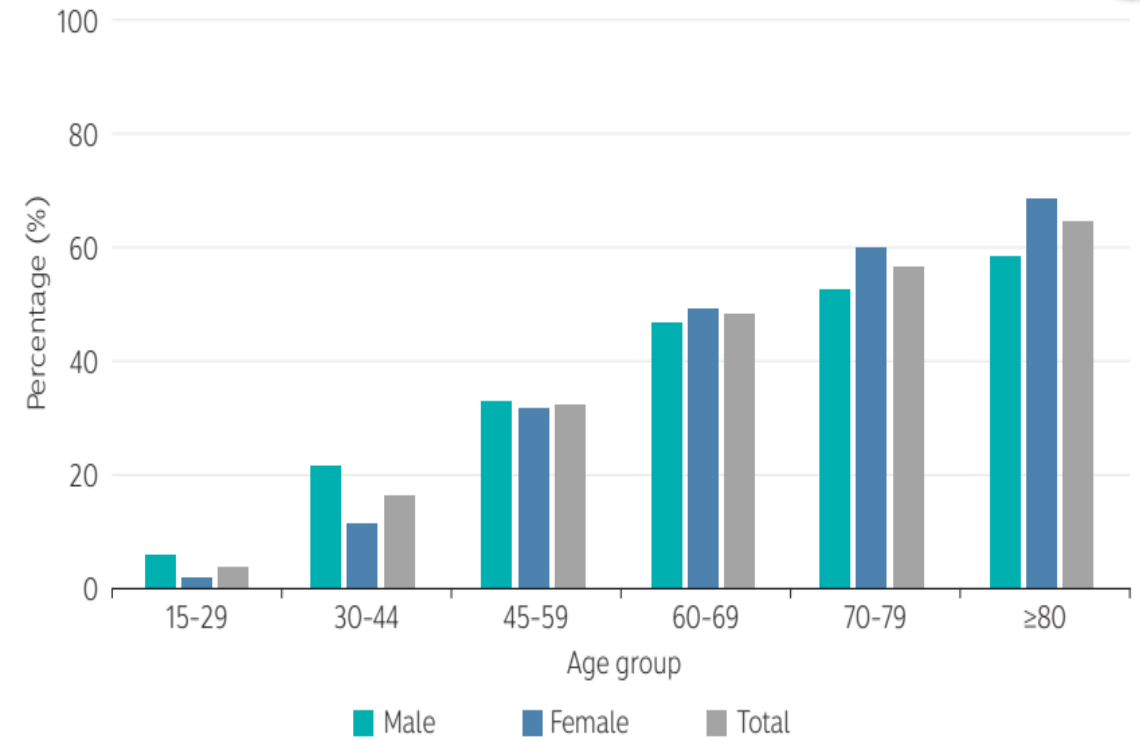
| Correlation test | Parametric? | Variables |
|---------------------------------|-------------|----------------------------------|
| Pearson's <i>r</i> | Yes | Interval/ratio variables |
| Spearman's <i>r</i> | No | Ordinal/interval/ratio variables |
| Chi square test of independence | No | Nominal/ordinal variables |

Fig. 7.5. Frequency of complications among patients with hypertension, 2012–2018, Thailand



Source: National Health Security Office, 2012–2018

Fig. 1.2. Prevalence of hypertension by age and sex, Thailand, 2014



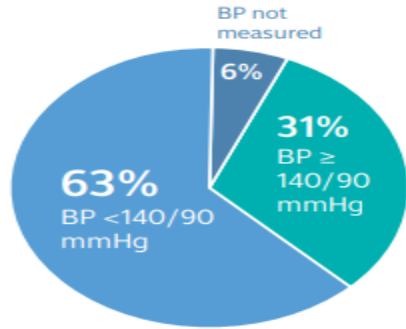
Source: National Health Examination Survey 2014

Source:
 HYPERTENSION CARE IN THAILAND - WHO |
 World Health ...<https://apps.who.int>

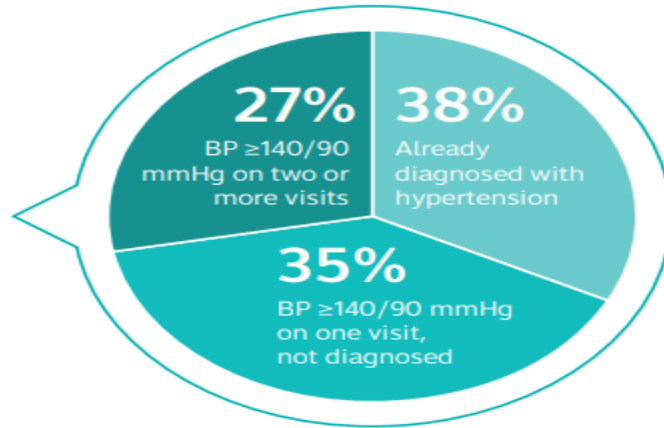
Data Presentation

Fig. 4.1. Blood pressure of adult Thais visiting an MoPH health facility one or more times during 2017, Thailand

Distribution of outpatients aged >15 years, by blood pressure measurement, 2017 (N=32 512 861)



Status of hypertension diagnosis among outpatients with raised blood pressure, 2017 (N=9 962 528)



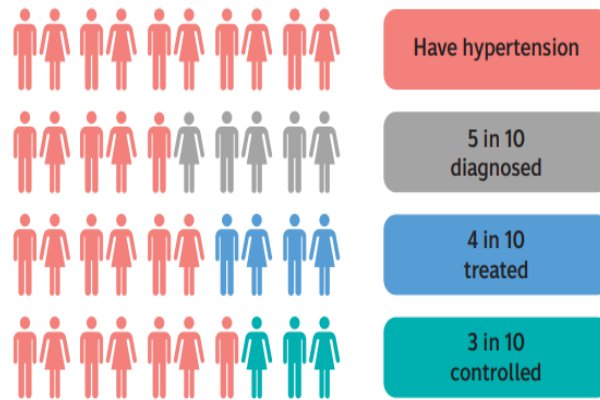
Source: Health Data Centre

Data Presentation



Source:
HYPERTENSION CARE IN THAILAND - WHO | World Health
...<https://apps.who.int>

Fig. 7.1. Only 3 out of 10 persons with hypertension have a controlled blood pressure



Source: Aekplakorn W, Hathaichanok P, Kanittha T, Warapone S. The 5th National Health Examination Survey, 2557 (2014). Nonthaburi: Health System Research Institute; 2016 (<https://www.hsri.or.th/researcher/research/new-release/detail/7711>, accessed 5 November 2019).

Table 3.1. Period for which medicines are dispensed by size of health facility and status of blood pressure control, Thailand, 2018

| | BP controlled | | BP uncontrolled | |
|---------------------|-----------------|---------------|-----------------|---------------|
| | Average (weeks) | Range (weeks) | Average (weeks) | Range (weeks) |
| Provincial hospital | 12 | 10–16 | 6 | 1–10 |
| District hospital | 10 | 6–12 | 3 | 2–4 |
| Health centre | 10 | 8–12 | 4 | 3–6 |

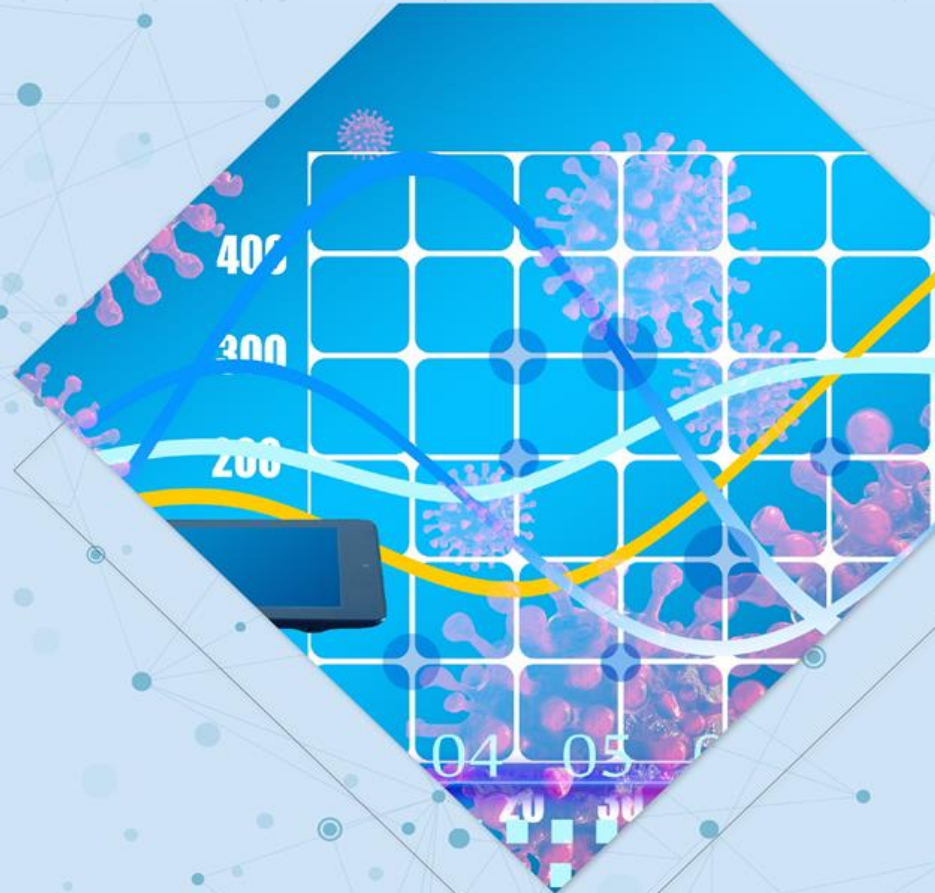
Source: Srivanichakorn S, Angkurawaranon C, Sanchaisuriya P, Theptien B-O, Trakoonmuang P. Rapid assessment of the health service system for hypertension care in Thailand. 2019.



CHAPTER 3

Biostatistics

(ชีวสถิติ)



Episode 1

1.2 

Wanpen Waelveerakup, Dr.P.H.
Email: wanpenw@webmail.npru.ac.th





The epidemiological measurements



TYPES OF EPIDEMIOLOGICAL MEASUREMENTS



1. Measures of disease frequency

EP.1



2. Measures of association

EP.2

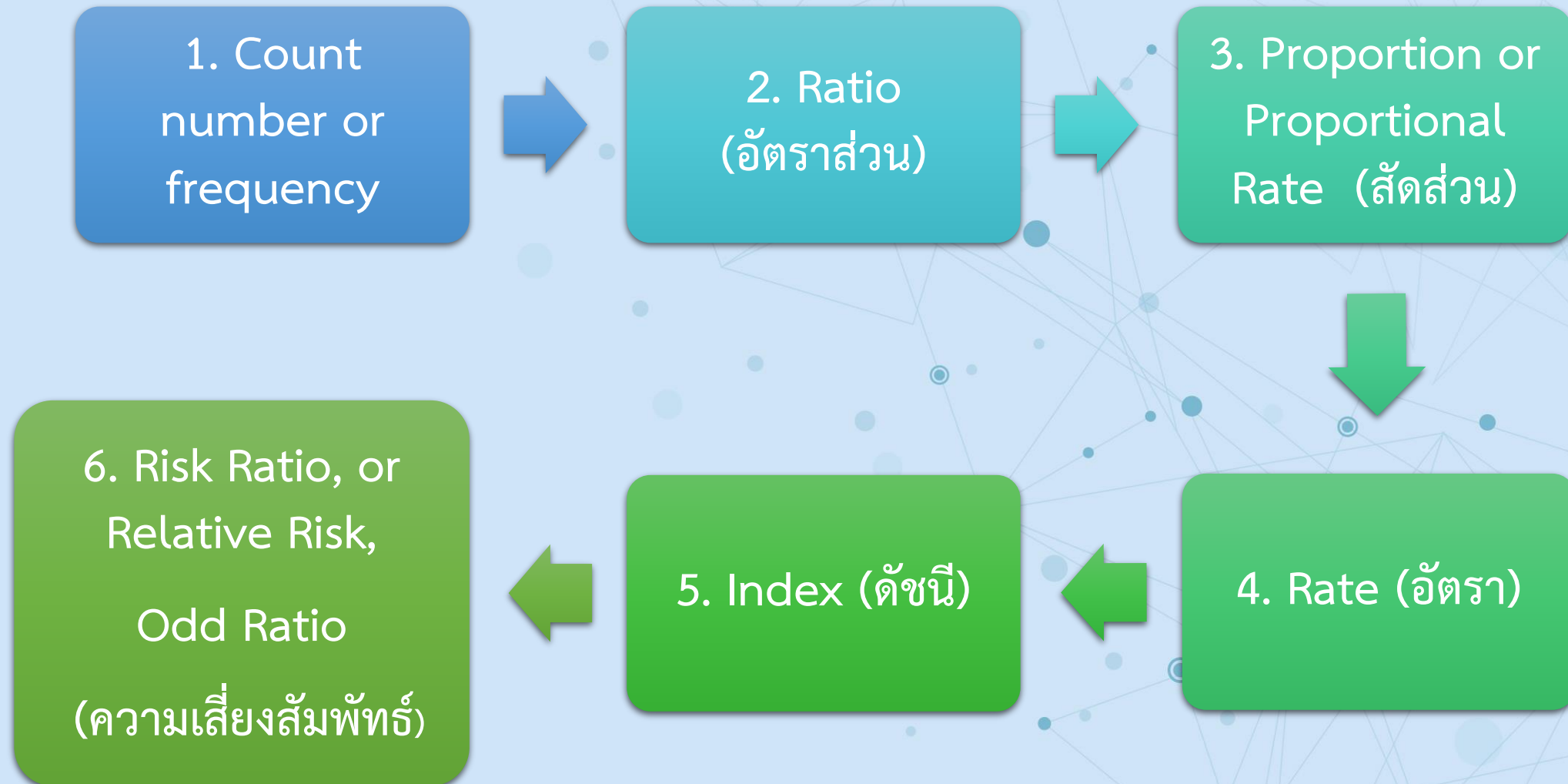


3. Measures of potential impact

EP.2



Basic tools for epidemiological measurements



Basic tools for epidemiological measurements



1. Count number or frequency

Counting the number of actual events in the population of interest or have some characteristics in common at a given area and during the study period.

2. Ratio (อัตราส่วน)

The relationship between two groups or amounts expresses how much bigger one is than the other

The ratio formula for any two quantities say a and b is given as,

$$\mathbf{a : b = a / b.}$$

Since a and b are individual amounts for two quantities, the total quantity combined is given as (a + b)

3. Proportion or Proportional Rate (สัดส่วน)

Proportion, in general, is referred to as a part, share, or number considered in comparative relation to a whole,

$$\text{(total value) Proportion} = \frac{a}{a+b+c+ \dots \dots \dots \text{(total)}}$$

$$\text{or Proportion} = \frac{a}{a+b+c+ \dots \dots \dots \text{(total)}} \times 100$$

Basic tools for epidemiological measurements

4. Rate (อัตรา)

$$\text{Rate} = \frac{a}{a+b} \times k$$

a = Numerator (เลขตัวตั้ง)

$a + b$ = Denominator (เลขตัวหาร)

The value of k (constant) can be 100, 1,000, 10,000, 100,000.

Rate

Definition: Frequency of events, that occur in a defined time period, divided by the average population of risk.

$$\text{Rate} = \frac{\text{Numerator}}{\text{Denominator}} \times \text{Constant multiplier}$$

$$\text{Crude death rate} = \frac{\text{Number of deaths (defined place and time period)}}{\text{Mid-period population (same place and population)}} \times 1000$$



5. Index (ดัชนี)

is an approximate value of the rate

➤ It is often used only when the true denominator cannot be found.

➤ The value of something else that is relevant must be used instead.

❑ Maternal death rate = 0.08/100 live births/year

❑ Death rate = 5.16/1,000 population/year



Relative Risk (RR)

A relative risk (RR), also called risk ratio, compares the risk of a health event (disease, injury, risk factor, or death) among one group with the risk among another group.

The formula for RR is:

$$\text{Relative Risk (RR)} = \frac{\text{Risk of disease (incidence proportion, attack rate) in group of exposed (primary interest)}}{\text{Risk of disease (incidence proportion, attack rate) in group of unexposed (comparison group)}}$$



Measures of disease frequency



Measurement of disease frequency (Measure of morbidity..การป่วย..)

1. Prevalence rate (อัตราความชุก)

is the proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time.

$$= \frac{\text{All new and pre-existing cases during a given time period}}{\text{Population during the same time period}} \times K$$

2. Incidence rate (อัตราอุบัติการณ์)

Number of new cases of disease or injury during specified period

$$= \frac{\text{Number of new cases of disease or injury during specified period}}{\text{Size of population at start of period}} \times K$$

- Incidence refers to the occurrence of new cases of disease or injury in a population over a specified period of time.

The value of k (constant) can be 100, 1000, 10,000

Prevalence Rate



❖ **Point prevalence** is a measure of the proportion of people in a population who have a disease or condition at a particular time, such as a particular month or date. It is like a snapshot of the disease in time.

❖ **Period prevalence** is the number of individuals identified as cases during a specified period of time, divided by the total number of people in that population.



1. Point prevalence

- Point prevalence measures the proportion of existing people with a disease in a defined population at a single point in time.

$$\text{Point Prevalence} = \frac{\text{No. of cases in a defined population at one point in time}}{\text{No. of persons in a defined population at the same point in time}}$$

x **K**

The value of k (constant) can be 100, 1000, 10,000

2. Period prevalence

- Prevalence is a useful measure to quantify the burden of disease in a population at a given point in time.

Incidence Rate

1. Cumulative Incidence (CI):

$$\equiv \frac{\text{Number of new cases of disease or injury during specified period}}{\text{Time each person was observed, totaled for all persons}} \times K$$

The value of k (constant) can be 100, 1000, 10,000

2. Incidence Density (ID) or Person-time Incidence

$$\text{Person-time rate} = \frac{\text{Number of new cases of disease or injury during specified period}}{\text{Sum of the person-time of the at-risk population (Person-time at risk)}} \times K$$

Incidence Rate



Source:
<https://www.cdc.gov/cels/dsepd/ss1978/lesson3/section2.html>

3. Attack Rate (อัตราโจมตี)

3.1 Primary Attack Rate

$$= \frac{\text{Number of new cases of disease or injury during specified period}}{\text{Size of population at start of period}} \times 100$$

3.2 Secondary Attack Rate

$$= \frac{\text{Number of cases among contacts of primary cases}^*}{\text{Total number of contacts}^{**}} \times 100$$

* กลุ่ม ผู้ป่วยที่อาจได้รับเชื้อจากกลุ่มผู้ป่วยปฐมภูมิ

** ไม่นับผู้ป่วยรายแรก (หรือชุดแรก) และรายที่มีภูมิคุ้มกัน

3. Attack Rate



| <u>Measure</u> | <u>Numerator</u> | <u>Denominator</u> |
|---|---|--------------------------------------|
| Primary attack rate Incidence proportion (or attack rate or risk) | Number of new cases of disease during specified time interval | Population at start of time interval |
| Secondary attack rate | Number of new cases among contacts | Total number of contacts |

The Population at Risk (ประชากรกลุ่มเสี่ยง)

The population at risk (ประชากรกลุ่มเสี่ยง)

The term population-at-risk refers to the group of individuals susceptible to the event of interest (e.g., infection, disease, death) at or during the time period of interest.

mid-year population



mid-year population calculation

1. Mid-year population (July 1) of a given year

2. National mid-year population estimates are made every 10 years.

3. Based on the calculation of data from the civil registration of the Ministry of Interior.
(กระทรวงมหาดไทย)

3.1 Mid-year population

= (population on January 1) + 1/2 of number of births - number of deaths + Number of people moving in - Number of people moving out) of the given year

3.3 Mid-year population

= 1/2 (population on December 31 of the previous year + population on December 31 of the given year)

3.2 Mid-year population

= (population on December 31) - 1/2 of number of births - number of deaths + Number of people moving in - Number of people moving out) of the given year

SUMMARIZE EPIDEMIOLOGICAL MEASUREMENTS



Overview of Biostatistics

1. Prevalence rate

Period prevalence

Point prevalence

1. Cumulative incidence (CI)
2. Incidence density (ID) or Person-time incidence
3. Attack Rate (อัตราโจมตี)
 - 3.1 Primary
 - 3.2 Secondary

2. Incidence rate

Measures of disease frequency



Email: wanpenw@webmail.npru.ac.th

