

# Statistical Application in Clinical Research For Medical Students

---

Phichayut Phinyo

MD, Dip(ClinEpidemio) Dip(ClinStat)

Center for Clinical Epidemiology and Clinical Statistics

Department of Family Medicine

Faculty of Medicine, Chiang Mai University



# Class objectives

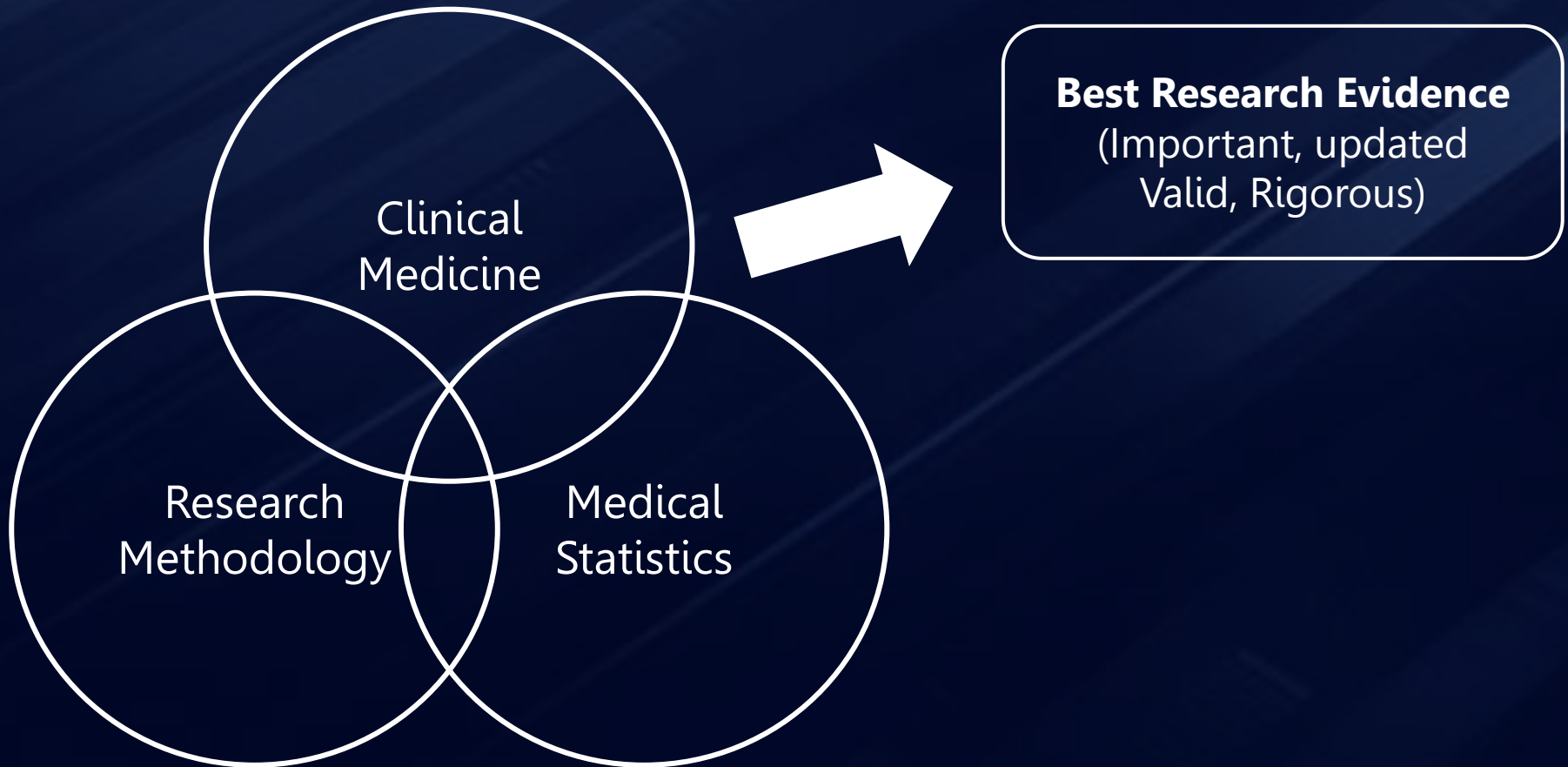
- After the end of this session, (hopefully) the students would be able to understand and correctly describe the concept of the following topics:

# Topics outline

- **Medical statistics I**
  - Statistical concepts in medical practice
  - Internal and external validity

# Clinical Epidemiology

Scientific investigations to provide sound evidence for clinical problems in practice

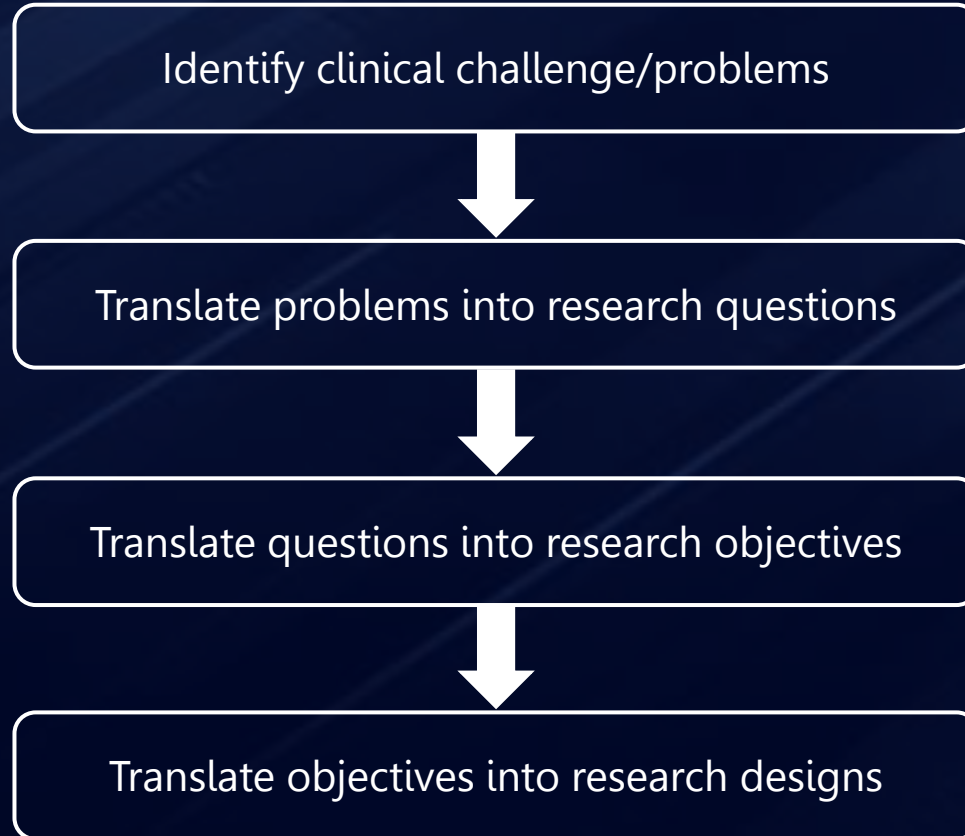


# Clinical Problems

- ศาสตร์ทาง**ระบาดวิทยาและสถิติ**ถูกนำมาใช้ในการออกแบบการศึกษาวิจัยเพื่อตอบโจทย์ปัญหาต่าง ๆ ในวงการแพทย์
- **Four Main Aspects in Clinical Medicine**
  - **D**iagnosis
  - **E**tiognosis
  - **P**rognosis
  - **T**herapeutic

# Clinical Epidemiology

- Basic Epidemiologic Thinking process



# Clinical Epidemiology

- Core backbone of research = DDO
  - Domain  
ศึกษาในใคร? จะนำไปใช้ในใคร?
  - Determinant (s)  
ตัวแปรต้น สิ่งกำหนด สิ่งที่น่าสนใจ
  - Outcome  
ผลลัพธ์ที่น่าสนใจ
  - Covariates  
(ตัวแปรอื่นที่เกี่ยวข้อง (ถ้ามี))



# Clinical Epidemiology

- Core backbone of research = DDO
  - Domain  
ศึกษาในใคร? จะนำไปใช้ในใคร?
  - Determinant (s)  
ตัวแปรต้น สิ่งกำหนด สิ่งที่น่าสนใจ
  - Outcome  
ผลลัพธ์ที่น่าสนใจ

คำถามวิจัยที่ดีต้องมีครบทั้ง 3 ประเด็น + ชัดเจน!



# Clinical Epidemiology

- Two main types of research question
  - Descriptive  
(บรรยาย หรือพรรณนา ไม่เปรียบเทียบ)
  - Hypothesis testing /comparative  
(มีการทดสอบหรือเปรียบเทียบ)

# Diagnosis

- How to categorize patient into groups?
  - Illness : Yes or No?
  - If Yes : Disease A or Disease B?

**Example:** COVID-19 or other infections (influenza) ?

*What are the clinical signs and symptoms that could help clinicians distinguish COVID-19 from influenza infection?*

What types of answer are we looking for?

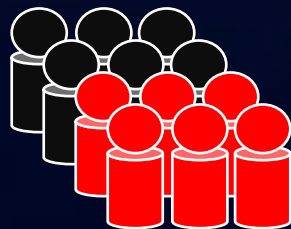
How To design the study?

# Diagnosis

*What are the clinical signs and symptoms that could help clinicians distinguish COVID-19 from influenza infection?*

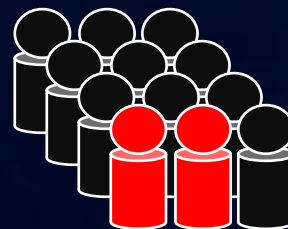
สมมติฐานของผู้วิจัย (research hypothesis)

- ผู้วิจัย "เชื่อว่า" น่าจะมีอาการหรืออาการแสดงบางอย่างที่สามารถแยกผู้ป่วยที่มาด้วยอาการเจ็บป่วยทางระบบทางเดินหายใจระหว่างเชื้อไวรัส COVID-19 และ influenza
- ความเชื่ออาจเกิดจากการสังเกต หรือหลักฐานในอดีต



6/12 (50%)  
มี anosmia

**Confirmed COVID-19**



2/12 (16%)  
มี anosmia

**Confirmed Influenza**

# Diagnosis

*What are the clinical signs and symptoms that could help clinicians distinguish COVID-19 from influenza infection?*

D

What is the study domain?

D

What are the study determinants?

O

What is the study outcome?

# Diagnosis

*What are the clinical signs and symptoms that could help clinicians distinguish COVID-19 from influenza infection?*

D

Patients with respiratory illness

D

Having anosmia

O

COVID-19 or influenza infection

# Diagnosis

*What are the clinical signs and symptoms that could help clinicians distinguish COVID-19 from influenza infection?*

D

Patients with respiratory illness

เราจะวางแผนเก็บรวบรวมข้อมูลผู้ป่วยกลุ่มนี้อย่างไรเพื่อทำวิจัย?

D

Having anosmia

จะเก็บข้อมูลตัวอย่างใด มีตัวแปรอื่น ๆ ที่ต้องเก็บเพิ่มอีกหรือไม่?

O

COVID-19 or influenza infection

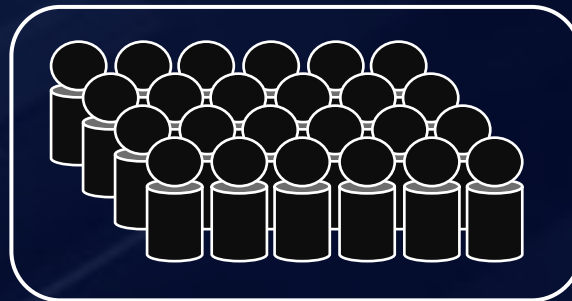
จะวัดผลลัพธ์เมื่อไหร่ อย่างไร?



# Diagnosis

**Example:** COVID-19 or other infections (influenza) ?

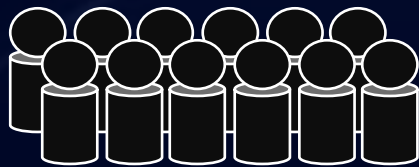
Fever = (yes/no)  
Cough = (yes/no)  
Dyspnea = (yes/no)  
**Anosmia = (yes/no)**  
Rhinorrhea = (yes/no)



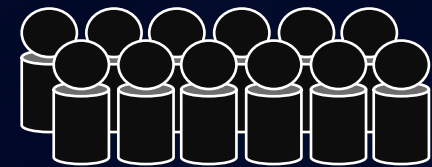
Patients who presented with respiratory illness

Viral testing to verify endpoint

**COVID-19 = (yes/no)**



**Confirmed COVID-19**



**Confirmed Influenza**



# Diagnosis


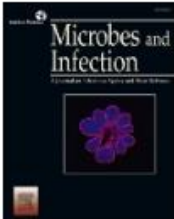
Example: COVID-19 or other infections (influenza) ?

ARTICLE IN PRESS

Microbes and Infection xxx (xxxx) xxx

---

Contents lists available at ScienceDirect

 **Microbes and Infection** 

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

---

## Clinical features of COVID-19 and influenza: a comparative study on Nord Franche-Comte cluster

Souheil Zayet <sup>a,\*</sup>, N'dri Juliette Kadiane-Oussou <sup>a</sup>, Quentin Lepiller <sup>b</sup>, Hajer Zahra <sup>c</sup>,  
Pierre-Yves Royer <sup>a</sup>, Lynda Toko <sup>a</sup>, Vincent Gendrin <sup>a</sup>, Timothée Klopfenstein <sup>a,\*\*</sup>

<sup>a</sup> Infectious Disease Department, Nord Franche-Comté Hospital, France  
<sup>b</sup> Virology Department, Centre-Hospitalier Universitaire de Besançon, France  
<sup>c</sup> Diabetology Department Nord Franche-Comté Hospital, France

<https://www.sciencedirect.com/science/article/pii/S1286457920300940>

# Diagnosis

**Example:** COVID-19 or other infections (influenza) ?

**Table 2**

Differences in symptoms between SARS-CoV-2 and Influenza virus A/B infected patients in Nord Franche-Comte Hospital.

Functional Signs	SARS-CoV-2 (n = 70)	Influenza A/B (n = 54)	p-(value)
<b>General symptoms</b>			
<b>Fever (Number, (%))</b>			
<b>Fever ≥ 38 (objective)</b>	53 (75.7)	50 (92.6)	<b>0.042</b>
<b>Feeling of fever</b>	13 (18.6)	3 (5.6)	0.056
<b>No fever neither feeling of fever</b>	4 (5.7)	1 (1.8)	0.386
<b>Highest temperature (mean, extremes, SD)</b>	38.7 [36.7–41.5] ±0.91	39 [36.2–41.4] ± 0.97	0.064
<b>Fatigue (Number, (%))</b>	65 (92.9)	47 (87)	0.362
<b>Pain symptoms</b>			
<b>Myalgia (Number, (%))</b>	41 (58.6)	38 (70.4)	0.192
<b>Arthralgia (Number, (%))</b>	38 (54.3)	36 (66.7)	0.198
<b>Headache (Number, (%))</b>			
<b>Total</b>	51 (72.9)	31 (57.4)	0.086
<b>Diffuse</b>	20 (28.6)	24 (44.4)	0.088
<b>Frontal</b>	18 (25.7)	5 (9.3)	<b>0.021</b>
<b>Others<sup>d</sup></b>	13 (17.2)	2 (3.7)	<b>0.013</b>
<b>Respiratory symptoms</b>			
<b>Cough (Number, (%))</b>	56 (80)	44 (81.5)	1
<b>Sputum production (Number, (%))</b>	20 (28.6)	28 (51.9)	<b>0.010</b>
<b>Sneezing (Number, (%))</b>	13 (18.6)	25 (46.3)	<b>0.001</b>
<b>Chest pain (Number, (%))</b>	18 (25.7)	10 (18.5)	0.391
<b>Hemoptysis (Number, (%))</b>	6 (8.6)	3 (5.6)	0.730
<b>Dyspnea (Number, (%))</b>	24 (34.3)	32 (59.3)	<b>0.007</b>
<b>Otorhinolaryngological symptoms</b>			
<b>Tinnitus (Number, (%))</b>	7 (10)	4 (4.7)	0.755
<b>Sore throat (Number, (%))</b>	14 (20)	24 (44.4)	<b>0.006</b>
<b>Hearing loss (Number, (%))</b>	4 (5.7)	4 (7.4)	0.727
<b>Dysgeusia (Number, (%))</b>	34 (48.6)	11 (20.4)	<b>0.001</b>
<b>Anosmia (Number, (%))</b>	37 (52.9)	9 (16.7)	<b>&lt;0.001</b>
<b>Rhinorrhea (Number, (%))</b>	34 (48.6)	30 (55.6)	0.4
<b>Nasal obstruction (Number, (%))</b>	13 (18.6)	19 (35.2)	0.08
<b>Epistaxis (Number, (%))</b>	3 (4.3)	3 (5.6)	1

Zayet et al., 2020

# Diagnosis

**Example:** COVID-19 or other infections (influenza) ?

**Table 2**

Differences in symptoms between SARS-CoV-2 and Influenza virus A/B infected patients in Nord Franche-Comte Hospital.

Functional Signs	SARS-CoV-2 (n = 70)	Influenza A/B (n = 54)	p-(value)
<b>General symptoms</b>			
<b>Fever (Number, (%))</b>			
<b>Fever ≥ 38 (objective)</b>	53 (75.7)	50 (92.6)	<b>0.042</b>
<b>Feeling of fever</b>	13 (18.6)	3 (5.6)	0.056
<b>No fever neither feeling of fever</b>	4 (5.7)	1 (1.8)	0.386
<b>Highest temperature (mean, extremes, SD)</b>	38.7 [36.7–41.5] ±0.91	39 [36.2–41.4] ± 0.97	0.064
<b>Fatigue (Number, (%))</b>	65 (92.9)	47 (87)	0.362
<b>Pain symptoms</b>			
<b>Myalgia (Number, (%))</b>	41 (58.6)	38 (70.4)	0.192
<b>Arthralgia (Number, (%))</b>	38 (54.3)	36 (66.7)	0.198
<b>Headache (Number, (%))</b>			
<b>Total</b>	51 (72.9)	31 (57.4)	0.086
<b>Diffuse</b>	20 (28.6)	24 (44.4)	0.088
<b>Frontal</b>	18 (25.7)	5 (9.3)	<b>0.021</b>
<b>Others<sup>a</sup></b>	13 (17.2)	2 (3.7)	<b>0.013</b>
<b>Respiratory symptoms</b>			
<b>Cough (Number, (%))</b>	56 (80)	44 (81.5)	1
<b>Sputum production (Number, (%))</b>	20 (28.6)	28 (51.9)	<b>0.010</b>
<b>Sneezing (Number, (%))</b>	13 (18.6)	25 (46.3)	<b>0.001</b>
<b>Chest pain (Number, (%))</b>	18 (25.7)	10 (18.5)	0.391
<b>Hemoptysis (Number, (%))</b>	6 (8.6)	3 (5.6)	0.730
<b>Dyspnea (Number, (%))</b>	24 (34.3)	32 (59.3)	<b>0.007</b>
<b>Otorhinolaryngological symptoms</b>			
<b>Tinnitus (Number, (%))</b>	7 (10)	4 (4.7)	0.755
<b>Sore throat (Number, (%))</b>	14 (20)	24 (44.4)	<b>0.006</b>
<b>Hearing loss (Number, (%))</b>	4 (5.7)	4 (7.4)	0.727
<b>Dysgeusia (Number, (%))</b>	34 (48.6)	11 (20.4)	<b>0.001</b>
<b>Anosmia (Number, (%))</b>	37 (52.9)	9 (16.7)	<b>&lt;0.001</b>
<b>Rhinorrhea (Number, (%))</b>	34 (48.6)	30 (55.6)	0.4
<b>Nasal obstruction (Number, (%))</b>	13 (18.6)	19 (35.2)	0.08
<b>Epistaxis (Number, (%))</b>	3 (4.3)	3 (5.6)	1

Zayet et al., 2020

# Diagnosis

**Example:** COVID-19 or other infections (influenza) ?

## Symptom Comparison

Symptoms	Allergies	Cold	Influenza (Flu)	Coronavirus COVID-19
Body Aches	Never	Often	Often	Sometimes
Cough	Sometimes	Often	Often	Often
Diarrhea/GI	Rare	Rare	Sometimes	Sometimes
Fatigue	Sometimes	Sometimes	Often	Often
Fever / Chills / Shaking	Never	Rare	Often	Often
Headache	Rare	Rare	Often	Sometimes
Loss of Taste or Smell	Never	Never	Never	Sometimes
Shortness of Breath or Difficulty Breathing	Rare	Rare	Rare	Often
Sneezing	Often	Often	Rare	Rare
Sore Throat	Rare	Often	Sometimes	Sometimes
Stuffy Nose	Often	Often	Sometimes	Rare

Yale  
NewHaven  
Health

Questions about COVID-19?  
**833-ASK-YNHH**

\*Symptoms may range from mild to severe depending on individual health and/or pre-existing conditions. Consult your physician with questions.



# Etiognosis

- To identify risk or etiologic factors of diseases?
  - Does having A causes B? (Explain)
  - What are the factors that cause B? (Explore)

**Example:** PM2.5 and Lung Cancer Incidence?

*Does living in an area with chronic exposure to PM2.5 increases the risk of lung cancer?*

What types of answer are we looking for?

How To design the study?

# Etiognosis

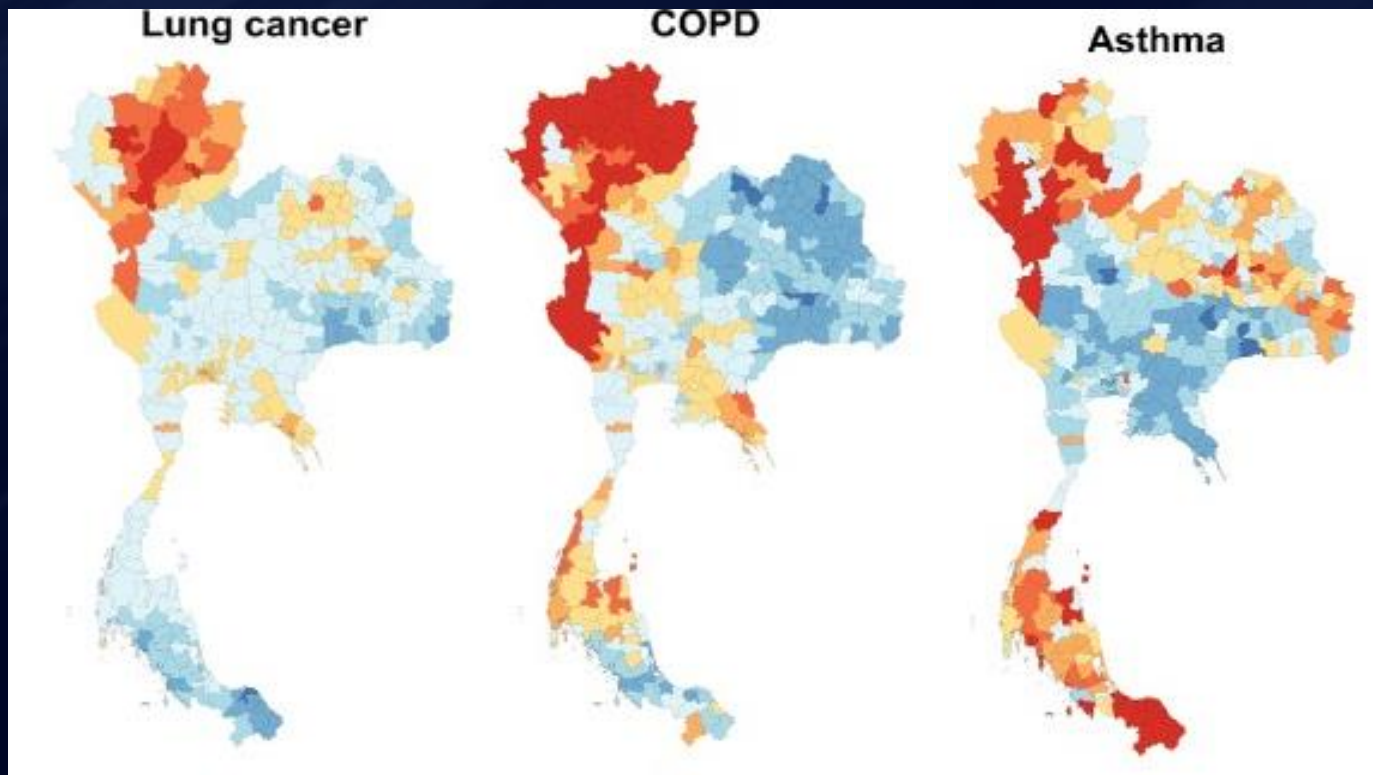
*Does living in an area with chronic exposure to PM2.5 increase the risk of lung cancer?*

## สมมติฐานของผู้วิจัย (research hypothesis)

- ผู้วิจัย “เชื่อว่า” การอยู่อาศัยในเขตพื้นที่หรือบริเวณที่มีการ expose ต่อฝุ่น PM 2.5 เป็นเวลานานน่าจะมีความสัมพันธ์ในเชิงเพิ่มความเสี่ยงต่อการเกิดมะเร็งปอด
- ความเชื่ออาจเกิดจากการสังเกต หรือหลักฐานในอดีต

# Etiognosis

*Does living in an area with chronic exposure to PM2.5 increase the risk of lung cancer?*



Aungkulanon, 2016



# Etiognosis

*Does living in an area with chronic exposure to PM2.5 increase the risk of lung cancer?*

D

What is the study domain?

D

What are the study determinants?

O

What is the study outcome?

# Etiognosis

*Does living in an area with chronic exposure to PM2.5 increase the risk of lung cancer?*

D

Thai population (general)

D

Chronic exposure to PM 2.5

O

Lung cancer incidence

# Etiognosis

*Does living in an area with chronic exposure to PM2.5 increase the risk of lung cancer?*

D

Thai population (general)

เราจะวางแผนเก็บรวบรวมข้อมูลคนกลุ่มนี้อย่างไรเพื่อทำวิจัย?

D

Chronic exposure to PM 2.5

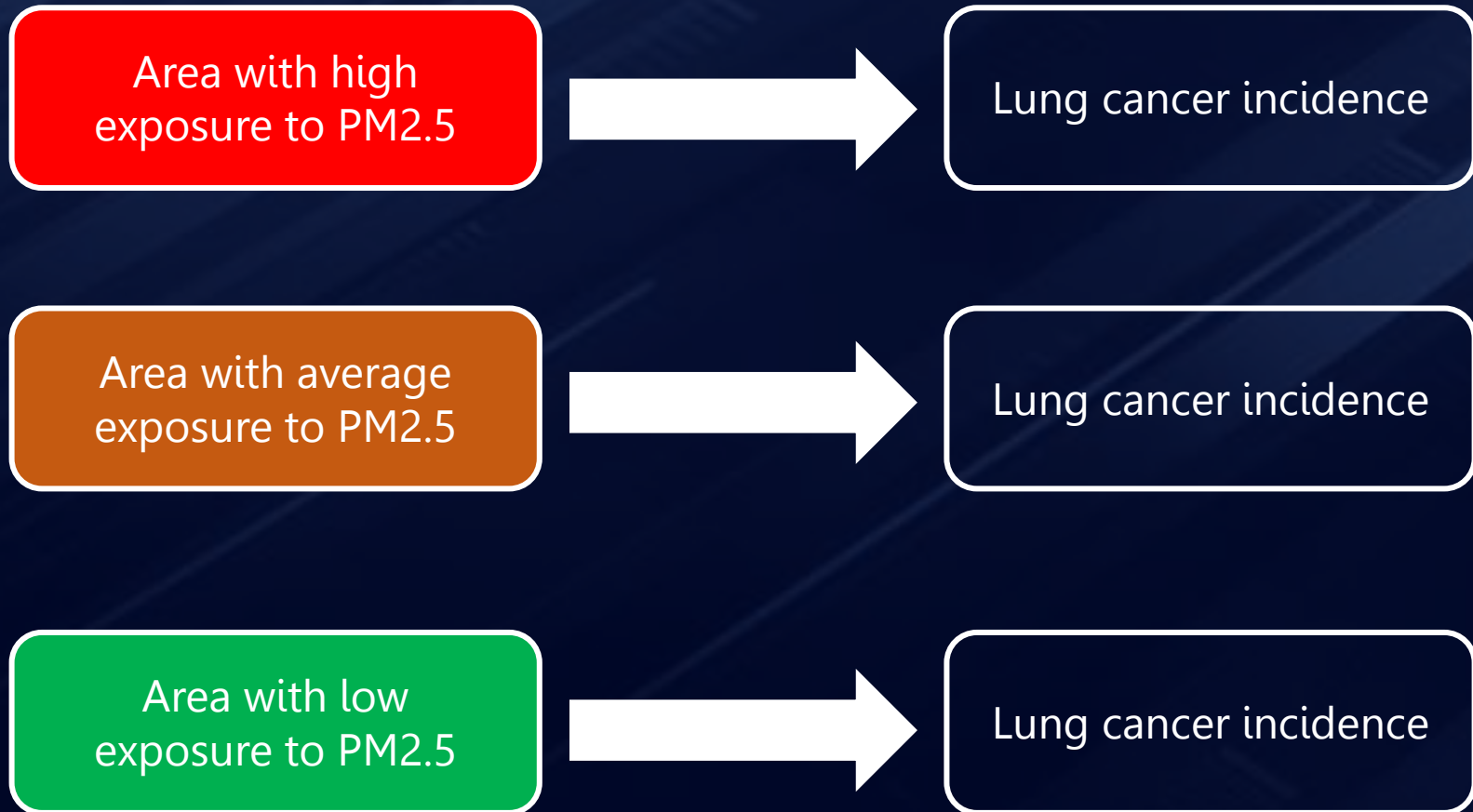
จะเก็บข้อมูลตัวอย่างใด มีตัวแปรอื่น ๆ ที่ต้องเก็บเพิ่มอีกหรือไม่?

O

Lung cancer incidence

จะวัดผลลัพธ์เมื่อไหร่ อย่างไร?

# Etiognosis



# Prognosis

- How to foresee the course of the diseases?

**Example:** Breast cancer with metastasis AND mortality

*What is the average survival of breast cancer patients with metastatic disease?*

*What factors influence the survival of metastatic breast cancer patients?*

What types of answer are we looking for?  
How To design the study?

# Prognosis

*What is the average survival of breast cancer patients with metastatic disease?*

สมมติฐานของผู้วิจัย (research hypothesis)

- ไม่มีสมมติฐาน ไม่มีวัตถุประสงค์เพื่อทดสอบความแตกต่าง ผู้วิจัยเพียงต้องการทราบการรอดชีพเฉลี่ยของผู้ป่วยในกลุ่มดังกล่าว

*What factors influence the survival of metastatic breast cancer patients?*

สมมติฐานของผู้วิจัย (research hypothesis)

- ผู้วิจัย "เชื่อว่า" มีปัจจัยหรือลักษณะบางอย่างของผู้ป่วยที่ส่งผลต่อการดำเนินโรคหรือการรอดชีพของผู้ป่วยได้



# Prognosis

*What factors influence the survival of metastatic breast cancer patients?*

D

What is the study domain?

D

What are the study determinants?

O

What is the study outcome?



# Prognosis

*What factors influence the survival of metastatic breast cancer patients?*

D

Metastatic breast cancer patients

D

Factors: demographic, clinical

O

Survival outcomes

# Prognosis

*What factors influence the survival of metastatic breast cancer patients?*

D

Metastatic breast cancer patients

เราจะวางแผนเก็บรวบรวมข้อมูลผู้ป่วยกลุ่มนี้อย่างไรเพื่อทำวิจัย?

D

Factors: demographic, clinical

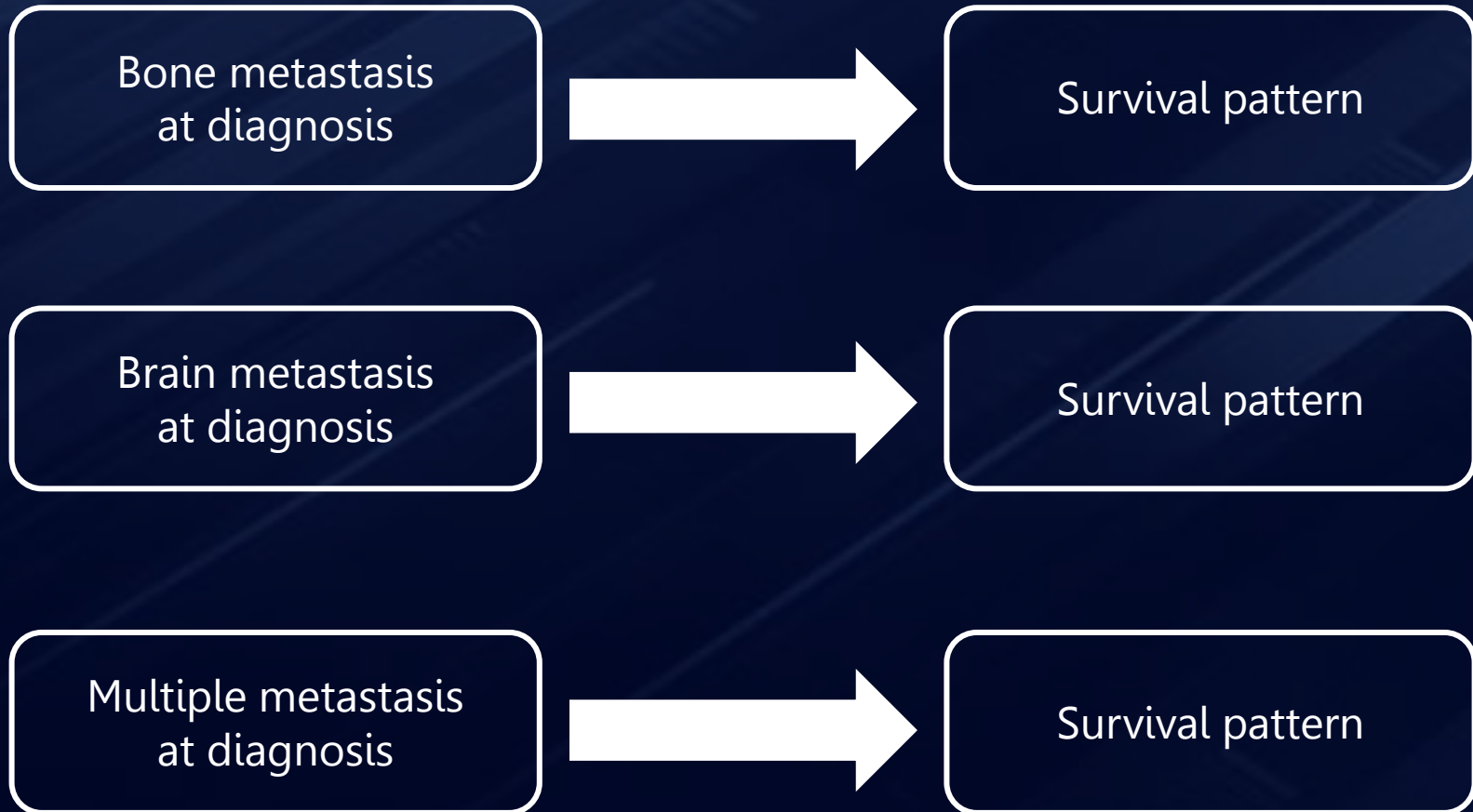
จะเก็บข้อมูลตัวอย่างไร?

O

Survival outcomes

จะวัดผลลัพธ์เมื่อไหร่ อย่างไร?

# Prognosis



# Therapy

- Does treatment A perform better than treatment B?

**Example:** Statin and Myocardial infarction in DM

*Does statin reduce the risk of myocardial infarction in patients with Diabetes Mellitus?*

What types of answer are we looking for?

How To design the study?

# Therapy

*Does statin reduce the risk of myocardial infarction in patients with Diabetes Mellitus?*

สมมติฐานของผู้วิจัย (research hypothesis)

- ผู้วิจัย “เชื่อว่า” การได้รับ primary prevention ด้วย simvastatin ในผู้ป่วยเบาหวานสามารถลดความเสี่ยงต่อการเกิดโรคกล้ามเนื้อหัวใจขาดเลือดที่ 5 ปีได้ เมื่อเทียบกับการไม่ได้รับ
- ความเชื่ออาจเกิดจากการสังเกต หรือหลักฐานในอดีต

# Therapy

*Does statin reduce the risk of myocardial infarction in patients with Diabetes Mellitus?*

D

What is the study domain?

D

What are the study determinants?

O

What is the study outcome?



# Therapy

*Does statin reduce the risk of myocardial infarction in patients with Diabetes Mellitus?*

D

Patients with Diabetes

เราจะวางแผนเก็บรวบรวมข้อมูลผู้ป่วยกลุ่มนี้อย่างไรเพื่อทำวิจัย?

D

Simvastatin vs. Placebo

จะทำการสุ่มเข้ากลุ่มอย่างไร?

O

Myocardial infarction

จะวัดผลลัพธ์เมื่อไหร่ อย่างไร?



# Therapy

*Does statin reduce the risk of myocardial infarction in patients with Diabetes Mellitus?*

D

Patients with Diabetes

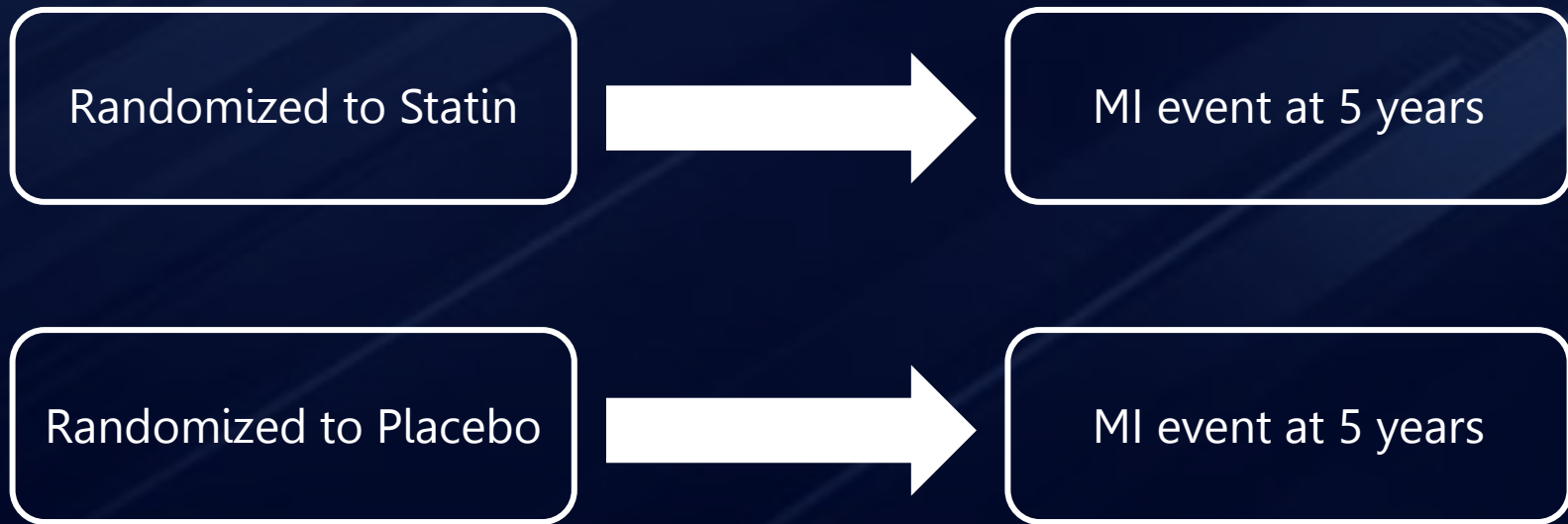
D

Simvastatin vs. Placebo

O

Myocardial infarction

# Therapy



# Therapy

**Example:** Statin and Myocardial infarction in DM

ARTICLES

## Articles

### MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20 536 high-risk individuals: a randomised placebo-controlled trial

*Heart Protection Study Collaborative Group\**

#### Summary

**Background** Throughout the usual LDL cholesterol range in Western populations, lower blood concentrations are associated with lower cardiovascular disease risk. In such populations, therefore, reducing LDL cholesterol may reduce the development of vascular disease, largely irrespective of initial cholesterol concentrations.

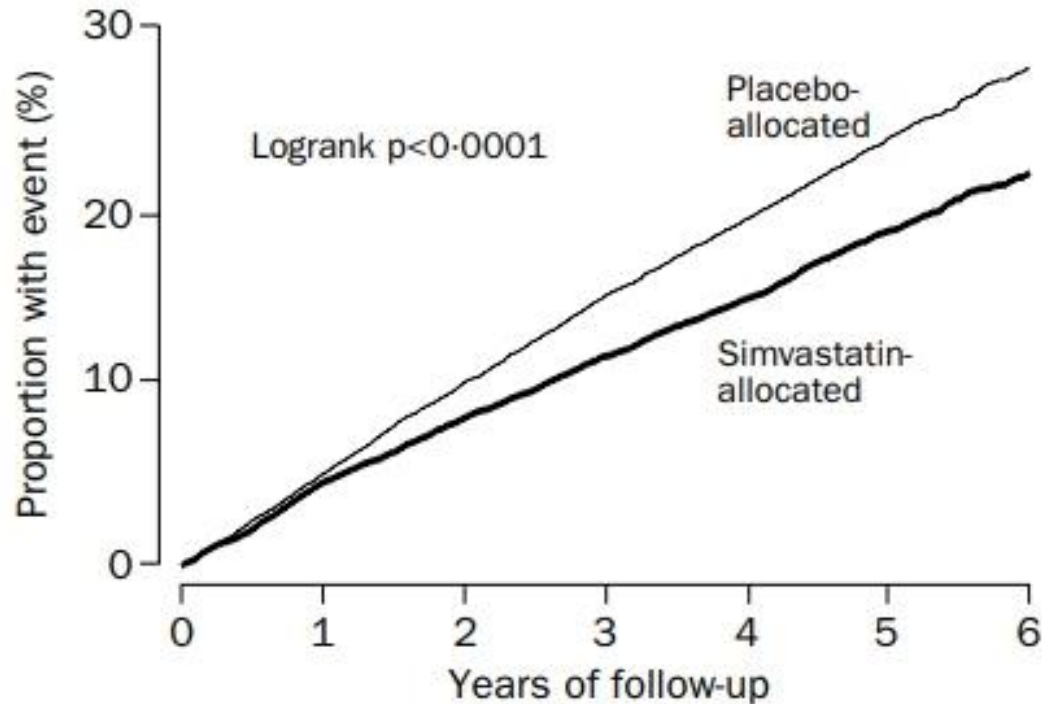
**Methods** 20 536 UK adults (aged 40–80 years) with coronary disease, other occlusive arterial disease, or diabetes were randomly allocated to receive 40 mg simvastatin daily (average compliance: 85%) or matching

participant studied, including: those without diagnosed coronary disease who had cerebrovascular disease, or had peripheral artery disease, or had diabetes; men and, separately, women; those aged either under or over 70 years at entry; and—most notably—even those who presented with LDL cholesterol below 3.0 mmol/L (116 mg/dL), or total cholesterol below 5.0 mmol/L (193 mg/dL). The benefits of simvastatin were additional to those of other cardioprotective treatments. The annual excess risk of myopathy with this regimen was about 0.01%. There were no significant adverse effects on cancer incidence or on hospitalisation for any other non-vascular cause.

HPS study 2002

# Therapy

**Example:** Statin and Myocardial infarction in DM

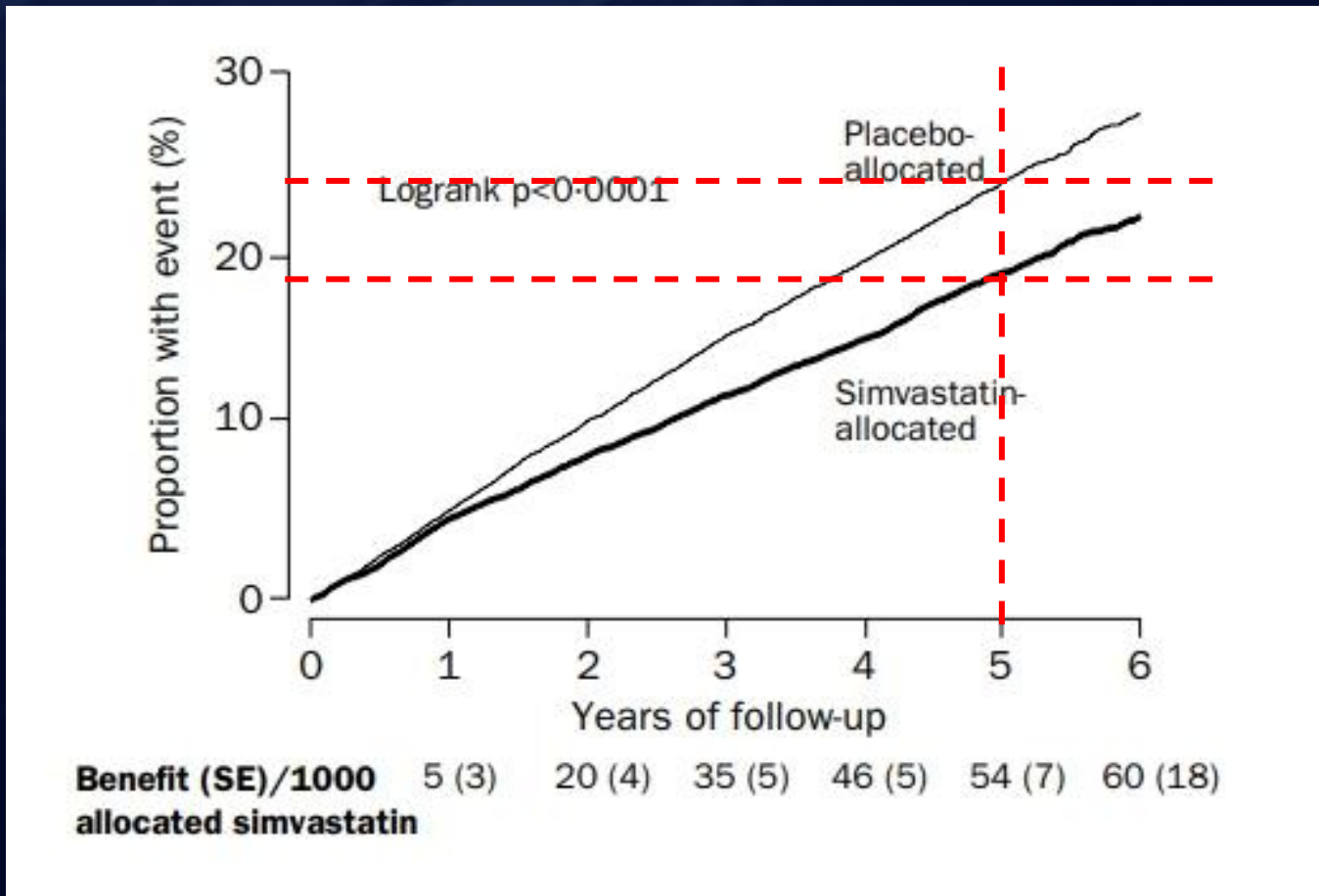


**Benefit (SE)/1000 allocated simvastatin** 5 (3) 20 (4) 35 (5) 46 (5) 54 (7) 60 (18)

HPS study 2002

# Therapy

**Example:** Statin and Myocardial infarction in DM



HPS study 2002



# Therapy

Identify optimal character of treatment for the patients

Example: CPR duration and Survival status

*Does the survival rates differ between long- and short-resuscitation procedure?*

*How long is a cardiopulmonary resuscitation procedure required after cardiac arrest?*

What types of answer are we looking for?  
15, 30, 60, or 120 minutes?

How To design the study?



# Therapy

*Does the survival rates differ between long- and short-resuscitation procedure?*

## สมมติฐานของผู้วิจัย (research hypothesis)

- ผู้วิจัย “เชื่อว่า” การทำ CPR เป็นเวลานานเกินไปอาจไม่มีความจำเป็นเนื่องจากสังเกตว่าผู้ที่ได้รับการ CPR เป็นเวลานานแม้จะตรวจพบการกลับมาทำงานของหัวใจและระบบไหลเวียน (ROSC) สุดท้ายส่วนใหญ่ไม่รอดชีพจากโรงพยาบาลหรือมักมีระบบประสาททำงานผิดปกติ ไม่ตื่น ไม่รู้สึกตัว
- ผู้วิจัยต้องการทราบเวลาที่เหมาะสมสำหรับการทำ CPR ถ้าหากเกินจุดเวลาใดที่โอกาสรอดชีพจากโรงพยาบาลหรือโอกาสที่จะกลับมามีระบบประสาทใกล้เคียงปกติน้อยลง การทำ CPR เกินกว่านั้นอาจไม่มีความจำเป็น

# Therapy

*Does the survival rates differ between long- and short-resuscitation procedure?*

D

What is the study domain?

D

What are the study determinants?

O

What is the study outcome?

# Therapy

*Does the survival rates differ between long- and short-resuscitation procedure?*

D

Cardiac arrest patients - CPR

เราจะวางแผนเก็บรวบรวมข้อมูลผู้ป่วยกลุ่มนี้อย่างไรเพื่อทำวิจัย?

D

Duration of CPR

จะทำการสุ่มเข้ากลุ่มอย่างไร?

O

Survival at discharge, Neuro status

จะวัดผลลัพธ์เมื่อไหร่ อย่างไร?

# Clinical Problems

- To create sound clinical evidence:  
the following should steps are crucial
  - Well formulated questions (hypothesis)
  - Proper study design
  - Proper data collection
  - Proper data analysis
  - Thus, data is the key!
- All of which will be provided in this course!