

Hội thảo Y học Thực chứng
Bệnh viện Hùng Vương

Đánh giá nghiên cứu quan sát

Professor Tuan V. Nguyen
Garvan Institute of Medical Research
University of New South Wales
Sydney – Australia

Mô hình nghiên cứu

- Thực nghiệm (can thiệp) và quan sát
- Nghiên cứu thực nghiệm
 - Phân nhóm ngẫu nhiên ?
 - RCT và quasi-randomized hoặc thử nghiệm tự nhiên (natural experiments)
- Nghiên cứu quan sát
 - Phân tích và mô tả
 - Phân tích
 - NC theo thời gian, bệnh chứng
 - Mô tả
 - Ca lâm sàng, NC cắt ngang

Phân biệt mô hình nghiên cứu

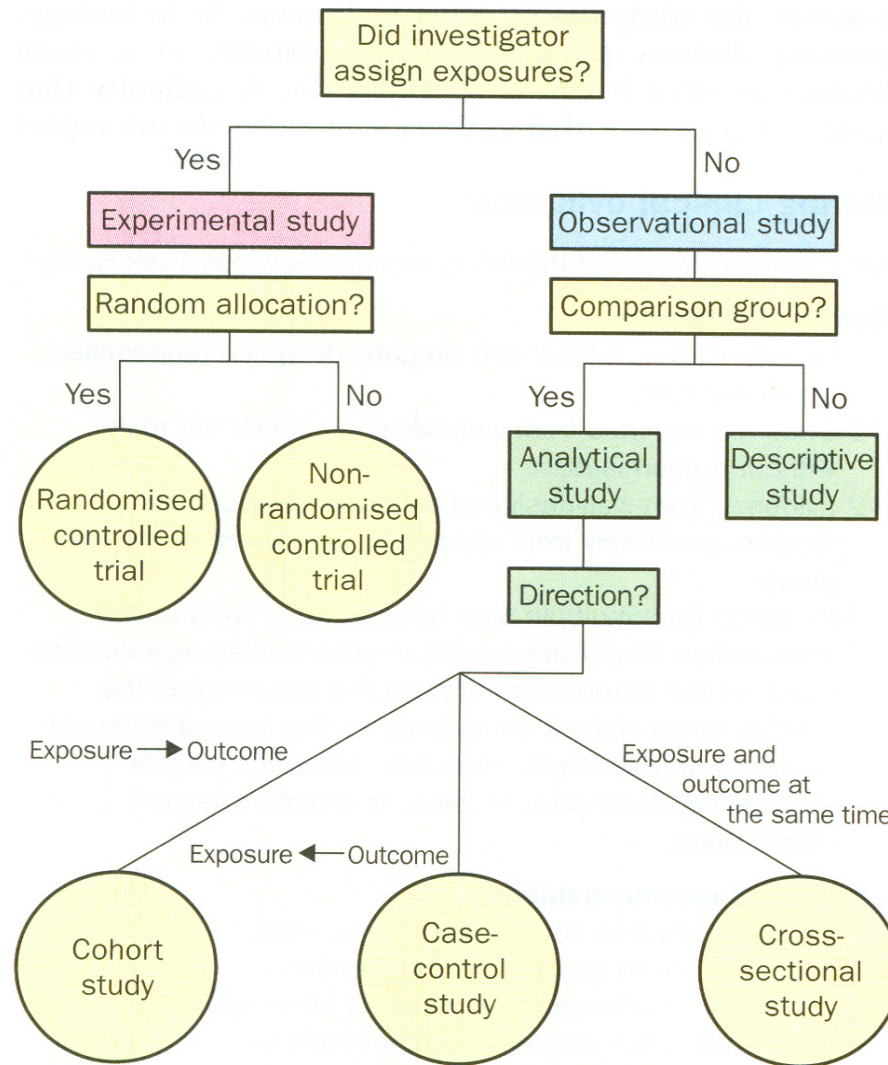


Figure 1: **Algorithm for classification of types of clinical research**

Exposure và outcome

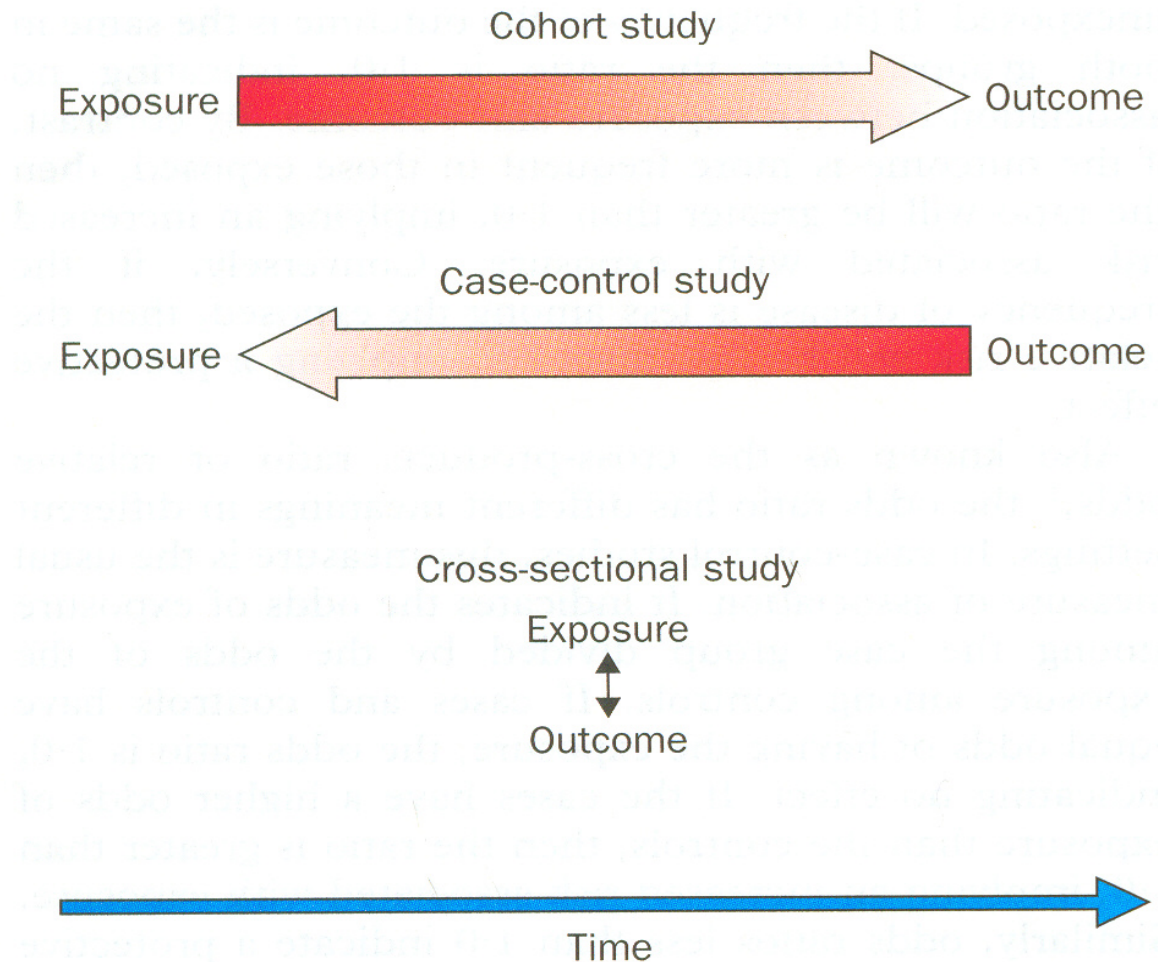
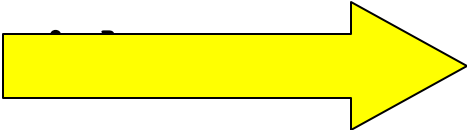


Figure 2: **Schematic diagram showing temporal direction of three study designs**

Kết quả: tính hợp lí (validity)

- **Internal validity (hợp lí nội tại):** kết quả có đáng tin cậy không?
- **External validity (hợp lí ngoại tại):** kết quả từ đối tượng nghiên cứu có thể khái quát hóa cho quần thể lớn hơn?

Kết quả có đáng tin cậy?

Nuôi chó/mèo trong  Giardiasis

OR = 7.3

Liên quan thật?

nhân quả
không phải nhân quả

Bias?

Confounding?

Chance?

Nội dung

- Khái niệm về bias
- Thể loại bias
- Phòng ngừa biases
- Cảnh giác

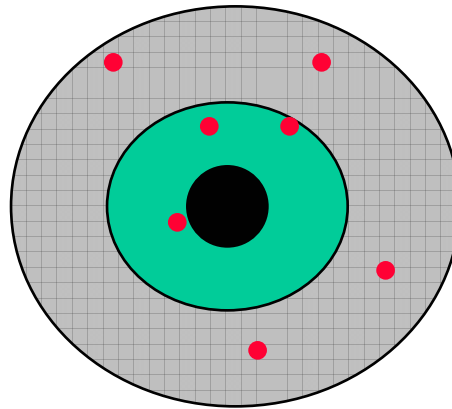
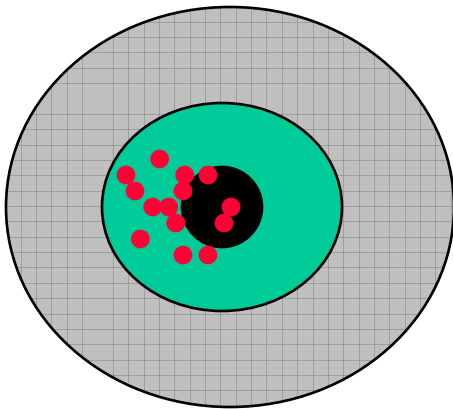
Errors (sai sót)

- Sai sót ngẫu nhiên (→ độ chính xác thấp)
Random error (→ low precision)
- Sai sót có hệ thống (→ giảm tính hợp lí)
Systematic error (→ low validity)

Chất lượng của một ước số (estimate)

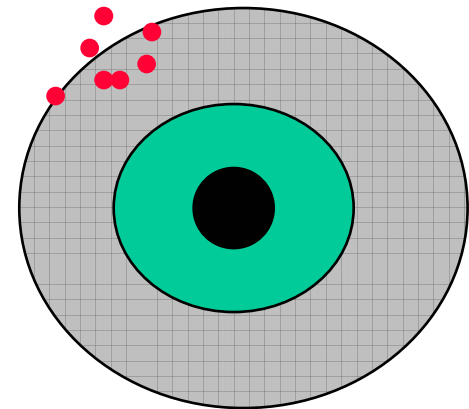
Không chính xác

**Chính xác
& hợp lí**



**Sai ngẫu
nhiên
(random
error) !**

**Chính xác
nhưng không
Hợp lí**



**Sai sót có hệ
thống (Bias) !**

Ví dụ

Đo chiều cao:

- Dùng thước dây (do nhiều người đo)
 - **thiếu chính xác**
 - error
- Thước bị dòn / co
 - **systematic error**
 - bias (không thể điều chỉnh được)

180
179
178
177
176
175
174
173

Định nghĩa bias

Bất cứ sai sót có hệ thống (systematic error) trong một nghiên cứu dịch tễ học có thể dẫn đến kết quả sai về mối liên hệ giữa exposure và disease

Dao động có hệ thống của *giá trị đo lường* so với *giá trị thật* (J. Last)

Systematic variation of measurements from the true value (J. Last)

Nguồn bias

- Sai sót trong quá trình thu thập dữ liệu, phân tích, diễn giải, công bố, và duyệt dữ liệu
- Có thể xảy ra trong bất cứ giai đoạn nào trong nghiên cứu (thiết kế, thực hiện, hay trình bày)
- Có thể xuất phát từ nhà nghiên cứu hay đối tượng nghiên cứu
- Xảy ra trong các nghiên cứu thực nghiệm và nghiên cứu quan sát

Loại bias

- **Bias trong chọn lựa đối tượng (Selection bias)**
- **Bias trong thông tin (Information bias)**
- **Confounding**

Selection Bias

Selection Bias

- Một loại “Systematic error” hệ quả từ việc chọn đối tượng nghiên cứu hay các yếu tố ảnh hưởng đến việc tham gia của đối tượng
 - Kết quả có thể khác với trị số thực nếu thu thập từ *toàn bộ* quần thể
- Xảy ra trong tất cả nghiên cứu
- Không thể điều chỉnh khi phân tích. Cần phải tránh hay giảm khả năng selection bias

Những loại *Selection Bias*

- Bias tình nguyện: Volunteer bias
- Bias mẫu: Sampling bias
- Bias lành mạnh: Healthy worker effect

Selection bias trong nghiên cứu bệnh chứng

Selection Bias – Nghiên cứu bệnh chứng

	<u>Cases</u> liver cirrhosis	<u>Controls A</u> trauma ward
Heavy alcohol use	80	40
Light/no alcohol use	20	60

OR=6

Bệnh: chọn từ khoa cấp cứu, bệnh nhân bị tai nạn xe cộ

Tính đại diện?

Selection Bias – Nghiên cứu bệnh chứng

	<u>Cases</u> liver cirrhosis	<u>Controls A</u> trauma ward	<u>Controls B</u> non-trauma
Heavy alcohol use	80	40	10
Light/no alcohol use	20	60	90

OR=6



Chọn từ khoa
cấp cứu

OR=36



Chọn từ các
khoa khác

Diagnostic Bias

Diagnostic đề cập đến việc chúng ta biết trước yếu tố phơi nhiễm

	Cases uterine cancer	Controls
Takes oral contraceptives	a 	b
Does not take oral contraceptives	c	d

- Dùng OC $\Rightarrow$ ra máu $\Rightarrow$ tăng nguy cơ phát hiện ung thư tử cung
- Ước tính cao hơn thực tế (Overestimation)
“a” $\Rightarrow$ overestimation of OR

Admission Bias

Nhóm bệnh có yếu tố nguy cơ có khả năng nhập viện cao hơn nhóm chứng

	Cases of lung cancer	Controls from surgical wards
Contact with asbestos	a ↑	b
No contact with asbestos	c	d

- Giáo sư, sếp bộ môn hô hấp, chuyên gia hàng đầu về phổi nhiễm asbestos, 145 bài báo khoa học trong lịch
- Ca bệnh ung thư phổi phổi nhiễm asbestos không đại diện cho các bệnh nhân ung thư phổi
- Overestimation of “a” $\Rightarrow$ overestimation of OR

Survival Bias

Chỉ có người còn sống (từ một bệnh nguy hiểm) tham gia vào công trình nghiên cứu

	Cases of haemorrhagic fever	Controls
Exposed to “lethal”	a 	b
Not exposed to “lethal”	c	d

- Phơi nhiễm yếu tố nguy cơ “lethal” dẫn đến chết sớm
- Underestimation of “a” $\Rightarrow$ underestimation of OR

Non-response Bias

Papanicolau test

	Cases of cervical cancer	Controls
Did not have test	a	b
Had test	c	d 
Total	1000	1060

- Nhóm chứng được chọn từ nữ ở nhà họ: 13000 nhà liên lạc $\Rightarrow$ 1060 tham gia
- Controls chủ yếu là nội trợ, khả năng tham gia xét nghiệm thấp hơn phụ nữ đi làm
- Underestimation of “d” $\Rightarrow$ underestimation of OR

Selection bias trong nghiên cứu theo thời gian (prospective study)

Non-response Bias (không trả lời, không tham gia)

Lung cancer

yes no

	yes	no	
Smoker	90	910	1000
Non-smoker	10	990	1000

$$RR = \frac{90}{1000} \div \frac{10}{1000} = 9$$

Non-response Bias (không trả lời, không tham gia)

Lung cancer
yes no

10% of smokers
dare to respond

	<u>Lung cancer</u>		
	yes	no	
Smoker	9	91	100
Non-smoker	10	990	1000

$$RR = \frac{9}{100} \div \frac{10}{1000} = 9$$

Non-response Bias (không trả lời, không tham gia)

	<u>Lung cancer</u>		
	yes	no	
Smoker	45	910	955
Non-smoker	10	990	1000

$$RR = \frac{45}{955} \div \frac{10}{1000} = 4.7$$

50% bệnh nhân hút thuốc lá
bỏ cuộc (không theo dõi được)

Loss to follow-up

- Bias do khác biệt về "loss to follow-up" giữa hai nhóm
- Ví dụ
 - Nghiên cứu bệnh trong nhóm di dân
 - Di dân thường về quê khi mắc bệnh
 - ⇒ lost to follow-up
 - ⇒ tỉ lệ mắc bệnh thấp hơn trong nhóm exposure (=di dân)

Tối thiểu hóa selection bias

- Xác định quần thể nghiên cứu cụ thể
- Xác định bệnh và chứng cụ thể qua tiêu chuẩn rành mạch
- Bệnh và chứng phải xuất phát từ một quần thể

Information bias (Bias thông tin)

Information Bias

- **Systematic error trong việc đo lường và thu thập thông tin / dữ liệu**
- **Khác biệt về tính chính xác của**
 - **exposure giữa nhóm chứng và nhóm bệnh**
 - **outcome giữa nhóm chứng và nhóm bệnh**
- **Đối tượng có thể được phân nhóm sai trong khi phân tích**

Misclassification (phân nhóm sai lạc)

Là một dạng "measurement error" dẫn đến phân nhóm exposure hay outcome sai

Non-differential

- Random error
- Không liên quan đến exposure hay outcome
- Không phải là bias
- Làm suy giảm mức độ tương quan

Differential

- Systematic error
- Không liên quan đến exposure hay outcome
- Bias
- Làm sai lạc mối tương quan

Hai loại information bias

- **Bias trong báo cáo (reporting bias)**
 - Nhớ sai -- Recall bias
- **Bias quan sát viên (Observer bias)**
 - Phỏng vấn -- Interviewer bias
 - Theo dõi -- Biased follow-up

Nhớ sai -- Recall Bias

Bệnh nhân nhớ về yếu tố phơi nhiễm khác với nhóm chứng

	Mothers of	
	Children with malformation	Controls
Took tobacco, alcohol, drugs	a ↑	b
Did not take	c	d

- Bà mẹ với con bị dị tật sẽ nhớ yếu tố phơi nhiễm trong quá khứ tốt hơn là những bà mẹ trong nhóm chứng với con khỏe mạnh
- **Overestimation of “a” $\Rightarrow$ overestimation of OR**

Phỏng vấn -- Interviewer Bias

Người phỏng vấn hỏi bệnh nhân và nhóm chứng khác về yếu tố phơi nhiễm

	Cases of listeriosis	Controls
Eats soft cheese	a 	b
Does not eat soft cheese	c	d

- Nhà nghiên cứu có thể đặt câu hỏi listeriosis về tiêu thụ cheese
- Overestimation of “a” $\Rightarrow$ overestimation of OR

Theo dõi -- Biased Follow-up

Trong nghiên cứu theo thời gian, nhóm không phơi nhiễm có thể ít có khả năng chẩn đoán hơn nhóm phơi nhiễm

- **Ví dụ**

- Cohort study nghiên cứu yếu tố nguy cơ về u trung biểu mô (mesothelioma)
- Khó chẩn đoán bằng histology
- Histologist có thể chẩn đoán mesothelioma nếu biết bệnh nhân từng phơi nhiễm asbestos

Non-differential misclassification

- **Phân nhóm sai không tùy thuộc vào trị số các biến khác**
 - Phân nhóm phơi nhiễm không liên quan gì đến bệnh đang nghiên cứu, hoặc
 - Phân nhóm bệnh không liên quan gì đến yếu tố phơi nhiễm
- **Hệ quả:**
 - Giảm mối tương quan (“bias towards the null”)

Non-differential misclassification

Cohort study: Alcohol $\Rightarrow$ laryngeal cancer

No misclassification	Incidence per mill	RR
1,000,000 drinkers	50	5.0
500,000 nondrinkers	10	

50% drinkers misclassified		
500,000 drinkers	50	1.7
1,000,000 “nondrinkers”	30	

Non-differential misclassification

Cohort study: Alcohol $\Rightarrow$ laryngeal cancer

No misclassification	Incidence per mill	RR
1,000,000 drinkers	50	5.0
500,000 nondrinkers	10	
50% drinkers & 33% nondrinkers misclassified		
666,667 “drinkers”	40	1.2
833,333 “nondrinkers”	34	

Bias trong nghiên cứu theo thời gian

- **Loss to follow up**
 - Nguồn chính trong nghiên cứu theo thời gian
 - Giả định rằng tất cả không mắc bệnh?
- **Ascertainment bias và interviewer bias**
 - Đáng quan tâm: Biết yếu tố nguy cơ có thể ảnh hưởng đến cách xác định bệnh / outcome
- **Non-response, refusals**
 - Không đáng quan tâm mấy: Bias xảy ra khi có liên quan đến yếu tố nguy cơ và outcome
- **Recall bias**
 - Không có vấn đề gì quan trọng: Phơi nhiễm đư75c xác định lúc tham gia nghiên cứu

Bias trong nghiên cứu theo thời gian

- **Ascertainment bias, participation bias, interviewer bias**
 - Exposure và outcome có thể đã xảy ra $\Rightarrow$ differential selection / interviewing
- **Recall bias**
 - Bệnh nhân có thể nhớ yếu tố phơi nhiễm khác với nhóm chứng

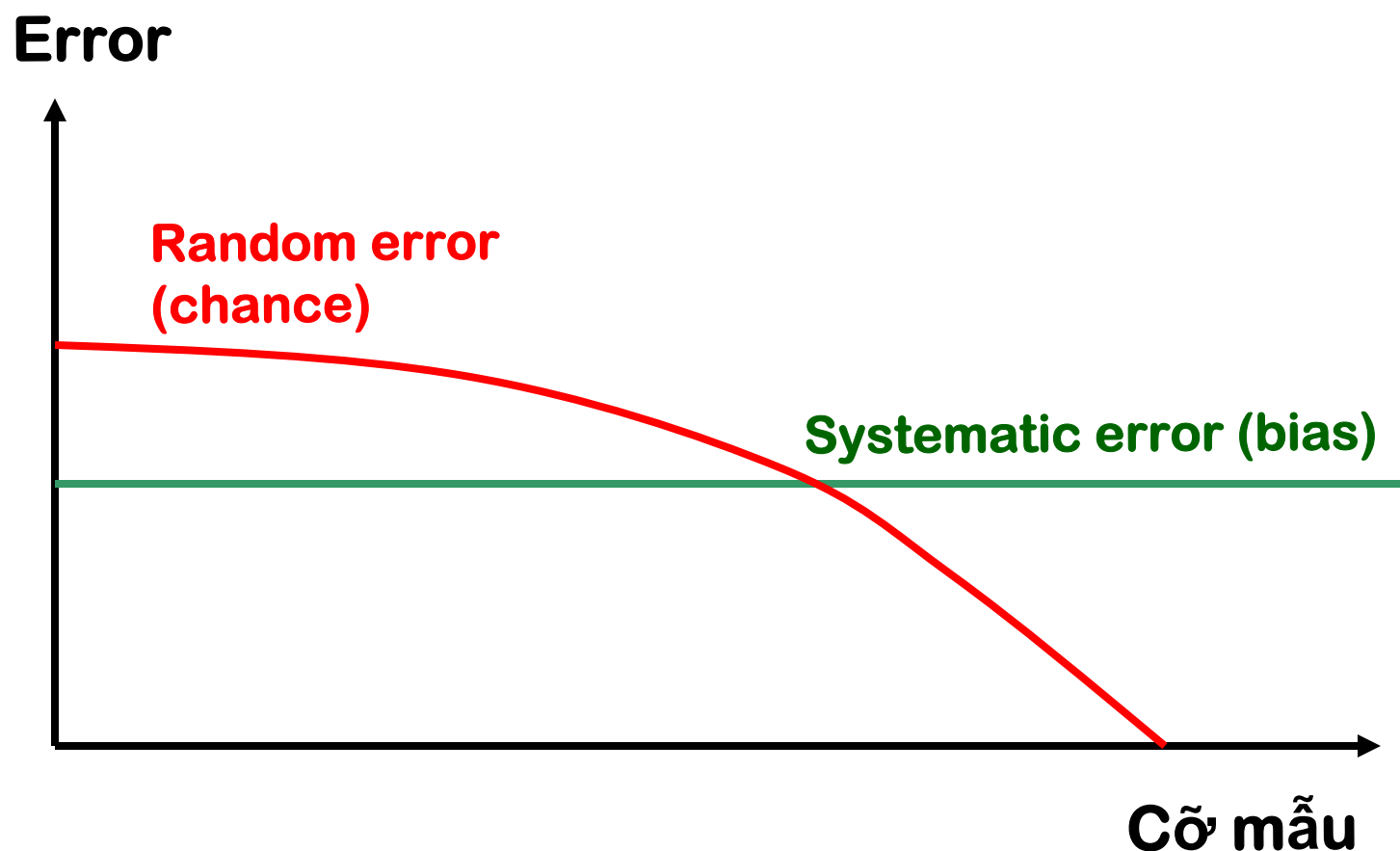
Giảm information bias

- Chuẩn hóa bộ câu hỏi hay phương tiện đo lường
- Phương tiện thu thập dữ liệu áp dụng cho cùng nhóm bệnh và nhóm chứng
- Dùng nhiều nguồn thông tin
 - Questionnaires
 - Thiết bị đo lường trực tiếp
 - Registries
 - Case records
- Dùng nhiều nhóm chứng

Bộ câu hỏi -- Questionnaire

- Câu hỏi “closed”, cụ thể và chính xác; tránh câu hỏi “open-ended”
- Dùng nhiều câu hỏi để thu thập dữ liệu cho một giả thuyết
- “Ngụy trang” câu hỏi về giả thuyết bằng những câu hỏi không liên quan nhau
- Thử nghiệm và hoàn thiện bộ câu hỏi trước khi áp dụng
- Chuẩn hóa người phỏng vấn, dùng mô hình (như muốn, chèn, li, tách) để hỏi các câu hỏi mang tính định lượng

Giảm sai sót trong nghiên cứu

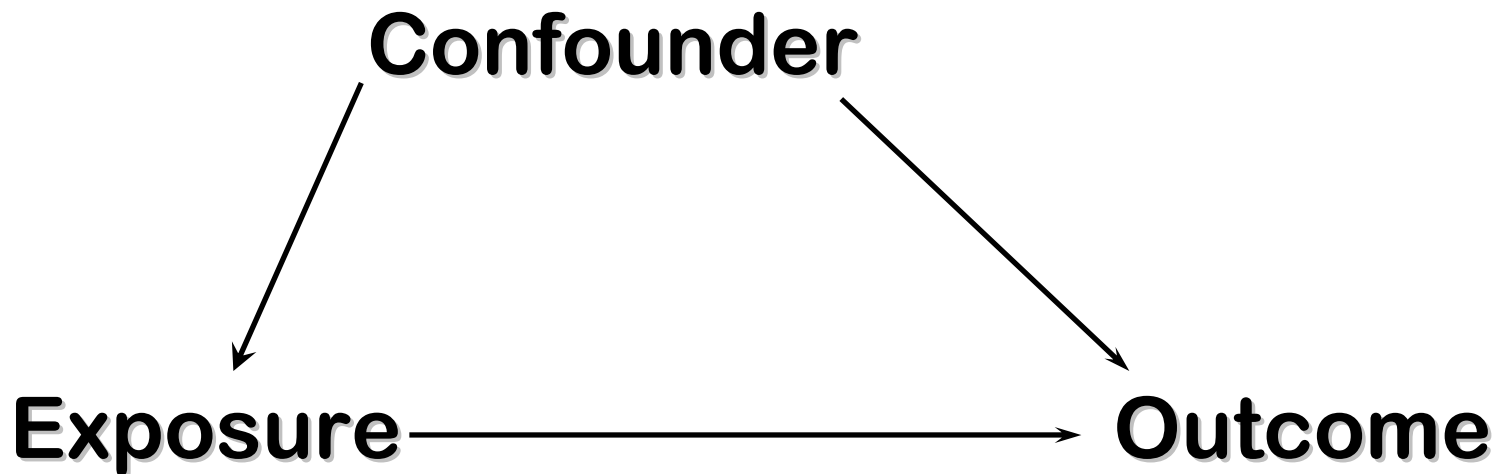


Source: Rothman, 2002

Yếu tố confounding

- Nhầm lẫn ảnh hưởng của một exposure (yếu tố phơi nhiễm) với một yếu tố ngoại tại (extraneous variable).
- Ước số của ảnh hưởng của exposure có thể do ảnh hưởng của covariate.

Điều kiện cần của confounder



Cả hai exposure và confounder đều có liên quan đến outcome.

Ngoài ra, yếu tố confounder có liên quan đến exposure.

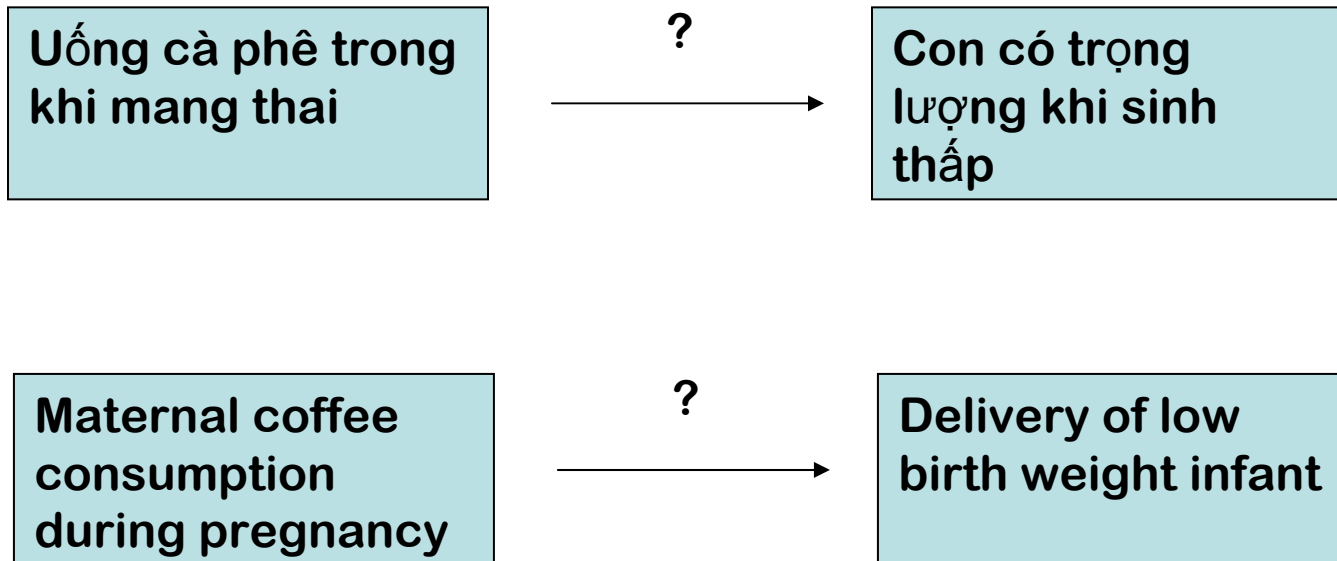
Nhưng yếu tố confounder không nằm trong “causal pathway”.

Confounding

Khác biệt với selection bias và information bias, vì nó có thể đánh giá và điều chỉnh trong khi phân tích

Differs from selection and information bias because it can be evaluated and controlled to some extent in the analysis phase of the study

Một ví dụ về confounding



Một ví dụ về confounding

	Low Birth Weight	Normal Birth Weight
Cà phê	170	96
Không cà phê	90	88

Chưa điều chỉnh:

$$OR = (170)(88) / (96)(90) = 1.73$$

Phân tích trong nhóm hút thuốc lá

	Low Birth Weight	Normal Birth Weight
Cà phê	160	16
Không cà phê	80	8

$$OR_{\text{smoker}} = (160)(8) / (16)(80) = 1.00$$

Phân tích trong nhóm không hút thuốc lá

	Low Birth Weight	Normal Birth Weight
Coffee	10	80
No Coffee	10	80

$$OR_{\text{non-smokers}} = (10)(80) / (10)(80) = 1.00$$

Bảng chứng về confounding

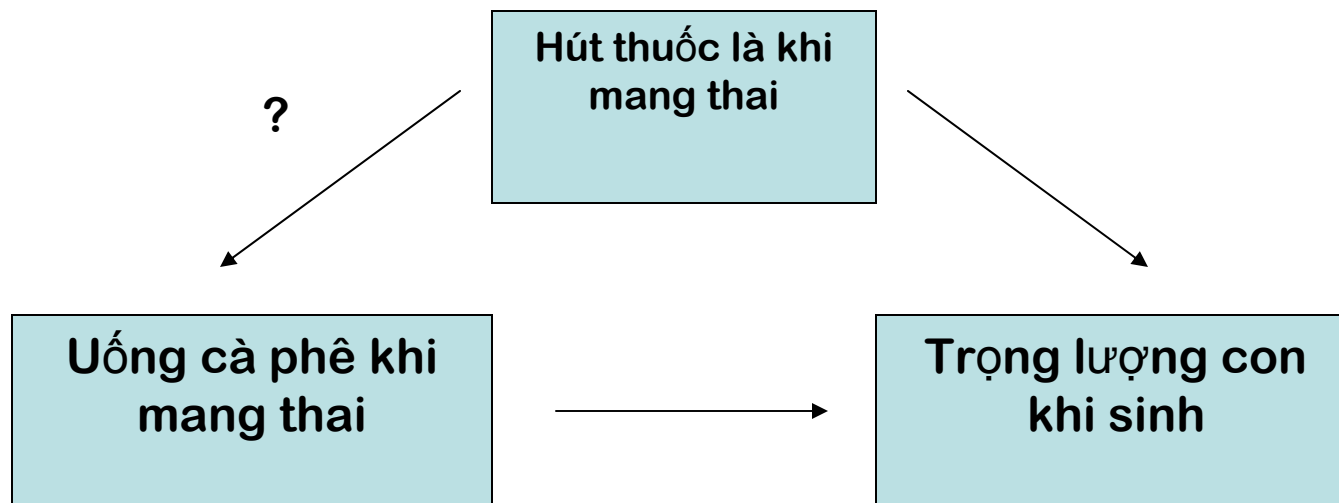
$$OR_{\text{crude}} = 1.73$$

$$OR_{\text{smokers}} = 1.00$$

$$OR_{\text{non-smokers}} = 1.00$$

Mối tương quan giữa uống cà phê và trọng lượng khi sinh bị “confounded” bởi hút thuốc lá. Trong thực tế, không có mối liên hệ nào giữa cà phê và trọng lượng khi sinh

Một ví dụ về confounding

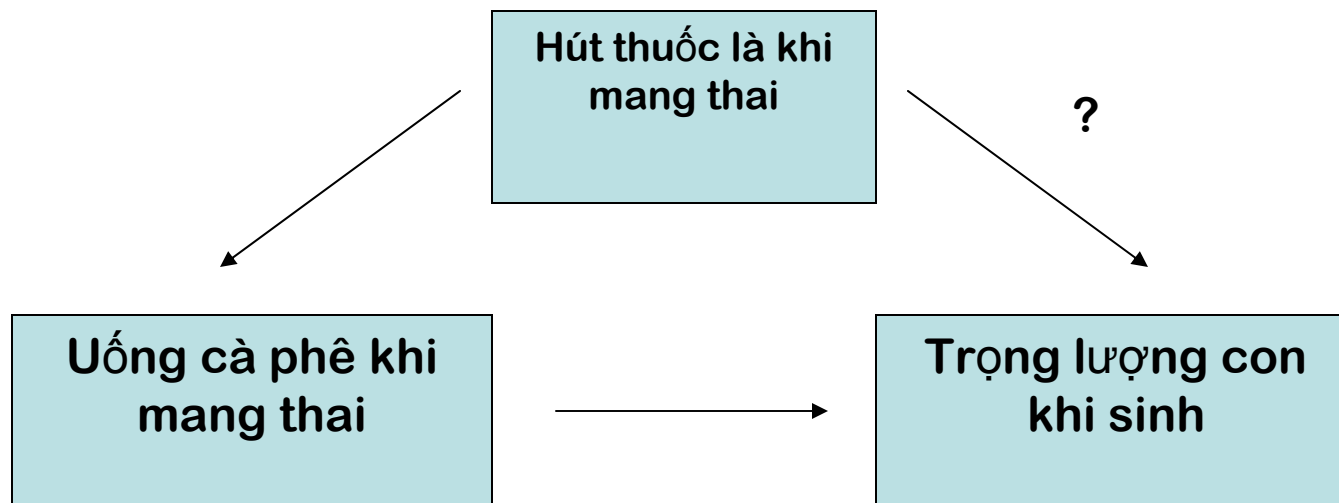


Ví dụ: cà phê và hút thuốc lá

	Smoking	No Smoking
Coffee	176	90
No Coffee	88	90

$$\text{Crude OR} = (176)(90) / (90)(88) = 2.00$$

Ví dụ: cà phê và hút thuốc lá



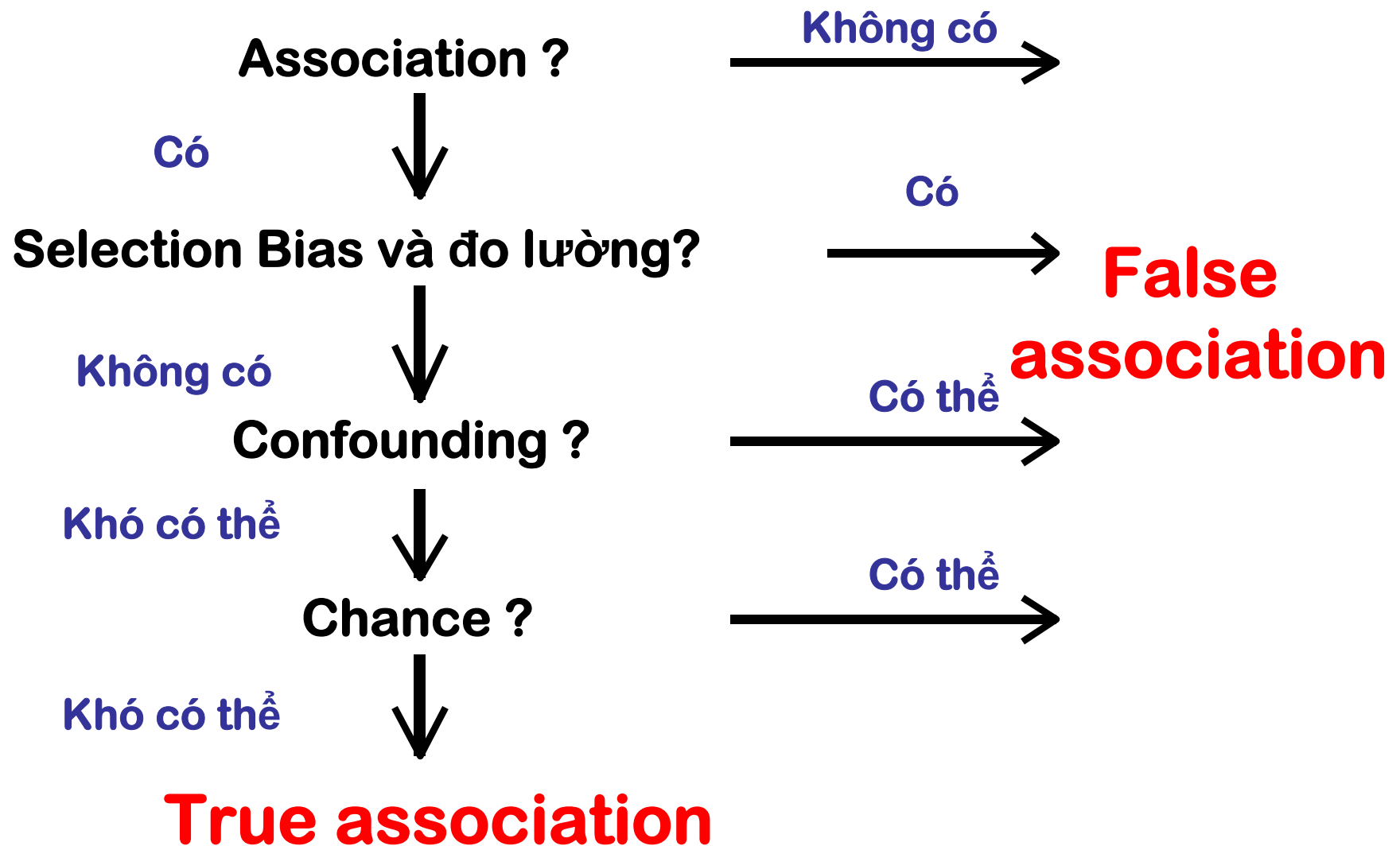
Hút thuốc lá và trọng lượng khi sinh

	Low Birth Weight	Normal Birth Weigh
Smoking	240	24
No Smoking	20	160

$$\text{Crude OR} = (240)(160) / (24)(20) = 80.0$$

Đánh giá mối tương quan

Epidemiologic association: true or false



Tóm lược

- **Hợp lí nội tại (Internal validity) và hợp lí ngoại tại (external validity):**
 - Thiết kế nghiên cứu
 - Phương pháp đo lường
 - Phân tích
- **Biases**
 - Selection bias
 - Information bias
 - Confounding bias

Further reading

1. Grimes DA, Schulz KF, Bias and Causal Associations in Observational Research, Lancet 2002; 359: 248-52
2. Sackett DL, Bias in Analytic Research, J Chron Dis 1979, Vol. 32, 51-6
3. Hill AB. The environment and disease: association or causation? Proc R Soc Med, 1965; 58:295-300
4. Hennekens CH, Epidemiology in Medicine, Lippincott Williams&Wilkins, 1987, first edition
5. Rothman KJ, Modern Epidemiology, Lippincott Williams&Wilkins, 1998, second edition
6. Giesecke J, Modern infectious disease epidemiology
7. Last JM, A Dictionary of Epidemiology, Oxford University press, 2001, fourth edition