

# Bài giảng 22:

## Ước tính cỡ mẫu: nguyên lí

**Nguyễn Văn Tuấn**

Viện nghiên cứu y khoa Garvan (Úc)

Đại học New South Wales và Đại học Công nghệ Sydney

Đại học Tôn Đức Thắng, Việt Nam

**Trường hợp 1:** Tạp chí *Nghiên cứu Y học* (2005, tập XX, số X, trang XX), trong bài “*Nghiên cứu đặc điểm lâm sàng ...*”, có đoạn viết về cỡ mẫu như sau:

“Thiết kế nghiên cứu: Đây là loại nghiên cứu mô tả đánh giá kết quả trước và sau mổ. Cỡ mẫu được tính theo công thức:

$$n = \frac{pqZ^2(1-\alpha/2)}{d^2}$$

Trong đó,  $p = 0,96$ ;  $q = 1-p$ ;  $Z(1-\alpha/2) = 1.96$ , khi  $\alpha = 0,05$ ;  $d$ : sai số ấn định trong nghiên cứu (5%). Từ công thức trên, tính được  $n = 31$ .” (*hết trích*)

**Trường hợp 2:** Cũng cùng số của tạp chí, trang 72, trong bài “*Thực trạng thiếu ăn ...*”, tác giả viết:

“Cỡ mẫu được tính theo công thức:

$$n = \frac{Z^2 p(100 - p)}{e^2}$$

Trong đó, Z = Giá trị tương ứng với mức tin cậy 95% là 1,96 làm tròn bằng 2); p là tỉ lệ suy dinh dưỡng thể nhẹ cân của [...] 2002 (31.1%); e là sai số cho phép thường bằng 0,05. Thay các giá trị vào công thức tính cỡ mẫu cần thiết là: 343 trẻ.” (*hết trích*)

**Trường hợp 3:** Cũng cùng tạp chí, trang 79, có bài “*Kiến thức, thực hành chăm sóc ...*”. Trong đó, tác giả viết:

“Cỡ mẫu: áp dụng công thức cỡ mẫu:

$$n = Z^2 (1 - \alpha / 2) \frac{pq}{d^2}$$

Với p là tỷ lệ bà mẹ được khám thai là đủ 3 lần là 60%, d là sai số cho phép (0.08) và được  $n = 144$ , lấy tròn 150.” (*hết trích*)

**Trường hợp 4:** Cũng cùng số trên tạp chí *Nghiên cứu Y học*, trang 84 có bài “*Kiến thức, thái độ và thực hành ...*”, trong phần phương pháp, tác giả mô tả như sau:

Đối tượng nghiên cứu: 195 bà mẹ sinh con trong thời gian từ ngày 1/12/2003 đến ngày 31/3/2003 tại 9 xã/thị trấn được chọn ngẫu nhiên đại diện chung cho quần thể nghiên cứu của huyện [...] được thực hiện theo phương pháp cắt ngang mô tả với công thức tính cỡ mẫu:

$$n = Z^2 (1 - \alpha / 2) \frac{pq}{d^2} = 192$$

(*hết trích*)

**Trường hợp 5:** *Tạp chí Y học TPHCM* (XX, tập XX, phụ bản X, trang XX) có bài “*Đặc điểm lâm sàng ...*”. Phần phương pháp có đoạn viết về cỡ mẫu như sau:

“Do một dấu hiệu lâm sàng đặc thù của bệnh S là dấu K, cỡ mẫu nghiên cứu nhằm xác định tỉ lệ trẻ có dấu K với sai số tương đối không quá 20% (tỉ lệ trẻ bệnh S có dấu hiệu K được ước đoán vào khoảng 60%). Do đó, cỡ mẫu cần thiết được tính theo công thức

$$N = Z_{1-\alpha/2}^2 \frac{(1-p)}{pe^2} = 3,84 \frac{0,4}{0,6 \times 0,2^2} = 64$$

với N: cỡ mẫu cần thiết cho nghiên cứu

Z<sub>x</sub>: Giá trị của phân phối chuẩn tại định vị x

α: Ngưỡng sai lầm loại I tương ứng với khoảng tin cậy 95%. Vậy α = 0,05

p: Tỉ lệ trẻ em có dấu hiệu K. Giả định p = 0,6

e: Sai số tương đối được chọn bằng 0,2.” (hết trích)

**Trường hợp 6:** Cùng tạp chí *Nghiên cứu Y học*, trang 97, trong bài “*Ảnh hưởng của mô hình ...*”, các tác giả ước tính cỡ mẫu như sau:

“Mẫu nghiên cứu

$$n_1 = n_2 = Z_{\alpha/2}^2 \frac{\left[ (1 - p_1) / p_1 \right] + \left[ (1 - p_0) / p_0 \right]}{\left[ \ln(1 - \varepsilon) \right]^2}$$

Cỡ mẫu: 250 phụ nữ, được tính theo công thức tính cỡ mẫu cho nghiên cứu can thiệp.

Trong đó:  $n_1, n_2$  : cỡ mẫu mỗi nhóm;  $p_1$  : tỷ lệ hiệu biết trong nhóm được can thiệp = 0,75,  $p_0$  tỷ lệ hiệu biết trong nhóm không được can thiệp = 0,35;  $\varepsilon$  là mức chính xác mong muốn, lấy 0.25RR;  $Z(a, b) = 1.96$ . Cỡ mẫu theo tính toán = 117,5 thực tế được nhân với 2 (hệ số thiết kế) và làm tròn 250.” (hết trích)

# Nguyên lí ước tính cỡ mẫu

- Vấn đề hiện hành
- Các "chất liệu" để ước tính cỡ mẫu



## The scandal of poor medical research

*We need less research, better research, and research done for the right reasons*

[Home](#) [Contents](#)

[Records](#)

[jameslindlibrary.org](http://jameslindlibrary.org)

[Key Passages](#) [Commentary](#) [Context](#)

**Bracken MB. Why animal studies are often poor predictors of human reactions to exposure.**

**Commentary on:** Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock S, Macleod M, Mionini LE, Javaram D, Khan KS (2007) Comparison of treatment effects between animal

Research article

[Open Access](#)

**Methodological quality of systematic reviews of animal studies: a survey of reviews of basic research**

Luciano E Mionini<sup>1</sup> and Khalid S Khan<sup>\*2</sup>

Essay

## Why Most Published Research Findings Are False

John P.A. Ioannidis

# Khác biệt có ý nghĩa thực tế và ý nghĩa thống kê

Outcome	Group A	Group B
Improved	9	18
No improved	21	12
Total	30	30
% improved	30	60

Chi-square: 5.4;  $P < 0.05$   
“Statistically significant”

Outcome	Group A	Group B
Improved	6	12
No improved	14	8
Total	20	20
% improved	30	60

Chi-square: 3.3;  $P > 0.05$   
“Statistically insignificant”

## Chúng ta cần gì?

# Những vấn đề trong thiết kế nghiên cứu

- Địa điểm / setting
- Đối tượng: tiêu chuẩn chọn và loại
- Mô hình: survey, prospective, experimental ...
- Đo lường
- Phân tích
- **Sample size / power**

**Bốn yếu tố để ước tính cỡ mẫu**

# Phương pháp ước tính cỡ mẫu

- Không có con số huyền thoại 30!
- Có rất nhiều phương pháp
- Áp dụng đúng phương pháp đòi hỏi suy nghĩ

# Bốn "chất liệu" để ước tính cỡ mẫu

- Sai sót loại I và II
- Biến outcome là gì?
- Độ dao động (độ lệch chuẩn) của outcome
- Mức độ ảnh hưởng

# 1. Sai sót loại I (type I error)

- Nghiên cứu tìm ảnh hưởng / mối liên quan (H)
- Sai sót I: xác suất tìm được H nhưng trong thực tế H không đúng
- Còn gọi là alpha ( $\alpha$ )
- Tương đương với "**dương tính giả**" (false positive)
  - Xác suất kết quả xét nghiệm dương tính, nhưng cá nhân không có bệnh

# Sai sót loại II (type II error)

- Sai sót II: xác suất không tìm được H, nhưng trong thực tế H tồn tại
- Còn gọi là alpha ( $\beta$ )
- Tương đương với "âm tính giả" (false negative)
  - Xác suất kết quả xét nghiệm âm tính, nhưng cá nhân có bệnh



# Power

- $\text{Power} = 1 - \beta$
- Xác suất tìm được H và trong thực tế H tồn tại
- Giống như "sensitivity" = độ nhạy
- Power = độ nhạy của nghiên cứu

# Chẩn đoán và nghiên cứu khoa học

## Chẩn đoán (diagnosis)

	Bệnh trạng	
	Có	Không
Xét nghiệm +ve	True +ve (sensitivity)	False +ve
-ve	False -ve	True -ve (Specificity)

## Nghiên cứu (research)

	Có ảnh hưởng	Không có ảnh hưởng
Test thống kê Có ý nghĩa	OK $1 - \beta$	Sai sót I $\alpha$
Không có	Sai sót II $\beta$	OK $1 - \alpha$

“Có ý nghĩa”:  $P < 0.05$

“Không có ý nghĩa”:  $P > 0,05$

# Các thông số sai sót I và II

- Thông thường chúng ta cho

$$\alpha = 0.05$$

$$\beta = 0.20 \text{ (Power} = 0.80\text{)}$$

# Mỗi giá trị alpha và beta gắn liền với z

- Z là chỉ số của phân bố chuẩn

$\alpha$	$Z_{\alpha}$ (One-sided)	$Z_{\alpha/2}$ (Two-sided)
0.20	0.84	1.28
0.10	1.28	1.64
0.05	1.64	1.96
0.01	2.33	2.81

$\beta$	$Z_{1-\beta}$
0.20	0.84
0.10	1.28
0.05	1.64
0.01	2.33

## 2. Biến outcome dạng gì?

- Biến liên tục
  - Huyết áp, mật độ xương, glucose, insulin, v.v.
- Biến phân loại
  - Sống / chết, bệnh / không bệnh

### 3. Độ dao động của biến outcome

- Nếu là biến liên tục: cần phải biết độ lệch chuẩn (standard deviation)
- Nếu là biến phân loại: cần phải biết tỉ lệ mắc bệnh (hay của biến cố)

# Ảnh hưởng của độ dao động: ví dụ

	True mean: 100 True SD: 15		True mean: 100 True SD: 35	
Sample size	Est. M	SD	Est. M	SD
10	98.0	11.0	108.9	32.2
50	100.4	13.6	95.3	41.4
100	101.3	14.4	99.1	35.5
200	99.9	15.2	100.3	33.2
500	99.8	15.3	98.9	33.8
1000	99.5	15.1	99.9	35.0
2000	99.7	15.0	99.9	34.7
10000	100.1	15.0	99.9	35.0
100000	100.0	15.0	100.0	35.0

## 4. Mức độ ảnh hưởng

- Effect size (ES)
- Một thông số rất quan trọng!
- Tùy thuộc vào mô hình nghiên cứu



## 4. Mức độ ảnh hưởng

- Nghiên cứu cắt ngang: trung bình / sai số (mean / marginal error)
- Nghiên cứu so sánh hai số trung bình ( $m_1$  và  $m_2$ )
- Nghiên cứu so sánh 2 tỉ lệ ( $p_1$  và  $p_2$ )

## 4 "chất liệu" để ước tính cỡ mẫu

- Không có 4 chất liệu trên = không thể ước tính cỡ mẫu
- Cần phải có tất cả 4 thông số

# Ước tính cỡ mẫu

- Phải đọc và **tham khảo nghiên cứu trước** để có thông số
- Phải **suy nghĩ** về biến phân tích (outcome)
- Phải chọn sai số có thể chấp nhận được

# Ước tính cỡ mẫu: 4 câu hỏi

- Tôi chấp nhận sai số I và II bao nhiêu?
- Biến outcome của tôi là gì?
- Độ dao động của biến outcome là bao nhiêu?
- Mức độ ảnh hưởng tôi muốn biết là bao nhiêu

# Ví dụ

- Nghiên cứu so sánh hiệu quả 2 thuật can thiệp A và B về mật độ xương
- Mô hình nghiên cứu: randomized controlled trial (RCT)
- 4 thông số:
  - Alpha = 5%, beta = 10%
  - Biến outcome: glucose
  - Độ lệch chuẩn: 0.12 g/cm<sup>2</sup>
  - Giả thuyết 2 nhóm khác biệt 5% mật độ xương