

Bài giảng 11a: So sánh 2 tỉ lệ bằng relative risk

Nguyễn Văn Tuấn

Garvan Institute of Medical Research, Australia

Đại học Tôn Đức Thắng, Việt Nam

Chúng ta đã học

- Hai tỉ lệ p_1 và p_2
- So sánh
 - Tính $d = p_1 - p_2$
 - Tính độ lệch chuẩn của d , gọi là s
 - Tính $z = d / s$
- Có hai cách khác để mô tả khác biệt
 - $RR = p_1 / p_2$
 - $OR = (p_1 / (1 - p_1)) / (p_2 / (1 - p_2))$

Nội dung

- Khái niệm *risk* (nguy cơ)
- Tỷ số nguy cơ (*relative risk* hay *risk ratio*, RR)
- Cách ước tính khoảng tin cậy 95% của RR
- Phân tích bằng R

Khái niệm "risk"

Định nghĩa *risk*

- Xác suất một biến cố bất lợi xảy ra trong một thời gian nhất định
- Ba yếu tố:
 - xác suất
 - biến cố bất lợi (adverse event)
 - thời gian

Ví dụ minh họa cho *risk*

- Một nhóm bệnh nhân gồm **100** người được điều trị. Trong thời gian 3 năm sau đó, có **15** người tử vong

risk of mortality – nguy cơ tử vong là

$$15/100 = 0.15 \text{ hay } 15\%$$

(có khi đề cập như là *incidence*)

Nguy cơ tử vong của 2 nhóm

- Nghiên cứu lâm sàng đối chứng ngẫu nhiên (Randomized controlled trial)
- Hai nhóm bệnh nhân, theo dõi 3 năm
- Tiêu chí đánh giá: tử vong

	Rx	Placebo
Số bệnh nhân	1065	1062
Tử vong	101	141

Nguy cơ

	Rx	Placebo
Số bệnh nhân	1065	1062
Tử vong	101	141
Nguy cơ tử vong	0.095	0.133
Kí hiệu	p_1	p_2

- **Câu hỏi nghiên cứu:** Thuốc có hiệu quả giảm tử vong?
- **Câu hỏi thống kê:** nguy cơ tử vong khác nhau?

Khái niệm *risk ratio*

Thuật ngữ

- Risk ratio
- Relative risk
- Incidence risk ratio

Trường hợp chung

	Rx	Control
Bệnh	a	b
Không bệnh	c	d
Cỡ mẫu	N_1	N_2

Nguy cơ

$$p_1 = \frac{a}{N_1}$$

$$p_2 = \frac{b}{N_2}$$

Hai cách để so sánh

- Hiệu số (absolute risk difference)

$$d = p_1 - p_2$$

- Tỷ số hai nguy cơ

$$RR = \frac{p_1}{p_2}$$

Ý nghĩa của RR

- p_1 : nguy cơ nhóm *được điều trị*, p_2 nguy cơ nhóm *không điều trị*

$$RR = p_1 / p_2$$

- Ý nghĩa:
 - Nếu $RR = 1$, hai nhóm có hiệu quả như nhau
 - Nếu $RR > 1$, thuốc có hại
 - Nếu $RR < 1$, thuốc có lợi

Ước tính khoảng tin cậy 95% của RR

- Nhưng RR chỉ là số *trung bình*
- Nếu lặp lại nghiên cứu, RR có thể khác
- Lặp lại lần nữa, RR lại khác
- Nếu lặp lại 100 lần, và nếu tất cả hay 95% RR đều dưới 1 thì đó là chứng cứ khoa học
- Vấn đề: làm sao ước tính khoảng tin cậy 95%?

Qui trình ước tính khoảng tin cậy

- Cần phải ước tính phương sai của RR
- Sau đó, ước tính độ lệch chuẩn (căn số bậc 2 của phương sai)
- Áp dụng lí thuyết phân bố chuẩn: khoảng tin cậy 95% bằng

trung bình $\pm 1.96 \times$ độ lệch chuẩn

Ước tính phương sai của RR

- RR là một tỉ số
- Rất khó ước tính phương sai của một tỉ số!

Tính phương sai của $\log(RR)$

Cách tính "gián tiếp" hay "vòng"

1. Hoán chuyển RR sang $\log(RR)$
2. Tính phương sai (và độ lệch chuẩn) của $\log(RR)$
3. Tính khoảng tin cậy 95% của $\log(RR)$
4. Hoán chuyển khoảng tin cậy $\log(RR)$ sang RR

Ví dụ cách tính khoảng tin cậy 95%

	Rx	Control
Bệnh	a	b
Không bệnh	c	d
Cỡ mẫu	N_1	N_2

Nguy cơ

$$p_1 = \frac{a}{N_1}$$

$$p_2 = \frac{b}{N_2}$$

- $RR = p_1 / p_2$
- $L = \log(RR) = \log(p_1) - \log(p_2)$

Ví dụ cách tính khoảng tin cậy 95%

	Rx	Control
Bệnh	a	b
Không bệnh	c	d
Cỡ mẫu	N_1	N_2

Nguy cơ

$$p_1 = \frac{a}{N_1}$$

$$p_2 = \frac{b}{N_2}$$

- $RR = p_1 / p_2$
- $L = \log(RR) = \log(p_1) - \log(p_2)$

- Phương sai của L:
$$V = \left(\frac{1}{a} - \frac{1}{N_1} \right) + \left(\frac{1}{b} - \frac{1}{N_2} \right)$$

Ví dụ cách tính khoảng tin cậy 95%

- $RR = p_1 / p_2$
- $L = \log(RR) = \log(p_1) - \log(p_2)$

- Phương sai của L:
$$V = \left(\frac{1}{a} - \frac{1}{N_1} \right) + \left(\frac{1}{b} - \frac{1}{N_2} \right)$$

- Độ lệch chuẩn của L: $S = V^{1/2}$
- Khoảng tin cậy 95% của L: $L \pm 1.96xS$
- Khoảng tin cậy 95% của RR:
 $\exp(L-1.96xS)$ đến $\exp(L+1.96xS)$

Phân tích với R

Tính toán "thủ công"

	Rx	Placebo
Tử vong	101	141
Số bệnh nhân	964	921
Số bệnh nhân	1065	1062
Tỉ lệ tử vong	0.0948	0.1327

- Tính p_1 và p_2
- $RR = p_1 / p_2$
- $L = \log(RR)$
- $V = 1/101 - 1/1065 + 1/141 - 1/1062$
- $S = \text{căn số bậc 2 của } V$
- KTC95: $L - 1.96S$ đến $L + 1.96S$
- KTC95% RR: $\exp(L - 1.96S)$ đến $\exp(L + 1.96S)$

```
a=101; b=141;
n1=1065; n2=1062
RR = (a/n1) / (b/n2)
L = log(RR)
S = sqrt(1/a-1/n1+1/b-1/n2)
L95 = exp(L-1.96*S)
U95 = exp(L+1.96*S)
cbind(RR, L95, U95)
```

Kết quả tính toán thủ công

```
> cbind(RR, L95, U95)
```

RR

L95

U95

```
[1,] 0.7142943 0.5613506 0.9089084
```

Dùng package **epitools**

```
library(epitools)
```

```
data = matrix(c(101, 964, 141, 921), byrow=T,  
nrow=2)
```

```
riskratio(data, rev="c")
```


Dùng package **epitools**

```
> riskratio(tab, rev="c")
```

Predictor	Outcome		
	Disease2	Disease1	Total
Exposed1	964	101	1065
Exposed2	921	141	1062
Total	1885	242	2127

Tại sao RR = 1.49?

(Trước đó RR = 0.71)

\$measure

risk ratio with 95% C.I.

Predictor	estimate	lower	upper
Exposed1	1.000000	NA	NA
Exposed2	1.399983	1.100226	1.78141

\$p.value

two-sided

Predictor	midp.exact	fisher.exact	chi.square
Exposed1	NA	NA	NA
Exposed2	0.005904708	0.006268449	0.005874976

Tóm lược

- RR = risk ratio hay relative risk
- Ước tính khoảng tin cậy 95% của RR là một cách đánh giá mức độ ảnh hưởng