

Lâm sàng thống kê

Ý nghĩa của odds ratio và relative risk

Nguyễn Văn Tuấn

Hỏi: “Nhiều bài báo nghiên cứu y học lúc thì báo cáo odds ratio, lúc thì relative risk. Em xin hỏi thầy hai chỉ số này khác nhau như thế nào và chỉ số nào đúng. Em thấy hình như không có sách nào ở Việt Nam giải thích rõ về hai chỉ số này.”

Đây là một thắc mắc không chỉ ở nước ta mà ... toàn thế giới. Rất nhiều nhà nghiên cứu, dù đã công bố báo cáo khoa học trên các tập san danh tiếng như New England Journal of Medicine, vẫn hiểu sai – có khi cực kỳ sai lầm – về hai chỉ số này. Tôi muốn mượn câu hỏi này để giải thích những khác biệt quan trọng của hai chỉ số odds ratio (OR) và relative risk (RR). Tạm thời tôi sẽ không dịch hai thuật ngữ này ở đây. Tôi hi vọng rằng qua bài viết ngắn này, bạn đọc sẽ nắm vững vấn đề hơn và phân biệt những khác biệt căn bản của hai chỉ số để không vấp phải những sai lầm như các nhà nghiên cứu y học Tây phương từng phạm phải.

OR và RR: cơ chế tính toán

OR và RR là hai chỉ số thống kê rất phổ biến và có ích trong dịch tễ học, vì cả hai chỉ số kiểm định mối liên hệ giữa một yếu tố nguy cơ và bệnh tật – một mục tiêu gần như căn bản của nghiên cứu y học hiện đại. Cơ chế tính toán của hai chỉ số này cực kỳ đơn giản.

RR: Nếu trong một nhóm n_1 bệnh nhân bị phơi nhiễm một yếu tố nguy cơ (chẳng hạn như hút thuốc lá), và sau khi theo dõi trong một thời gian, có k_1 bệnh nhân mắc bệnh, thì tỉ lệ mắc bệnh đơn giản là $\frac{k_1}{n_1}$ (hãy gọi tỉ lệ này là p_1). Nếu một nhóm n_2 bệnh nhân khác không bị phơi nhiễm yếu tố nguy cơ đó và trong cùng thời gian theo dõi có k_2 bệnh nhân mắc bệnh, thì tỉ lệ mắc bệnh là $p_2 = \frac{k_2}{n_2}$. Do đó, RR đơn giản là tỉ số của hai tỉ lệ:

$$RR = \frac{p_1}{p_2} \quad [1]$$

Nếu $RR > 1$ (hay $p_1 > p_2$), chúng ta có thể phát biểu rằng yếu tố nguy cơ làm tăng khả năng mắc bệnh; nếu $RR = 1$ (tức là $p_1 = p_2$), chúng ta có thể nói rằng không có mối liên hệ nào giữa yếu tố nguy cơ và khả năng mắc bệnh; và nếu $RR < 1$ (tức $p_1 < p_2$),

chúng ta có bằng chứng để thể phát biểu rằng yếu tố nguy cơ có thể làm giảm khả năng mắc bệnh.

OR: Thay vì sử dụng tỉ lệ phát sinh p để đo lường khả năng mắc bệnh, thống kê cung cấp cho chúng ta một chỉ số khác: đó là *odd*. Odd được định nghĩa là tỉ số của hai xác suất. Nếu p là xác suất mắc bệnh, thì $1-p$ là xác suất sự kiện không mắc bệnh. Theo đó, odd được định nghĩa bằng:

$$odd = \frac{p}{1-p} \quad [2]$$

Như vậy, nếu $odd > 1$, khả năng mắc bệnh cao hơn khả năng không mắc bệnh; nếu $odd = 1$ thì điều này cũng có nghĩa là khả năng bằng với khả năng không mắc bệnh; và nếu $odd < 1$, chúng ta có thể nói khả năng mắc bệnh thấp hơn khả năng không mắc bệnh.

Với định nghĩa này, chúng ta quay lại với ví dụ vừa trình bày trong mục RR. Odd mắc bệnh trong nhóm bị phơi nhiễm yếu tố nguy cơ là: $odd_1 = \frac{p_1}{1-p_1}$, và odd mắc bệnh trong nhóm không bị phơi nhiễm yếu tố nguy cơ là: $odd_2 = \frac{p_2}{1-p_2}$. Từ đó, OR được định nghĩa là tỉ số của hai odds:

$$OR = \frac{p_1 / (1-p_1)}{p_2 / (1-p_2)} \quad [3]$$

(Hay nếu ai thích đại số, OR có thể viết lại “gọn” hơn: $OR = \frac{p_1}{1-p_1} \times \frac{1-p_2}{p_2}$).

Mối liên hệ giữa RR và OR. Qua công thức [1], [2] và [3], chúng ta có thể thấy OR và RR có một mối liên hệ số học. Nếu ai thích đại số, có thể viết lại công thức RR như là một hàm số của OR (hay ngược lại), nhưng ở đây, tôi chỉ muốn lưu ý một điểm quan trọng có liên quan đến việc diễn dịch RR và OR.

Nhìn vào công thức [2] của odd, chúng ta dễ dàng thấy nếu tỉ lệ mắc bệnh p thấp (chẳng hạn như 0.001 hay 0.01 – tức 0.1% hay 1%), thì $odd = p$. Chẳng hạn như nếu $p = 0.01$, thì $1-p = 0.99$, và do đó $odd = 0.01 / 0.99 = 0.010101$. Quay lại với công thức [3], nếu p_1 hay p_2 thấp, thì OR có thể viết như sau:

$$OR = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} \approx \frac{p_1}{p_2} \quad [4]$$

Nói cách khác, nếu tỉ lệ mắc bệnh thấp, thì OR gần bằng với RR. Nhưng nếu tỉ lệ mắc bệnh cao (chẳng hạn như trên 10%) thì chỉ số OR cũng cao hơn chỉ số RR.

Có thể làm một vài tính toán để thấy sự khác biệt giữa RR và OR qua bảng số liệu sau đây (Bảng 1). Với những trường hợp tỉ lệ mắc bệnh dưới 5%, OR và RR không khác nhau đáng kể. Nhưng nếu tỉ lệ mắc bệnh cao hơn 10%, thì OR thường ước tính RR cao hơn thực tế.

Bảng 1. So sánh RR và OR với nhiều tỉ lệ khác nhau (số liệu mô phỏng)

Trường hợp	Tỉ lệ mắc bệnh		Odd mắc bệnh		So sánh giữa RR và OR	
	Nhóm 1 (p_1)	Nhóm 2 (p_2)	Nhóm 1 (odd_1)	Nhóm 2 (odd_2)	RR	OR
1	0.001	0.003	0.002	0.003	3	3.01
2	0.01	0.03	0.01	0.03	3	3.06
3	0.02	0.06	0.02	0.06	3	3.13
4	0.05	0.15	0.05	0.18	3	3.35
5	0.10	0.30	0.11	0.43	3	3.86
6	0.15	0.45	0.18	0.82	3	4.64
7	0.20	0.60	0.25	1.50	3	6.00
8	0.25	0.75	0.33	3.00	3	9.00
9	0.30	0.90	0.43	9.00	3	21.0
10	0.33	0.99	0.49	99.0	3	2101.0

Chú ý: Bảng trên đây được mô phỏng sao cho RR = 3 để chứng minh rằng OR ước tính độ ảnh hưởng cao hơn so với thực tế.

RR và OR: ứng dụng

Ví dụ 1: truy tìm ung thư vú. Chương trình truy tìm ung thư vú được khuyến khích như là một phương cách y tế công cộng nhằm giảm tỉ lệ tử vong từ bệnh này ở phụ nữ. Một nhóm nghiên cứu ở Thụy Điển tiến hành một nghiên cứu lâm sàng đối chứng ngẫu nhiên, mà trong đó họ tuyển các phụ nữ tuổi 50 trở lên, và chia thành 2 nhóm: nhóm A gồm 66103 phụ nữ được chụp X quang (tức mammography) thường xuyên (cứ mỗi năm một lần), và nhóm B gồm 66105 phụ nữ không chụp X quang mà chỉ theo dõi bình thường (tức nhóm đối chứng). Sau 5 năm, nhóm A có 183 người chết vì ung thư vú và nhóm B có 177 người chết. Số liệu được trình bày trong Bảng 2 sau đây:

Bảng 2: Truy tìm ung thư vú và tử vong

Nhóm	Tổng số đối tượng tham gia	Số tử vong
A – X quang	66,103	183
B – Đối chứng	66,105	177

Với số liệu này, chúng ta có thể thấy tỉ lệ tử vong trong nhóm A là $p_A = 183/66103 = 0.002768$ và nhóm B là $p_B = 177/66105 = 0.002678$. Từ đó, RR có thể ước tính bằng công thức [1] như sau:

$$RR = \frac{0.002768}{0.002678} = 1.034$$

Nhưng thay vì áp dụng RR, chúng ta cũng có thể áp dụng OR. Odd tử vong trong nhóm chụp mammography là: $odd_A = \frac{p_A}{1-p_A} = \frac{0.002768}{1-0.002768} = 0.002776$, và nhóm đối

chứng: $odd_B = \frac{p_B}{1-p_B} = \frac{0.002678}{1-0.002678} = 0.002685$. Do đó, OR là:

$$RR = \frac{odd_A}{odd_B} = \frac{0.002776}{0.002685} = 1.034$$

Như vậy, OR bằng RR. Nhưng cách diễn dịch của OR khác với RR. Bởi vì đơn vị của RR là tỉ lệ tử vong, cho nên chúng ta có thể nói rằng nhóm chụp mammography thường xuyên có tỉ lệ tử vong cao hơn nhóm đối chứng khoảng 3.4%. Nhưng đơn vị của OR là odd, cho nên chúng ta không thể phát biểu về “tỉ lệ tử vong”, mà chỉ có thể phát biểu rằng “khả năng” tử vong của nhóm A cao hơn nhóm B khoảng 3.4%. Ở đây, vì tỉ lệ tử vong thấp, cho nên như công thức [4] cho thấy hai chỉ số này giống nhau, và trong thực tế chúng ta có thể diễn dịch một OR như là một RR.

Cách phân biệt trên có vẻ máy móc và lí thuyết, nhưng quan trọng. Để thấy rõ nguy hiểm trong cách diễn dịch OR, tôi sẽ trình bày một ví dụ sau đây:

Ví dụ 2: sắc tộc và tỉ lệ thông tim (cardiac catheterization). Tập san New England Journal of Medicine số ra ngày 25/2/1999 (tập 349; trang 618-626) công bố một nghiên cứu rất thú vị về ảnh hưởng của sắc tộc đến tỉ lệ thông tim. Trong nghiên cứu này, các nhà nghiên cứu mượn một số diễn viên điện ảnh người da trắng và da đen đóng vai bệnh nhân. Các diễn viên được chỉ cách trình bày các triệu chứng và bệnh trạng cẩn thận và đầy đủ, nhưng giống nhau. Họ thu hình các diễn viên vào video; chọn ngẫu nhiên 720 bác sĩ chuyên khoa tim người da trắng, cho họ xem các video này, và hỏi “ai cần được thông tim?”. Kết quả cho thấy 90.6% bác sĩ đề nghị các bệnh nhân da trắng nên

được thông tim, nhưng tỉ lệ này cho bệnh nhân da đen chỉ 84.7%. Một phần của kết quả có thể tóm lược trong Bảng 3 sau đây:

Bảng 3: Sắc tộc và tỉ lệ thông tim

Nhóm	Số bác sĩ đề nghị thông tim	Số bác sĩ không đề nghị thông tim
w - Bệnh nhân da trắng	652	68
b - Bệnh nhân da đen	610	110

Các nhà nghiên cứu kết luận rằng tỉ lệ bệnh nhân da đen được thông tim thấp hơn tỉ lệ ở bệnh nhân da trắng đến 40%. Sau khi nghiên cứu này công bố, giới truyền thông rầm rộ bàn về kết quả và ý nghĩa của nghiên cứu. Không cần nói ra, cũng có thể đoán được trong dư âm và tình trạng kì thị chủng tộc ở Mỹ còn kéo dài, những nhóm đấu tranh chống kì thị chủng tộc lấy kết quả này để làm bằng chứng tố cáo rằng các bác sĩ da trắng kì thị bệnh nhân da đen. Ý nghĩa còn sâu xa hơn: sự kì thị này có thể dẫn đến tử vong. Nói cách khác, có người diễn dịch rằng đây là một sự giết người có ý đồ!

Nhưng rất tiếc là con số 40% đó đã được diễn dịch cực kì sai. Không những diễn dịch sai mà cách tính toán cũng sai. Để hiểu tại sao cách diễn dịch đó sai, chúng ta hãy bắt đầu bằng cách tính OR của các tác giả. Odd thông tim trong nhóm bệnh nhân da trắng là: $odd_w = 652/68 = 9.59$. Tính tương tự, odd thông tim trong nhóm bệnh nhân da đen là: $odd_b = 610/110 = 5.545$. Do đó, OR giữa nhóm da đen và da trắng là:

$$OR = \frac{odd_b}{odd_w} = \frac{5.545}{9.59} = 0.578$$

(Chú ý, các nhà nghiên cứu tính số chẵn, nên OR là khoảng 0.60).

Nhưng nếu thay vì tính bằng odd, chúng ta có thể tính bằng tỉ lệ đề nghị thông tim cho hai nhóm. Tỉ lệ này cho nhóm bệnh nhân da trắng là: $p_w = 652/720 = 0.906$, và nhóm bệnh nhân da đen: $p_b = 610/720 = 0.847$. Và RR có thể tính từ hai tỉ lệ này:

$$RR = \frac{p_b}{p_w} = \frac{0.847}{0.906} = 0.935$$

Đến đây, chúng ta thấy rõ rằng, tỉ lệ đề nghị thông tim trong nhóm bệnh nhân da đen thấp hơn tỉ lệ trong nhóm bệnh nhân da trắng là 6.5% (lấy 1 trừ cho 0.935 và nhân kết quả cho 100), chứ không phải 40% như các tác giả và giới truyền thông tuyên bố.

Tại sao có sự khác biệt? Tại vì các tác giả và giới truyền thông nhầm lẫn rằng OR là RR. Trong trường hợp này, OR không phải là một chỉ số thích hợp để phân tích số liệu, bởi vì số tỉ lệ quá cao (84.7% và 90.6%), và vì tỉ lệ quá cao, cho nên OR ước tính RR quá cao hơn thực tế.

Thật ra, ở đây cách gọi “RR” cũng không chính xác. RR chỉ sử dụng cho tỉ lệ phát sinh (incidence), nhưng trong trường hợp này không có tỉ lệ phát sinh, mà là tỉ lệ lưu hành (prevalence). Do đó, thuật ngữ chính xác để mô tả 0.935 là **prevalence ratio (PR)**. (Đây là một đề tài khác mà tôi hi vọng sẽ có dịp quay lại để bàn thêm).

Điều ngạc nhiên là sai sót này lại hiện diện ngay trên giấy trắng mực đen của một tập san y học vào hàng số 1 trên thế giới!

Vấn đề diễn dịch OR

RR là tỉ số của hai tỉ lệ, và tỉ lệ thì chúng ta có thể hiểu được khá dễ dàng. Nếu nói tỉ lệ mắc bệnh 3%, chúng ta nghĩ ngay đến 3 trong 100 người mắc bệnh. Vì thế, vấn đề diễn dịch RR khá dễ dàng. Nếu $RR = 2$, chúng ta có thể nói rằng tỉ lệ tăng gấp 2 lần. Ai cũng hiểu được mà không cần phải hỏi thêm.

OR là tỉ số của hai odds. Odd phản ánh “khả năng” mắc bệnh. $Odd = 2$ có nghĩa là khả năng mắc bệnh cao hơn khả năng không mắc bệnh 2 lần. Khó hiểu. Odd đã khó hiểu thì tỉ số của hai odds (hay hai khả năng) lại càng là một đo lường khó hiểu hơn vì nó quá chung chung, khó cảm nhận được. Thật ra, một người bình thường (kể cả giới bác sĩ) khó có thể hiểu chính xác nghĩa của OR. Chúng ta biết $OR = 2$ không hẳn có cùng nghĩa với $RR = 2$. Chính vì thế mà gần đây có “phong trào xét lại” OR trên các tập san y học quốc tế. Nhiều nhà nghiên cứu, dịch tễ học và thống kê học kêu gọi bỏ OR!

Nhưng bất cứ đo lường nào cũng lợi thế và khiếm khuyết. RR, dù dễ diễn dịch cũng có khiếm khuyết của nó. Lấy ví dụ đơn giản: nếu tỉ lệ mắc bệnh ung thư trong nhóm A là 1% và nhóm B là 3%, chúng ta dễ dàng thấy $RR = 3$. Nhưng thay vì nói mắc bệnh, chúng ta lật ngược lại vấn đề “không mắc bệnh”: chúng ta có tỉ lệ cho nhóm A là 99% so với nhóm B là 97%, và như thế $RR = 0.97 / 0.99 = 0.98$, tức là tỉ lệ không mắc bệnh trong nhóm B thấp hơn nhóm A khoảng 2%. (Nhưng nếu dùng “mắc bệnh”, nhóm A mắc bệnh nhiều hơn nhóm B đến 3 lần!) Nói cách khác, RR có thể thiếu tính nhất quán (consistency).

Nhưng OR thì không có vấn đề thiếu nhất quán. Trong ví dụ trên, nếu lấy chỉ số là “mắc bệnh” làm so sánh, OR là 3.06. Nhưng nếu lấy “không mắc bệnh” làm chỉ số so sánh, thì OR vẫn là 3.06 (bạn đọc có thể kiểm tra con số này). Trong toán thống kê,

người ta gọi đặc tính của OR là symmetric (đối xứng), còn đặc tính của RR là asymmetric (bất đối xứng).

OR, RR và thể loại nghiên cứu

Một khác biệt cơ bản nữa giữa RR và OR là sự tùy thuộc vào thể loại nghiên cứu. Nói một cách ngắn gọn, RR chỉ có thể ước tính từ nghiên cứu xuôi thời gian (cohort prospective study), nhưng OR thì có thể ước tính từ tất cả thể loại nghiên cứu, nhưng chủ yếu là nghiên cứu bệnh – chứng.

Giả sử chúng ta muốn tìm hiểu mối liên hệ giữa phơi nhiễm chất độc màu da cam (Agent Orange – AO) và bệnh ung thư. Một cách nghiên cứu qui mô là tuyển chọn [ngẫu nhiên] một nhóm đối tượng, sau đó phân nhóm dựa vào tiền sử có bị phơi nhiễm độc chất hay không. Sau đó, theo dõi cả hai nhóm đối tượng một thời gian (chẳng hạn như 5 năm) và ghi nhận số người bị ung thư. Kết quả của nghiên cứu như thế có thể tóm lược trong **Bảng 4** sau đây. Trong số 1000 người được thăm định bị phơi nhiễm lúc ban đầu, có 20 người (hay 2%) bị ung thư trong thời gian theo dõi; trong số 10,000 người không bị phơi nhiễm AO, có 100 người (tức 1%) bị ung thư sau đó. Như vậy, $RR = 0.02/0.01 = 2$. Nhưng nếu tính bằng odd thì $OR = 2.02$. Hai chỉ số này không khác nhau đáng kể.

Bảng 4. Một nghiên cứu xuôi thời gian (giả tưởng)

Nhóm	Ung thư	Không ung thư	Tổng số
Phơi nhiễm AO	20	980	1000
Không phơi nhiễm AO	100	9900	10000

Nhưng theo dõi đối tượng một thời gian dài thường rất tốn kém. Một phương pháp nghiên cứu khác cũng có thể đáp ứng mục đích tìm hiểu mối liên hệ giữa AO và ung thư, nhưng cần ít đối tượng hơn và không cần theo dõi một thời gian dài: đó là nghiên cứu bệnh – chứng. **Bảng 5** dưới đây trình bày kết quả một nghiên cứu (giả tưởng) như thế. Trong nghiên cứu này, chúng ta chọn 100 bệnh nhân ung thư và 100 đối tượng không bị ung thư, nhưng hai nhóm này tương đương nhau về các yếu tố nguy cơ. Sau đó, chúng ta tìm hiểu qua hồ sơ bệnh lý (hay phỏng vấn) trong mỗi nhóm có bao nhiêu người bị phơi nhiễm độc chất. Nói cách khác, đây là một nghiên cứu “ngược thời gian” (so với nghiên cứu “xuôi thời gian” như trình bày trong Bảng 4. Kết quả nghiên cứu bệnh chứng này được trình bày như sau:

Bảng 5. Một nghiên cứu bệnh - chứng (giả tưởng)

Nhóm	Ung thư	Không ung thư
------	---------	---------------

Phơi nhiễm AO	10	5
Không phơi nhiễm AO	90	95
Tổng số	100	100

Trong nhóm bệnh nhân, có 10 người (hay 10%) từng bị phơi nhiễm AO; và trong nhóm không ung thư số đối tượng từng bị phơi nhiễm là 5 người (hay 5%). Ở đây, chúng ta không thể tính tỉ lệ phát sinh bệnh (incidence), bởi vì số lượng bệnh nhân và đối chứng đã được xác định trước. Vì không thể ước tính tỉ lệ phát sinh, nghiên cứu bệnh chứng không cho phép chúng ta ước tính RR. Tuy nhiên, chúng ta có thể tính OR, và OR trong trường hợp này là một ước tính chỉ số RR.

Số liệu **Bảng 5** cho thấy odd bị phơi nhiễm trong nhóm bệnh nhân là: $10/90 = 0.1111$, và nhóm đối chứng: 0.05263 . Do đó, $OR = 0.1111 / 0.05263 = 2.11$. Thật ra, có thể tính đơn giản hơn bằng công thức “giao chéo”: $OR = \frac{10 \times 95}{90 \times 5} = 2.11$.

Điểm chính để phân biệt hai hình thức nghiên cứu này là phương pháp chọn mẫu. Với nghiên cứu xuôi thời gian, chúng ta xác định số lượng đối tượng theo yếu tố nguy cơ ngay từ đầu, và số lượng bệnh phát sinh là một số ghi nhận. Ngược lại, với nghiên cứu ngược thời gian, chúng ta xác định số lượng bệnh nhân và đối tượng ngay từ đầu, và số lượng phơi nhiễm yếu tố nguy cơ là số ghi nhận.

Tuy kết quả nghiên cứu của hai thể loại nghiên cứu được trình bày rất giống nhau: hai cột và hai dòng (2x2 table), nhưng “câu chuyện” đằng sau của các số liệu này rất khác nhau. Không am hiểu câu chuyện đằng sau của một bảng số liệu rất dễ dàng sai lầm trong khi phân tích!

Tóm tắt

Để kết thúc bài viết này, tôi xin nhấn mạnh một số điểm chính:

- Cả hai RR và OR đều là những chỉ số phản ánh độ tương quan giữa một yếu tố nguy cơ và bệnh; nhưng RR mới là chỉ số chúng ta cần biết (còn OR chỉ là ước số của RR).
- Vì OR tùy thuộc vào tỉ lệ hiện hành (và khi tỉ lệ hiện hành của bệnh cao – như trên 10% – thì OR thường cao hơn so với thực tế), cho nên các nghiên cứu này nên sử dụng prevalence ratio.

- Các nghiên cứu bệnh chứng (case-control study) chỉ có thể cung cấp OR, chứ không phải RR. Các nghiên cứu cắt ngang có thể cung cấp OR, chứ không phải RR.
 - Các nghiên cứu xuôi thời gian (prospective study) cung cấp cả OR và RR, nhưng RR là chỉ số có ý nghĩa và dễ diễn dịch.
 - Tuy nhiên, RR không nhất quán như OR.
-