

“Đo lường” bệnh tật: Tỉ số, tỉ lệ, tỉ suất, v.v...

Nguyễn Văn Tuấn

Một bạn đọc viết thư hỏi như sau: “*Em là một Bs hiện đang công tác tại Trung tâm Y tế dự phòng, trong thời gian qua em đã theo dõi rất kỹ các bài viết nào của thầy trên ykhoanet. Em đã học được ở thầy rất nhiều vấn đề, đặc biệt là phương pháp dịch tễ học ... Thưa thầy, hiện nay, Bộ Y tế đã soạn xong một bộ tài liệu, đang đưa vào sử dụng, có 3 quyển, [...]. Qua đọc quyển 1, ở phần nội dung tỉ số, tỉ lệ và tỉ suất, em thấy các tác giả viết về tỉ lệ, tỉ suất rất khác so với hồi em học với Giáo sư Dương Đình Thiện (ở bộ môn VSMT - dịch tễ, trường ĐHY Hà Nội trước kia, nay đã về hưu). Em nhờ thầy trình bày giúp em rõ về tỉ lệ, tỉ suất trong dịch tễ học. Em xin cảm ơn và kính chúc sức khỏe thầy. (Em gửi cho thầy tóm tắt tài liệu quyển 1 mà các tác giả soạn theo file gửi kèm)”*

Thành thật cảm ơn bạn (tôi gọi thế vì không biết bạn đọc là anh hay chị) về lời khích lệ, và cảm ơn một số định nghĩa và cách tính của BYT mà bạn gửi. Tôi nghĩ dù sao đi nữa thì đây vẫn là một điều tốt. Phòng chống dịch cần phải có thông tin và phương pháp. Tôi nghĩ tập tài liệu sẽ giúp đỡ cho những người trực tiếp với công việc này.

Tôi sẽ không bình luận chi tiết những chỉ dẫn về cách tính các chỉ số dịch tễ học trong tài liệu. Cảm tưởng chung của tôi là khó hiểu, một phần vì thiếu minh họa, một phần do công thức khá rối rắm với hằng số 10^n , và một phần do ... trình độ tiếng Việt của tôi có hạn. Ngoài ra, hình như tác giả nhầm lẫn giữa “tỉ suất mật độ” và “tỉ suất mới mắc”, vì thật ra hai chỉ số này là một (không có gì khác nhau cả)! Do đó, tôi chỉ mô tả những gì tôi hiểu về các chỉ số này để bạn so sánh và rút ra kết luận.

Trong các nghiên cứu trên quần thể (population), không chỉ riêng y khoa, chúng ta có khá nhiều chỉ số để đánh giá qui mô và tầm quan trọng của một sự kiện trong một quần thể hay cho một cá nhân. Sự phổ biến của các chỉ số này còn tùy thuộc vào mức độ chuyên sâu và tinh vi của nghiên cứu, nhưng ở mức đơn giản nhất, có 3 chỉ số quan trọng như sau:

- Tỉ số
- Tỉ lệ
- Tỉ suất

Đặc điểm chung của các chỉ số này là chúng đều là phân số. Điều này có nghĩa là mỗi chỉ số đều có tử số và mẫu số. Những khác biệt về ý nghĩa giữa các chỉ số này chính là phần mẫu số.

Tỉ số, tỉ lệ, và tỉ suất là những danh từ chung. Khi “du nhập” vào dịch tễ học, chúng ta có những thuật ngữ “đặc hiệu” hơn là tỉ suất / tỉ lệ phát sinh (incidence), tỉ lệ hiện hành (prevalence), và rủi ro (hazards). Trong phần sau đây, tôi sẽ giải thích cụ thể từng chỉ số.

1. Các chỉ số chung

1.1 Tỉ số (ratio)

Tỉ số là một phân số mà tử số có thể không có liên hệ gì với mẫu số. Tỉ số có giá trị từ 0 đến vô hạn ($0 \rightarrow \infty$).

Ví dụ 1: Chẳng hạn như trong một quần thể gồm có 193 người (95 nam và 98 nữ), chúng ta hay viết mô tả tình trạng này bằng tỉ số giới tính (sex ratio) nam:nữ như $95/98 = 0.97$. Tỉ số nam:nữ bằng 0.97 có nghĩa là cứ 100 nữ thì có 97 nam. Chỉ số này cho chúng ta biết rằng nam ít hơn nữ (vì nếu hai giới bằng nhau thì tỉ số phải bằng 1).

Các chỉ số khác như tỉ trọng khối cơ thể (body mass index hay BMI) hay mật độ xương (bone mineral density) cũng là một dạng của tỉ số. Các đo lường quan trọng trong dịch tễ học như risk ratio, odds ratio, hazards ratio, v.v... cũng là những tỉ số.

1.2. Tỉ lệ (proportion)

Tỉ lệ là một dạng tỉ số, nhưng cấu trúc của tỉ lệ khác với tỉ số ở phần tử số. Trong tỉ số, tử số có thể khác đơn vị đo lường với mẫu số (như trường hợp BMI), nhưng với tỉ lệ thì tử số và mẫu số phải cùng đơn vị đo lường. Tử số của một tỉ lệ là một phần của mẫu số. Nói cách khác, tử số của một tỉ lệ lúc nào cũng thấp hơn hay bằng mẫu số. Tỉ lệ có thể nhân cho 100 để thành số phần trăm (**percent**). Chính vì thế mà giá trị của tỉ lệ chỉ dao động từ 0 đến 1 (hay nếu diễn tả bằng phần trăm thì dao động từ 0 đến 100%).

Ví dụ 2: Trong đợt dịch tả bộc phát hồi tháng 3 vừa qua, tính đến ngày 13/4/2008 có 2490 người mắc bệnh tiêu chảy cấp tính, trong số này có 377 người nhiễm vi khuẩn tả *V. cholerae*. Do đó, tỉ lệ nhiễm khuẩn tả trong quần thể những người tiêu chảy cấp tính là $377 / 2490 = 0.151$ hay 15.1%.

Qua cách tính trên, chúng ta dễ dàng thấy rằng tỉ lệ thực chất là một xác suất. Trong ví dụ trên, chúng ta cũng có thể diễn giải rằng xác suất người bị tiêu chảy cấp tính nhiễm khuẩn tả là 0.15 (tức trong số 100 người tiêu chảy cấp tính, chúng ta kì vọng có khoảng 15 người bị nhiễm vi khuẩn tả).

1.3. Tỉ suất (rate)

Tỉ suất cũng là một dạng của tỉ số hay tỉ lệ. Trong kinh tế học, người ta hay nói đến tỉ suất lời (*interest rate*), và nó có yếu tố thời gian trong đó. Do đó, thuật ngữ **rate (tỉ suất)** có một khác biệt quan trọng với tỉ lệ: đó là mẫu số của tỉ suất có thể bao gồm yếu tố *thời gian* và số đối tượng, nhưng mẫu số của tỉ lệ thì không có thời gian. Do đó, giá trị của tỉ suất có thể đi từ 0 đến vô hạn.

2. Ứng dụng vào dịch tễ học

Như đề cập trong phần đầu, tỉ số, tỉ lệ và tỉ suất là những danh từ chung, những khái niệm của nghiên cứu quần thể. Dịch tễ học ứng dụng những khái niệm này để xây dựng thành một số chỉ số liên quan đến qui mô và yếu tố nguy cơ bệnh tật. Hai chỉ số phổ biến nhất là:

- Tỉ lệ lưu hành
- Tỉ suất / tỉ lệ phát sinh

Cần phải phân biệt hai khái niệm này: lưu hành (prevalence) và phát sinh (incidence). Lưu hành đề cập đến hiện tại, còn phát sinh đề cập đến những ca bệnh mới xảy ra trong một khoảng thời gian nhất định. Khi nói đến tỉ lệ lưu hành 10% chúng ta chỉ biết hiện nay cứ 100 người thì có 10 người mắc bệnh, nhưng không biết bệnh phát sinh từ lúc nào. Khi nói tỉ lệ phát sinh 10% trong 1 năm, chúng ta biết rằng nếu theo dõi 100 người trong vòng 1 năm thì sẽ có 10 người mắc bệnh.

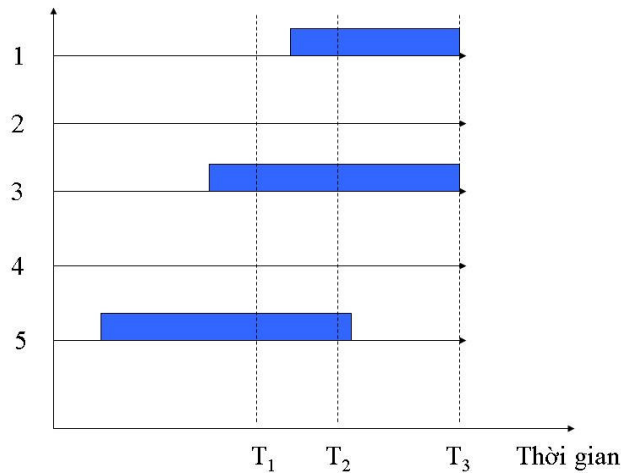
2.1 Tỉ lệ lưu hành (prevalence hay prevalence proportion)

Trong dịch tễ học, **tỉ lệ lưu hành (prevalence)**, như tên gọi, chính là một dạng của tỉ lệ. Tỉ lệ lưu hành phản ánh số ca bệnh *hiện lưu hành* trong một quần thể. Cách tính tỉ lệ này rất đơn giản: lấy số ca người mắc bệnh trong một quần thể chia cho tổng dân số trong quần thể đó. Tỉ lệ lưu hành phản ánh *qui mô* của một vấn đề y tế, nhưng không cho chúng ta biết về bệnh căn học (etiology).

Ví dụ 3: Năm 2001, trong một điều tra dịch tễ học tại Thành phố Hồ Chí Minh trên 2932 đối tượng, các nhà nghiên cứu phát hiện 111 người mắc bệnh tiểu đường. Do đó, tỉ lệ lưu hành bệnh tiểu đường là $111 / 2932 = 0.038$ hay 3.8%.

Cần phải phân biệt hai loại tỉ lệ lưu hành: **tỉ lệ lưu hành tại một thời điểm (point prevalence)** và **tỉ lệ lưu hành trong một thời gian (period prevalence)**. Để hiểu hai chỉ số này, chúng ta có thể xem qua ví dụ sau đây.

Ví dụ 4: Một quần thể gồm có 5 đối tượng. Đối tượng 1, 3, và 5 mắc bệnh; đối tượng 2 và 4 không mắc bệnh. Biểu đồ thanh (màu xanh) mô tả thời gian mắc bệnh của từng đối tượng (xem biểu đồ dưới đây).



Nếu chúng ta tiến hành điều tra vào *thời điểm* T_1 , có 2 bệnh nhân trong số 5 đối tượng, và do đó, tỉ lệ lưu hành là $2/5 = 0.4$ hay 40%. Nếu chúng ta tiến hành điều tra vào thời điểm T_2 , có 3 bệnh nhân trong số 5 đối tượng, và tỉ lệ lưu hành là $3/5 = 0.6$ hay 60%. Đây là *tỉ lệ lưu hành tại một thời điểm*.

Nhưng nếu chúng ta tính từ thời điểm T_1 đến T_3 thì trong số 5 đối tượng trong quần thể này, có tất cả 3 người mắc bệnh, và do đó tỉ lệ lưu hành trong một thời gian này là 60%.

2.2 Tỉ suất phát sinh (incidence rate)

Tỉ suất phát sinh có khi còn gọi là **incidence density (mật độ phát sinh)**. Vì là “mật độ” nên chúng ta có thể suy đoán được mẫu số là một số có yếu tố thời gian hay không gian. Thật vậy, theo định nghĩa thông thường, tỉ suất phát sinh có nghĩa là số ca mới mắc bệnh chia cho số đối tượng được theo dõi trong một thời gian. Định nghĩa này có thể hiểu qua ví dụ sau đây.

Ví dụ 5: Chúng ta theo dõi 10 bệnh nhân từ năm 1989 đến ngày 31/12/2001. Trong thời gian này chúng ta ghi nhận 3 người mắc bệnh tả, kết quả từng bệnh nhân như sau: Bệnh nhân 1 được theo dõi từ ngày 12/1/1989 đến 31/12/2001, tức 12.95 năm, và không bị bệnh tả. Bệnh nhân thứ 2 có thời gian theo dõi là 7.61 năm, tính từ 14/8/1990 đến 21/3/1998 khi bị bệnh. *Vân vân.*

Bệnh nhân	Ngày tham gia nghiên cứu	Ngày bị bệnh tả	Thời gian theo dõi, tính đến ngày khóa sổ 31/12/2001 (năm)
1	12/1/1989		12.95

2	14/08/1990	21/03/1998	7.61
3	27/03/1989		12.77
4	01/02/1989	17/03/1999	10.13
5	07/09/1989		12.32
6	21/11/1989		12.12
7	02/12/1990	01/05/1992	1.41
8	07/04/1989		12.74
9	08/01/1990		11.99
10	30/07/1989		12.43
Tổng cộng			106.49

Như đề cập trên, chúng ta phát hiện 3 ca bệnh tả, và đó là tử số của tỉ suất phát sinh. Về mẫu số, chúng ta thấy tổng cộng thời gian theo dõi của 10 bệnh nhân là 106.49 năm. Nhưng nên nhớ là chúng ta theo dõi 10 người, cho nên đơn vị phải là **năm-người** hay nói theo thuật ngữ tiếng Anh là **person-years**. Tổng thời gian theo dõi là 106.49 năm-người hay person-years.

Tỉ suất mắc bệnh tả, do đó, được tính bằng cách lấy số ca bệnh chia cho số năm-người. Cụ thể hơn: $3 / 106.49 = 0.0282$. Nói cách khác, tỉ suất mắc bệnh là 2.82% năm-người. Con số này còn có thể diễn giải một cách khác nữa, *nhưng tôi sẽ để dành cho bạn đọc suy nghĩ!*

2.3 Tỉ lệ phát sinh (incidence hay incidence proportion)

Một lợi điểm nhưng cũng là khiếm khuyết của tỉ suất phát sinh là mẫu số do hỗn hợp của hai đơn vị số đối tượng và thời gian nên đôi khi cũng khó hiểu. Chẳng hạn như khi nói tỉ suất phát sinh 2.8 trên 100 người-năm, chúng ta không biết đó là số lần mắc bệnh của một cá nhân được theo dõi 100 năm, hay 2 người được theo dõi mỗi người 50 năm. Do đó, dịch tễ học đề ra một chỉ số khác tạm dịch là *tỉ lệ phát sinh*.

Mẫu số của tỉ lệ phát sinh không phải là người-năm mà chỉ là số đối tượng theo dõi. Vì là tỉ lệ, cho nên mẫu số của tỉ lệ phát sinh là một phần của mẫu số. Do đó, tỉ lệ phát sinh được định nghĩa là số ca mắc bệnh trong một khoảng thời gian chia cho số đối tượng được theo dõi trong khoảng thời gian đó.

Chú ý rằng tỉ lệ phát sinh còn đôi khi được đề cập đến như là **tỉ lệ tấn công (attack rate)**. Trong y văn, hai thuật ngữ này (*tỉ suất phát sinh* và *tỉ lệ phát sinh*) thỉnh thoảng vẫn được sử dụng qua lại mà không phân biệt mẫu số. Tỉ suất phát sinh hay tỉ lệ phát sinh có thể ước tính cho hai nhóm và so sánh cho ra tỉ số nguy cơ (relative risk). Do đó, tỉ suất có giá trị khoa học là cung cấp cho chúng ta một vài thông tin về bệnh căn học.

Trong ví dụ trên, chúng ta có 10 bệnh nhân, và qua theo dõi trung bình 10.65 năm (lấy 106.49 chia cho 10), có 3 ca mắc bệnh tử. Do đó, tỉ lệ phát sinh trong vòng 10.65 năm là: $3 / 10 = 0.30$ hay 30%.

2.4. Mối liên hệ giữa tỉ suất phát sinh và tỉ lệ lưu hành

Để thấy mối liên hệ giữa tỉ suất phát sinh và tỉ lệ lưu hành, chúng ta cần đến vài kí hiệu. Gọi N là số đối tượng trong một quần thể được theo dõi trong thời gian từ năm t_0 đến năm t_1 , và trong thời gian đó chúng ta phát hiện có X người phát sinh bệnh. Giả dụ rằng tổng số người-năm toàn bộ quần thể là T .

Gọi P là tỉ lệ lưu hành, và theo định nghĩa trong thời gian t_0 đến năm t_1 , P có thể ước tính qua công thức:

$$P = X / N$$

Gọi I là tỉ suất phát sinh, chúng ta có thể ước tính I như sau:

$$I = X / T$$

Vì tổng số người-năm là T , và chúng ta có N đối tượng, cho nên *thời gian theo dõi trung bình* (kí hiệu D) của quần thể là:

$$D = T / N$$

Qua vài thao tác đại số có thể thấy ngay rằng:

$$P = I \times T / N = I \times D$$

Điều này có nghĩa là tỉ lệ lưu hành là tích số của tỉ suất phát sinh và thời gian theo dõi trung bình.

Ví dụ 5 (tiếp tục): Trong ví dụ trên, chúng ta theo dõi 10 đối tượng với tổng số năm-người theo dõi là 106.49, do đó, thời gian theo dõi trung bình cho mỗi đối tượng là $D = 106.49 / 10 = 10.65$ năm. Chúng ta cũng biết rằng tỉ suất phát sinh là $I = 3 / 106.49 = 0.0282$. Do đó, tỉ lệ lưu hành trong thời gian đó là: $P = 0.0282 \times 10.65 = 0.30$ hay là 30%. Thật vậy, trong số 10 người theo dõi trong thời gian đề cập, có 3 người mắc bệnh, và do đó tỉ lệ lưu hành là 30%. Cần nói thêm rằng tỉ lệ lưu hành một thời gian này còn được gọi là tỉ lệ phát sinh tích lũy (**cumulative incidence**).

Bảng sau đây sẽ tổng hợp các khác biệt, đặc tính và ứng dụng của tỉ số, tỉ lệ, và tỉ suất:

Đặc tính	Tỉ số (ratio)	Tỉ lệ (proportion)	Tỉ suất (rate)
----------	------------------	-----------------------	-------------------

Toán học	Phân số	Phân số	Phân số
Tử số và mẫu số	Không hẳn phải có cùng đơn vị	Phải có cùng đơn vị	Không hẳn phải có cùng đơn vị
Tử số (trong văn cảnh dịch tễ học)	Có thể bất cứ số nào	Số ca bệnh	Số ca bệnh
Mẫu số (trong văn cảnh dịch tễ học)	Có thể bất cứ số nào	Tổng số đối tượng trong quần thể	Tổng số người-thời gian (“thời gian” có thể là năm, tháng, v.v...)
Giới hạn của giá trị	Từ 0 đến vô hạn	Từ 0 đến 1 (hay 100)	Từ 0 đến vô hạn
Chỉ số dịch tễ học	Tỉ số giới tính, BMI, mật độ xương, mật độ dân số, v.v...	Tỉ lệ lưu hành (prevalence), kể cả tỉ lệ lưu hành tích lũy hay một thời điểm	Tỉ lệ phát sinh (incidence), tỉ lệ tấn công (attack rate), tỉ suất phát sinh (incidence rate), v.v...
Ý nghĩa	Đơn thuần mô tả	Qui mô của bệnh	Mối liên hệ giữa yếu tố nguy cơ và bệnh (bệnh căn học)

Ngoài các chỉ số trên, còn một số chỉ số khác cũng phản ánh qui mô bệnh tật như lifetime risk (nguy cơ mắc bệnh trọn đời), hazards rate (tỉ suất nguy cơ), v.v... cũng có khi sử dụng trong các nghiên cứu dịch tễ học thuộc dạng phân tích. Ngoài ra, các chỉ số mang tính đo lường khác [tôi chưa biết dịch sang tiếng Việt là gì] như attributable risk fraction, etiologic fraction, potential years of life lost, v.v... cũng có khi sử dụng cho các công trình dịch tễ học mang tính phân tích. Tuy nhiên, cách tính các chỉ số này cũng phức tạp, và đôi khi đòi hỏi phải có phần mềm máy tính mới thực hiện được.

Cổ nhiên, trên đây chỉ là các chỉ số cực kì đơn giản trong dịch tễ học. Các chỉ số này chỉ là số *trung bình*, chưa phản ánh độ dao động mẫu (sampling variation). Cách tính độ dao động mẫu có phần tương đối phức tạp hơn nhưng đó là một đề tài cho câu hỏi sau. :-)

Có thể nói rằng bộ môn dịch tễ học được hình thành từ ba ngành học chính: y học, dân số học, và thống kê học. Y học cung cấp cho chúng ta những khái niệm bệnh lí, bệnh cảnh, đặc tính lâm sàng từ cá nhân để từ đó suy luận cho một quần thể. Dân số học cung cấp cho chúng ta những khái niệm quần thể như tỉ lệ, tỉ suất, và tỉ số. Thống kê học cung cấp cho chúng ta những khái niệm xác suất, mô hình toán học, và mô hình suy luận khoa học (như kiểm định giả thuyết và kiểm định ý nghĩa thống kê). Do đó, dịch tễ học thực chất là một liên ngành, và đòi hỏi người thực hành phải nắm vững cả ba lĩnh vực học thuật trên. Dịch tễ học không thể là “khoa học” nếu không có đo lường. Để kết thúc bài trả lời này, tôi mượn lời của ông William Thomson (còn gọi là Lord Kelvin hay Bá tước Kelvin), người phát triển nhiệt kế Kelvin (độ K) như sau: “*Tôi thường nói rằng khi bạn*

có thể đo lường một sự kiện và diễn giải sự kiện đó bằng con số, bạn biết đôi điều về sự kiện. Nếu bạn không thể diễn tả một sự kiện bằng con số thì kiến thức của bạn sẽ nghèo nàn đi. Không đo lường được, bạn khó có thể làm cho khoa học tiến triển”. Trong khoa học “cân đo đong đếm” (như ông bà ta hay nói) rất quan trọng. Có thể nói không ngoa rằng bất cứ ngành nào mà không có “cân đo đong đếm” thì chưa hẳn là ngành khoa học!

Chú thích:

Bạn đọc có thể tìm hiểu thêm qua Từ điển Dịch tễ học trực tuyến ở đây: <http://www.swintons.net/jonathan/Academic/glossary.html>. Cuốn này thật ra không hẳn tốt. Cuốn sách tham khảo chính là “A Dictionary of Epidemiology” của John M. Last. Cuốn này rất mỏng, nhưng ... khá đắt tiền!

Sau đây là định nghĩa và chỉ dẫn của Bộ Y tế (Dự án hỗ trợ hệ thống y tế dự phòng) mà bạn đọc cho biết (chỉ trích dẫn phần liên quan) để các bạn tham khảo.

II. Phân tích số liệu giám sát

Phân tích số liệu kịp thời là việc làm quan trọng để có thể biến thông tin thành hành động.

1. Các chỉ số:

1.1. Khái niệm: Chỉ số là số đo giúp cho việc lượng hóa và so sánh những sự thay đổi. Sự thay đổi có thể được thể hiện theo chiều hướng (tăng hay giảm), mức độ (ít hay nhiều) và phạm vi (rộng hay hẹp).

1.2. Các dạng thức của chỉ số:

1.2.1. **Tần số:** Biểu thị số lần xuất hiện của một quan sát nào đó. Ví dụ: số người mang vi rút cúm gây bệnh H5N1 khi kiểm tra bệnh phẩm.

1.2.2. **Tần số cộng dồn:** thường được sử dụng khi trình bày trong bảng. Tần số cộng dồn của một ô nào đó bằng tần số của chính ô đó cộng với tần số của các ô trước.

1.2.3. **Tần số tuyệt đối:** chính là tần số thực của một quan sát và không phụ thuộc vào cỡ mẫu lớn hay nhỏ. Tần số tuyệt đối được biểu thị bằng giá trị tuyệt đối của sự kiện.

1.2.4. **Tần số tương đối hay còn gọi là tần suất:** biểu thị của tần số trong một mối tương quan với cỡ mẫu. Tần số tương đối được biểu thị bằng giá trị tương đối của sự kiện, bao gồm tỷ số, tỷ lệ và tỷ suất.

1.3. Tỷ số, tỷ lệ và tỷ suất:

Tỷ số, tỷ lệ và tỷ suất là những công cụ quan trọng dùng để đo tần số của bệnh. Các thước đo tần số này được dùng cho các biến phân loại (ví dụ như sống hoặc chết, nam hoặc nữ). Ba thước đo có cùng công thức:

$$X/Y \times 10^n$$

trong đó x và y là những số lượng được so sánh và 10^n là hằng số.

1.3.1. **Tỷ số:** là biểu thị mối tương quan của hai đại lượng. Hai đại lượng này có thể liên quan với nhau hoặc hoàn toàn độc lập với nhau. Tỷ lệ và tỷ suất là những trường hợp riêng của tỷ số.

Ví dụ:

$$\text{Mật độ dân số} = \frac{\text{Số dân trung bình của một khu vực trong khoảng thời gian xác định (người)}}{\text{Diện tích của khu vực đó (km}^2\text{)}}$$

1.3.2. **Tỷ lệ:** là một phân số trong đó tử số là một phần của mẫu số và có cùng đơn vị đo

như nhau, [x/y, trong đó x là một phần của y]

Ví dụ:

$$\text{Tỷ lệ nam} = \frac{\text{Số nam của một khu vực trong khoảng thời gian xác định}}{\text{Tổng dân số (nam+nữ) của khu vực đó trong cùng thời gian}}$$

- Tỷ lệ phần trăm: Có cùng công thức như tỷ lệ, nhưng được nhân với 100 (x/y X 100).

1.3.3. *Tỷ suất*: là phân số dùng để đo lường xác suất xảy ra của một sự kiện đặc biệt, trong đó tử số là các sự kiện (sinh, chết, tai biến, bệnh tật,...) và mẫu số là số lượng cá thể có trong các sự kiện đó (dân số chung, số trẻ em dưới 5 tuổi, số phụ nữ trong độ tuổi sinh đẻ,...) trong một khoảng thời gian nhất định. Tỷ suất thường để xác định mức độ biến động của các hiện tượng trong một khoảng thời gian xác định.

Tỷ suất trong dịch tễ học thường chứa 3 phần: tần suất mắc bệnh, mẫu của quần thể và thời gian xảy ra sự kiện.

Không giống như những con số, các tỷ suất có thể giúp hình dung ra nguy cơ mắc bệnh của con người trong quần thể. Tỷ suất thường được dùng để so sánh sự xuất hiện bệnh giữa các vùng, nhóm tuổi, giới tính hoặc nhóm dân tộc khác nhau.

$$\text{Tỷ suất} = \frac{\text{Số sự kiện xảy ra trong khoảng thời gian xác định thuộc một khu vực}}{\text{Số lượng trung bình cá thể có khả năng sinh "sự kiện" đó trong khu vực/thời gian}} \times k \quad (k=10^n)$$

Ví dụ: *Tỷ suất tử vong ca bệnh (Case Fatality Rate - CFR)* còn gọi là tỷ lệ chết trên mắc là tỷ lệ thường dùng. Cách tính tỷ suất tử vong ca bệnh như sau:

- Tính tổng số trường hợp tử vong.
- Chia tổng số trường hợp tử vong cho tổng số trường hợp bệnh ghi nhận được (cả tử vong và không tử vong).
- Nhân kết quả với 100.

Bảng Tỷ suất tử vong ca bệnh của bệnh sởi theo nhóm tuổi năm 2006

Nhóm tuổi	Số mắc	Số chết	Tỷ suất tử vong ca bệnh
0-4 tuổi	40	4	10%
5-14 tuổi	9	1	11%
Từ 15 tuổi trở lên	1	0	0%
Không rõ tuổi	28	0	0%
Tổng cộng	78	5	6%

- Tỷ suất tử vong ca bệnh giúp:
 - o Chỉ ra mức độ trầm trọng của bệnh trạng.
 - o Xác định những mầm bệnh mới có độc tính cao hơn hoặc kháng thuốc.
 - o Chỉ ra chất lượng chăm sóc y tế kém hoặc tình trạng chưa được chăm sóc
 - o Chỉ ra những vấn đề trong quản lý ca bệnh. So sánh chất lượng quản lý ca bệnh giữa những vùng khác nhau.
- Các chương trình y tế dự phòng có thể tác động đến tỷ suất tử vong ca bệnh bằng cách đảm bảo rằng ca bệnh được phát hiện ngay và được quản lý với chất lượng tốt. Một số khuyến nghị về công tác phòng chống những bệnh cụ thể coi việc làm giảm tỷ suất tử vong ca bệnh là tiêu chí đánh giá tính hiệu quả của việc xử lý dịch (ví dụ như CFR thấp)

hơn 1% trong vụ dịch tả).

Các tài liệu trong y học sử dụng các thuật ngữ tỷ số, tỷ lệ và tỷ suất không nhất quán như là trong định nghĩa của toán học. Điều quan trọng là phải hiểu cách đo lường, tính toán và ý nghĩa của các thuật ngữ này.

1.5. Đo lường mắc bệnh:

1.5.1. Tỷ suất hiện mắc: Đo lường tần số hiện mắc bệnh của một quần thể ở một điểm nhất định của thời gian (tỷ lệ hiện mắc tại một thời điểm) hay ở một khoảng thời gian nhất định (tỷ lệ hiện mắc trong một khoảng thời gian). Tỷ suất hiện mắc thường ký hiệu là P (prevalence)

$$P = \frac{\text{Tổng số bệnh nhân hiện mắc của một bệnh tại một thời điểm nhất định hay trong một khoảng thời gian nhất định}}{\text{Tổng số người ở quần thể có nguy cơ}} \times k \quad (k=10^n)$$

- ý nghĩa của tỷ suất hiện mắc:

- Tỷ suất hiện mắc được xác định trong các cuộc điều tra ngang giám sát tình hình sức khỏe của cộng đồng, đánh giá một hiện tượng mắc hàng loạt. Nếu có thể đối chiếu đỉnh cao nhất của nguy cơ (xảy ra trong quá khứ) với đỉnh cao của tỷ lệ mới mắc thì có thể ước lượng khoảng thời kỳ ủ bệnh hoặc thời kỳ tiềm tàng của bệnh, thời gian tiếp xúc, và cùng với các thông tin khác về dịch tễ, có thể cho ta đánh giá một cách logic quá trình mắc bệnh hàng loạt đó và áp dụng những phương pháp hợp lý và hữu hiệu trong giám sát bệnh hàng loạt.
- Tỷ suất hiện mắc được sử dụng để đánh giá tình trạng sức khỏe của một quần thể, là cơ sở để lập kế hoạch đáp ứng nhu cầu chăm sóc sức khỏe của quần thể (như số cán bộ y tế, số giường bệnh, cơ sở thuốc điều trị).

1.5.2. Tỷ suất mới mắc: Đo lường tần số mới mắc bệnh của một quần thể trong một khoảng thời gian nào đó. Tỷ suất mới mắc thường có ký hiệu là I (incidence)

Tỷ suất mới mắc tích lũy (CI):

$$CI = \frac{\text{Tổng số trường hợp mới mắc bệnh của một quần thể trong một khoảng thời gian xác định}}{\text{Tổng số người có nguy cơ mắc bệnh của quần thể đó trong thời gian đó}}$$

Tỷ suất mật độ mới mắc (ID): là tỷ số của số trường hợp mới mắc của một bệnh trong một thời gian nhất định, được chia cho thời gian-người có nguy cơ.

$$ID = \frac{\text{Tổng số trường hợp mới mắc bệnh của một quần thể trong một khoảng thời gian xác định}}{\text{Tổng thời gian có nguy cơ theo dõi được của tất cả các cá thể thuộc quần thể đó}}$$

- Tỷ suất mật độ mới mắc thường là một phép đo chính xác hơn tỷ suất mới mắc tích lũy, bởi vì nó quan tâm đến thời gian nguy cơ của từng cá thể.

- *Tỷ lệ tấn công:* Là một dạng đặc biệt của tỷ suất mới mắc để thấy được phạm vi và mức độ dân số bị ảnh hưởng, thường được sử dụng trong điều tra, nghiên cứu các vụ dịch. Tỷ lệ tấn công thường được ký hiệu là AR (attack rate).

Công thức:

$$AR = \frac{\text{Số mắc mới của quần thể trong một khoảng thời gian}}{\text{Dân số nguy cơ vào lúc đầu khoảng thời gian}} \times 100 \quad (1000)$$

Ví dụ:

Bảng minh họa tỷ lệ tấn công của bệnh sốt xuất huyết theo vùng:

Quận/huyện	Dân số	Số mắc	Tỷ lệ tấn công/1
------------	--------	--------	------------------

Thủy Nguyên	271.728	1.370	5	
Ngô Quyền	196.188	1.400	7	
An Hải	193.997	870	4	
Vĩnh Bảo	191.719	450	2	
An Lão	182.481	400	2	
Khác	133.000	0	0	
Tổng cộng	1.169.113	4.490	4	

- ý nghĩa của tỷ suất mới mắc:

- Tỷ suất mới mắc được xác định trong các nghiên cứu thuần tập tương lai, trong đó một nhóm cá thể có nguy cơ nhưng chưa mắc bệnh được theo dõi trong một thời gian để xác định nguy cơ phát triển bệnh
- Tỷ suất mới mắc có ích cho việc đánh giá nguy cơ phát triển bệnh theo thời gian do tiếp xúc với yếu tố nguy cơ đó.
- Tỷ suất mới mắc cho phép đánh giá hiệu quả của các biện pháp can thiệp đã được áp dụng.

1.5.3. Liên quan giữa tỷ suất hiện mắc P và tỷ suất mới mắc I:

- Tỷ suất hiện mắc P phụ thuộc chủ yếu vào tỷ suất mới mắc I và thời gian mắc bệnh trung bình:

$$P = I \times D$$

(Trong đó D là thời gian mắc bệnh trung bình, còn gọi là bệnh kỳ, đo bằng ngày, tuần, tháng hoặc năm).

- Sự liên quan này giúp chúng ta thấy một điều quan trọng: nếu muốn giảm tỷ lệ hiện mắc thì có thể thực hiện biện pháp:

- Giảm số mới mắc (chống dịch hữu hiệu, như bảo vệ khỏi cảm nhiễm, cắt đứt đường truyền nhiễm, không để xuất hiện trường hợp mắc bệnh mới, có biện pháp phòng bệnh đặc hiệu) và hoặc
- Giảm thời gian mắc bệnh trung bình (có biện pháp điều trị tốt, rút ngắn thời gian điều trị, tăng cường sức khỏe nhân dân).