

# Ngộ nhận về tính đại diện trong nghiên cứu khoa học

Hầu hết những ai làm nghiên cứu lâm sàng ở Việt Nam đều từng trải qua các hội đồng xét duyệt đề tài hay nghiệm thu, và tuyệt đại đa số đều nhận được lời phê bình là nghiên cứu không mang tính đại diện. Những người thốt ra phê bình này thường làm về y tế công cộng, chứ không phải nghiên cứu y học và nghiên cứu lâm sàng. Lời phê bình này hàm ý nói kết quả nghiên cứu không có giá trị khoa học, và có khi nó được dùng như là một cách bác bỏ nỗ lực của nhà nghiên cứu. Nhưng khi được hỏi "tính đại diện" có nghĩa là gì, thì mỗi người giải thích một cách khác nhau, càng làm cho nhà nghiên cứu thiếu kinh nghiệm thêm lúng túng. Trong note này tôi sẽ giải thích tại sao một nghiên cứu y học nghiêm chỉnh không nhất thiết phải có mẫu mang tính đại diện.

## Những cách hiểu về khái niệm "đại diện"

Một trong những nguyên lý cổ điển về nghiên cứu y học được giảng dạy ở Việt Nam là mẫu nghiên cứu phải mang tính đại diện. Nói theo tiếng Anh là *representativeness*. Theo nguyên lý này, mẫu mang tính đại diện thì kết quả nghiên cứu mới có giá trị khoa học, dù nhiều người không định nghĩa rõ "giá trị khoa học" có nghĩa gì trong thực tế. Nhưng từ suy nghĩ đó, đối với những người đã qua đào tạo hay chịu ảnh hưởng bởi nguyên lý cổ điển đó đòi hỏi nghiên cứu phải mang tính "đại diện" cho quần thể mà mẫu được lấy từ đó. Nhưng khi được hỏi khái niệm "đại diện" có nghĩa là gì, thì hình như mỗi người hiểu một cách khác nhau. Qua thăm dò ý kiến và đọc những phê bình trong các đề cương nghiên cứu khoa học, tôi thấy có 4 cách hiểu về tính đại diện.

### 1. Đại diện mang ý nghĩa thống kê

Một nghiên cứu y học thường lấy mẫu (sample) từ một quần thể (population). Từ một quần thể, chúng ta có thể lấy rất nhiều mẫu khác nhau, và dĩ nhiên mỗi lần lấy mẫu thì mỗi lần chúng ta có kết quả hơi khác. Nếu quần thể gồm 100 người có chiều cao 160 cm và độ lệch chuẩn 6.5 cm, thì nếu chúng ta lấy mẫu ngẫu nhiên 10 người từ quần thể đó thì chiều cao trung bình của 10 người đó có thể dao động trong khoảng 157 đến 162 cm. Ở đây, tính đại diện có nghĩa là phân bố chiều cao của mẫu nghiên cứu giống như phân bố chiều cao của quần thể.

### 2. Đại về một đặc điểm lâm sàng

Đối với nhiều người, cách hiểu về tính đại diện có vẻ đơn giản hơn, và đó là kì vọng rằng mẫu nghiên cứu có những đặc điểm giống như đặc điểm của quần thể. Theo cách hiểu này, nếu quần thể có 51% dân số là phụ nữ, 30% nam giới hút thuốc lá, 15% người có thu nhập cao, 1% người với tuổi 20, v.v. thì mẫu nghiên cứu cũng phải có những đặc điểm tương tự.

### **3. Đại diện về kết quả**

Một cách hiểu khác là đại diện tính có nghĩa là kết quả, ví dụ như tỉ lệ mắc bệnh, trong mẫu nghiên cứu gần bằng với tỉ lệ của quần thể. Nói cách khác, nếu trong quần thể có 10% người mắc bệnh tiểu đường, thì mẫu nghiên cứu cũng có khoảng 10% người mắc bệnh tiểu đường. Nếu mẫu nghiên cứu có 20% người mắc bệnh thì được xem là không mang tính đại diện.

### **4. Đại diện về phân bố địa lí**

Một cách hiểu khác mà nhiều người khi đề cập đến tính đại diện hàm ý nói đến tính địa lí và đặc điểm của đối tượng nghiên cứu. Rất nhiều người nghĩ rằng một nghiên cứu y học, đặc biệt là dịch tễ học, phải mang tính đại diện địa lí, có nghĩa là chọn ngẫu nhiên trong cộng đồng. Họ nghĩ rằng để đảm bảo tính đại diện của mẫu nghiên cứu, phải chọn mỗi tỉnh thành, chọn ngẫu nhiên vài quận/huyện, mỗi quận/huyện chọn ngẫu nhiên vài phường/xã, mỗi phường/xã chọn ngẫu nhiên vài ấp, v.v. Nói cách khác, nếu Quận I có dân số chiếm 10% dân số của Thành phố, thì mẫu nghiên cứu cũng nên có khoảng 10% cá nhân đến từ Quận I.

Nhưng rất tiếc là 4 cách hiểu về đại diện là 4 ngộ nhận về nghiên cứu khoa học. Cách hiểu thứ nhất là một sự hiểu lầm giữa đại diện mang tính thống kê và đại diện mang tính khoa học, và tôi sẽ bàn thêm dưới đây. Một cách ngắn gọn trong suy luận khoa học, chúng ta không nhất thiết phải cần đến đại diện thống kê.

Cách hiểu thứ hai và thứ ba là rất phản khoa học, bởi vì để kiểm định một giả thuyết khoa học nhà khoa học không cần chọn đối tượng mang tính đại diện, mà có thể chọn một nhóm đối tượng với những đặc điểm đặc biệt. Chẳng hạn như để kiểm định giả thuyết về ảnh hưởng của thiếu vận động đến loãng xương, chúng ta có thể chọn bệnh nhân đột quỵ và những phi hành gia.

Cách hiểu thứ tư là một hiểu lầm về điều tra dân số, lấy ý kiến quần chúng và nghiên cứu khoa học. Khi lấy ý kiến trong công chúng, người ta có thể [chỉ "có thể" thôi] cần đến mẫu đại diện, nhưng nghiên cứu khoa học không phải là điều tra xã hội. Ngay cả điều tra xã hội, kết quả là một tỉ lệ (ví dụ như) ủng hộ ứng viên nào đó cũng có thể là ... vô nghĩa. Cần phải nói thêm rằng nếu phải lấy mẫu đại diện địa lí thì có thể nói rằng 100% các nghiên cứu y học đều vô giá trị! Nếu phải đáp ứng đại diện địa

lí thì tất cả những nghiên cứu nổi tiếng như Framingham, Rotterdam (Hà Lan), CaMOS (Canada), Dubbo Osteoporosis Epidemiology Study (của tôi), tất cả các công trình nghiên cứu lâm sàng ngẫu nhiên (RCT), v.v. đều không có giá khoa học?! Do đó, những ngộ nhận này tuy đáng tiếc, nhưng lại rất nguy hiểm, vì làm chùn bước rất nhiều người muốn làm nghiên cứu khoa học có ý nghĩa. Những hiểu lầm này có thể xuất phát từ sự ngộ nhận về suy luận thống kê và suy luận khoa học mà tôi giải thích dưới đây.

## Suy luận thống kê và suy luận khoa học

Để hiểu tại sao khái niệm đại diện không cần thiết trong nghiên cứu khoa học, chúng ta cần phải biết mục tiêu thật sự của một nghiên cứu khoa học. Ở đây, tôi không bàn đến những cuộc điều tra xã hội để biết ý kiến của công chúng về một vấn đề nào đó (ví dụ như bầu cử) hay để biết tỉ lệ mắc bệnh trong quần thể, vì đó không phải là nghiên cứu khoa học. Thường, các nghiên cứu khoa học có mục tiêu chính là kiểm định một giả thuyết khoa học. Giả thuyết khoa học thường mang tính tiên lượng, giống như qui luật, trong thế giới tự nhiên. Do đó, mục tiêu chính của khoa học là đi tìm một qui luật của thế giới tự nhiên (law of nature).

Qui luật này có thể thể hiện qua hiện tượng chúng ta gọi là *liên quan* (1). Ý tưởng "liên quan" ở đây có thể là mối liên hệ nhân quả (causal relationship) hay mối tương quan thống kê (correlation). Dĩ nhiên, tương quan thống kê chỉ là điều kiện cần, chứ chưa đủ, để nói lên mối liên hệ nhân quả. Chẳng hạn như để kiểm định giả thuyết thiếu vận động thể lực dẫn đến tăng nguy cơ loãng xương, nhà nghiên cứu có thể đặt mục tiêu nghiên cứu là tìm mối liên quan giữa vận động thể lực và loãng xương. Vận động thể lực có thể được đo bằng chỉ số tiêu hao năng lượng (biến liên tục), nhưng cũng có thể so sánh giữa nhóm bất động và nhóm bình thường. Đó là nghiên cứu tìm mối liên quan.

Khía cạnh quan trọng nhất trong nghiên cứu khoa học là kết quả (qui luật quan sát trong nghiên cứu) phù hợp với thực tế của **qui luật tự nhiên**. Chẳng hạn như trong nghiên cứu về mối liên quan giữa vận động thể lực và loãng xương, chúng ta có thể thiết kế nghiên cứu bằng cách so sánh mật độ xương của các phi hành gia và người bình thường, và nếu kết quả cho thấy phi hành gia có mật độ xương thấp hơn người bình thường, thì đó là một qui luật, một mối liên quan. Mối liên quan đó đúng hay sai không phải tùy thuộc vào tính đại diện của đối tượng nghiên cứu, mà phụ thuộc vào kiến thức sinh học về xương và những nghiên cứu trước đây trong y văn.

Từ những cân nhắc trên, chúng ta có thể phân biệt khái niệm suy luận thống kê (**statistical inference**) và suy luận khoa học (**scientific inference**). Suy luận thống kê dựa trên kết quả phân tích trong một mẫu nghiên cứu để phát biểu về quần thể. Những kết quả này bao gồm các chỉ số kiểm định thống kê (như t-test, Ki bình

phương, mô hình hồi qui tuyến tính) kèm theo trị số P và khoảng tin cậy 95%. Suy luận khoa học liên quan đến khái quát hoá mối liên quan quan sát từ nghiên cứu đến thế giới tự nhiên. Suy luận khoa học dựa trên kết quả suy luận thống kê, nhưng đó chỉ là điều kiện cần, chứ chưa đủ, vì ngoài thống kê ra, chúng ta cần có kiến thức chuyên ngành trước khi phát biểu một suy luận khoa học.

### **Hợp lí nội tại và hợp lí ngoại tại**

Khi nói một nghiên cứu khoa học có giá trị, người ta nghĩ đến hai tính hợp lí: hợp lí nội tại (internal validity) và hợp lí ngoại tại (external validity). Hiểu một cách đơn giản, hợp lí ngoại tại là đảm bảo "chất lượng" (quality assurance) của nghiên cứu. Trong thực tế, **hợp lí nội tại** có nghĩa là chọn phương pháp đo lường chính xác và có độ tin cậy cao, là sử dụng phương pháp phân tích tốt và thích hợp, là kiểm soát và giảm thiểu tối đa các yếu tố nhiễu, để kết quả nghiên cứu có tính hợp lí. Chẳng hạn như nghiên cứu về loãng xương mà dùng phương pháp siêu âm thì có nghĩa là tính hợp lí nội tại chưa cao.

**Hợp lí ngoại tại** (còn gọi là khái quát hoá) là mức độ mà kết quả của nghiên cứu có thể áp dụng, hoặc liên quan, hoặc khái quát hoá cho một quần thể lớn hơn hoặc cho những người không tham gia vào công trình nghiên cứu. Khái quát hoá theo cách hiểu này không phải chỉ đơn thuần là dựa vào thống kê (mẫu đại diện) mà *quan trọng hơn* là còn phải dựa vào kiến thức chuyên ngành và các nghiên cứu trong quá khứ.

*Trong suy luận khoa học, tính hợp lí nội tại đóng vai trò quan trọng nhất, còn hợp lí ngoại tại chỉ là phụ.* Nếu một nghiên cứu mà phương pháp đo lường không chuẩn mực thì cũng chẳng có giá trị gì về hợp lí ngoại tại. Chẳng hạn như nghiên cứu về béo phì mà chỉ đơn thuần dựa vào các chỉ số như chiều cao và cân nặng (hay body mass index -- BMI) thì giá trị khoa học và tính hợp lí nội tại không cao (vì phương pháp chuẩn là đo lường tỉ trọng mỡ trong cơ thể). Do đó, xin nhắc lại: nếu một nghiên cứu thiếu tính hợp lí nội tại thì kết quả cũng không có ý nghĩa hợp lí ngoại tại. Đảm bảo tính hợp lí nội tại là rất quan trọng trong bất cứ công trình nghiên cứu y học nào.

### **Tại sao nghiên cứu y học không cần mang tính đại diện?**

Nghiên cứu khoa học không cần mẫu nghiên cứu phải mang tính đại diện, bởi vì mục tiêu chính là suy luận khoa học. Để đánh giá sự ảnh hưởng của yếu tố di truyền đến một bệnh lí, chúng ta có thể nghiên cứu những thành viên trong gia đình hay những cặp sinh đôi, và cách chọn hoàn toàn không cần phải mang tính đại diện địa lí hay bất cứ chỉ số lâm sàng nào. Để đánh giá hiệu quả của một loại thuốc điều trị bệnh loãng xương, chúng ta có thể thực hiện một nghiên cứu lâm sàng đối chứng

ngẫu nhiên (RCT), và bệnh nhân được tuyển chọn theo hàng chục tiêu chuẩn thu nhận và tiêu chuẩn loại trừ. Nghiên cứu RCT không cần đến tính đại diện của mẫu nghiên cứu. Ngay cả công trình vĩ đại UKBiobank với mục tiêu tư vấn cho chính sách y tế ở Anh, với gần nửa triệu người tham gia, hoàn toàn không mang tính đại diện. Những nghiên cứu như thế đòi hỏi tính hợp lý nội tại rất cao.

Cần nói thêm rằng đối với các nghiên cứu khoa học thì các phát hiện phải mang tính tái lập (repeatability). Tính tái lập cực kì quan trọng, vì nó phân biệt giữa khoa học thật và khoa học dỏm. Tái lập không chỉ quần thể này mà còn ở quần thể khác, và thường không tùy thuộc vào thời gian. Khi chúng ta phát hiện người hút thuốc lá có nguy cơ tử vong cao, thì mối liên quan này gần như được lặp lại ở bất cứ quần thể lớn nào và ở thời điểm nào cũng thế. Đó chính là tính đại diện. Và, tính đại diện này không phải do chọn đối tượng nghiên cứu mang tính đại diện theo nghĩa xác suất, càng không có gì liên quan đến đại diện địa lí.

Tính đại diện khoa học hay tái lập, một cách lí tưởng, không tùy thuộc vào thời gian và không gian. Trong y học, chúng ta đã thấy mối liên quan giữa cholesterol và bệnh tim mạch xuất phát từ một nghiên cứu không mang tính đại diện địa lí bên Mỹ, nhưng kết quả đó áp dụng cho toàn thế giới. Một nghiên cứu về ảnh hưởng của di truyền đến xương ở Việt Nam cũng được lặp lại ở tất cả các nhóm khác trên thế giới. Nếu giả dụ như mẫu nghiên cứu "đại diện" theo cách hiểu đó thì mối liên quan mà chúng ta quan sát được có lẽ không áp dụng cho bất cứ nhóm nào, nam hay nữ, người giàu hay người nghèo. Cái mối liên quan giữa  $x$  và  $y$ , ví dụ như thể hiện qua tỉ số nguy cơ, đã được trọng số hoá và trung bình hoá từ nhiều nhóm nhỏ, nên nó không thể khái quát hoá cho nhóm nào cả. Sự thật này nói lên một lần nữa, trong nghiên cứu kiểm định giả thuyết khoa học, chúng ta không cần chọn mẫu đại diện.

Nhưng có trường hợp tính đại diện có thể cần thiết. Trung cầu dân ý, điều tra xã hội để tìm hiểu qui mô của một bệnh (như tiểu đường) hay thói quen (như hút thuốc lá, uống rượu bia) rất khó có thể xem là nghiên cứu, vì không có giả thuyết đằng sau. Ở đây, chúng ta không nói những điều tra như thế là không quan trọng; thật ra, kết quả của các điều tra đó rất quan trọng cho việc hoạch định chính sách. Những dạng điều tra xã hội này có thể đòi hỏi phải có mẫu đại diện. Xin nhắc lại: "có thể" mà thôi, chứ không phải đúng.

Tuy nhiên, nếu mẫu điều tra mang tính đại diện theo nghĩa thứ hai (tức các đặc điểm của mẫu nghiên cứu giống như đặc điểm của quần thể) thì kết quả của một nghiên cứu như thế sẽ không có tính hợp lý ngoại tại. Chẳng hạn như nếu một mẫu điều tra đại diện về cấu trúc độ tuổi trong dân số, và phát hiện tỉ lệ mắc bệnh là 10%, thì con số này không đại diện cho cái gì cả. Tại sao? Tại vì con số 10% là số trung bình của nhiều độ tuổi, và tỉ lệ này khác biệt giữa các độ tuổi. Ở đây, chúng ta thấy cái gọi là tính "đại diện" là rất vô nghĩa!

Một ảo tưởng (và cũng là sai lầm) khác là cho rằng nếu mẫu nghiên cứu đại diện cho quần thể về vài đặc điểm nào đó thì kết quả hoàn toàn đại diện cho quần thể. Những kết quả điều tra như thế (tỉ lệ mắc bệnh) chỉ giá trị trong vòng một thời gian nhất định, và phụ thuộc vào không gian. Nói cách khác, nghiên cứu tìm hiểu tỉ lệ của một bệnh trong cộng đồng cũng chỉ có giá trị ngắn hạn ngay cả được thực hiện trong một quần thể cụ thể, và không thể khái quát hoá thậm chí cho quần thể khác. Do đó, tính đại diện ở đây cũng rất ư vô nghĩa.

Khái niệm đại diện theo tỉ lệ càng vô lí. Theo khái niệm và cách hiểu này, nếu quần thể có 5% người tuổi 70-79 thì mẫu cũng nên có 10% người trong độ tuổi đó. Đây là một hiểu lầm tai hại. Một nghiên cứu nghiêm chỉnh với cỡ mẫu gần bằng nhau cho từng nhóm tuổi có độ nhạy cao hơn nghiên cứu đảm bảo tỉ lệ của mẫu nghiên cứu và quần thể bằng nhau. Thử tưởng tượng, nếu chúng ta có 3 nhóm tuổi 50-59, 60-69, và 70-79, nghiên cứu với cỡ mẫu bằng nhau (ví dụ như 100, 100, 100 cho 3 nhóm tuổi) có độ nhạy cao hơn là theo tỉ lệ (mà kết quả có thể là 100, 60, 40). Do đó, tính đại diện cho mỗi nhóm là rất ư phản tác dụng.

Nhưng tính đại diện bị hiểu lầm khá nhiều bởi những người thiếu kinh nghiệm về nghiên cứu khoa học và các nhà dịch tễ học hiểu chưa đến vấn đề. Ở nước ngoài, thỉnh thoảng cũng có những phê bình sai về tính đại diện, vì người phát biểu hiểu sai. Tuy nhiên, họ không dám bài bác kết quả nghiên cứu, có lẽ một phần là do kinh nghiệm thực tế nên họ cũng dè dặt trong phê phán. Nhưng ở Việt Nam thì xuất phát từ ngộ nhận và hiểu sai về khái niệm đại diện nên những phê phán cũng sai và gây tác hại cho không ít nhà nghiên cứu chân chính.

Khái niệm "đại diện" theo cách hiểu 1-4 như trên đã được bàn luận trong y văn từ hơn 30 năm nay. Mới đây, một số nhà dịch tễ học lừng danh trên thế giới tranh luận trên tập san *International Journal of Epidemiology* (2-7), và ngoại trừ 1 người phản đối yếu ớt, tất cả các nhà dịch tễ học còn lại đều đồng ý rằng tính đại diện không cần cho nghiên cứu y học nhằm kiểm định một giả thuyết (8) liên quan đến một qui luật tự nhiên. Bất cứ ai làm nghiên cứu khoa học với giả thuyết khoa học đều chẳng ngạc nhiên về sự đồng ý đó, bởi vì tính đại diện theo cách hiểu 1-4 là không cần thiết và có thể gây tác hại cho sự tiến bộ của y học.

### **Chú thích:**

(1) Cần nói thêm rằng chữ "liên quan" ở đây là tương đương với thuật ngữ "association" trong tiếng Anh. Thuật ngữ *association* trong nghiên cứu khoa học dùng để chỉ những hiện tượng thống kê học như *correlation* (tương quan), *relationship* (liên hệ thường là nhân quả), *linkage* (trong di truyền học), và khác biệt

(*difference*). Một mối liên quan có thể là nhân quả hay không nhân quả. Khi tỉ lệ mắc bệnh ở nam và nữ khác nhau, chúng ta cũng có thể nói có "liên quan" giữa giới tính và nguy cơ mắc bệnh. Các chỉ số thống kê như tỉ số odds, tỉ số nguy cơ, tất cả đều phản ánh mối liên quan. Do đó, liên quan dùng ở đây là một **cách nói cấp cao** hay "**high level language**" (chứ không phải chỉ đơn thuần theo nghĩa thống kê).

2. Rothman KJ, Gallacher JEJ, Hatch E. Why representativeness should be avoided. *Int J Epidemiol* 2013;42:1012-4.

3. Elwood M. Commentary: On representativeness. *Int J Epidemiol* 2013;42:1014-5.

4. Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years – has the time come to move one? *Int J Epidemiol* 2013;42:1016-7.

5. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and should be avoided. *Int J Epidemiol* 2013;42:1018-22.

6. Ebrahim S, Smith GD. Commentary: Should we always deliberately be non-representative? *Int J Epidemiol* 2013;42:1022-6.

7. Rothman KJ, Gallacher JEJ, Hatch E. Rebuttal: When it comes to scientific inference, sometimes a cigar is just a cigar. *Int J Epidemiol* 2013;42:1026-8.

8. Trích dẫn vài bình luận của các chuyên gia:

*"Representativeness is time and place specific, and will therefore be a historical concept."*

*"We apply knowledge of genetics from fruit flies to humans because the biological relationships are generalizable although the individuals studied are not."*

*"Representativeness is not desirable even if the goal is to study effect-measure modification."*