

Tổng quan

NGHIÊN CỨU DỊCH TỄ HỌC TRONG BỆNH VIỆN: NHỮNG VẤN ĐỀ CẦN LƯU Ý VỀ THIẾT KẾ VÀ DIỄN GIẢI

Nguyễn Văn Tuấn^{1,2}

1. Khoa Y, Đại học New South Wales, Úc

2. Trung tâm Công nghệ Y tế, Đại học Công nghệ Sydney, Úc

* Tác giả liên hệ: Nguyễn Văn Tuấn ✉ tuan.nguyen@unsw.edu.au

TÓM TẮT: Phần lớn các nghiên cứu dịch tễ học và nghiên cứu lâm sàng quan sát tại Việt Nam được thực hiện trên bệnh nhân nhập viện và thường không có nhóm chứng. Mặc dù có tính thực tiễn và thuận lợi về mặt dữ liệu, các nghiên cứu này thường được sử dụng để trả lời những câu hỏi vượt quá khả năng suy luận của thiết kế. Trong bài này tôi phân tích các hạn chế phương pháp của nghiên cứu dịch tễ học dựa trên bệnh nhân trong bệnh viện, đặc biệt trong trường hợp không có nhóm chứng, trong bối cảnh nghiên cứu y học tại Việt Nam. Dựa trên phân tích khái niệm và nguyên lý dịch tễ học, bài viết thảo luận các nguồn sai lệch thường gặp trong nghiên cứu bệnh viện, bao gồm sai lệch chọn mẫu, sai lệch đồng căn (collider bias), thiếu nhóm chứng, và hiện tượng hồi quy về trung bình. Bệnh nhân nhập viện là một quần thể đã qua chọn lọc, chịu ảnh hưởng không chỉ bởi mức độ bệnh mà còn bởi khả năng tiếp cận y tế, tuyến chuyển viện và các yếu tố xã hội. Việc điều kiện hóa phân tích trên nhóm đã nhập viện có thể làm lệch mối liên quan giữa phơi nhiễm và bệnh, tạo ra các mối liên quan giả tạo hoặc làm mờ mối quan hệ nhân quả thực sự. Khi thiếu nhóm chứng, các mối liên quan quan sát dễ bị diễn giải sai như yếu tố nguy cơ ở mức quần thể. Các nghiên cứu trước-sau trong bệnh viện còn dễ bị ảnh hưởng bởi hiện tượng hồi quy về trung bình, dẫn đến đánh giá sai hiệu quả can thiệp. Các mô hình tiên lượng xây dựng từ dữ liệu nội viện có thể vận hành tốt trong bối cảnh huấn luyện nhưng khó khái quát ra ngoài bệnh viện. Nghiên cứu dựa trên bệnh nhân nhập viện vẫn có giá trị quan trọng trong mô tả bệnh nặng, tiên lượng nội viện và đánh giá quy trình chăm sóc. Tuy nhiên, phạm vi suy luận của thiết kế này là có giới hạn. Việc nhận diện rõ các giới hạn phương pháp và gắn chặt diễn giải kết quả với điều kiện chọn mẫu là cần thiết để tránh những kết luận sai lệch và nâng cao chất lượng nghiên cứu y học tại Việt Nam.

Từ khóa: Sai lệch chọn mẫu, Sai lệch đồng căn, Nghiên cứu dựa trên bệnh viện

EPIDEMIOLOGICAL RESEARCH IN HOSPITAL: KEY CONSIDERATIONS FOR DESIGN AND INTERPRETATION

Nguyen Van Tuan

ABSTRACT: Most epidemiological and observational clinical studies in Vietnam are conducted on hospitalized patients, frequently lacking a control group. Despite their practicality and data accessibility, these studies are often utilized to address questions that exceed the inferential capacity of their design. In this article, I analyze the methodological limitations of hospital-based epidemiological research, particularly in the absence of a control group, within the context of medical research in Vietnam. Drawing on conceptual analysis and epidemiological principles, the article discusses common sources of bias in hospital-based studies, including selection bias, collider bias, the lack of control groups, and the regression toward the mean phenomenon. Inpatients represent a highly selected population, influenced not only by disease severity but also by healthcare accessibility, referral pathways, and socio-economic factors. Conditioning the analysis on a hospitalized group can distort the association between exposure and disease, creating spurious correlations or masking true causal relationships. Without a control group, observed associations are easily misinterpreted as population-level risk factors. Furthermore, before-after studies in hospitals are susceptible to regression toward the mean, leading to inaccurate evaluations of intervention efficacy. Prognostic models developed from inpatient data may perform well in training environments but often fail to generalize to broader settings. While hospital-based research remains vital for describing severe illness, inpatient prognosis, and evaluating care processes, its inferential scope is inherently limited. Clearly identifying these methodological boundaries and anchoring interpretations to sampling conditions is essential to avoid biased conclusions and to enhance the quality of medical research in Vietnam.

Keywords: Selection bias, Collider bias, Hospital-based research

Tóm tắt thông điệp chính

1. Bệnh nhân nhập viện là quần thể đã qua chọn lọc. Bệnh nhân nhập viện không đại diện cho cộng đồng chung; họ thường bệnh nặng hơn và chịu ảnh hưởng của khả năng tiếp cận y tế, chuyển tuyến và yếu tố xã hội. Các ước lượng từ dữ liệu này chỉ phản ánh mối liên quan trong nhóm đã nhập viện.
2. Sai lệch chọn mẫu và sai lệch đồng căn là hệ quả của thiết kế, không phải lỗi phân tích. Việc điều kiện hóa phân tích trên nhóm đã nhập viện có thể tạo ra hoặc làm sai lệch các mối liên quan giữa yếu tố phơi nhiễm và bệnh (kết cục). Các mô hình thống kê hay machine learning không thể "sửa" được sai lệch này.
3. Thiếu nhóm chứng đồng nghĩa với thiếu phản thực. Không có nhóm chứng, nghiên cứu không thể phân biệt đâu là yếu tố làm tăng nguy cơ mắc bệnh, đâu là yếu tố làm tăng khả năng nhập viện, và đâu chỉ là dấu hiệu của mức độ bệnh trong bệnh viện.
4. Nghiên cứu trước-sau trong bệnh viện dễ bị hồi quy về trung bình. Sự cải thiện quan sát được sau nhập viện có thể phản ánh xu hướng thống kê tự nhiên hơn là hiệu quả thực sự của can thiệp, nếu không có nhóm so sánh phù hợp.
5. Mô hình tiên lượng nội viện khó khái quát ra ngoài bệnh viện. Các mô hình xây dựng từ dữ liệu bệnh viện học cả đặc điểm bệnh sinh lẫn "dấu vân tay" của hệ thống y tế; vì vậy, chúng thường hoạt động tốt nội viện nhưng suy giảm hiệu năng khi áp dụng ở bối cảnh khác.

1. DẪN NHẬP

Trong nhiều năm qua, nghiên cứu khoa học tại Việt Nam phát triển nhanh chóng [1], đặc biệt trong các bệnh viện tuyến trung ương và tuyến tỉnh. Trong nghiên cứu y học, nhiều đề tài thường bắt đầu ở khoa lâm sàng, dữ liệu được thu thập từ bệnh nhân nội trú, và phạm vi nghiên cứu hiếm khi vượt ra ngoài cổng bệnh viện. Điều này diễn ra quen thuộc đến mức nhiều người xem đó là lẽ tự nhiên của nghiên cứu y học, ít khi đặt câu hỏi vì sao lại như vậy.

Ở bệnh viện, đặc biệt là bệnh viện giảng dạy, bệnh nhân tập trung, dữ liệu lâm sàng phong phú, xét nghiệm sẵn có, và các bác sĩ / nhà nghiên cứu vừa điều trị vừa làm nghiên cứu khoa học. Bệnh nhân đã ở đó, bác sĩ gặp họ mỗi ngày, và việc trao đổi, giải thích hay thu thập thông tin diễn ra một cách liền mạch trong dòng chảy của chăm sóc y tế. Không cần tìm kiếm đối tượng, không cần xây dựng mạng lưới ngoài cộng đồng, không cần theo dõi từ xa. Nghiên cứu, trong trường hợp này, gần như phát sinh từ công việc lâm sàng thường nhật.

Song song với sự hiện diện của bệnh nhân là kho dữ liệu lâm sàng ngày càng phong phú. Bệnh án điện tử, kết quả xét nghiệm, hình ảnh học, chẩn đoán và diễn tiến điều trị được ghi nhận như một phần tất yếu của thực hành y khoa hiện đại. Đối với nhà nghiên cứu, nhất là trong bối cảnh nguồn lực còn hạn chế, việc có sẵn dữ liệu là một lợi thế lớn. Nghiên cứu trong bệnh viện cho cảm giác "dữ liệu đã có rồi, chỉ cần phân tích", trong khi nghiên cứu cộng đồng đòi hỏi nhiều công sức hơn, từ thiết kế khung chọn mẫu đến tổ chức thu thập và theo dõi dài hạn.

Yếu tố đạo đức và hậu cần cũng góp phần làm cho nghiên cứu trong bệnh viện trở thành lựa chọn quen thuộc. Trong môi trường lâm sàng, việc xin ý kiến chấp thuận thường thuận lợi hơn, vì bệnh nhân đã có mối quan hệ sẵn có với bệnh viện. Việc tái khám, liên lạc hay cập nhật kết cục diễn ra trong một không gian quen thuộc, ít phụ thuộc vào những yếu tố khó kiểm soát ngoài cộng đồng. Ngay cả hội đồng đạo đức cũng thường cảm thấy "yên tâm" hơn với những nghiên cứu diễn ra trong phạm vi bệnh viện, nơi các quy trình chăm sóc và giám sát đã được thiết lập.

Một lý do khác, tuy ít được nói ra nhưng rất thực, là cảm giác gần gũi với thực hành lâm sàng. Kết quả rút ra từ bệnh nhân đang điều trị thường mang lại ấn tượng rằng chúng áp dụng được ngay. Khi một yếu tố nào đó liên quan đến biến chứng, tử vong hay thời gian nằm viện, bác sĩ dễ dàng liên hệ với những ca bệnh mình đang chăm sóc.

Sự gần gũi này tạo ra cảm giác giá trị tức thời, đôi khi mạnh hơn những cân nhắc trừu tượng về thiết kế nghiên cứu hay tính đại diện của quần thể.

Khi nghiên cứu bệnh nhân trong bệnh viện trở thành lựa chọn mặc định, ranh giới giữa điều “dễ làm” và điều “nên làm” bắt đầu bị xóa mờ. Những mối liên quan quan sát được trong một quần thể nội viện hẹp dễ dàng được diễn giải như yếu tố nguy cơ chung; các mô hình tiên lượng nội viện có thể được kỳ vọng vượt ra ngoài bối cảnh mà chúng được xây dựng. Sự thuận tiện về dữ liệu và hậu cần, nếu không được đặt trong một khung phương pháp rõ ràng, có thể khiến nhà nghiên cứu vô tình yêu cầu thiết kế nghiên cứu trả lời những câu hỏi mà nó không có khả năng trả lời.

Trong phần tiếp theo, tôi sẽ tập trung phân tích các vấn đề phương pháp phổ biến của nghiên cứu dịch tễ học dựa trên bệnh nhân nhập viện không có nhóm chứng, bao gồm sai lệch chọn mẫu, sai lệch đồng căn, sự thiếu vắng so sánh phản thực, và hiện tượng hồi quy về trung bình. Đây không phải là những lỗi kỹ thuật có thể sửa chữa bằng phân tích thống kê tinh vi hơn hay AI, mà là những hệ quả trực tiếp của cách thiết kế và chọn đối tượng nghiên cứu ngay từ đầu [2].

2. SAI LỆCH CHỌN MẪU (SELECTION BIAS)

Trong nghiên cứu y học, có một giả định ngầm thường được chấp nhận mà ít khi được nói ra: những gì nhà nghiên cứu quan sát được trong bệnh viện có thể đại diện, ít nhất, cho những gì đang diễn ra ngoài cộng đồng. Giả định này không hẳn là sai trong tất cả trường hợp, nhưng nó mong manh. Chính từ giả định mong manh đó, sai lệch chọn mẫu bắt đầu hình thành và len lỏi vào nhiều kết luận tưởng như rất hợp lý.

Bệnh nhân nhập viện, trước hết, không phải là một mẫu ngẫu nhiên từ quần thể. Họ thường là những người bệnh nặng hơn, lớn tuổi hơn, có nhiều bệnh đi kèm hơn, và trong nhiều hệ thống y tế, bao gồm cả Việt Nam, họ thường thuộc nhóm có điều kiện kinh tế – xã hội bất lợi hơn. Việc họ có mặt trong bệnh viện đã phản ánh một chuỗi lựa chọn và hoàn cảnh: từ mức độ bệnh, khả năng tiếp cận chăm sóc ban đầu, cho đến quyết định chuyển tuyến của bác sĩ. Vì vậy, ngay từ thời điểm nhà nghiên cứu bắt đầu thu thập dữ liệu, quần thể nghiên cứu đã khác xa với quần thể mà ta thường ngầm nghĩ đến khi dùng những từ như “nguy cơ” hay “yếu tố liên quan”.

Hậu quả trực tiếp của điều này là tính đại diện bị suy giảm, đôi khi rất nghiêm trọng. Những ước lượng rút ra từ mẫu bệnh viện, dù được phân tích cẩn thận đến đâu, cũng chỉ phản ánh mối liên quan trong một nhóm người đã được chọn lọc sẵn. Quên hay bỏ qua sự thật đó và diễn giải kết quả như thể chúng áp dụng cho toàn bộ cộng đồng, sai lệch chọn mẫu bắt đầu biến thành sai lệch suy luận.

Có thể nghĩ đến những nghiên cứu về biến chứng của bệnh tiểu đường dựa trên bệnh nhân nhập viện như là những ca minh họa cho sai lệch chọn mẫu. Trong nhóm này, nhà nghiên cứu thường thấy tỷ lệ kiểm soát đường huyết kém rất cao, cùng với nhiều bệnh kèm như cao huyết áp, suy thận hay bệnh tim mạch. Nếu từ đó nhà nghiên cứu kết luận rằng một số đặc điểm lâm sàng hoặc hành vi nào đó là “yếu tố nguy cơ” của biến chứng, thì câu hỏi cần được đặt ra là: nguy cơ đối với ai? Đối với tất cả người bệnh tiểu đường ngoài cộng đồng, hay chỉ đối với nhóm đã đủ nặng hoặc đủ khó khăn để phải nhập viện?

Trong thực tế, nhóm người tiểu đường không nhập viện (những người kiểm soát bệnh tương đối tốt) được theo dõi thường xuyên ở tuyến cơ sở, hoặc đơn giản là chưa gặp biến chứng nặng, gần như vắng mặt hoàn toàn trong dữ liệu. Khi một nửa bức tranh bị che khuất, bức tranh còn lại khó có thể được xem là đại diện.

Sai lệch chọn mẫu còn tinh vi hơn khi nó làm méo mó phân bố của các yếu tố phơi nhiễm (exposures) [3]. Ở những bệnh nhân nhập viện, một số yếu tố xuất hiện rất thường xuyên, không phải vì chúng thực sự phổ biến trong cộng đồng, mà vì chúng làm tăng khả năng nhập viện. Chẳng hạn, kiểm soát đường huyết kém, tuân thủ điều trị thấp, hoặc điều kiện sống khó khăn có thể xuất hiện dày đặc trong bệnh viện công, đơn giản vì những yếu tố này đẩy người bệnh đến ngưỡng cần nhập viện. Khi phân tích dữ liệu, nhà nghiên cứu dễ có cảm giác rằng đây là những yếu tố phổ biến hoặc quan trọng, trong khi thực tế chúng chỉ phản ánh logic chọn lọc của hệ thống y tế.

3. SAI LỆCH ĐỒNG CĂN

Nếu sai lệch chọn mẫu liên quan đến câu hỏi *ai được đưa vào nghiên cứu*, thì sai lệch đồng căn liên quan đến một vấn đề tinh vi hơn: *điều gì xảy ra với các mối liên quan khi nhà nghiên cứu chỉ phân tích trong số những người đã được chọn lọc đó*. Hai dạng sai lệch này thường xuất hiện đồng thời trong nghiên cứu bệnh viện, nhưng chúng không đồng nhất và không thể thay thế cho nhau.

Trong nghiên cứu dựa trên bệnh nhân nhập viện, nhà nghiên cứu luôn bắt đầu từ một điều kiện chung: tất cả đối tượng đều đã vào viện. Điều kiện này tương chừng hiển nhiên, nhưng chính nó lại tạo ra một “điểm giao” nơi nhiều yếu tố khác nhau cùng tác động. Nhập viện không phải là một sự kiện ngẫu nhiên; nó là kết quả của mức độ bệnh, khả năng tiếp cận y tế, quyết định của bác sĩ, hoàn cảnh kinh tế – xã hội, và nhiều yếu tố khác cùng hội tụ. Trong ngôn ngữ của dịch tễ học, nhập viện trở thành một *đồng căn* (collider).

Khi nhà nghiên cứu giới hạn phân tích vào nhóm đã đi qua điểm giao này, các mối liên quan giữa phơi nhiễm và kết cục có thể bị bóp méo: những liên hệ vốn không tồn tại ngoài cộng đồng có thể xuất hiện trong bệnh viện, và những liên hệ có thật lại có thể bị làm mờ hoặc đảo chiều. Hiện tượng này đã được Joseph Berkson mô tả từ năm 1946 [4] và thường được gọi là Berkson’s bias, nhưng trong thực hành hiện đại, nó xuất hiện dưới nhiều hình thức tinh vi hơn.

Để thấy rõ điều này, hãy tưởng tượng một nghiên cứu tìm hiểu mối liên quan giữa hút thuốc và mức độ nặng của viêm phổi trong số bệnh nhân nhập viện. Trong cộng đồng, hút thuốc làm tăng nguy cơ viêm phổi và nhiều bệnh hô hấp khác. Điều này không gây tranh cãi. Tuy nhiên, trong bệnh viện, bức tranh có thể khác: những người hút thuốc nặng thường có nhiều bệnh nền, nguy cơ tử vong sớm cao, hoặc đã có các biến chứng tim phổi từ trước. Một số người trong số họ có thể không sống đủ lâu để nhập viện vì viêm phổi, hoặc khi vào viện thì đã ở trong tình trạng rất nặng và được phân loại khác.

Do đó, khi chỉ phân tích nhóm đã nhập viện, nhà nghiên cứu có thể quan sát thấy rằng tỷ lệ hút thuốc ở nhóm viêm phổi nặng không cao như mong đợi, thậm chí thấp hơn ở nhóm nhẹ. Nếu không ý thức về sai lệch đồng căn, nhà nghiên cứu có thể bối rối, hoặc tệ hơn, suy diễn rằng hút thuốc “không liên quan” hoặc “có vẻ bảo vệ” viêm phổi. Thực ra, mối liên quan đã bị sai lệch ngay từ lúc nhà nghiên cứu đặt điều kiện “đã nhập viện”. Nhập viện, trong trường hợp này, là điểm giao của cả hút thuốc lẫn mức độ nặng của bệnh, và việc điều kiện hóa trên điểm giao đó đã tạo ra một mối liên quan giả tạo.

Điều đáng nói là sai lệch đồng căn không biến mất khi nhà nghiên cứu dùng hồi quy đa biến, mô hình phức tạp hay machine learning. Nó không phải là vấn đề thiếu điều chỉnh, mà là vấn đề *điều chỉnh sai chỗ*. Khi thiết kế nghiên cứu đã khóa chặt đối tượng trong một không gian chọn lọc như bệnh viện, thì mọi phân tích sau đó đều vận hành trong không gian méo mó đó. Không có thuật toán nào có thể ‘mở khóa’ lại thông tin đã bị loại bỏ ngay từ đầu.

Trong bối cảnh Việt Nam, nơi bệnh viện vẫn là trung tâm của cả chăm sóc lẫn nghiên cứu, sai lệch đồng căn đặc biệt dễ xảy ra nhưng lại ít khi được gọi tên. Nhiều nghiên cứu quan sát thấy những mối liên quan lạ, nghịch lý, hoặc trái với hiểu biết sinh học, nhưng thay vì đặt câu hỏi về điều kiện chọn mẫu, nhà nghiên cứu lại tìm cách giải thích bằng cơ chế mới, yếu tố mới, hoặc giả thuyết mới. Cách tiếp cận này vô tình làm sai lệch phương hướng khoa học.

4. VẤN ĐỀ NGHIÊN CỨU THIẾU NHÓM CHỨNG

Trong dịch tễ học, có một ý niệm tương chừng rất trừu tượng nhưng lại chi phối toàn bộ cách nhà nghiên cứu suy luận: *phản thực* (counterfactual), hay nói một cách giản dị hơn, là câu hỏi “*điều gì sẽ xảy ra nếu mọi thứ khác giữ nguyên, chỉ thay đổi một yếu tố?*”. Vì không thể quay ngược thời gian hay đặt cùng một con người vào hai hoàn cảnh khác nhau, khoa học buộc phải dùng nhóm chứng như một cách xấp xỉ thực tế cho phản thực đó. Nhóm chứng không phải là chi tiết kỹ thuật phụ trợ; nó là nền móng để mọi so sánh có ý nghĩa khoa học.

Khi nghiên cứu trong bệnh viện không có nhóm chứng, nền móng phản thực biến mất. Nhà nghiên cứu chỉ còn lại một tập hợp bệnh nhân, và tất cả phân tích đều diễn ra bên trong tập hợp đó. Không có mốc so sánh bên ngoài, không có bối cảnh để đặt câu hỏi “so với ai?”, và vì vậy, rất dễ nhầm lẫn giữa mối liên quan nhân quả.

Trong những nghiên cứu như vậy, tần suất của một phơi nhiễm nào đó (như cao tuổi, bệnh nền, chỉ số sinh học bất thường) không có điểm tựa để đánh giá. Nhà nghiên cứu thấy nó xuất hiện nhiều trong bệnh viện, nhưng không biết nó có thực sự phổ biến hơn ngoài cộng đồng hay không. Nhà nghiên cứu biết rằng trong số những người đã nhập viện, yếu tố đó liên quan đến kết cục xấu hơn, nhưng không biết liệu chính yếu tố đó có làm tăng nguy cơ mắc bệnh hay chỉ làm tăng khả năng phải nhập viện. Khi không có nhóm chứng, hai câu hỏi này dễ bị trộn lẫn, và sự trộn lẫn đó dẫn đến sai lệch trong diễn giải.

Một ví dụ điển hình là các nghiên cứu về COVID-19 trong giai đoạn đầu của đại dịch [5]. Nhiều công trình nghiên cứu dựa trên bệnh nhân nhập viện đã xác định các yếu tố nguy cơ của COVID-19 nặng như tuổi cao, béo phì, tiểu đường hay bệnh tim mạch. Những kết quả này nhìn qua có vẻ hiển nhiên, nhưng câu hỏi quan trọng hơn lại ít khi được đặt ra: những yếu tố này làm tăng nguy cơ diễn tiến nặng sau khi đã nhập viện, hay làm tăng nguy cơ phải nhập viện ngay từ đầu?

Nếu chỉ nhìn vào nhóm bệnh nhân nội trú, nhà nghiên cứu chỉ có thể trả lời câu hỏi thứ nhất, và ngay cả câu trả lời đó cũng chỉ đúng trong điều kiện đã nhập viện. Có thể một yếu tố nào đó dự đoán tử vong rất tốt trong bệnh viện, nhưng đồng thời lại làm giảm khả năng nhập viện vì những người mang yếu tố đó ít khi đến được bệnh viện, hoặc được quản lý sớm ở tuyến dưới. Không có nhóm chứng, sự khác biệt này hoàn toàn không thể được nhận diện.

Chính ở đây, khái niệm “yếu tố nguy cơ” bắt đầu trở nên ... trôi nổi. Trong bối cảnh bệnh viện, một phơi nhiễm hay yếu tố có thể là dấu hiệu của mức độ nặng, của cơ may tiếp cận y tế, hoặc của ngưỡng nhập viện, chứ không nhất thiết là yếu tố quyết định bệnh. Tuy nhiên, khi thuật ngữ “risk factor” (yếu tố nguy cơ) được sử dụng mà không kèm theo điều kiện “trong số bệnh nhân đã nhập viện”, nó mang theo một hàm ý ngầm rằng kết quả đó có giá trị rộng hơn, ở mức quần thể. Đây là bước trượt rất nhỏ về ngôn ngữ, nhưng lại là bước trượt lớn về suy luận.

5. HỒI QUY VỀ SỐ TRUNG BÌNH (REGRESSION TOWARD THE MEAN)

Nếu việc thiếu nhóm chứng làm mất đi khả năng so sánh giữa các nhóm người khác nhau, thì hiện tượng hồi quy về trung bình [6] làm méo mó việc so sánh *giữa các thời điểm trên cùng một nhóm người*. Hai vấn đề này thường cùng tồn tại trong nghiên cứu bệnh viện, nhưng chúng đại diện cho hai dạng sai lệch khác nhau trong suy luận.

Trong bối cảnh bệnh viện, bệnh nhân hiếm khi được đưa vào nghiên cứu khi tình trạng còn ổn định. Việc nhập viện thường xảy ra đúng vào thời điểm bệnh đạt đến mức độ nặng, hoặc khi các chỉ số sinh học vượt xa giá trị trung bình. Chính việc chọn thời điểm quan sát này tạo điều kiện cho hiện tượng hồi quy về trung bình: *những giá trị cực đoan có xu hướng tiến gần hơn về mức trung bình ở các lần đo sau, ngay cả khi không có bất kỳ can thiệp nào* [2,6]

Khi sự cải thiện tự nhiên này trùng với quá trình điều trị nội viện, nhà nghiên cứu rất dễ diễn giải nhầm rằng can thiệp là nguyên nhân chính của thay đổi quan sát được. Không có nhóm chứng hoặc nhóm so sánh phù hợp, sự phân biệt giữa hiệu quả điều trị và xu hướng thống kê tự nhiên trở nên bất khả thi.

Chẳng hạn như bệnh nhân tiểu đường nhập viện vì đường huyết rất cao, và lúc nhập viện, đường huyết hoặc HbA1c ở mức khiến cả bệnh nhân lẫn bác sĩ đều lo lắng. Sau vài ngày điều trị nội viện, các chỉ số này giảm xuống rõ rệt. Một nghiên cứu trước-sau (before-after study), nếu chỉ dựa vào hai thời điểm đo này, rất dễ kết luận rằng phác đồ điều trị trong bệnh viện có hiệu quả rõ ràng. Nhưng câu hỏi cần được đặt ra là: bao nhiêu phần của sự “cải thiện” đó đến từ can thiệp, và bao nhiêu phần đơn giản là do

bệnh nhân được chọn vào nghiên cứu đúng vào lúc chỉ số của họ đang ở mức cực đoan?

Không có nhóm chứng, câu hỏi này không thể được trả lời. Nhà nghiên cứu không biết đường huyết của những bệnh nhân tương tự sẽ diễn tiến ra sao nếu không nhập viện, hoặc nếu chỉ được theo dõi mà không thay đổi điều trị. Nhà nghiên cứu chỉ thấy rằng từ một điểm rất cao, giá trị đi xuống, và não bộ con người (vốn thích tìm nguyên nhân) nhanh chóng gán sự đi xuống đó cho can thiệp đang diễn ra. Hồi quy về trung bình, trong trường hợp này, bị che khuất bởi trực giác lâm sàng.

Vấn đề trở nên rõ nét hơn trong các phân tích trước-sau được thực hiện hoàn toàn trong bệnh viện. Những nghiên cứu này thường được thiết kế với ý định tốt: đánh giá hiệu quả của một biện pháp, một quy trình, hoặc một phác đồ mới. Nhưng khi đối tượng được chọn là những bệnh nhân nhập viện vì tình trạng xấu nhất của họ, thì chính tiêu chí chọn mẫu đã đảm bảo rằng, trung bình, tình trạng của họ sẽ tốt lên theo thời gian! Sự cải thiện, dù có thật, không thể được tách bạch khỏi xu hướng tự nhiên của dữ liệu.

Trong thực hành lâm sàng, điều này dễ bị bỏ qua vì nó phù hợp với kỳ vọng. Bác sĩ điều trị kỳ vọng bệnh nhân cải thiện, và khi điều đó xảy ra, bác sĩ có xu hướng xem đó là bằng chứng cho quyết định của mình. Nhưng nghiên cứu khoa học đòi hỏi một khoảng lùi. Nó buộc nhà nghiên cứu và bác sĩ phải chấp nhận khả năng rằng một phần – đôi khi là phần lớn – của sự cải thiện quan sát được không đến từ can thiệp, mà từ chính cách nhà nghiên cứu chọn thời điểm quan sát.

Hồi quy về trung bình không chỉ giới hạn ở đường huyết. Nó xuất hiện trong huyết áp, chức năng thận, chỉ số viêm, thậm chí cả các thang điểm lâm sàng. Bất cứ khi nào bệnh nhân được đưa vào nghiên cứu vì một chỉ số "quá xấu", nhà nghiên cứu đều đang đứng trước nguy cơ diễn giải nhầm. Và nguy cơ này tăng lên đáng kể khi nghiên cứu thiếu nhóm đối chứng, hoặc khi mọi so sánh chỉ xoay quanh hai thời điểm trước và sau nhập viện.

Điều đáng nói là hồi quy về trung bình không phải là lỗi thống kê có thể chỉnh sửa bằng mô hình phức tạp; nó là hệ quả trực tiếp của thiết kế và cách chọn đối tượng nghiên cứu. Không có thuật toán nào có thể phân biệt đâu là cải thiện tự nhiên, đâu là hiệu quả điều trị, nếu dữ liệu ban đầu không cho phép sự phân biệt đó. Khi thiết kế đã gắn chặt việc quan sát với thời điểm bệnh nặng nhất, mọi phân tích sau đó đều phải gánh theo hệ quả này.

Hồi quy về trung bình còn là một lời nhắc về sự khiêm tốn trong suy luận. Khi một kết quả trông quá đẹp, quá phù hợp với kỳ vọng, khoa học buộc nhà nghiên cứu phải tự hỏi: nếu không làm gì cả, liệu điều này có xảy ra không? Khi câu hỏi đó chưa được trả lời, mọi kết luận về hiệu quả đều cần được giữ ở mức dè dặt. Không phải vì bi quan, mà vì nhà nghiên cứu hiểu rằng dữ liệu, giống như bệnh nhân, luôn mang theo lịch sử của chính nó.

6. XÂY DỰNG MÔ HÌNH TIỀN LƯỢNG TỪ BỆNH NHÂN NHẬP VIỆN

Những vấn đề phương pháp đã phân tích ở trên không chỉ ảnh hưởng đến việc xác định mối liên quan hay yếu tố nguy cơ, mà còn liên quan đến các mô hình tiên lượng được xây dựng từ dữ liệu bệnh viện.

Trong những năm gần đây, cùng với sự phát triển của hồ sơ bệnh án điện tử và năng lực tính toán, việc xây dựng mô hình tiên lượng từ dữ liệu bệnh viện trở nên ngày càng phổ biến. Các mô hình này hứa hẹn giúp bác sĩ dự đoán tử vong, biến chứng, thời gian nằm viện hay nhu cầu hồi sức ngay từ lúc bệnh nhân nhập viện. Trên bề mặt, đây là một hướng đi hợp lý: dữ liệu sẵn có, kết cục rõ ràng, và giá trị ứng dụng dường như rất gần với thực hành lâm sàng.

Tuy nhiên, cần nhấn mạnh rằng các mô hình này được huấn luyện hoàn toàn trong một không gian đã qua chọn lọc: không gian của những bệnh nhân đủ điều kiện để nhập viện. Do đó, chúng không chỉ học các đặc điểm bệnh sinh, mà còn vô tình học cả những quy luật ngầm của hệ thống y tế, từ tiêu chí nhập viện, mô hình chuyển tuyến, cho đến khả năng tiếp cận chăm sóc. Khi mang các mô hình này ra ngoài bối cảnh mà chúng được xây dựng, hiệu năng suy giảm không phải là điều bất ngờ, mà là hệ quả tất

yếu của thiết kế dữ liệu ban đầu.

Lý do không nằm ở thuật toán, mà nằm ở thiết kế. Dữ liệu bệnh viện không chỉ phản ánh bệnh sinh, mà còn phản ánh cách hệ thống y tế vận hành. Một mô hình học từ dữ liệu này sẽ vô tình học cả những quy luật ngầm của nhập viện: ai thường được chuyển tuyến, ai có bảo hiểm, ai đến muộn, ai được làm nhiều xét nghiệm. Những yếu tố đó có thể giúp mô hình dự đoán rất tốt kết cục nội viện, nhưng lại không có nhiều ý nghĩa khi bối cảnh thay đổi.

Một ví dụ dễ hình dung là mô hình tiên lượng tử vong ở bệnh nhân viêm phổi nhập viện. Mô hình có thể sử dụng tuổi, SpO₂, CRP, chức năng thận và một số bệnh nền để dự đoán nguy cơ tử vong trong bệnh viện. Trong dữ liệu huấn luyện, mô hình hoạt động xuất sắc: phân biệt rõ nhóm nguy cơ cao và thấp, đường cong ROC đẹp, độ chính xác cao. Nhưng nếu đem mô hình này áp dụng ở tuyến cơ sở để sàng lọc bệnh nhân viêm phổi trong cộng đồng, kết quả sẽ rất khác. Nhiều người có SpO₂ thấp hay CRP cao ngoài cộng đồng chưa chắc đã nhập viện, trong khi những người nhập viện trong dữ liệu gốc đã trải qua một quá trình chọn lọc rất riêng.

Không có nhóm chứng, mô hình khó có thể phân biệt đâu là yếu tố dự đoán bệnh, đâu là yếu tố dự đoán nhập viện, và đâu là yếu tố dự đoán kết cục sau khi đã nhập viện. Cao tuổi, chẳng hạn, có thể là yếu tố nguy cơ của tử vong nội viện, nhưng đồng thời cũng là yếu tố làm tăng khả năng được nhập viện ngay từ đầu. Khi cả ba vai trò này bị trộn lẫn trong cùng một mô hình, kết quả của mô hình tiên lượng trở nên phụ thuộc chặt chẽ vào bối cảnh huấn luyện, và mất đi khả năng đi xa hơn.

Vấn đề này thể hiện rõ nhất ở hiệu chuẩn mô hình. Một mô hình có thể dự đoán đúng thứ tự nguy cơ trong bệnh viện, nhưng lại dự đoán sai xác suất tuyệt đối khi áp dụng ở nơi khác. Nguy cơ 20% trong bệnh viện không còn là 20% ở tuyến cơ sở hay trong cộng đồng, không phải vì mô hình 'tê', mà vì quần thể đã thay đổi. Nguồn dữ liệu khác nhau dẫn đến cùng một đầu vào mang ý nghĩa khác nhau. Đây không phải là thất bại của thống kê học, mà là hệ quả tất yếu của việc dùng dữ liệu hẹp để trả lời câu hỏi rộng.

Trong thực tế, nhiều mô hình tiên lượng được công bố mà không có ngoại kiểm (external validation). Khi có kiểm định, hiệu năng thường giảm đáng kể, đôi khi đến mức không còn giá trị thực hành. Hiện tượng này thường được gọi là 'quá lạc quan' (optimism), nhưng cách gọi đó chưa phản ánh hết vấn đề. Mô hình không quá lạc quan; nó chỉ đang làm rất tốt công việc mà nó được xây dựng trong một không gian rất cụ thể.

Trong bối cảnh nghiên cứu tại Việt Nam, nơi dữ liệu bệnh viện là nguồn chủ đạo cho cả nghiên cứu và ứng dụng, việc xây dựng mô hình tiên lượng tử vong bệnh nhân nhập viện là điều khó tránh. Vấn đề không nằm ở việc có nên xây dựng hay không, mà ở việc nhà nghiên cứu kỳ vọng gì ở những mô hình này. Nếu mục tiêu là hỗ trợ quyết định trong bệnh viện, chúng có thể rất hữu ích. Nhưng nếu được diễn giải như công cụ dự đoán nguy cơ cộng đồng, hay như bằng chứng cho yếu tố căn nguyên của bệnh, thì đó là một bước đi quá xa.

Sau cùng, mô hình tiên lượng cũng giống như bất kỳ công cụ lâm sàng nào khác: giá trị của nó phụ thuộc vào bối cảnh sử dụng. Một mô hình sinh ra từ bệnh viện sẽ mang theo 'dấu vân tay' của bệnh viện đó. Nhận ra dấu vân tay này không làm mô hình kém đi, mà giúp nó được dùng đúng chỗ, đúng việc. Trong khoa học, hiểu rõ giới hạn của công cụ đôi khi quan trọng không kém việc cải tiến chính công cụ đó.

7. KẾT LUẬN

Những phân tích trong bài này không nhằm phủ nhận vai trò của nghiên cứu dịch tễ học và nghiên cứu quan sát dựa trên bệnh nhân trong bệnh viện. Trong thực tế, với điều kiện nguồn lực còn hạn chế, mô hình nghiên cứu này đã và đang đóng góp đáng kể vào việc mô tả đặc điểm bệnh nặng, đánh giá kết cục nội viện, cải thiện quy trình chăm sóc, và xây dựng các mô hình tiên lượng phục vụ thực hành lâm sàng. Ở những mục tiêu đó, dữ liệu bệnh viện không chỉ phù hợp mà đôi khi còn là lựa chọn khả thi duy nhất.

Vấn đề cốt lõi nằm ở chỗ thiết kế nghiên cứu thường bị yêu cầu trả lời những câu hỏi vượt quá năng lực suy luận vốn có của nó. Khi dữ liệu được thu thập trong một quần

thể đã qua chọn lọc có hệ thống, không có nhóm chứng, và bị điều kiện hóa bởi việc nhập viện, các khái niệm như “yếu tố nguy cơ”, “liên quan nhân quả” hay “hiệu quả can thiệp” dễ bị sử dụng một cách không thận trọng. Sai lệch chọn mẫu, sai lệch đồng căn, hồi quy về trung bình và thiếu nhóm chứng (phản thực) không phải là những lỗi kỹ thuật có thể giải quyết bằng phân tích tinh vi hơn, mà là hệ quả trực tiếp của thiết kế nghiên cứu ngay từ đầu.

Trong bối cảnh đó, điều quan trọng không phải là hạn chế hay né tránh nghiên cứu trong bệnh viện, mà là đặt nó vào đúng vị trí phương pháp luận. Nghiên cứu dựa trên bệnh nhân nhập viện có thể trả lời tốt các câu hỏi về tiên lượng nội viện và vận hành hệ thống chăm sóc y tế, nhưng không đủ cơ sở để suy luận về nguy cơ ở mức quần thể hay về quan hệ nhân quả nếu thiếu thiết kế so sánh phù hợp. Việc diễn giải kết quả cần luôn gắn chặt với điều kiện chọn mẫu, thay vì được “mở rộng” bằng ngôn ngữ hoặc kỳ vọng vượt ra ngoài dữ liệu.

Bài học nền tảng của dịch tễ học: thiết kế nghiên cứu quan trọng hơn kỹ thuật phân tích, và nhóm chứng là nền móng của suy luận khoa học. Bài học này vẫn còn nguyên giá trị, ngay cả trong kỷ nguyên dữ liệu lớn và thông minh nhân tạo. Nếu không nhận dạng giới hạn này, nghiên cứu y học rất dễ tạo ra những kết quả có vẻ hợp lý và thuyết phục, nhưng lại mang theo nguy cơ dẫn dắt sai lầm cả về khoa học lẫn chính sách y tế. Ngược lại, khi các giới hạn được nhận diện và tôn trọng, nghiên cứu trong bệnh viện không những không bị giảm giá trị, mà còn trở nên trung thực hơn, hữu ích hơn, và đóng góp đúng chỗ vào tiến trình tích lũy tri thức y học.

TÀI LIỆU THAM KHẢO

- [1] Nguyen TV, Pham L. Scientific output and its relationship to knowledge economy: an analysis of ASEAN countries. *Scientometrics* 2011;89:107-117.
- [2] Nguyen TV. Common methodological issues and suggested solutions in bone research. *Osteoporosis and Sarcopenia* 2020;6:161-167.
- [3] Hernán MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-625.
- [4] Berkson J. Limitations of the Application of Fourfold Table Analysis to Hospital Data. *Biometrics Bulletin* 1946;2(3):47-53.
- [5] Haddad, C., Bou Malhab, S., Sacre, H. & Salameh, P. Smoking and COVID-19: A scoping review. *Tob. Use Insights* 2021 Feb 15;14:1179173X21994612.
- [6] Bland JM, Altman DG. Regression towards the mean *BMJ* 1994 Jun 4;308(6942):1499.