

Những tiêu chí cần thiết để đánh giá bằng chứng y học

Nguyễn Văn Tuấn

(trích từ sách “*Hai mặt sáng tối của y học hiện đại*”

Nhà xuất bản Trẻ 2004)

Tin tức y học ngày nay không chỉ được lưu truyền trong các tập san khoa học chuyên môn, mà còn xuất hiện trong các hệ thống truyền thông đại chúng. Hầu như hàng tuần, thậm chí hàng ngày, báo chí, đài truyền thanh, hay đài truyền hình đều có những bản tin về y học. Một phần lớn những bản tin này tường thuật những phát hiện mới trong nghiên cứu y khoa, đặc biệt là những mối liên hệ giữa bệnh tật và thực phẩm, hay môi trường, hay hóa chất mà chúng ta dùng hàng ngày. Trong nhiều trường hợp những bản tin này mâu thuẫn nhau một cách đến ngớ ngẩn. Có lúc chúng ta được cho biết rằng uống cà phê có hại cho sức khỏe; lúc khác, chúng ta được báo rằng cà phê có lợi cho sức khỏe. Có nghiên cứu kết luận rằng dùng điện thoại di động có thể gây ra ung thư; nhưng sau đó lại có nghiên cứu khác cho rằng điện thoại di động rất an toàn. Có thời gian, phương pháp thay thế kích thích tố (còn gọi là HRT, hay hormone replacement therapy) được xem là “thần dược,” nhưng gần đây, giới y khoa cho chúng ta biết HRT là thủ phạm gây ra ung thư vú và có hại nhiều hơn có lợi [1].

Những thông tin mâu thuẫn như thế tăng theo cấp số nhân và cường độ xuất hiện của những bản tin như thế cao đến nỗi người ta khó mà phân biệt được đâu là thật và đâu là giả. Hậu quả là những bản tin như thế làm cho nhiều người hoang mang, và có phản ứng không thích hợp, và có khi nguy hiểm. Chẳng hạn như trường hợp một phụ nữ người Mỹ, sau khi nghe tin chất silicon dùng để làm lớn vú có thể gây ra ung thư, chị này bèn dùng dao cạo tự mình giải phẫu lấy túi silicon ra [2]! Một người phụ nữ khác sau khi được biết qua nghe tin tức là những quả táo được xử lí vệ sinh bằng hóa chất có thể gây ra ung thư, bà ta bèn gọi điện cho cảnh sát chặn chiếc xe bus đang trên đường chở con bà đến trường học để lấy quả táo trong cặp của đứa bé [3].

Hai câu hỏi được đặt ra là những bản tin như thế có đáng được tin cậy hay không, và bằng cách nào để xác nhận một mối liên hệ giữa bệnh tật và một nhân tố? Ngày nay, trên thế giới có khoảng một ngàn tập san y khoa, mỗi năm công bố hơn trăm ngàn bài báo y học. Với một lượng thông tin khổng lồ như thế, chắc chắn không phải thông tin nào cũng có giá trị như nhau. Do đó, có thể nói rằng không phải kết quả của bất cứ nghiên cứu nào đều có độ tin cậy như nhau. Những

ngiên cứu không có mục tiêu rõ ràng, phương pháp thiết kế nghiên cứu sai, cách thu thập và phân tích dữ kiện không thích hợp, diễn dịch dữ kiện ngoài phạm vi cho phép, v.v... rất phổ biến trong các tập san y khoa chuyên môn, nhưng lại ít khi nào được giới truyền thông quan tâm đến, và do đó, có thể nói phần lớn những bản tin y khoa có chất lượng khoa học không cao.

Mức độ tin cậy của bằng chứng y học

Do đó, đối với công chúng, vấn đề được đặt ra phải dùng những tiêu chí nào để phân biệt các nghiên cứu có chất lượng cao với các nghiên cứu có chất lượng thấp. Cố nhiên, cơ bản của vấn đề là mức độ tin cậy của các bằng chứng y học được rút ra từ những nghiên cứu. Nếu một bản tin báo cho chúng ta biết ăn nhiều chất béo sẽ làm tăng nguy cơ bị bệnh tim mạch, câu hỏi mà chúng ta cần phải đặt ra là “Phát hiện này có độ tin cậy cỡ nào?” Để trả lời câu hỏi này, người ta xếp loại bằng chứng nghiên cứu y học thành 6 nhóm (tính từ thấp nhất, hay có độ tin cậy thấp, đến loại có độ tin cậy cao nhất) như sau:

Loại một, là những **ý kiến cá nhân của các bác sĩ hay nhà khoa học**. Những ý kiến cá nhân của những người này, bất kể họ giữ chức vụ cao hay thấp trong khoa bảng như giáo sư, giám đốc, viện trưởng, v.v... là loại bằng chứng có giá trị thấp nhất. Thực ra, đứng trên phương diện khoa học mà nói, những ý kiến cá nhân, kinh nghiệm cá nhân không có một giá trị nào đáng kể, nếu chúng không được đi kèm theo những dẫn chứng khoa học đáng tin cậy.

Bằng chứng khoa học có giá trị thấp vào hàng thứ hai là những **ngiên cứu trên thú vật**, hay trong ống nghiệm. Đây là các nghiên cứu mà đối tượng không phải là con người mà là những động vật (thường là chuột và thỏ), hay những tế bào được chiết ra từ động vật đó. Hai động vật này tuy có cấu trúc di truyền giống con người, nhưng môi trường sinh sống và quá trình tiến hóa khác nhau, cũng như hệ thống nội tiết rất khác với con người. Do đó, các kết quả nghiên cứu trong chuột và thỏ hay các động vật khác không thể áp dụng cho con người được.

Bằng chứng có giá trị thấp vào hàng thứ ba là những **báo cáo về các trường hợp lâm sàng** (tức là những “case reports”). Trong nhiều tập san y học, bác sĩ và nhà khoa học thường báo cáo những trường hợp bệnh nhân cá biệt và “hấp dẫn”, tức là những trường hợp bệnh khó chẩn đoán, hay những trường hợp bệnh lí chưa được mô tả trong sách giáo khoa, những ca giải phẫu ngoạn mục, v.v... Giá trị của những trường hợp này chủ yếu là ở khía cạnh giáo dục y học.

Kết quả khảo cứu trong các trường hợp đó không có giá trị khái quát hóa đến một cộng đồng.

Những bằng chứng có giá trị khoa học “trung bình” là kết quả của những **nghiên cứu bệnh án và đối chứng** (còn gọi là “case-control studies”). Đây là một phương pháp nghiên cứu mà trong đó các nhà nghiên cứu so sánh một nhóm bệnh nhân với một căn bệnh cụ thể và một nhóm người bình thường, không bị bệnh (gọi chung là nhóm “control” hay “đối chứng”). Chẳng hạn như để nghiên cứu mối liên hệ giữa mật độ xương và ung thư vú, nhà nghiên cứu có thể chọn 50 phụ nữ bị ung thư vú, và cứ mỗi trường hợp ung thư vú, nhà nghiên cứu chọn một (hay nhiều hơn) phụ nữ không bị ung thư vú nhưng có cùng độ tuổi, cùng tuổi có kinh, cùng tuổi mãn kinh, tương đương các đặc tính lâm sàng, v.v... Sau đó, nhà nghiên cứu đo lường và so sánh mật độ xương của hai nhóm phụ nữ. Nếu có sự khác biệt về mật độ xương thì sự khác biệt đó là do ảnh hưởng của ung thư chứ không phải do ảnh hưởng của tuổi tác hay các đặc tính lâm sàng. Kết quả của các nghiên cứu loại này có độ tin cậy trung bình, và có giá trị chủ yếu trong việc xác định một (hay nhiều) yếu tố dẫn đến căn bệnh.

Trong các bằng chứng có giá trị trung bình này phải kể đến các kết quả được rút ra từ những **nghiên cứu thiết diện** (còn gọi là “cross-sectional studies”). Trong các nghiên cứu này, một nhóm đối tượng được chọn một cách ngẫu nhiên, dựa trên một số tiêu chí chung, ở một thời điểm xác định. Nhà nghiên cứu thu thập các dữ kiện lâm sàng hay yếu tố được cho là có khả năng gây ra căn bệnh. Sau đó, nhà nghiên cứu sẽ so sánh các yếu tố này giữa những người đã bị bệnh và những người chưa bị bệnh.

Cao hơn các nghiên cứu bệnh án và đối chứng hay nghiên cứu thiết diện là những **nghiên cứu viễn cảnh** (thường được gọi là “Prospective cohort studies”). Khác với nghiên cứu thiết diện, nghiên cứu viễn cảnh theo dõi đối tượng trong một thời gian tương đối dài, nhưng không chữa trị. Trong nghiên cứu viễn cảnh, các nhà nghiên cứu chọn một nhóm đối tượng (thường có qui mô hàng ngàn người) và đo lường các yếu tố lâm sàng (chẳng hạn như áp xuất máu, hay mức độ cholesterol) mà nhà nghiên cứu nghĩ là có thể liên quan đến căn bệnh (như bệnh tim chẳng hạn). Nhóm đối tượng này sẽ được theo dõi trong một thời gian dài, và trong thời gian đó nhà nghiên cứu quan sát tỉ lệ phát bệnh tim trên một ngàn cá thể (chẳng hạn). Tỉ lệ phát bệnh này sẽ được phân tích với các yếu tố lâm sàng đã được đo lường trước khi bệnh nhân mắc bệnh, và qua đó xác định mối liên hệ giữa yếu tố đó và căn bệnh.

Thử nghiệm lâm sàng đối chứng (hay “Controlled clinical trials”) là những nghiên cứu có giá trị khoa học cao hơn các nghiên cứu viễn cảnh một bậc. Đây là một loại nghiên cứu rất tốn kém, và là một thước đo chuẩn về sự hiệu nghiệm của các loại dược phẩm, hay thuật chữa trị. Trong nghiên cứu này, nhà nghiên cứu cẩn thận tuyển chọn một số bệnh nhân theo những tiêu chuẩn đã định trước; rồi ngẫu nhiên chia thành hai (hay nhiều hơn hai) nhóm: nhóm 1 sẽ không được điều trị (gọi là nhóm đối chứng); nhóm 2 được điều trị bằng một loại thuốc mà người ta nghĩ là sẽ có hiệu nghiệm. Nhóm hai này có thể được chia thành nhiều nhóm dựa theo liều lượng của thuốc. Hai nhóm bệnh nhân sẽ được theo dõi trong một thời gian, và trong thời gian đó, nhà nghiên cứu sẽ quan sát những biến chuyển sinh hóa liên quan bệnh tật. Sau khi thời gian theo dõi kết thúc, nhà nghiên cứu phân tích và so sánh mối tương quan giữa các thuốc (và liều lượng thuốc), các yếu tố sinh hóa, và quá trình bệnh lý giữa hai nhóm bệnh nhân để đi đến một kết luận về mức độ hiệu nghiệm của thuốc.

Chẳng hạn như để biết xem thuốc Alendronate có hiệu quả trong việc phòng chống gãy xương hay không, các nhà nghiên cứu chọn ra một nhóm bệnh nhân phụ nữ (có thể lên đến 1000 người) bị chứng loãng xương, ở độ tuổi 60 đến 70, chưa từng bị gãy xương, và không sống trong các nhà dưỡng lão, v.v... Bệnh nhân sẽ được phân chia một cách ngẫu nhiên thành hai nhóm: phân nửa sẽ được điều trị bằng thuốc alendronate, và phân nửa sẽ không được điều trị bằng thuốc nào cả. Các nhà nghiên cứu có thể theo dõi hai nhóm bệnh nhân trong vòng 2 năm (hay lâu hơn), và quan sát số bệnh nhân bị gãy xương trong thời gian theo dõi. Sau khi mãn thời hạn theo dõi, tỉ lệ gãy xương giữa hai nhóm sẽ được phân tích và so sánh để xác định mức độ mà thuốc alendronate có thể phòng chống gãy xương. Đây là một cách so sánh và xác định sự hiệu nghiệm của một thuật chữa trị khách quan nhất và khoa học nhất. Bất cứ một thuật chữa trị nào nếu chưa qua cách so sánh và thử nghiệm lâm sàng đối chứng này đều phải được xem là còn trong vòng nghi vấn.

Và cuối cùng, bằng chứng y học có giá trị cao nhất là những kết quả **phân tích khuynh tâm** (còn gọi là “meta-analysis”). Đây là một nghiên cứu trên nhiều nghiên cứu. Trong thực tế, đây là một sự tổng hợp và hệ thống hóa kết quả từ nhiều nghiên cứu khác nhau, như tổng kết các thử nghiệm lâm sàng đối chứng. Trong nghiên cứu này, nhà nghiên cứu ứng dụng các phương pháp thống kê phức tạp (gọi là phân tích khuynh tâm) để phân tích tất cả các kết quả từ các nghiên cứu độc lập trên thế giới, và qua đó đưa đến một kết luận về một vấn đề y học. Chẳng hạn như để xác định một cách dứt khoát xem ăn chay có đem lại lợi ích trong việc phòng ngừa bệnh tim, nhà nghiên cứu có thể thu thập tất cả các kết quả nghiên cứu

về ăn chay và bệnh tim trên thế giới, và sau đó là phân loại và phân tích kết quả để đi đến một kết luận tổng hợp. Vì tính tổng hợp và hệ thống hóa cao của phân tích khuynh tâm, nên kết quả của nghiên cứu này được xem là đáng tin cậy nhất trong tất cả các loại nghiên cứu y học.

Bốn tiêu chí để xác định một mối liên hệ

Xác định độ tin cậy của một nghiên cứu tuy quan trọng, nhưng đó chỉ mới là một khía cạnh, một bước đầu trong việc đánh giá mối liên hệ giữa một yếu tố và bệnh tật. Thực vậy, một vấn đề khác và quan trọng hơn là bằng cách nào để biết một mối liên hệ giữa một yếu tố và bệnh tật được báo cáo trong các tập san khoa học chuyên môn là thực hay giả. Đây là một vấn đề phức tạp, đòi hỏi nhiều cân nhắc, nhưng dùng các tiêu chí của dịch tễ học và thống kê học, một mối liên hệ chỉ có ý nghĩa nếu nó đáp ứng bốn điều kiện sau đây:

Thứ nhất là sự **nhất quán** giữa các cuộc nghiên cứu. Đây là một tiêu chí quan trọng trong hoạt động nghiên cứu khoa học. Mỗi công trình nghiên cứu y học chỉ dựa trên một số đối tượng thường được chọn một cách ngẫu nhiên (như một số bệnh nhân ung thư chẳng hạn), ở một địa điểm cụ thể và một thời gian nhất định. Nếu một nghiên cứu độc lập khác cũng dựa vào những đối tượng, cùng địa điểm, và cùng thời gian, cũng như cùng phương pháp mà cho ra kết quả tương tự thì đó là một dấu hiệu của sự nhất quán. Tuy nhiên, nếu cuộc nghiên cứu lặp lại trong một nhóm đối tượng khác ở một địa điểm khác hay thời gian theo dõi ngắn/dài hơn có thể cho ra kết quả khác nhau. Do đó, nếu mối liên hệ giữa một yếu tố và bệnh tật là thật thì kết quả của nhiều nghiên cứu độc lập với nhau phải nhất quán. Chẳng hạn như nếu có 100 nghiên cứu về dioxin và dị tật bẩm sinh, thì ít ra phải có 80% nghiên cứu cho thấy cùng kết quả thì mới có thể đi đến một kết luận.

Thứ hai là loại bỏ **yếu tố ngẫu nhiên**. Như đề cập trên, vì phần lớn – nếu không muốn nói là tất cả – nghiên cứu y khoa dựa vào chọn mẫu đối tượng (samples), chứ không nghiên cứu trên toàn bộ đối tượng. Chẳng hạn như để ước đoán tỉ lệ người bị bệnh tiểu đường trong một cộng đồng gồm 1000 người, nhà nghiên cứu có thể chọn một cách ngẫu nhiên gồm 200 người, và tìm hiểu xem trong số 200 người này có bao nhiêu người bị tiểu đường. Tỉ lệ này chỉ là một ước đoán số người bị tiểu đường cho cộng đồng đó. Vì ước đoán dựa vào mẫu, nên nó có thể thay đổi tùy theo số lượng mẫu được chọn: tỉ lệ ước đoán dựa vào một mẫu 200 người có thể khác với tỉ lệ ước đoán dựa vào một mẫu 100 người. Thực tế này có nghĩa là khi chọn mẫu (sampling) thì các yếu tố ngẫu nhiên có thể ảnh

hưởng để kết quả nghiên cứu, dù kết quả đó là so sánh hay ước tính, và loại bỏ yếu tố ngẫu nhiên là một vấn đề của thống kê học. Một mối liên hệ giữa một yếu tố và bệnh tật chỉ có ý nghĩa khoa học nếu các yếu tố ngẫu nhiên đã được loại trừ khỏi mối liên hệ đó.

Thứ ba là mối liên hệ phải tuân theo một **hàm số liều lượng**. Nếu một hóa chất có ảnh hưởng đến bệnh tật, thì mối liên hệ đó phải biến chuyển một cách có hệ thống với liều lượng của hóa chất. Chẳng hạn như khi chứng minh sự ảnh hưởng của dioxin đến xác suất sinh con trai, các nhà nghiên cứu Ý không những chỉ trình bày dữ kiện cho thấy những đàn ông bị nhiễm dioxin có xác suất sinh con trai thấp hơn những đàn ông không bị nhiễm dioxin, mà còn phải trình bày dữ kiện cho thấy nồng độ tích tụ dioxin trong máu càng cao càng có ảnh hưởng lớn đến tỷ lệ sinh con trai. Chẳng hạn như trong các người cha có nồng độ dioxin từ 61 tới 117 ppt, tỷ lệ sinh con trai là 47%, và những người có nồng độ dioxin từ 118 trở lên, tỷ lệ nam giới giảm xuống còn chỉ 40% [4].

Thứ tư là các cơ chế sinh học cơ bản của mối **liên hệ** phải được thiết lập. Chiều sâu về nghiên cứu y học có thể được chia làm ba tầng (từ thấp nhất đến cao nhất): nghiên cứu cơ bản, nghiên cứu lâm sàng, và nghiên cứu dịch tễ học. Nghiên cứu dịch tễ học dựa vào đối tượng là một cộng đồng dân số lớn; nghiên cứu lâm sàng dựa vào đối tượng là những bệnh nhân được tuyển chọn theo những tiêu chuẩn cụ thể; và nghiên cứu cơ bản là những phân tích ở mức độ tế bào, tìm ra cơ chế của bệnh tật. Chẳng hạn như nếu mối liên hệ giữa dioxin và dị tật bẩm sinh được xác định bằng nghiên cứu lâm sàng thì bước kế tiếp là phải xác định bằng khoa học cơ bản, có nghĩa là phải làm thí nghiệm để xác định cơ chế về mối liên hệ này, để biết chắc chắn rằng dioxin, mà không phải là một yếu tố khác, gây ra dị tật bẩm sinh. Trong nhiều trường hợp, nghiên cứu khoa học cơ bản phải đi trước nghiên cứu lâm sàng để xác định mối liên hệ giữa một độc chất và hậu quả bệnh tật. Tất cả những liên hệ giữa một nhân tố và bệnh tật được quan sát từ các nghiên cứu dịch tễ học và nghiên cứu lâm sàng sẽ không có ý nghĩa khoa học nếu các mối liên hệ này chưa được “xác nhận” bằng các kết quả của nghiên cứu cơ bản.

Do đó, trực diện với một bản tin về y học, người tiếp nhận cần phải đặt một số câu hỏi thông thường như: nghiên cứu này được tiến hành ở đâu, do ai làm, đối tượng trong nghiên cứu thuộc độ tuổi và giới tính nào, có bệnh gì, thời gian theo dõi bao lâu, các dữ kiện được phân tích như thế nào, cách diễn dịch có phù hợp với dữ kiện hay không, và quan trọng là nghiên cứu đã được công bố chưa, ở đâu. Tất cả các câu hỏi này sẽ giúp cho người tiếp nhận thông tin quyết định xem các

kết quả nghiên cứu có giá trị gì hay không, và nếu có thì chúng có áp dụng cho cá nhân mình hay không.

Một điều mà người tiếp nhận thông tin cần phải nhận thức rõ là mỗi con người chúng ta là một cá nhân độc đáo, một tác phẩm nghệ thuật, chứ không phải là một con số trung bình. Mặt khác, những con số khoa học trong các công trình nghiên cứu thường là những con số trung bình. Điều đó có nghĩa là kết quả của một nghiên cứu không thể nào ứng dụng trực tiếp cho mọi cá nhân trong cộng đồng.

Khoa học y khoa hiện đại ngày nay cực kỳ đa dạng về chất lượng, và có thể nói rằng phần lớn những nghiên cứu xuất hiện trên mặt báo chí và hệ thống truyền thông đại chúng có chất lượng không cao. Hàng tháng, hàng ngàn tập san y học công bố hàng vạn bài báo y học chuyên môn. Nếu chỉ giới hạn trong số 90 tập san lớn, hàng năm cũng có hơn 15 ngàn bài báo y học. Với một lượng thông tin khổng lồ và đa dạng như thế, người đọc cần phải có những kỹ năng xử lý thông tin cần thiết. Nhưng phân biệt những kết quả đáng tin cậy với những kết quả không đáng tin cậy không phải là một điều dễ dàng. Những tiêu chí trên đây hi vọng sẽ giúp cho những người không làm trong khoa học, hay không quen với y học, một phương tiện để phân biệt và tiếp nhận thông tin y khoa một cách tích cực hơn.

Ngày nay với cuộc cách mạng thông tin càng ngày càng đem lại cho bệnh nhân một cơ hội và cũng là một lợi ích mà họ chưa bao giờ có được trong vài thập niên trước đây: đó là sự tiếp cận với thông tin y học. Chỉ cần một computer nối vào liên mạng (internet), bất cứ ai cũng có thể tìm tất cả các tài liệu khoa học liên quan đến bất cứ bệnh gì qua một thư viện y khoa đại chúng vĩ đại do Viện Y tế Quốc gia Hoa Kỳ quản lý [5]. Qua tiếp cận với những thông tin y học, bệnh nhân (hay thân nhân bệnh nhân) có thể đóng vai trò chủ động cùng làm việc với bác sĩ mình để đi đến một quyết định tối ưu trong việc điều trị bệnh cho chính mình, thay vì tùy thuộc hoàn toàn vào sự định đoạt của người thầy thuốc như ngày xưa.

Chú thích:

[1] “Risks and benefits of estrogen plus progestin in healthy postmenopausal women”, [Tập san] Journal of the American Medical Association, số ra ngày 17 tháng Bảy, 2002; trang 321-333/

[2] “Woman removes own breast implant with razor blade,” Agence France Presse April 17, 1992.

[3] “A morbid fear of illness makes America trash good food and common sense,” LA Times March 10, 1989.

[4] “Paternal concentration of dioxin and sex ratio of offspring,” Tập san Lancet, số 355, ra ngày 27 tháng 5, 2000, trang 1858-1863.

[5] Địa chỉ trang nhà của Thư viện Y khoa Quốc gia Mỹ:
<http://www.ncbi.nlm.nih.gov>