



XÂY DỰNG BỘ CÔNG CỤ ĐÁNH GIÁ NĂNG LỰC SUY LUẬN THỐNG KÊ TRONG Y HỌC CỦA SINH VIÊN

Trần Thúy Hiền*, Hoàng Minh Vũ, Hoàng Thanh Hải

Trường Đại học Y – Dược, Đại học Huế, 6 Ngô Quyền, Thành phố Huế, Việt Nam

Lê Phước Sơn

Trường Đại học Luật, Đại học Huế, Đường Võ Văn Kiệt, Thành phố Huế, Việt Nam

Tác giả liên hệ: **Trần Thúy Hiền** <tthien@huemed-univ.edu.vn >

(Ngày nhận bài: 29-9-2022; Ngày chấp nhận đăng: 28-12-2022)

Tóm tắt. Trong chương trình đào tạo y khoa dựa trên năng lực, mục tiêu dạy học thống kê là hướng đến phát triển cho sinh viên y khoa năng lực suy luận thống kê đặc thù trong lĩnh vực này, do vậy đánh giá cũng phải hướng đến mục tiêu đó. Bài viết này có mục tiêu nghiên cứu vận dụng lý thuyết đánh giá trong giáo dục toán để xây dựng các thang đánh giá năng lực suy luận thống kê trong y học của sinh viên khi giải quyết vấn đề thực tế. Trên cơ sở đó, chúng tôi tiếp tục nghiên cứu thiết kế bộ công cụ đánh giá năng lực suy luận thống kê trong y học của sinh viên và tiến hành thực nghiệm. Phân tích thể hiện bài làm của sinh viên cho chúng tôi những đánh giá tổng quan về năng lực suy luận thống kê trong y học của sinh viên và bước đầu cho thấy tính hiệu quả của bộ công cụ trong việc đánh giá năng lực này của sinh viên.

Từ khóa. Đánh giá, suy luận thống kê, suy luận thống kê trong y học, sinh viên y khoa.

DESIGNING A TOOLKIT FOR ASSESSING STUDENTS' STATISTICAL REASONING IN MEDICINE

Tran Thuy Hien*, Hoang Minh Vu, Hoang Thanh Hai

University of Medicine and Pharmacy, Hue University, 6 Ngo Quyen St, Hue, Vietnam

Le Phuoc Son

University of Law, Hue University, Vo Van Kiet St., Hue, Viet Nam.

* Correspondence to **Tran Thuy Hien** <tthien@huemed-univ.edu.vn >

(Received: September 29, 2022; Accepted: December 28, 2022)

Abstract. In a competency-based medical training program, the goal of teaching medical statistics to students is to develop students' statistical reasoning competency. Therefore, assessment must also aim at achieving that goal. We have studied and applied assessment theory in mathematics education to build a framework for assessing students' statistical reasoning competency in medicine in solving real-life problems. We continue to study and design a toolkit to assess students' statistical reasoning ability and conduct experiments. Analysis of student work gives us an overview of medical students' statistical reasoning competency and initially shows the toolkit's effectiveness in assessing this capacity of medical students.

Keywords. Assessment, statistical reasoning, statistical reasoning competency in medicine, medical students.

1. Mở đầu

Đánh giá là yếu tố quan trọng trong quá trình dạy học, đánh giá, dưới nhiều hình thức, có thể tác động tích cực đối với quá trình dạy học [7]. Theo Guskey [2], đánh giá là một yếu tố quan trọng trong nỗ lực của chúng ta để cải thiện giáo dục. Tuy nhiên, nếu sử dụng đánh giá chỉ như là một phương tiện để xếp hạng các trường hay xếp loại sinh viên thì chúng ta sẽ bỏ lỡ những lợi ích mạnh mẽ nhất của đánh giá. Qua khảo sát về thực tế việc kiểm tra, đánh giá trong dạy học Thống kê y học (TKYH), chúng tôi nhận thấy việc đánh giá được sử dụng chủ yếu với mục đích để cho điểm, để xếp loại, để cung cấp thông tin phản hồi định kì về việc học tập của sinh viên (SV). Những bài tập trong bài kiểm tra, bài thi cuối học phần TKYH của SV dưới hình thức tự luận mới đơn thuần chỉ ở mức tái hiện, nhớ lại, thể hiện khả năng SV có thể thực hiện các tính toán và áp dụng các quy trình một cách chính xác, mà chưa thực sự phản ánh tốt cách SV tư duy, suy luận, áp dụng kiến thức và năng lực suy luận thống kê trong y học của SV [4]. Để sử dụng các đánh giá trong lớp học nhằm cải tiến việc dạy và học, giáo viên cần thay đổi cách tiếp cận của mình về đánh giá, đánh giá phải căn cứ vào mục tiêu, chương trình và các chuẩn mực cụ thể, không phụ thuộc vào chủ quan của người dạy [2].

Chương trình đào tạo y khoa theo hướng tích hợp, dựa trên năng lực đòi hỏi dạy học TKYH phải đổi mới nhằm nâng cao năng lực suy luận thống kê cho SV khi giải quyết các vấn đề trong thực hành nghề nghiệp. Việc nghiên cứu xây dựng bản đồ mục tiêu môn học, các thang đánh giá năng lực suy luận thống kê trong y học cụ thể sẽ là một căn cứ quan trọng giúp chúng tôi nghiên cứu xây dựng bộ công cụ đánh giá năng lực suy luận thống kê trong y học của SV. Mục đích nghiên cứu là đưa ra được đánh giá tổng quan về năng lực suy luận thống kê trong y học của SV và đề xuất được một bộ công cụ đánh giá hữu ích nhằm phát huy vai trò của đánh giá trong việc cải tiến chất lượng dạy và học TKYH cho SV ngành y.

2. Cơ sở lý thuyết

Dựa trên định nghĩa về năng lực của tổ chức Hợp tác và Phát triển kinh tế (OECD) [9], định nghĩa Hiểu biết toán của PISA [11], mô hình suy luận thống kê của Jones và cộng sự [6], Mooney [8] và quan điểm về quá trình phát triển tri thức khoa học của Riffenburgh [14], chúng tôi cho rằng có ba quá trình TKYH then chốt thể hiện năng lực suy luận thống kê, bao gồm quá trình Mô tả, Giải thích và Dự đoán. Theo đó, “Năng lực suy luận thống kê trong y học có thể xem là khả năng thực hiện một cách hiệu quả và khoa học các nhiệm vụ trong quá trình TKYH, bao gồm Mô tả, Giải thích, Dự đoán” [3].

Dựa trên phân loại tư duy Bloom sửa đổi (Anderson và cộng sự) [1], Thứ bậc nhiệm vụ đánh giá toán (phân loại MATH) (Smith và cộng sự) [15] kết hợp phân loại Hiểu biết toán trong chương trình đánh giá học sinh quốc tế (PISA) của OECD [10], chúng tôi đã xây dựng thang đánh giá tổng quát năng lực suy luận thống kê của SV y khoa [5]. Đánh giá lĩnh vực Hiểu biết toán của PISA [11] liên quan đến Nội dung toán học (các ý tưởng bao quát), Quá trình toán học (xác định bởi các năng lực toán học) và Bối cảnh trong đó toán học được sử dụng. Tính không chắc chắn là một trong bốn ý tưởng bao quát, được đề xuất hai chủ đề liên quan là dữ liệu và cơ hội, tương ứng với các chủ đề toán xác suất và thống kê. Như vậy, theo một khía cạnh nào đó thì Hiểu biết toán bao hàm năng lực suy luận thống kê. Dựa trên quan điểm này, chúng tôi cũng đã đề xuất một mô hình đánh giá năng lực suy luận thống kê trong y học của SV khi giải quyết vấn đề thực tế (Hình 1).



Hình 1. Mô hình đánh giá năng lực suy luận thống kê trong y học

Trong mô hình này, đánh giá liên quan đến ba khía cạnh là Nội dung TKYH (bao gồm các ý tưởng cơ bản được sử dụng để giải quyết vấn đề), Quá trình TKYH thể hiện năng lực suy luận thống kê trong y học (xác định bởi ba thành tố là khả năng Mô tả, Giải thích, Dự đoán) và Bối cảnh lâm sàng. Trong Bối cảnh thực tế liên quan đến lâm sàng y học gắn liền với cuộc sống

nghề nghiệp, SV phải đưa ra quyết định về kiến thức nào có thể liên quan, quy trình hay thủ tục nào sẽ dẫn đến giải pháp khả thi và sử dụng suy luận thống kê ở các mức độ khác nhau. *Nội dung* TKYH nên tập trung vào các ý tưởng lớn của TKYH bao gồm Dữ liệu; Phân phối; Xu hướng; Biến thiên; Mô hình; Mối liên quan; Mẫu và lấy mẫu; và Suy diễn thống kê. Dựa trên mô hình đánh giá và thang đánh giá tổng quát năng lực suy luận thống kê của SV y khoa, chúng tôi đã xây dựng thang đánh giá năng lực suy luận thống kê trong y học của SV khi giải quyết vấn đề thực tế. Năng lực này được mô tả theo ba cụm năng lực nhận thức như trong Bảng 1.

Bảng 1. Ma trận hai chiều về năng lực suy luận thống kê trong y học và các mức độ nhận thức

| Cụm năng lực nhận thức Năng lực suy luận thống kê trong y học | <i>Tái tạo</i> | | <i>Liên kết</i> | | <i>Phản ánh</i> | |
|---|----------------|-------|-----------------|-------|-----------------|-------|
| | Mức 1 | Mức 2 | Mức 3 | Mức 4 | Mức 5 | Mức 6 |
| <i>Mô tả</i> | – | – | – | – | – | – |
| <i>Giải thích</i> | – | – | – | – | – | – |
| <i>Dự đoán</i> | – | – | – | – | – | – |

Các mức năng lực nhận thức của mỗi thành tố của suy luận thống kê được mô tả trong ba thang đánh giá cụ thể [4].

3. Xây dựng bộ công cụ đánh giá năng lực suy luận thống kê trong y học

3.1. Thiết kế các vấn đề có bối cảnh lâm sàng y học

Thang đánh giá năng lực tổng quát và các năng lực thành phần như *Mô tả*, *Giải thích*, *Dự đoán* được mô tả cụ thể là một căn cứ quan trọng để chúng tôi xây dựng các câu hỏi đánh giá. Căn cứ tiếp theo để xây dựng các câu hỏi đánh giá là Bản đồ mục tiêu môn học, với 5 mục tiêu môn học và 29 mục tiêu tương ứng 7 bài giảng, trong đó liên quan đến TKYH có 21 mục tiêu. Theo thang đánh giá năng lực suy luận thống kê trong y học [4], mỗi mục tiêu được xác định tương ứng với một cụm năng lực suy luận thống kê. Dựa trên các thang đánh giá, bản đồ mục tiêu môn học và yêu cầu về mặt kĩ thuật khi biên soạn từng loại câu hỏi, chúng tôi đã xây dựng một bộ các câu hỏi gồm câu hỏi trắc nghiệm khách quan (TNKQ) và tự luận (Tlu) theo 14 vấn đề có bối cảnh lâm sàng y học. Chúng tôi sử dụng dạng câu hỏi nhiều lựa chọn (NLC), câu hỏi có trả lời đóng hay kết thúc đóng (TL đóng) và câu hỏi có trả lời mở hay kết thúc mở (TL mở).

3.2. Thiết kế ma trận để kiểm tra đánh giá năng lực suy luận thống kê trong y học

Nghiên cứu về cấu trúc đánh giá của PISA cho thấy đánh giá Hiểu biết toán của PISA thông qua một sự kết hợp các loại câu hỏi: câu hỏi NLC, câu hỏi TL đóng và câu hỏi TL mở. Số lượng câu hỏi của ba loại này gần như bằng nhau. Theo kinh nghiệm phát triển và sử dụng câu hỏi đánh giá cho các kì PISA, loại câu hỏi NLC nhìn chung được xem là phù hợp nhất để đánh

giá các cụm năng lực tái tạo và liên kết. Để đánh giá các mức độ tư duy cao hơn, những loại câu hỏi như câu hỏi TL đóng hay câu hỏi TL mở sẽ thường được ưa thích hơn. Nét chính của câu hỏi TL mở là chúng cho phép người học trình bày khả năng của mình bằng cách cung cấp các lời giải khác nhau ở nhiều mức độ khác nhau của tính phức tạp toán học. Để đảm bảo theo một trọng số nào đó được đặt ra (chẳng hạn trọng số điểm) thì tỉ lệ phần trăm (%) của các câu hỏi ở cụm *Tái tạo* : *Liên kết* : *Phản ánh* tương ứng khoảng là 25 : 50 : 25 là phù hợp [10]. Việc vận dụng tỉ lệ này của đánh giá PISA để xây dựng ma trận đề kiểm tra cũng tương đối phù hợp với phân bố tỉ lệ các mục tiêu bài giảng mà chúng tôi đã xây dựng cho dạy học TKYH.

Chúng tôi lựa chọn xây dựng ma trận đề kiểm tra với sự kết hợp cả hai hình thức TNKQ (bao gồm những câu hỏi NLC) và TLu (bao gồm những câu hỏi TL đóng và câu hỏi TL mở). Tỉ lệ phần trăm (%) về trọng số điểm của các câu hỏi ở cụm *Tái tạo* : *Liên kết* : *Phản ánh* tương ứng khoảng là 25 : 50 : 25. Chúng tôi thực hiện quy trình 5 bước như sau:

Bước 1. Xác định thời gian dành cho từng phần và trọng số điểm tương ứng. Theo đặc thù môn Toán, tỉ trọng điểm thích hợp giữa hình thức TNKQ và TLu nên là 3:7; 4:6 hoặc 5:5 [13]. Căn cứ trên thời gian kiểm tra để xác định thời gian tương ứng cho mỗi hình thức.

Bước 2. Liệt kê các mục tiêu cần kiểm tra (mục tiêu môn học, bài giảng), xác định năng lực suy luận thống kê cần đánh giá (đánh giá tổng quát hay năng lực cụ thể *Mô tả*, *Giải thích*, *Dự đoán*).

Bước 3. Xác định trọng số điểm tương ứng với mỗi mục tiêu, nội dung đánh giá (mục tiêu môn học, bài giảng). Việc xác định này căn cứ chủ yếu vào số tiết qui định trong phân phối chương trình và tầm quan trọng của mục tiêu trong chương trình. Trọng số điểm của các cấp độ tư duy của các câu hỏi ở cụm *Tái tạo* : *Liên kết* : *Phản ánh* tương ứng khoảng 2,5 : 5,0 : 2,5.

Bước 4. Xác định trọng số điểm của mỗi chuẩn (mỗi ô trong bảng là các chuẩn chương trình cần kiểm tra) và số lượng câu hỏi tương thích với trọng số điểm của mỗi chuẩn. Việc xác định này căn cứ nguyên tắc: mỗi câu hỏi TNKQ có trọng số như nhau và mỗi câu hỏi TLu có trọng số phù hợp với thời gian dự định hoàn thành nó và mức độ phức tạp về tư duy dự định đánh giá.

Bước 5. Đánh giá lại ma trận đề kiểm tra đã xây dựng và điều chỉnh nếu cần thiết.

Ví dụ 1. Chúng tôi đề xuất một ma trận đề kiểm tra để đánh giá thành tố *Mô tả* của năng lực suy luận thống kê trong y học khi SV học xong chủ đề lý thuyết mẫu (Ma trận 1, dùng trong đánh giá quá trình) và một ma trận đề kiểm tra để đánh giá năng lực suy luận thống kê khi kết thúc học phần TKYH (Ma trận 2, dùng trong đánh giá tổng kết). Bảng 2 là ma trận đề kiểm tra đánh giá năng lực suy luận thống kê trong y học (Ma trận 2). Tỉ trọng điểm giữa hình thức TNKQ và TLu là 4 : 6. Thời gian kiểm tra là 60 phút, tương ứng với 24 phút cho TNKQ (4 điểm) và 36 phút cho TLu (6 điểm). Bộ câu hỏi gồm 10 TNKQ và 5 TLu (4 TL đóng, 1 TL mở), trong đó

mỗi câu hỏi TNKQ đều có trọng số điểm như nhau là 0,4 và thời gian trung bình hoàn thành là khoảng 2,4 phút; câu hỏi TLu được đánh số phù hợp với thời gian dự định hoàn thành và mức độ phức tạp về tư duy dự định đánh giá, trong đó TLu có trọng số điểm là 1 tương ứng thời gian trung bình hoàn thành khoảng 6 phút. Trọng số điểm tương ứng với mỗi mục tiêu môn học từ 3 đến 5 là 2,6 : 5 : 2,4. Trọng số điểm tương ứng mỗi mức độ tư duy *Tái tạo* : *Liên kết* : *Phản ánh* là 2,4 : 5,1 : 2,5 nghĩa là 75% ở mức tái tạo và liên kết, 25% ở mức phản ánh, tỉ lệ này là phù hợp với tỉ lệ các mục tiêu đã đề ra.

Bảng 2. Ma trận đề kiểm tra đánh giá năng lực suy luận thống kê trong y học (Ma trận 2)

| MT môn học | <i>Tái tạo</i> | | <i>Liên kết</i> | | <i>Phản ánh</i> | | Tổng | | |
|------------|----------------|-----|-----------------|---------|-----------------|---------|-------------|---------|-------|
| | TNKQ | TLu | TNKQ | TLu | TNKQ | TLu | TNKQ | TLu | Điểm |
| MT 3 | 3 | | 1 | 1 (1) | | | 4 (1,6) | 1 (1,0) | (2,6) |
| MT 4 | 3 | | 2 | 1 (1,5) | | 1 (1,5) | 5 (2) | 2 (3) | (5) |
| MT 5 | | | 1 | 1 (1) | | 1 (1) | 1 (0,4) | 2 (2) | (2,4) |
| Tổng | 6 (2,4) | | 4 (1,6) | 3 (3,5) | | 2 (2,5) | 10 (4) | 5 (6) | (10) |
| | (2,4) | | (5,1) | | (2,5) | | | | |

(Ở mỗi ô: chữ số ở bên ngoài dấu ngoặc là số lượng câu hỏi, chữ số ở bên trong dấu ngoặc là trọng số điểm)

3.3. Thiết kế bộ câu hỏi đánh giá theo ma trận đề kiểm tra

Căn cứ trên ma trận đề kiểm tra để xác định số lượng câu hỏi và trọng số điểm tương ứng với mỗi mục tiêu bài giảng. Đây cũng là một căn cứ để xác định phạm vi đánh giá của mỗi câu hỏi, mức độ phức tạp và thời gian dự kiến hoàn thành câu hỏi. Câu hỏi được biên soạn tương ứng từng mục tiêu bài giảng, phù hợp với từng mức độ suy luận thống kê như đã xác định ở các thang đánh giá năng lực suy luận thống kê trong y học. Chẳng hạn, đối với ma trận 2, bộ câu hỏi của bài kiểm tra thứ hai (Test 2) được biên soạn tương ứng các mục tiêu bài giảng, các thành tố năng lực suy luận thống kê trong y học là *Mô tả*, *Giải thích*, *Dự đoán*, và theo một trọng số điểm. Bảng 3 là bộ 15 câu hỏi của bài kiểm tra Test 2 được chọn từ các câu hỏi đã biên soạn theo 14 vấn đề có bối cảnh lâm sàng y học.

Bảng 3. Mô tả bộ câu hỏi của bài kiểm tra Test 2

| Stt | VẤN ĐỀ | Câu hỏi tương ứng | Loại câu hỏi | Cụm năng lực | Mức SL TKYH | Điểm | MT bài giảng |
|-----|---|-------------------|--------------|--------------|-------------|------|--------------|
| 1 | Bệnh mạch vành | Câu hỏi 1 | MCQ | Tái tạo | 2 | 0,4 | MTBG 3.2 |
| 2 | Hàm lượng Protein | Câu hỏi 1 | MCQ | Tái tạo | 2 | 0,4 | MTBG 3.3 |
| 3 | | Câu hỏi 2 | MCQ | Liên kết | 3 | 0,4 | MTBG 3.5 |
| 4 | Nồng độ Hemoglobin | Câu hỏi 1 | MCQ | Tái tạo | 2 | 0,4 | MTBG 3.4 |
| 5 | Nồng độ men ALT | Câu hỏi 1 | TL đóng | Liên kết | 4 | 1,0 | MTBG 3.5 |
| 6 | Hàm lượng Cholesterol | Câu hỏi 1 | MCQ | Tái tạo | 2 | 0,4 | MTBG 4.1 |
| 7 | Nồng độ Beta-Crosslaps | Câu hỏi 1 | MCQ | Tái tạo | 2 | 0,4 | MTBG 4.2 |
| 8 | | Câu hỏi 2 | TL đóng | Liên kết | 4 | 1,5 | MTBG 5.2 |
| 9 | Điều trị cai nghiện ma túy | Câu hỏi 1 | MCQ | Liên kết | 3 | 0,4 | MTBG 6.4 |
| 10 | Phân bố nhóm máu | Câu hỏi 1 | MCQ | Tái tạo | 2 | 0,4 | MTBG 6.4 |
| 11 | | Câu hỏi 2 | MCQ | Liên kết | 3 | 0,4 | MTBG 6.4 |
| 12 | Bệnh ung thư tuyến tiền liệt | Câu hỏi 1 | TL đóng | Phản ánh | 5 | 1,5 | MTBG 6.4 |
| 13 | Bề dày lớp Nội trung mạc động mạch cảnh | Câu hỏi 1 | TL đóng | Liên kết | 4 | 1,0 | MTBG 7.2 |
| 14 | Liên quan giữa Glucose và Hormone | Câu hỏi 1 | MCQ | Liên kết | 3 | 0,4 | MTBG 7.3 |
| 15 | | Câu hỏi 3 | TL mở | Phản ánh | 5 | 1,0 | MTBG 7.4 |

Ví dụ 2. Câu hỏi 15 trong bài kiểm tra Test 2

Câu 15. (M5): Nghiên cứu sự phụ thuộc của hàm lượng Glucose (đường huyết) và Hormone (nội tiết tố). Hàm lượng Glucose (đơn vị mmol/L) và Hormone trong máu (đơn vị IU/mL) của một mẫu 35 bệnh nhân được xác định và lưu trong file Data.sav. Xử lý dữ liệu, kết quả của mô hình hồi qui tuyến tính thu được: $\text{Hormone} = -0,44 \cdot \text{Glucose} + 5,77$; $R^2 = 0,56$; test t của hệ số góc (độ dốc) có $\text{Sig.} = 10^{-4}$.

Từ kết quả thu được của mẫu số liệu, hãy viết một báo cáo trong đó đưa ra những nhận định có cơ sở cho nghiên cứu trên. Từ đó, đưa ra một ví dụ về dự đoán nồng độ Hormone của một bệnh nhân khi biết chỉ số đường huyết của bệnh nhân đó.

Phân tích. Loại câu hỏi: TL mở. Nội dung TKYH: Phân tích hồi qui tuyến tính đơn biến. Năng lực suy luận thống kê y học: *Dự đoán*.

Bối cảnh lâm sàng: Câu hỏi xây dựng trên bối cảnh nghiên cứu mối liên quan giữa hàm lượng Glucose và Hormone. Glucose máu là một chỉ số quan trọng giúp bác sĩ đánh giá khả năng kiểm soát đường huyết của cơ thể bệnh nhân, nhờ đó có thể xác định bệnh nhân có mắc các bệnh lý liên quan tới đường huyết (bệnh tiểu đường) hay không, cùng với đó còn có thể đánh giá xem bệnh nhân mắc bệnh tiểu đường có đáp ứng được các phương pháp điều trị đang áp dụng hay không. Hormone là những chất được sản xuất bởi các tuyến nội tiết có tác dụng to lớn đối với các quá trình của cơ thể. Chúng ảnh hưởng đến sự tăng trưởng và phát triển, tâm trạng, chức năng tình dục, sinh sản và trao đổi chất.

Cụm năng lực: *Phản ánh* – Mức 5. Câu hỏi trên được đưa ra nhằm đánh giá thành tố Dự đoán của năng lực suy luận thống kê trong y học của SV liên quan đến phân tích hồi qui tuyến tính đơn biến. Câu hỏi này phù hợp với mức 5 được mô tả trong thang đánh giá năng lực suy luận thống kê trong y học đã được chúng tôi xây dựng (thang đánh giá suy luận thống kê trong y học và thang đánh giá thành tố Phản ánh). Dựa vào những thông tin đã cho, viết báo cáo cho nghiên cứu này nghĩa là mô tả có lý giải một cách đầy đủ về mối liên hệ giữa hàm lượng glucose và hormone, SV cần có một sự hiểu biết sâu sắc về các khái niệm thống kê trong phân tích hồi qui tuyến tính đơn biến, phải biết tích hợp các khái niệm và áp dụng sự thấu hiểu suy luận để kiểm chứng, lý giải các thông tin và đưa ra những nhận định có cơ sở.

Khi viết báo cáo, sinh viên được yêu cầu đọc và giải thích trong một bối cảnh để xác định có tồn tại hay không mối liên hệ giữa glucose và yếu tố hormone, mức độ liên kết như thế nào. Mô hình hồi qui tuyến tính đơn biến có phù hợp để mô tả mối liên hệ giữa glucose và hormone, sự thay đổi của glucose có ý nghĩa giải thích cho nồng độ hormone hay không. Hàm hồi qui mẫu là một ước lượng của hàm hồi qui tổng thể từ mẫu dữ liệu, có ý nghĩa giải thích và tiên đoán như thế nào về mối liên hệ giữa hai biến. Sau khi đánh giá mức độ phù hợp của mô hình hồi qui mẫu thông qua kết quả của thủ tục kiểm định hệ số góc (độ dốc) và hệ số xác định R^2 , đưa ra được một ví dụ cụ thể minh họa về ý nghĩa của mô hình hồi qui cho phép dự đoán giá

trị của kết quả phát sinh từ một chỉ số lâm sàng cụ thể về biến glucose. Sử dụng suy luận một cách linh hoạt, thiết lập giải thích và giao tiếp có hiệu quả các kết quả của quá trình này.

Đây là những câu hỏi có trả lời mở và có nhiều cách trả lời khác nhau. Sinh viên thành công là SV biết khai thác đầy đủ ba thông tin đã cho: Với thông tin đã cho " $R^2 = 0,56$ ", đây là hệ số xác định của mô hình hồi qui, hệ số xác định đánh giá độ phù hợp của hàm hồi qui mẫu với tập dữ liệu, chỉ ra rằng trong 100% của toàn bộ các sai lệch của tập dữ liệu so với trung bình thì có 56% là do glucose giải thích, còn 44% là do sai số ngẫu nhiên và do các yếu tố khác (nếu có) mà ta không đưa vào trong mô hình này để xem xét. Với thông tin "test t của hệ số góc (độ dốc) có Sig. = 10^{-4} ", cho thấy độ dốc b khác 0 là có ý nghĩa thống kê ($p = 10^{-4} < 0,01$), nghĩa là có mối liên hệ giữa glucose và hormone, mô hình hồi qui tuyến tính là phù hợp để mô tả về mối liên hệ này. Với thông tin "Hormone = $-0,44 \cdot \text{Glucose} + 5,77$ " là hàm hồi qui mẫu, là một ước lượng của hàm hồi qui tổng thể, trong đó một ước lượng là $b = -0,44$ và $a = 5,77$, cho thấy mối tương quan nghịch chiều giữa glucose và hormone, trong một giới hạn phù hợp của hàm lượng glucose, khi glucose tăng thêm 1 (mmol/L) thì hormone giảm 0,44 (IU/mL). Hàm hồi qui mẫu cho phép dự đoán nồng độ hormone trong máu đối với bệnh nhân có hàm lượng glucose cụ thể, chẳng hạn đối với bệnh nhân có chỉ số đường huyết là 6,5 (mmol/L) thì dự đoán nồng độ hormone trong máu là 2,91 (IU/mL).

3.4. Thiết kế biểu điểm chấm bài kiểm tra

Biểu điểm chấm bài kiểm tra đã được xây dựng chi tiết ngay từ khi lập ma trận để kiểm tra. Căn cứ vào ma trận để kiểm tra, chúng tôi xác định thang điểm tương ứng các mức suy luận thống kê trong y học với thang điểm toàn bài kiểm tra. Bảng 4 mô tả thang điểm tương ứng các mức suy luận của ma trận 2.

Bảng 4. Thang điểm tương ứng mức Suy luận thống kê trong y học của ma trận 2

| Cụm | Mức suy luận thống kê trong y học | Mức điểm |
|-----------------|-----------------------------------|-----------|
| <i>Phản ánh</i> | 5, 6 | (7,5; 10] |
| <i>Liên kết</i> | 4 | (4; 7,5] |
| | 3 | (2,4; 4] |
| <i>Tái tạo</i> | 1, 2 | (0; 2,4] |

3.5. Thiết kế đề kiểm tra thực hành

Ngoài việc xây dựng các bài kiểm tra đánh giá năng lực suy luận thống kê trong y học của SV khi giải quyết vấn đề thực tế chỉ dùng giấy bút (như bài kiểm tra Test 1, Test 2), chúng tôi cũng đã xây dựng các bài kiểm tra đánh giá năng lực suy luận thống kê trong y học của SV thông qua việc các em thực hành xử lý, phân tích dữ liệu bằng phần mềm thống kê SPSS (Statistical Package for Social Sciences). Đối với bài kiểm tra này, các em được yêu cầu thực hành trên máy

tính với phần mềm SPSS và trình bày bài làm trên giấy. Bài Test thực hành nhằm khảo sát để đưa ra những đánh giá về năng lực suy luận thống kê trong y học của SV khi giải quyết vấn đề thực tế với sự hỗ trợ của phần mềm SPSS, bao gồm *Mô tả*, *Giải thích*, *Dự đoán* ở các mức tái tạo, liên kết và phân ánh.

4. Khảo sát đánh giá năng lực suy luận thống kê trong y học của sinh viên

Phương pháp khảo sát được chúng tôi sử dụng nhằm tìm hiểu năng lực suy luận thống kê của SV y khoa. Chúng tôi tiến hành thực nghiệm với bộ công cụ đánh giá đã xây dựng, thu thập kết quả bài làm của SV, phân tích kết quả và rút ra kết luận. Chúng tôi sử dụng phương pháp phân tích định lượng kết hợp với định tính. Thực nghiệm trên $N = 106$ SV y khoa năm thứ nhất, học kỳ 1, năm học 2020-2021, thuộc chương trình đào tạo y khoa dựa trên năng lực của trường Đại học Y - Dược, Đại học Huế. Những SV này vừa học xong những kiến thức liên quan chủ đề Lý thuyết mẫu trong học phần TKYH. Các em đã nắm bắt tốt các khái niệm, công thức TKYH liên quan biến ngẫu nhiên, định lý giới hạn trung tâm, tổng thể, mẫu ngẫu nhiên, lấy mẫu, các loại bảng, biểu đồ, các thống kê mẫu, một số phân phối thông dụng và phân phối lấy mẫu. Chúng tôi thu thập dữ liệu bài làm của SV đối với bài Test 1. Sau khi các em học xong phần thực hành TKYH, thu thập bài làm của các em đối với bài kiểm tra Test_thuchanh. Bài Test_thuchanh thực hiện tại phòng thực hành máy tính, trang thiết bị đảm bảo mỗi SV có một máy tính đã cài đặt sẵn phần mềm SPSS. Sau khi kết thúc học phần, thu thập dữ liệu bài làm của SV đối với bài Test 2. Kết quả thực nghiệm được phân tích định lượng bằng phần mềm SPSS (IBM SPSS Statistics 23).

5. Kết quả và bàn luận

Kết quả xử lý điểm bài kiểm tra Test 1, Test 2 cho chúng tôi xác định tỉ lệ phần trăm (%) SV đạt các mức suy luận thống kê. Trong phạm vi bài báo này, chúng tôi trình bày cách phân tích kết quả thực nghiệm đối với bài Test 2. Bảng 5 thể hiện tỉ lệ (%) SV đạt được các mức suy luận thống kê trong y học đối với bài Test 2.

Bảng 5. Tỉ lệ (%) sinh viên đạt các mức Suy luận thống kê trong y học đối với Test 2

| Cụm | Mức suy luận thống kê | Mức điểm | Tỉ lệ (%) ($N = 106$) |
|-----------------|-----------------------|-----------|-------------------------|
| <i>Phản ánh</i> | 5, 6 | (7,5; 10] | 13,2 |
| <i>Liên kết</i> | 4 | (4; 7,5] | 30,2 |
| | 3 | (2,4; 4] | 48,1 |
| <i>Tái tạo</i> | 1, 2 | (0; 2,4] | 8,5 |

Từ Bảng 5 cho thấy đối với SV được khảo sát, chỉ có 8,5% SV đạt được mức năng lực thấp là mức tái tạo và 56,6% sinh viên đạt được các mức 1, 2, 3. Sinh viên chủ yếu đạt năng lực ở các

mức 3, 4 (liên kết) là 78,3%, các mức cao 5, 6 (phản ánh) vẫn còn thấp, chỉ có 13,2%. Mức độ suy luận thống kê y học của SV là tương đối đồng đều, đa số các em đều được trang bị tốt các kỹ năng thống kê cơ bản. Đạt được mức năng lực tái tạo bao gồm truy xuất kiến thức, truy hồi lại công thức, sử dụng các quy trình quen thuộc là không khó khăn với hầu hết SV. Tuy nhiên, SV khó để có thể đạt được mức năng lực phản ánh, các mức năng lực cao đòi hỏi SV phải biết vận dụng sáng tạo các kiến thức đã biết để làm nên cái mới, không quen thuộc với những dạng đã biết hay vận dụng suy luận thống kê để giải quyết các vấn đề thực tế y học. Để thấy rõ hơn về năng lực suy luận thống kê y học của SV, chúng tôi phân tích câu trả lời của SV đối với từng câu hỏi và nhận thấy đa phần SV có câu trả lời tốt đối với những câu hỏi ở mức tái tạo, đối với 6 câu hỏi NLC ở mức tái tạo (Câu 1, 2, 4, 6, 7, 10) trong Test 2, thống kê được tỉ lệ (%) SV trả lời đúng với mỗi câu hỏi ở mức này đều đạt từ 86,5% trở lên. Tuy nhiên, đối với hai câu hỏi ở mức phản ánh, tỉ lệ SV đạt điểm tối đa là 8,5% đối với câu 12 và chỉ có 1,9% đối với câu hỏi 15.

Ví dụ 3. Xem xét câu trả lời của SV đối với câu hỏi 15 trong bài kiểm tra Test 2. Thống kê kết quả trả lời của SV đối với câu hỏi này được thể hiện trong Bảng 6.

Bảng 6. Kết quả trả lời tương ứng câu hỏi 15 của Test 2

| | Số điểm | | |
|---------------------|---------|------|-----|
| | 0 | 0,5 | 1,0 |
| Tỉ lệ (%) (N = 106) | 79,2 | 18,9 | 1,9 |

Kết quả bài làm của SV cho thấy, chỉ có 2 SV thành công chiếm tỉ lệ 1,9%. Hầu hết SV thất bại ở câu hỏi này, có 79,2% SV bỏ trống, hoặc có trả lời nhưng không đúng phần nào của câu hỏi, hoặc đưa ra được ví dụ về dự đoán nồng độ hormone với bệnh nhân có chỉ số đường huyết cụ thể nhưng trước đó không giải thích được mô hình hồi qui tuyến tính là phù hợp để mô tả mối liên hệ giữa glucose và hormone. Ngoài 2 SV đạt điểm tối đa cho câu hỏi này, còn lại hầu hết SV mới chỉ dừng lại ở mức tính toán, cho một giá trị glucose cụ thể và thay vào hàm hồi qui mẫu để xác định nồng độ hormone; tuy nhiên không có SV nào đưa ra được kết luận và giải thích đúng dựa vào thông tin “test t của hệ số góc có Sig. = 10^{-4} ”. Hình 2, Hình 3 là bài làm của SV1, SV2 đối với câu hỏi này.

Cho glucose = 5 (mmol/l) thì
 Hormone = $0,44 \times 5 + 5,77$
 = 3,57 (IU/L)

Hình 2. Trả lời của SV1 đối với câu hỏi 15 của Test 1

SV1 đưa ra được ví dụ về dự đoán nồng độ hormone với bệnh nhân có chỉ số đường huyết cụ thể nhưng trước đó chưa xét xem mô hình hồi qui tuyến tính có phù hợp để mô tả mối

liên hệ giữa glucose và hormone hay không, SV1 mới chỉ dừng lại ở mức tính toán, chưa thực hiện suy luận thống kê, chưa đạt được mức suy luận thống kê trong y học.

Do $R^2 = 0,56$; $sig = 10^{-4} < 0,05$ nên Hormone
và Glucose có mối liên quan với nhau, trong
đó Glucose có ý nghĩa giải thích được 56% cho
Hormone
Giá số 1 bệnh nhân có Glucose = 6 (mmol/l).
Thì dự đoán Hormone của bệnh nhân này là
 $Hormone = -0,44 \times 6 + 5,77 = 3,13$ (IU/ml)

Hình 3. Trả lời của SV2 đối với câu hỏi 15 của Test 1

SV2 đưa ra được ví dụ về dự đoán về nồng độ hormone với bệnh nhân có chỉ số đường huyết cụ thể, trước đó có xem xét và đánh giá có mối liên quan giữa glucose và hormone dựa vào hệ số xác định $R^2 = 0,56$, tuy nhiên chưa đưa ra giải thích dựa vào thông tin “test t của hệ số góc có $Sig. = 10^{-4}$ ”, chưa giải thích đầy đủ để kết luận mô hình hồi qui tuyến tính là phù hợp để mô tả mối liên hệ giữa glucose và hormone. SV2 đạt được mức 3 thuộc cụm năng lực liên kết theo thang đánh giá năng lực Dự đoán.

Chúng tôi đã tiến hành thực nghiệm và đánh giá thực nghiệm dựa trên bộ công cụ đánh giá năng lực suy luận thống kê trong y học đã xây dựng (bài kiểm tra Test 1, Test 2, Test_thuchanh). Qua phân tích kết quả bài làm của SV được khảo sát, chúng tôi có được những đánh giá tổng quan về năng lực suy luận thống kê trong y học của SV khi giải quyết vấn đề thực tế: SV chủ yếu đạt năng lực ở mức độ 1, 2, 3. Tỷ lệ sinh viên đạt năng lực ở mức độ cao 4, 5, 6 vẫn còn rất thấp. Nghĩa là, đa phần các em đã đạt được kết quả tốt về các kỹ năng, hiểu biết thống kê cơ bản (thể hiện ở mức độ 1, 2, 3), đó nền tảng cho việc phát triển tư duy thống kê trong y học nhưng rõ ràng việc dạy học TKYH vẫn chưa phát huy được tiềm năng đó. Sinh viên mới chỉ giải quyết được những vấn đề đã được mô hình hóa toán học một phần hay toàn bộ. Khi gặp những vấn đề có lời văn theo bối cảnh thực tế y học, nhiều sinh viên đã không chuyển thể thành công các mô hình toán để giải, đặc biệt gặp khó khăn với những câu hỏi TL mở, khả năng sáng tạo nên các phương án hiệu quả để giải quyết vấn đề mới lạ còn hạn chế, thiếu linh hoạt trong việc lên phương án để giải quyết vấn đề thực tế y học không quen thuộc.

Kết luận

Chúng tôi đã đề xuất việc thiết kế bộ công cụ đánh giá năng lực suy luận thống kê trong y học của SV, bộ công cụ đánh giá đã giúp chúng tôi đưa ra được những đánh giá tổng quan về năng lực suy luận thống kê trong y học của SV được khảo sát, các em có nền tảng tốt về kỹ năng toán học cơ bản, hiểu biết thống kê cơ bản, đạt mức cao về các kỹ năng liên quan đến áp dụng các quy tắc, quy trình tính toán thống kê, đó là nền tảng cho việc phát triển tư duy thống kê trong y học. Điều

đó có nghĩa dạy học TKYH cần đổi mới để phát huy tiềm năng đó của SV, để thúc đẩy SV sử dụng các kỹ năng thống kê cơ bản trong giải quyết vấn đề thực tế y học yêu cầu suy luận thống kê ở mức cao. Dạy học TKYH ở các trường ĐHY – Dược phải có sự đổi mới thực sự và đồng bộ, từ mục tiêu, chương trình, nội dung, giáo trình, phương pháp dạy học và phương pháp đánh giá. Đổi mới theo hướng tích hợp, tăng cường sự lồng ghép, nối kết TKYH với y học cơ sở, lâm sàng và nghiên cứu y học. Dạy học TKYH không nên chỉ dừng lại ở việc SV nhận biết, thông hiểu các khái niệm hay áp dụng các thủ tục, quy trình và thực hiện các tính toán thống kê, mà quan trọng hơn là SV phải ứng dụng được kiến thức TKYH để giải quyết vấn đề thực tế trong thực hành nghề nghiệp.

TÀI LIỆU THAM KHẢO

1. Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. *New York: Longman*, 352.
2. Guskey, T. R. (2003). How classroom assessments can improve learning, *Educational Leadership*, 60(5).
3. Trần Thúy Hiền (2019). Đánh giá năng lực suy luận thống kê y học của sinh viên y khoa từ quan điểm đào tạo vì cuộc sống nghề nghiệp. *Tạp chí Khoa học, Trường Đại học Sư phạm TP Hồ Chí Minh*, 11(16), tr. 309–322.
4. Trần Thúy Hiền (2022). Luận án Tiến sĩ *Đánh giá năng lực suy luận thống kê y học của sinh viên y khoa khi giải quyết vấn đề thực tế*. Trường Đại học Sư phạm TP Hồ Chí Minh.
5. Trần Thúy Hiền, Lê Phước Sơn (2019). Vận dụng phân loại tư duy Bloom và phân loại tư duy MATH để đánh giá mức độ suy luận thống kê y học của sinh viên ngành y. *Tạp chí Giáo dục, Bộ Giáo dục và Đào tạo*, Số 447, kì 1, 2/2019, tr. 43–49.
6. Jones, G. A., Langrall, C. W., Thornton, C. A., Mooney, E. S., Wares, A., Jones, M. R., Perry, B., Putt, I. J., & Nisbet, S. (2001). Using students' statistical thinking to inform instruction. *Journal of Mathematical Behavior*, 20(1), pp. 109–144.
7. Lewy, A. (1990). *Formative and Summative Evaluation*, Publisher at Pergamon Press.
8. Mooney, E. S. (2002). A Framework for Characterizing Middle School Students' Statistical Thinking. *Mathematical Thinking and Learning*, 4(1), pp. 23–63.
9. OECD. (2002). Definition and Selection of Competencies (DeSeCo): Theoretical and Conceptual Foundation. *OECD, Paris, France*.
10. OECD. (2009). The PISA 2009 Assessment Framework–Key competencies in reading, mathematics and science. *OECD, Paris, France*.

11. OECD. (2009b). Learning Mathematics for Life – a view perspective from PISA. *OECD, Paris, France*.
12. Nguyễn Thị Lan Phương (2007). *Đánh giá và thẩm định trong dạy học toán*. Bài giảng dành cho học viên cao học, Viện chiến lược và Chương trình giáo dục, Hà Nội.
13. Nguyễn Thị Lan Phương (2016). *Đánh giá và thẩm định trong dạy học toán*. Bài giảng dành cho học viên cao học, Viện chiến lược và Chương trình giáo dục, Hà Nội.
14. Riffenburgh, R. (2012). *Statistics in Medicine*. Academic Press.
15. Smith, G., Wood, L., Coupland, M., Stephenson, B., Crawford, K. & Ball, G. (1996). Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal of Mathematical Education in Science and Technology* 27(1), pp. 65–77.
16. Susan, M. C., & David, B. S. (2001). *Constructing Written Test Questions For the Basic and Clinical Sciences* (Third edition). National Board of Medical Examiners, Printed in the United States of America.
17. Trần Vui (2014). *Giải quyết vấn đề thực tế trong dạy học toán*. Nxb. Đại học Huế.