

DOI: 10.59715/pntjimp.1.2.8

## Lựa chọn phương pháp lượng giá phù hợp cho đầu ra học tập mong đợi (ELO) trong giáo dục Y khoa

Nguyễn Đức Quỳnh Vy<sup>1</sup>, Nguyễn Dũng Tuấn<sup>2</sup>

<sup>1</sup>Phòng Khảo thí, Trường Đại học Y khoa Phạm Ngọc Thạch, Thành phố Hồ Chí Minh.

<sup>2</sup>Phòng Khảo thí, Khoa Khoa học cơ bản - Y học cơ sở, Trường Đại học Y khoa Phạm Ngọc Thạch, TP. Hồ Chí Minh.

### Tóm tắt

Theo xu hướng đổi mới giáo dục y khoa thế giới và quy định của pháp luật hiện hành, chương trình đào tạo y khoa tại Việt Nam phải công bố chuẩn đầu ra của chương trình đào tạo. Đo lường kết quả học tập hay còn gọi là lượng giá để xác định việc người học đạt được chuẩn đầu ra là bắt buộc trong quá trình xây dựng và thực hiện chương trình đào tạo theo chuẩn đầu ra. Đầu ra của đào tạo y khoa là những năng lực đặc trưng nên cần những công cụ lượng giá phù hợp. Việc lựa chọn phương pháp lượng giá cần xét đến tính giá trị, tính tin cậy, tính khả thi của phương pháp. Trắc nghiệm nhiều lựa chọn là phương pháp được sử dụng phổ biến vì có sự khách quan và tính khả thi, tuy nhiên phương pháp này không có tính giá trị đối với việc lượng giá các chuẩn đầu ra lĩnh vực kỹ năng. OSCE là phương pháp có nhiều lợi thế khi lượng giá kỹ năng y khoa nhưng hạn chế khi lượng giá năng lực của người học trong bối cảnh thực tế, các phương pháp lượng giá tại nơi làm việc (Workplace Based Assessment) bổ khuyết cho OSCE ở khía cạnh này. Bài viết trình bày một cách tổng quan về các nguyên tắc và các phương pháp lượng giá phù hợp với giáo dục y khoa.

**Từ khóa:** Chuẩn đầu ra học tập; Lượng giá; Giáo dục y khoa.

### Abstract

#### A literature review on the assessment methods in outcome - based medical education

Outcome - based Education (OBE) has become regulated in Vietnam. OBE represents a clear and effective option for transforming medical education. Assessment in Outcome - based medical education gives the evidence that the learning objectives were achieved. Different characteristics of assessment tools need to be considered: validity, reliability, educational impact, feasibility. The major challenge in the choice of assessment method is the lack of comprehension about effective use of different assessment methods. Therefore, the basic aim of this review is to highlight the principle of assessment in medical education. The review also addresses the essential characteristics of various assessment tools that have been used in medical education. The data was collected from cross sectional studies, review articles, books on medical education, guidelines for assessment. Medline, Pubmed, Google Scholar were used to extract the data.

**Keywords:** Assessment Tools, Medical Education, Outcome Based Education

**Ngày nhận bài:**

15/02/2022

**Ngày phân biện:**

20/3/2022

**Ngày đăng bài:**

20/4/2022

**Tác giả liên hệ:**

Nguyễn Đức Quỳnh Vy

**Email:**

bsquynhvy@pnt.edu.vn

**ĐT:** 0937667627

## 1. ĐẶT VẤN ĐỀ

Giáo dục đại học theo chuẩn đầu ra là xu hướng đổi mới của giáo dục y khoa trên thế giới [1] và tại VN được thể hiện qua các văn bản của Bộ Giáo dục và đào tạo. Các trường đào tạo y khoa mặc dù là ngành đặc thù cũng không thể đặt mình bên ngoài xu hướng và các quy chuẩn pháp luật. Về phía đơn vị quản lý chuyên môn, Bộ Y tế đã ban hành văn bản quy định chuẩn năng lực nghề nghiệp để làm căn cứ tham khảo cho các trường đào tạo khối ngành sức khỏe tham khảo xây dựng chương trình của mình [2], các chương trình đào tạo bác sĩ y khoa qua đó đã công bố chuẩn đầu ra chương trình đào tạo.

Lợi ích của giáo dục theo chuẩn đầu ra trong giáo dục y khoa là: (1) Kiểm soát chất lượng; (2) Tạo ra một khung để lượng giá năng lực sinh viên, người được đào tạo trong các chuyên khoa, và bác sĩ; (3) Trao quyền cho sinh viên và giảng viên; (4) Tạo ra các ưu tiên để tránh việc quá tải thông tin; (5) Hỗ trợ cho sự hợp tác giữa các bậc đào tạo và các ngành đào tạo [1]. Việc xây dựng chương trình giảng dạy y khoa theo truyền thống thường là cách tiếp cận kế hoạch xuôi chiều thời gian bằng việc xây dựng nội dung khóa học đi trước và các phương pháp dạy - học tiếp theo. Tuy nhiên, để đạt được những lợi ích đã kể ra của chương trình đào tạo theo chuẩn đầu ra, cần nỗ lực rất nhiều trong đó cần cách tiếp cận “khác truyền thống” trong quá trình xây dựng chương trình đào tạo như: xây dựng một kế hoạch ngược dòng từ những kết quả định trước, kết quả định trước này phải căn cứ theo yêu cầu của các bên liên quan đến sản phẩm đào tạo, đổi mới phương pháp lượng giá theo hướng lượng giá năng lực, đổi mới phương pháp dạy - học hướng về việc học là trung tâm. Bắt đầu bằng việc mô tả những gì người học làm được vào lúc kết thúc chương trình đào tạo, người dạy sẽ chọn lựa các phương pháp đo lường kết quả học tập cần thiết và thiết kế các hoạt động dạy - học để giúp người học đạt được kết quả đó. Hoạt động đo lường đánh giá kết quả học tập đóng vai trò quan trọng để đảm bảo đạt được các mục đích của chương trình đào tạo theo chuẩn đầu ra, do đó các bên liên quan trong

chương trình đào tạo cần sử dụng các công cụ lượng giá một cách phù hợp, hiệu quả. Đối với giảng viên y khoa, khó khăn trong việc lựa chọn và sử dụng các phương pháp lượng giá là do họ không được trang bị đầy đủ các kiến thức cơ bản về lượng giá một cách có hệ thống cùng những hiểu biết về cách áp dụng các phương pháp lượng giá cụ thể vào công việc đào tạo hàng ngày.

Mục tiêu: Bài tổng quan này trình bày cơ sở lý luận về đo lường đầu ra học tập và đặc điểm của các phương pháp lượng giá được sử dụng trong giáo dục y khoa dựa trên các bằng chứng hiện có.

Dữ liệu được thu thập từ các nghiên cứu cắt ngang, bài báo tổng quan, sách chuyên khảo về giáo dục y khoa và hướng dẫn về lượng giá trong khoảng thời gian từ 1960 đến 2020. Sử dụng từ khóa: Assessment Tools, Medical Education, Outcome Based Education.

## TỔNG QUAN

### 1. CHUẨN ĐẦU RA/ ĐẦU RA HỌC TẬP MONG ĐỢI VÀ YÊU CẦU LIÊN KẾT CÁC THÀNH TỐ TRONG CHƯƠNG TRÌNH ĐÀO TẠO THEO CHUẨN ĐẦU RA.

#### 1.1 Chuẩn đầu ra/ Đầu ra học tập mong đợi

Chuẩn đầu ra là quy định về nội dung kiến thức chuyên môn; kỹ năng thực hành, nhận thức công nghệ và giải quyết vấn đề; công việc mà người học có thể đảm nhận sau khi tốt nghiệp và các yêu cầu đặc thù khác đối với từng trình độ, ngành đào tạo (Bộ Giáo dục và đào tạo, 2010) [3]. Trong một văn bản khác của Bộ Giáo dục và đào tạo ban hành năm 2015, chuẩn đầu ra là yêu cầu tối thiểu về kiến thức, kỹ năng, thái độ, trách nhiệm nghề nghiệp mà người học đạt được sau khi hoàn thành chương trình đào tạo, được cơ sở đào tạo cam kết với người học, xã hội và công bố công khai cùng với các điều kiện đảm bảo thực hiện [4].

Những danh từ tương đương với Chuẩn đầu ra được sử dụng trong văn bản của các cơ sở giáo dục là: Đầu ra học tập mong đợi/ dự kiến (Expected Learning Outcomes, Program Learning Outcomes, Student Learning

Outcomes). Chuẩn đầu ra cũng được phân chia thành các cấp độ: Cấp trường/viện (Institutional Learning Outcomes), Cấp chương trình đào tạo (Program Learning Outcomes), Cấp môn học/khóa học (Course Learning Outcomes), Bài học (Lesson Learning Outcome)

Trong các văn bản cấp cơ sở giáo dục, nhìn chung chuẩn đầu ra của chương trình đào tạo nhóm ngành khoa học sức khỏe được trình bày theo 2 hình thức: (1) các nhóm năng lực [5]; (2) các mục tiêu về kiến thức, kỹ năng, thái độ [6], [7].

Dù được trình bày theo hình thức khác nhau, chuẩn đầu ra có chung những ý nghĩa như sau [8]:

- Đối với người dạy: đây là các kiến thức, kỹ năng và thái độ mà nhà trường cần trang bị cho người học và mong đợi người học đạt được

- Đối với người học: đây là các kiến thức, kỹ năng và thái độ mà người học cần đạt được khi tốt nghiệp.

- Đối với nhà tuyển dụng và xã hội: chuẩn đầu ra là sự cam kết về chất lượng đạt được của người tốt nghiệp chương trình đào tạo, là tuyên bố trách nhiệm về chất lượng giáo dục đào tạo của nhà trường để xã hội giám sát, phản biện.

### **1.2. Khái niệm và yêu cầu liên kết các thành tố trong chương trình đào tạo theo chuẩn đầu ra.**

3 thành tố cơ bản trong quá trình dạy học phải được mô tả đầy đủ và tổ chức hợp lý trong chương trình đào tạo là: (1) Chuẩn đầu ra; (2) Phương pháp lượng giá; (3) Phương pháp dạy - học. John Biggs là tác giả nghiên cứu và đề xuất khái niệm “Thiết kế đồng bộ” (Constructive Alignment) giữa 3 thành tố trên [9]. “Thiết kế đồng bộ” cung cấp một khung hướng dẫn để điều chỉnh việc giảng dạy và lượng giá nhằm giải quyết việc đạt được các chuẩn đầu ra học tập [10]. Các hoạt động mà “Thiết kế đồng bộ” mô tả theo trình tự thời gian bao gồm:

1) Mô tả các kết quả học tập dự kiến cho đơn vị sử dụng.

2) Tạo môi trường học tập bằng cách thiết kế các hoạt động dạy/học tương ứng với kết quả học tập mong muốn.

3) Lựa chọn phương pháp lượng giá phù hợp để cung cấp thông tin về mức độ người học đáp

ứng chuẩn đầu ra, mức độ này được so sánh với một số mốc nhất định.

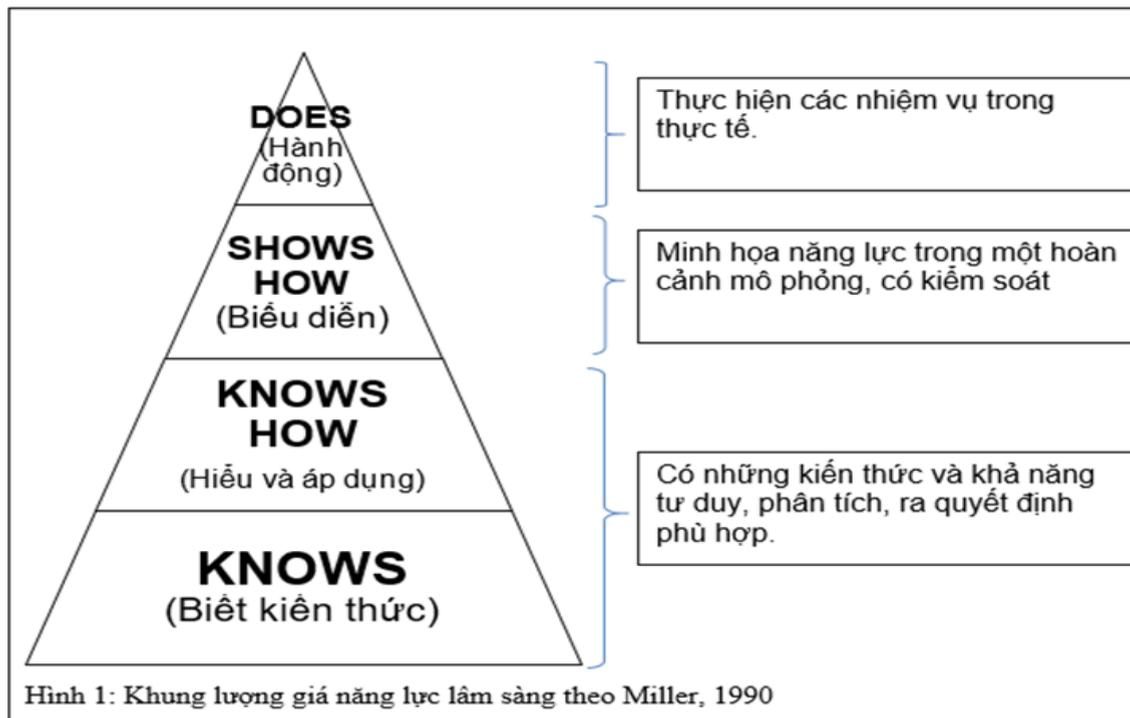
4) Chuyển những kết quả lượng giá này thành điểm cuối cùng.

Như vậy lượng giá vừa có tác dụng là thước đo kết quả học tập của người học và so sánh với chuẩn đầu ra, vừa có tác dụng thông tin cho người dạy và người học về hiệu quả của hoạt động dạy - học phục vụ việc điều chỉnh phương pháp dạy - học. Vleuten (1996) đã đề nghị việc lựa chọn phương pháp lượng giá phải tham khảo các đặc tính nội tại của phương pháp là: tính giá trị, tính tin cậy, tác động giáo dục, sự chấp nhận của các bên liên quan, và chi phí [11].

## **2. LƯỢNG GIÁ NĂNG LỰC TRONG BỐI CẢNH GIÁO DỤC Y KHOA**

Tác giả Miller đề xuất một hệ thống phân bậc lượng giá năng lực đến nay vẫn được ứng dụng rộng khắp trong giáo dục y khoa trên toàn thế giới. Hệ thống gồm 4 bậc, mô tả sự phát triển của năng lực trên cả 3 lĩnh vực: kiến thức, kỹ năng, thái độ và đề xuất các phương pháp phù hợp để lượng giá người học ở mỗi bậc [12].

Ở tầng thấp nhất của hình tháp là hai bậc phát triển năng lực về nhận thức bao gồm kiến thức, và cách sử dụng kiến thức (Knows; Know Hows) được tác giả đề nghị dùng các phương pháp lượng giá bằng văn bản (Written Assessment) để đo lường kiến thức của người học, với các câu hỏi mức độ ghi nhớ cho bậc Kiến thức - Knows; và mức độ giải quyết vấn đề cho bậc Biết Cách - Know Hows. Tầng thứ ba mô tả người học có thể minh họa cách giải quyết vấn đề y khoa trong một môi trường mô phỏng có kiểm soát (Show Hows), phương pháp lượng giá người học ở mức năng lực này thiên về các phương pháp thực hành, trình diễn (Performance Assessment). Tầng cao nhất mô tả năng lực của người học có thể bước vào giai đoạn hành nghề vì phải ứng dụng toàn diện kiến thức, kỹ năng, thái độ với một tỷ lệ hợp lý để giải quyết các tình huống thực tế do đó các phương pháp lượng giá được đề xuất là lượng giá dựa trên môi trường làm việc (Workplace Based Assessment).



### 3. MÔ TẢ CÁC PHƯƠNG PHÁP LƯỢNG GIÁ VÀ CÁCH ÁP DỤNG TRONG GIÁO DỤC Y KHOA

Thử thách đối với các nhà giáo dục y khoa là lựa chọn phương pháp lượng giá cho chương trình đào tạo, dù đã có những tiêu chuẩn được liệt kê ở các phần trên. Khó khăn thường do việc thiếu thông tin về những lựa chọn có thể và ưu khuyết của từng phương pháp lượng giá đang được sử dụng trong đào tạo y khoa. Lịch sử giáo dục y khoa trải dài từ truyền thống sử dụng các bài luận dài và đánh giá người học một cách chủ quan không được tiêu chuẩn hoá đến các phương pháp khách quan và tiêu chuẩn hóa như Trắc nghiệm khách quan nhiều lựa chọn Multiple Choice Question (MCQ); Thi lâm sàng cấu trúc theo mục tiêu Objective Structure Clinical Examination (OSCE)... Sau đây là các phương pháp được sử dụng rộng rãi bởi các hội đồng đánh giá quốc gia, các trường y khoa trên thế giới cùng các khuyến cáo về cách sử dụng.

Lượng giá kiến thức và áp dụng kiến thức:

#### 3.1. Trắc nghiệm khách quan nhiều lựa chọn

Câu trắc nghiệm nhiều lựa chọn được sử dụng rộng rãi trong giáo dục y khoa [13]. Câu trắc nghiệm nhiều lựa chọn là một phương pháp lượng giá khách quan, trong đó người trả lời được yêu cầu chọn một câu trả lời đúng (single best answer) trong một danh sách ít nhất

là 3 lựa chọn. Trong quá khứ, Hội đồng đánh giá y khoa Hoa Kỳ (National Board of Medical Examiners) phân loại các hình thức câu trắc nghiệm thành các phụ nhóm như Type A, Type B, Type X, Type R... [14], tuy vậy các phân loại này dường như ít có giá trị ứng dụng và đã được lược đi trong phiên bản hướng dẫn mới xuất bản năm 2020. MCQ thông thường bao gồm các thành phần sau:

- Phần thân (Stem)
- Câu hỏi dẫn (Lead-in)
- Danh sách các lựa chọn (Options) (thông thường từ 3 đến 5 lựa chọn), trong đó có: Các câu mỗi hay câu bẫy (Distractors); Đáp án (Key)

Phương pháp này có tính giá trị vì thường bao phủ được lượng lớn kiến thức trong một khoảng thời gian tương đối ngắn, phương pháp cũng được ghi nhận là khách quan và độ tin cậy cao [15]. Nhưng MCQ cũng dễ mắc lỗi có cấu trúc kém khi chỉ đánh giá mức độ nhớ lại kiến thức, hoặc mắc lỗi kỹ thuật khiến câu trả lời dễ đoán hoặc câu hỏi khó hiểu. Hội đồng đánh giá y khoa Hoa Kỳ, Châu Âu đều cung cấp một bộ hướng dẫn chi tiết để giúp các nhà giáo dục y khoa viết câu hỏi trắc nghiệm [16]. Các hướng dẫn và nghiên cứu đều chỉ ra sự cần thiết sử dụng các thông tin của một tình huống lâm sàng vào câu hỏi để tăng tính giá trị trong việc đo lường các mức nhận thức cao theo thang nhận

thức Bloom [17]. Không có sự suy giảm về tính giá trị và tin cậy trong việc giảm số lựa chọn từ 5 thành 4 hay 3 lựa chọn. [18], [19]. Sử dụng 3 lựa chọn có thể giúp đưa thêm nhiều câu trắc nghiệm vào đề thi qua đó giúp tăng tính giá trị.

**3.2. Câu hỏi ghép cặp - Extended Matching Items/ Extended Matching Questions (EMQ)**

EMQ gồm một danh sách các đáp án và một bộ câu hỏi tình huống, xoay quanh một chủ đề. Ở mỗi câu hỏi/ tình huống thí sinh chọn một hoặc nhiều đáp án trong danh sách cho sẵn. Câu EMQ bao gồm 4 phần: Chủ đề (Theme); Danh sách các lựa chọn (Option list); Câu dẫn (Lead-in); Các tình huống liên quan đến chủ đề (Item stems). Các tình huống sử dụng trong EMQ là các tình huống lâm sàng trải dài trong các chủ đề như chẩn đoán, xét nghiệm, điều trị, quản lý bệnh nhân..., danh sách các lựa chọn có thể từ 6 đến 20 lựa chọn [16]. Tuy vậy 8 lựa chọn là số lượng phù hợp để tiết kiệm thời gian mà vẫn đảm bảo chất lượng bộ câu hỏi [20].

**3.3. Câu hỏi ngắn và Tự luận cải tiến - Short Answer Questions and Modified Essay Questions (SAQ/MEQ)**

Short Answer Questions (SAQ): câu hỏi ngắn là những câu hỏi được trả lời bằng hình thức liệt kê ngắn gọn bằng vài từ. SAQ có thể đạt mức độ bao phủ nội dung rộng bằng cách hỏi một số câu hỏi quan trọng và phân bố toàn diện về nhiều khía cạnh cùng một chủ đề. Việc chấm điểm dễ dàng hơn bài viết tiểu luận dài vì câu trả lời cụ thể và ngắn gọn tuy vậy vẫn tốn thời gian chấm thi hơn MCQ. Độ tin cậy được cải thiện nếu dùng 2 người chấm [21].

Modified Essay Question (MEQ): tự luận cải tiến là dạng những câu hỏi ngắn cấu trúc tuần tự được xây dựng xung quanh một chủ đề

cụ thể hoặc một tình huống lâm sàng. Các chuỗi thông tin được trình bày theo trình tự thời gian giống với tình huống thực. MEQ có thể được thiết kế để kiểm tra khả năng giải quyết vấn đề và ra quyết định của người học. Một số vấn đề cần lưu ý với định dạng này: thứ nhất các câu hỏi sau có thể cung cấp gợi ý cho câu trả lời trước đặc biệt là với câu hỏi dựa trên tình huống lâm sàng; thứ hai, khi các câu hỏi được kết nối với nhau, có một khả năng người học sẽ bị trừ điểm nhiều lần vì cùng một lỗi; thứ ba, không thể thay đổi thứ tự xuất hiện của câu hỏi. Khan và cộng sự kết luận rằng MEQ không hiệu quả hơn MCQ trong việc lượng giá các mức độ cao của nhận thức và cũng dễ mắc lỗi khi soạn câu hỏi [22]. Nếu không có đủ nguồn lực để huấn luyện người ra đề MEQ đạt chất lượng thì MCQ là một lựa chọn hiệu quả hơn khi cần lượng giá các mức độ cao của nhận thức [23].

**3.4. Script Concordance Test - SCT**

SCT cho phép kiểm tra suy luận lâm sàng trong những tình huống không chắc chắn. Trong thực tế bác sĩ đối diện với những tình huống phải đặt ra nhiều giả thiết và phải ra quyết định mỗi khi xuất hiện thêm dữ kiện mới. Cách tiếp cận này dựa trên ba nguyên tắc: (1) thí sinh được thử thách bằng một tình huống thực tế trong đó có nhiều hơn một lựa chọn phù hợp, (2) định dạng trả lời là thang đo kiểu Likert, phản ánh cách xử lý dữ kiện liên quan trong các tình huống giải quyết vấn đề, theo lý thuyết kịch bản, (3) và cho điểm dựa trên phương pháp cho điểm tổng hợp để tính đến sự biến đổi của các quá trình lập luận giữa các chuyên gia. Định dạng này có một số dấu hiệu về tính giá trị và tin cậy tốt để đánh giá khả năng lý luận lâm sàng [24].

**Bảng 1:** Câu hỏi SCT mẫu (theo Wan S.H.) [25]

| Tình huống lâm sàng:                             |                       |   |                                 |    |   |    | bụng cấp dữ dội |                       |
|--|-----------------------|---|---------------------------------|----|---|----|-----------------|-----------------------|
| A: Bệnh nhân nữ 22 tuổi đến phòng cấp cứu vì đau |                       |   |                                 |    |   |    |                 |                       |
|  | Nếu bạn nghĩ đến:     | Sau đó bạn nhận được kết quả:               | Thì giả thiết ban đầu trở thành |    |   |    |                 |                       |
| 1  | Thai ngoài tử cung vỡ | Beta HCG âm tính                            | A                               | B  | C | D  | E               | -2: Không còn phù hợp |
|  |                       |   |                                 | -1 | 0 | +1 | +2              | -1: Ít phù hợp        |
| 2  | Viêm ruột thừa cấp    | Khám bụng nhận thấy dấu phản ứng và co cứng | A                               | B  | C | D  | E               | 0: Không thay đổi     |
|  |                       |   |                                 | -1 | 0 | +1 | +2              |                       |
| 3  | Viêm đường mật cấp    | Thân nhiệt đo được 36.8 độ C                | A                               | B  | C | D  | E               | +1: Phù hợp hơn       |
|  |                       |   |                                 | -1 | 0 | +1 | +2              | +2: Rất phù hợp       |

### **Lượng giá mức độ thực hiện:**

#### **3.5. Ca lâm sàng đầy đủ - Long case**

Ca lâm sàng đầy đủ là một định dạng thi truyền thống được sử dụng nhiều trong giáo dục y tế đại học và sau đại học trong một thời gian dài. Thí sinh được giao nhiệm vụ tiếp xúc trực tiếp với bệnh nhân trong vòng 30 - 60 phút, nhằm khai thác bệnh sử và thăm khám mà không có sự quan sát của giám khảo. Sau đó, người học trình bày kết quả và thi vấn đáp với các giám khảo trong vòng 20 - 30 phút xoay quanh ca bệnh vừa tiếp cận. Cách cho điểm thường không cấu trúc. Mặc dù các vấn đề của bệnh nhân rất chân thực nhưng việc mỗi thí sinh nhận ca bệnh khác nhau ảnh hưởng đến mức độ chuẩn hóa của kỳ thi [26]. Ca lâm sàng đầy đủ có lợi thế trong việc đo lường năng lực khám lâm sàng nhờ tiếp xúc với bệnh nhân thật, nhưng lượng giá bằng ca lâm sàng đầy đủ gặp hạn chế lớn về tính giá trị và tính tin cậy [27]. Do một ca lâm sàng không thể đo lường được một khoảng rộng lớn nhiều lĩnh vực lâm sàng nên cách thi này có tính giá trị không cao. Phương pháp này đã ngừng sử dụng tại các nước Bắc Mỹ [15] tuy nhiên các tác giả trên thế giới nhìn chung khuyến nghị dùng ca lâm sàng đầy đủ để lượng giá quá trình vì giá trị giáo dục của phương pháp là khuyến khích học tập dựa trên các tình huống bệnh nhân thật.[28]. Trong bối cảnh Việt Nam, việc sử dụng ca lâm sàng đầy đủ phục vụ cho lượng giá quá trình là phù hợp vì tính khả thi, giá trị giáo dục khi được sử dụng đúng bằng cách tăng tính chuẩn hóa của bảng chấm.

Để cải thiện các hạn chế của Long case, các nhà giáo dục đã cải biên theo các hướng: tăng số lượng ca bệnh mà thí sinh cần tiếp xúc (Ca lâm sàng từng phần - Short case); bổ sung tính chuẩn hóa và cấu trúc trong bảng chấm (Ca lâm sàng đầy đủ cấu trúc theo mục tiêu - Objective Structured Long Case Record/OSLER); vừa bổ sung tính cấu trúc, vừa thêm yêu cầu giám khảo quan sát thí sinh và tăng số lượng ca bệnh (Bài tập đánh giá lâm sàng thu gọn - Mini Clinical Evaluation Exercise)

#### **3.6. Ca lâm sàng từng phần - Short case**

Thí sinh được yêu cầu tiếp xúc với nhiều bệnh nhân mà mỗi ca chỉ tập trung vào một vài nhóm kỹ năng xác định trước. Sau đó đến phần

thảo luận chẩn đoán với hai giám khảo cho điểm độc lập, thường là cách cho điểm không cấu trúc. Phương pháp này cải thiện tính giá trị nhưng không có tính tin cậy vì cách chấm điểm không cấu trúc và không chuẩn hóa [26].

#### **3.7. Thi ca lâm sàng đầy đủ cấu trúc theo mục tiêu - Objective Structured Long Case Record (OSLER)**

OSLER sử dụng bảng điểm gồm 10 tiêu chí để phân tích và đánh giá năng lực người học và có sự quan sát của giám khảo khi thí sinh khai thác bệnh sử và thăm khám bệnh nhân. Việc chấm điểm dựa trên 10 tiêu chí để cải thiện tính khách quan, giá trị và độ tin cậy của Long case.

Mười tiêu chí đánh giá năng lực thí sinh bao gồm:

Bốn tiêu chí trong khai thác bệnh sử

1. Thời gian khai thác bệnh sử, tính rõ ràng, đầy đủ của bệnh sử.
2. Kỹ năng giao tiếp, thái độ với bệnh nhân.
3. Cách tiếp cận bệnh nhân có hệ thống
4. Hình thành tình tiết bệnh sử hợp lý theo thời gian.

Ba tiêu chí trong kỹ năng thăm khám

1. Cách tiếp cận thăm khám bệnh nhân có hệ thống
2. Kỹ thuật khám
3. Nhận diện các triệu chứng thực thể chính xác

Ba tiêu chí còn lại về quản lý bệnh nhân bao gồm:

1. Xây dựng các xét nghiệm cận lâm sàng thích hợp theo một trình tự hợp lý.
2. Hướng điều trị, mục tiêu điều trị cho bệnh nhân.
3. Sự nhạy bén trong lâm sàng: Mục này dựa trên chín mục trước để đánh giá khả năng xác định và giải quyết vấn đề của thí sinh.

Kết quả điểm được ghi nhận như sau

- P+ (rất tốt / xuất sắc): 60 - 80% mong đợi.
- P (đạt): 50 - 55% mong đợi.
- P- (dưới đạt): dưới 50% mong đợi.

Giám khảo cũng được yêu cầu ghi lại độ khó của ca lâm sàng

- Ca chuẩn (Standard case): có 1 vấn đề.
- Ca khó (Difficult case): BN có từ 2- 3 vấn đề.
- Ca rất khó (Very difficult case): trên 3 vấn đề. [29]

OSLER là phương pháp lượng giá có tính giá trị đối với nguyên tắc thực hành y khoa “bệnh nhân là trung tâm”, tận dụng được ưu điểm thực tế của Ca lâm sàng đầy đủ và bổ sung các khuyết điểm khi yêu cầu giám khảo quan sát đầy đủ quá trình thực hiện của thí sinh, cải thiện tính cấu trúc và tiêu chuẩn hóa việc chấm điểm. Tuy vậy, Wass, V. (2001) đề xuất rằng cần sử dụng 10 bệnh nhân thật không được chuẩn hoá với thời gian thi trung bình 3 giờ 30 phút mới đảm bảo tính tin cậy của bài kiểm tra năng lực lâm sàng [30] - điều này rõ ràng giảm tính khả thi và chấp nhận đối với phương pháp.

### **3.8. Kiểm tra lâm sàng có cấu trúc theo mục tiêu - Objective Structured Clinical Examination (OSCE)**

Kiểm tra lâm sàng có cấu trúc theo mục tiêu được định nghĩa bởi Harden (1988) là một “Cách tiếp cận trong việc lượng giá kỹ năng lâm sàng trong đó năng lực thành phần được đo một cách có kế hoạch, được cấu trúc tốt và chú ý đến mục tiêu học tập cần kiểm tra” [31]. Khan KZ (2013) đề xuất định nghĩa mở rộng cho OSCE như sau: “Một công cụ đánh giá dựa trên mục tiêu học tập và tiêu chuẩn hoá, trong đó thí sinh di chuyển qua một loạt các trạm liên mạch, giới hạn thời gian với mục đích là lượng giá các hoạt động chuyên môn trong môi trường mô phỏng. Tại mỗi trạm, thí sinh được lượng giá bằng cách cho các điểm trên các tiêu chuẩn được chuẩn hoá và thực hiện bởi các người đánh giá đã qua đào tạo” [32]. OSCE lần đầu tiên được mô tả vào năm 1975, khi tác giả Harden và cộng sự đã xuất bản một bài báo trên Tạp chí Y khoa Anh về việc đánh giá năng lực lâm sàng bằng cách sử dụng OSCE [33]. Harden đã tổ chức OSCE đầu tiên tại bệnh viện gồm 18 trạm để kiểm tra và 2 trạm nghỉ, mỗi trạm dài 4 phút rưỡi, và 30 giây di chuyển. OSCE được xác định là phương pháp thích hợp để lượng giá năng lực ở mức SHOWS HOW theo thang Miller [32]. Các yếu tố ảnh hưởng đến tính tin cậy của OSCE bao gồm [32]: Số lượng trạm [34], Rubric tiêu chuẩn tính điểm, Người đánh giá được đào tạo, Bệnh nhân chuẩn.

### **3.9. Bài tập đánh giá lâm sàng thu gọn - Mini Clinical Evaluation Exercise (Mini - CEX)**

Mini - CEX là phương pháp đánh giá do Hội đồng Nội khoa Hoa Kỳ (American Board of Internal Medicine) phát triển để đánh giá năng lực cốt lõi của bác sĩ nội trú. Phương pháp này được thực hiện bằng cách quan sát trực tiếp cuộc gặp gỡ thực tế giữa người học với bệnh nhân và cung cấp phản hồi. Các kỹ năng được đánh giá bao gồm: Kỹ năng phỏng vấn (medical interviewing skills); Kỹ năng khám lâm sàng (physical examination skills); Tính chuyên nghiệp (humanistic qualities/professionalism); Đánh giá lâm sàng (clinical judgment); Kỹ năng tư vấn (counselling skills); Kỹ năng tổ chức và tính hiệu quả (organization and efficiency); Năng lực lâm sàng tổng thể (overall clinical competency) [35]. Giám khảo sẽ quan sát trực tiếp cuộc gặp gỡ giữa người học - bệnh nhân tại bất kỳ cơ sở y tế nào. Thời gian cho mỗi cuộc gặp gỡ là 15 phút, người học sẽ tiến hành nhiệm vụ cụ thể tại thời điểm lượng giá, có thể là khai thác tiền sử, bệnh sử, khám lâm sàng, ra kế hoạch chẩn đoán và điều trị, thực hiện một phương pháp chăm sóc và nhận phản hồi từ giám khảo trong khoảng 5 - 10 phút. Mỗi người học dự kiến được lượng giá trong 4 - 6 cuộc gặp với bệnh nhân trong suốt học phần để đánh giá được toàn diện năng lực. Phương pháp này dần được mở rộng từ giáo dục sau đại học đến giáo dục đại học với sự gia tăng thời gian cho 1 lần gặp bệnh nhân và phản hồi lên 30 - 45 phút [36].

### **3.10. Portfolio và các dạng báo cáo thực tập**

Portfolio hay còn gọi là Hồ sơ năng lực là một tài liệu tích hợp các minh chứng về thành tích của một cá nhân và những suy tư về thành tích đó. Học tập dựa trên Hồ sơ năng lực (Portfolio - based learning) được Snadden đề cập như là một phương pháp đáp ứng với lý thuyết học tập ở người trưởng thành [37], do đó phương pháp này có được các ưu điểm mà các phương pháp khác không có được như: ghi nhận và khuyến khích việc tự học, là một phần không thể thiếu của tính chuyên nghiệp của thực hành y khoa; bằng chứng thu thập từ nhiều nguồn giúp hợp nhất toàn bộ quá trình học tập, kết nối từ lý thuyết đến thực hành; có thể dùng được cho lượng giá quá trình và lượng giá kết thúc [38].

Các trở ngại hiện nay của phương pháp này là tính tiêu chuẩn trong việc cho điểm, Challis lưu ý các tiêu chí chấm và các tiêu chuẩn giá trị của bằng chứng phải rõ ràng và công khai giữa người học và người chấm [38].

Các dạng báo cáo thực tập (Logbook) là các tập tài liệu về kinh nghiệm của người học, gần giống như portfolio. Tuy vậy chỉ giới hạn trong một số lĩnh vực chuyên môn hẹp hơn là toàn diện và không bao gồm các suy tư cá nhân về việc học. Báo cáo thực tập có thể là báo cáo ca bệnh, báo cáo thủ thuật, báo cáo nhiệm vụ, thông thường chỉ mang tính liệt kê hơn là phản ánh năng lực của người học. Báo cáo thực tập ghi nhận số lượng bệnh nhân và thủ thuật người học đã làm mà điều này không có liên hệ chắc chắn đến năng lực thực hành lâm sàng của người học do đó qua báo cáo thực tập điều chính xác nhất được phản ánh là những gì người học đã không thực hành [39].

#### 4. KẾT LUẬN

Lượng giá là công cụ cho phép người học chứng minh họ đạt được chuẩn đầu ra học tập mong đợi của chương trình đào tạo, nhưng không có công cụ duy nhất đáp ứng được tất cả các chuẩn đó. Để sử dụng công cụ lượng giá phù hợp mục đích, nhà giáo dục cần phải hiểu rõ các tiêu chuẩn lựa chọn dựa trên thuộc tính công cụ và mục đích lượng giá. Có 2 nhóm công cụ lượng giá chính theo khung lượng giá lâm sàng Miller: lượng giá kiến thức và lượng giá mức độ thực hiện. Theo thời gian, một số công cụ lượng giá đã bị loại bỏ hoặc ít sử dụng như các bài thi tự luận truyền thống, ca lâm sàng đầy đủ bên cạnh đó xuất hiện các công cụ mới như OSCE, OSPE, Mini - CEX, các công cụ lượng giá mới này giành được sự chú ý và ưa chuộng do có nhiều nghiên cứu khẳng định tính giá trị và tin cậy. Không có công cụ nào tỏ ra ưu thế vượt trội, việc sử dụng phải tùy thuộc vào mục đích lượng giá, thời điểm lượng giá và các đặc tính của công cụ. Để có thể lượng giá toàn bộ chuẩn đầu ra, cần lưu ý sử dụng lượng giá quá trình kết hợp với lượng giá kết thúc, và sử dụng nhiều loại công cụ phối hợp với nhau một cách thích hợp.

#### TÀI LIỆU THAM KHẢO

1. Ronald M. Harden. 23 Jul 2015, Why Outcome - based education (OBE) is an important development in medical education from: Routledge International Handbook of Medical Education, The cultural politics of education, development and childhood Routledge Accessed on: 27 Apr 2022 <https://www.routledgehandbooks.com/doi/10.4324/9780203066201.ch3>
2. Bộ Y tế (2015). Quyết định số 1854/QĐ-BYT ngày 18/05/2015 về việc phê duyệt tài liệu “Chuẩn năng lực cơ bản của bác sĩ đa khoa”
3. Bộ Giáo dục và Đào tạo (2010). Công văn số 2196/BGDĐT-GDĐH ngày 22/04/2010 về việc hướng dẫn xây dựng và công bố chuẩn đầu ra ngành đào tạo.
4. Bộ Giáo dục và Đào tạo (2015). Thông tư số 07/2015/TT-BGDĐT ngày 16/04/2015 về Quy định khối kiến thức tối thiểu, yêu cầu về năng lực mà người học đạt được sau khi tốt nghiệp đối với mỗi trình độ đào tạo của giáo dục đại học và quy trình xây dựng, thẩm định, ban hành chương trình đào tạo trình độ đại học, thạc sĩ, tiến sĩ.
5. Trường Đại học Y Dược Hải Phòng, Chuẩn đầu ra Chương trình đào tạo y khoa dựa trên năng lực, Ban hành kèm Quyết định số 622A/QĐ-YDHP, ngày 04 tháng 7 năm 2018 của Hiệu trưởng Trường Đại học Y dược Hải Phòng
6. <https://med.duytan.edu.vn/vi-vn/dao-tao/chuan-dau-ra-nganh-bac-si-da-khoa/> (truy cập ngày 18/04/2022)
7. [https://admin.ump.edu.vn/uploads/ckeditor/files/Y/Chu%e1%ba%a9n%20%20c4%91%e1%ba%a7u%20ra%20BS%20Y%20khoa%202020\\_1.pdf](https://admin.ump.edu.vn/uploads/ckeditor/files/Y/Chu%e1%ba%a9n%20%20c4%91%e1%ba%a7u%20ra%20BS%20Y%20khoa%202020_1.pdf) (truy cập ngày 18/04/2022)
8. Bộ Giáo dục và Đào tạo (2021). Thông tư số 17/2021/TT- BGDĐT về việc Quy định về chuẩn chương trình đào tạo; xây dựng, thẩm định và ban hành chương trình đào tạo các trình độ của giáo dục đại học
9. Biggs, J. B. (2003). Aligning Teaching and Assessment to Curriculum Objectives. Imaginative Curriculum Project, LTSN Generic Centre.
10. Biggs, J (2014). Constructive Alignment

- in university teaching. *HERDSA Review of Higher Education* Vol. 1, pp 5-22.
11. Van Der Vleuten, C.P.M. The assessment of professional competence: Developments, research and practical implications. *Adv Health Sci Educ* 1, 41-67 (1996). <https://doi.org/10.1007/BF00596229>
  12. Miller, G E The assessment of clinical skills/competence/performance, *Academic Medicine*: September 1990 - Volume 65 - Issue 9 - p S63-7
  13. Moss E. Multiple choice questions: their value as an assessment tool. *Curr Opin Anaesthesiol.* 2001;14(6):661-666. doi:10.1097/00001503-200112000-00011
  14. Case, Susan & Swanson, David. (2002). *Constructing Written Test Questions For the Basic and Clinical Sciences*. National Board of Medical Examiners.
  15. Wass, V., Van der Vleuten, C., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *Lancet* (London, England), 357(9260), 945-949. [https://doi.org/10.1016/S0140-6736\(00\)04221-5](https://doi.org/10.1016/S0140-6736(00)04221-5)
  16. NBME Item-writing guide. *Constructing Written Test Questions for the Health Sciences*. National Board of Medical Examiners, Philadelphia; 2020.
  17. Salam, A., Yousuf, R., & Bakar, S. M. A. (2020). Multiple Choice Questions in Medical Education: How to Construct High Quality Questions. *International Journal of Human and Health Sciences (IJHHS)*, 4(2), 79. <https://doi.org/10.31344/ijhhs.v4i2.180>
  18. Vyas R, Supe A. Multiple choice questions: a literature review on the optimal number of options. *Natl Med J India.* 2008 May-Jun;21(3):130-3. PMID: 19004145.
  19. Tarrant M, Ware J. A comparison of the psychometric properties of three- and four-option multiple - choice questions in nursing assessments. *Nurse Educ Today.* 2010 Aug;30(6):539-43. doi: 10.1016/j.nedt.2009.11.002. Epub 2010 Jan 6. PMID: 20053488.
  20. Swanson, David B.; Holtzman, Kathleen Z.; Allbee, Krista; Clauser, Brian E. *Psychometric Characteristics and Response Times for Content - Parallel Extended - Matching and One - Best - Answer Items in Relation to Number of Options*, *Academic Medicine*: October 2006 - Volume 81 - Issue 10 - p S52-S55 doi: 10.1097/01.ACM.0000236518.87708.9d
  21. Schuwirth, L. W., & van der Vleuten, C. P. (2004). Different written assessment methods: what can be said about their strengths and weaknesses?. *Medical education*, 38(9), 974-979. <https://doi.org/10.1111/j.1365-2929.2004.01916.x>
  22. Khan MU, Aljarallah BM. Evaluation of Modified Essay Questions (MEQ) and Multiple Choice Questions (MCQ) as a tool for Assessing the Cognitive Skills of Undergraduate Medical Students. *Int J Health Sci (Qassim).* 2011 Jan;5(1):39-43. PMID: 22489228; PMCID: PMC3312767.
  23. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Med Educ.* 2007 Nov 28;7:49. doi: 10.1186/1472-6920-7-49. PMID: 18045500; PMCID: PMC2148038.
  24. Charlin, B., & van der Vleuten, C. (2004). Standardized Assessment of Reasoning in Contexts of Uncertainty: The Script Concordance Approach. *Evaluation & the Health Professions*, 27(3), 304-319. <https://doi.org/10.1177/0163278704267043>
  25. Wan S. H. (2015). Using the script concordance test to assess clinical reasoning skills in undergraduate and postgraduate medicine. *Hong Kong medical journal = Xianggang yi xue za zhi*, 21(5), 455-461. <https://doi.org/10.12809/hkmj154572>
  26. Ponnampuruma, G. G., Karunathilake, I. M., McAleer, S., & Davis, M. H. (2009). The long case and its modifications: a literature review. *Medical education*, 43(10), 936-941. <https://doi.org/10.1111/j.1365-2923.2009.03448.x>
  27. Norcini J. J. (2002). The death of the long case?. *BMJ (Clinical research ed.)*, 324(7334), 408-409. <https://doi.org/10.1136/bmj.324.7334.408>
  28. Wass, V., & van der Vleuten, C. (2004). The long case. *Medical education*, 38(11),

- 1176-1180. <https://doi.org/10.1111/j.1365-2929.2004.01985.x>
29. Gleeson F. 1997. Assessment of Clinical Competence using the Objective Structured Long Examination Record (OSLER), AMEE Medical Education Guide No. 9. *Med Teach* 19:7-14
30. Wass, V., Jones, R., & Van der Vleuten, C. (2001). Standardized or real patients to test clinical competence? The long case revisited. *Medical education*, 35(4), 321-325. <https://doi.org/10.1046/j.1365-2923.2001.00928.x>
31. Harden R. M. (1988). What is an OSCE?. *Medical teacher*, 10(1), 19-22. <https://doi.org/10.3109/01421598809019321>
32. Khan, K. Z., Ramachandran, S., Gaunt, K., & Pushkar, P. (2013). The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: an historical and theoretical perspective. *Medical teacher*, 35(9), e1437-e1446. <https://doi.org/10.3109/0142159X.2013.818634>
33. Harden, R M et al. "Assessment of clinical competence using objective structured examination." *British medical journal* vol. 1,5955 (1975): 447-51. doi:10.1136/bmj.1.5955.447
34. Roberts, C., Newble, D., Jolly, B., Reed, M., & Hampton, K. (2006). Assuring the quality of high-stakes undergraduate assessments of clinical competence. *Medical teacher*, 28(6), 535-543. <https://doi.org/10.1080/01421590600711187>
35. Norcini, J., Blank, L., Arnold, G. K., & Kimball, H. R. (1995). The mini-CEX (clinical evaluation exercise): a preliminary investigation. *Annals of internal medicine*, 123(10), 795-799. <https://doi.org/10.7326/0003-4819-123-10-199511150-00008>
36. Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical teacher*, 29(9), 855-871. <https://doi.org/10.1080/01421590701775453>
37. Snadden, David. (1998). The use of portfolio learning in medical education. *Medical Teacher*. 20. 192-199. 10.1080/01421599880904.
38. Challis, M. (1999). AMEE Medical Education Guide No. 11 (revised): Portfolio-based learning and assessment in medical education. *Medical teacher*, 21(4), 370-386.
39. Shumway, J. M., Harden, R. M., & Association for Medical Education in Europe (2003). AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Medical teacher*, 25(6), 569-584. <https://doi.org/10.1080/0142159032000151907>