

# **ÁP DỤNG PHƯƠNG PHÁP THIẾT LẬP CHUẨN ĐẬU NEDELSKY TRONG LƯỢNG GIÁ CUỐI KỲ HỌC PHẦN SẢN PHỤ KHOA CHO SINH VIÊN Y KHOA TẠI TRƯỜNG ĐẠI HỌC Y KHOA PHẠM NGỌC THẠCH**

## ***Tác giả***

1. Bùi Hiếu Anh, Bộ môn Nhi Khoa, Đại học Y khoa Phạm Ngọc Thạch

Điện thoại: 038.262.0804, Email: [bhanh1142@gmail.com](mailto:bhanh1142@gmail.com)

2. Nguyễn Xuân Vũ, Bộ môn Sản Phụ Khoa, Đại học Y khoa Phạm Ngọc Thạch

## **TÓM TẮT**

**Mục tiêu:** Thiết lập quy chuẩn để xác định điểm đậu hay còn gọi là Thiết lập chuẩn đậu (standard setting) là nền tảng của sự lượng giá dựa trên năng lực trong giáo dục y khoa. Các chứng cứ trong y văn cho thấy thực hành lấy ngưỡng cắt là chuẩn 50% câu đúng để phân biệt nhóm đậu và nhóm rớt không có tính biện hộ được (defensible). Mục tiêu của nghiên cứu này là nhấn mạnh phương pháp khoa học để thiết lập chuẩn đậu, đặc biệt là phương pháp Nedelsky trong những bối cảnh giới hạn về nguồn lực (thiếu các chuyên gia Subject Matter Expert - SME), ứng dụng trong lượng giá cuối kỳ học phần Sản Phụ Khoa cho sinh viên Y khoa tại Trường Đại học Y khoa Phạm Ngọc Thạch.

**Phương pháp:** Đầu tiên, thiết kế và tiến hành quy trình gồm bốn bước áp dụng phương pháp Nedelsky để xác lập điểm đậu cho Đề thi cuối kỳ Học phần Sản Phụ Khoa cho sinh viên Y khoa năm thứ 4 ngày 3/12/2022, trong đó có hội đồng 3 người đánh giá là các chuyên gia có kinh nghiệm chuyên môn và lượng giá. Sau khi có điểm đậu theo phương pháp Nedelsky, chúng tôi tiến hành nghiên cứu cắt ngang mô tả trên 808 thí sinh tham gia kỳ thi để tính tỉ lệ đậu và tỉ lệ rớt kỳ thi. Ngoài ra, chúng tôi so sánh tỉ lệ rớt giữa phương pháp tuyệt đối Nedelsky và phương pháp tương đối Cohen.

**Kết quả:** Sau khi tiến hành quy trình phương pháp Nedelsky, chúng tôi thu được kết quả độc lập của 3 người đánh giá cho số câu bẫy mà nhóm MCC có thể loại trừ được từ đó tính Ngưỡng đậu tối thiểu (Minimal Passing Level – MLP) cho từng câu hỏi trong 50 câu của đề thi và tính trung bình cho toàn bộ đề thi của cả 3 chuyên gia là 65%, tương ứng với 33 câu đúng là đậu. Trong 808 thí sinh tham dự kỳ thi thì theo điểm đậu của phương pháp Nedelsky có 183 thí sinh đậu với tỉ lệ đậu là 23% và tỉ lệ rớt là 77% tương ứng với 625 thí sinh. So sánh với điểm đậu được tính toán theo phương pháp tương đối Cohen là 55%, tỉ lệ rớt chỉ chiếm 51%; thấp hơn đáng kể và có ý nghĩa thống kê so với tỉ lệ rớt của phương pháp Nedelsky. Với 209 thí sinh (chiếm ¼ số thí sinh tham gia kỳ thi) được đậu nếu xét theo phương pháp Cohen thì lại trở thành rớt theo phương pháp Nedelsky vì chưa đủ năng lực. Kết quả này đe dọa nghiêm trọng đến quá trình lượng giá dựa trên năng lực, vốn là cốt lõi trong giáo dục y khoa nhằm đào tạo ra các nhân viên y tế đủ chuyên môn để hành nghề.

**Kết luận:** Phương pháp thiết lập chuẩn đậu tương đối như phương pháp Cohen không còn phù hợp để phân định người có năng lực và người không có năng lực trong giáo dục y học. Nói cách khác, thiết lập chuẩn đậu nếu không dựa trên các phương pháp khoa học có tính biện

hộ được (defensible) sẽ dẫn đến hậu quả thất bại trong việc đánh rớt các thí sinh không có năng lực – “Failure to Fail”. Cần ứng dụng các phương pháp thiết lập chuẩn đầu phù hợp hơn dựa trên các chứng cứ khoa học và có tính có thể biện hộ (defensible) như phương pháp Nedelsky trong bối cảnh chưa có các chuyên gia SME, trong thời gian xây dựng hội đồng các SME và nghiên cứu ứng dụng các phương pháp dựa trên tiêu chuẩn như Angoff hay Ebel với sự tham gia của các SME.

**Từ khóa:** thiết lập chuẩn đầu, phương pháp Nedelsky, phương pháp Cohen, lượng giá dựa trên năng lực

## **ABSTRACT**

*APPLICATION OF THE DEFENSIBLE STANDARD SETTING METHOD - NEDELSKY METHOD FOR SUMMATIVE ASSESSMENT OF OBSTETRICS AND GYNECOLOGY COURSE FOR MEDICAL STUDENTS IN PHAM NGOC THACH UNIVERSITY OF MEDICINE*

**Objectives:** Defensible standard setting is the cornerstone of competency-based assessment. Evidence from the literature showed that the practice of taking a 50% mark as the cut off point differentiating passes and failures is not defensible. The objective of this paper is to highlight the standard setting methodology, particularly the Nedelsky’s method in the context of limited resources, in determining the pass score in the assessment of Obstetrics and Gynecology Course of undergraduate medical programme at the Pham Ngoc Thach University of Medicine.

**Method:** First, we designed and implemented a four-step process applying the Nedelsky method to establish the passing score for the Obstetrics and Gynecology Final Exam for fourth-year medical students on December 3, 2022. This process involved a panel of three assessors who were experienced experts in the field. Following the establishment of the passing score using the Nedelsky method, we conducted a cross-sectional descriptive study involving 808 candidates who participated in the exam to calculate the pass rate and fail rate for the examination. Additionally, we compared the fail rates between the absolute Nedelsky method and the relative Cohen method.

**Results:** After conducting the Nedelsky method procedure, we obtained independent results from three assessors regarding the number of distractors that the MCC team could eliminate. From these numbers, we calculated the Minimal Passing Level (MLP) for each question out of the 50 questions on the exam. The average MLP determined by all three experts was 65%, which corresponds to 33 correct answers as the passing score. Out of the 808 candidates who participated in the exam, 183 passed according to the Nedelsky method, resulting in a pass rate of 23% and a fail rate of 77%, which translates to 625 candidates. Comparatively, when we calculated the passing score using the Cohen method, it was set at 55%, resulting in a pass rate of only 51%. This pass rate was significantly lower and statistically different from the pass rate determined by the Nedelsky method. Notably, 209 candidates (25% of the total candidates) who passed according to the Cohen method would have failed according

to the Nedelsky method due to inadequate competence. These findings pose a significant threat to the competency-based assessment process, which is the core of medical education aimed at producing adequately skilled healthcare professionals.

**Conclusion:** The relative standard setting method like the Cohen method is no longer suitable for distinguishing between individuals with competence and those without competence in medical education. In other words, setting a passing standard that is not based on scientifically defensible methods can lead to the consequence of failing to identify incompetent candidates - a phenomenon often referred to as "Failure to Fail". It is essential to apply the defensible standard setting methods based on scientific evidence, such as the Nedelsky method, in situations where subject matter experts (SMEs) are not available during the standard setting process. This approach can be particularly valuable during the period of establishing an SME panel and conducting research on the application of criterion-referenced methods like Angoff or Ebel with the involvement of SMEs.

**Key words:** standard setting, Nedelsky's method, Cohen's method, competency-based assessment

## ĐẶT VẤN ĐỀ

Thiết lập quy chuẩn để xác định điểm đậu hay còn gọi là Thiết lập chuẩn đậu (standard setting) là quy trình xác định điểm đậu (passing score) có giá trị phân biệt giữa người có năng lực và người không có năng lực. Chuẩn đậu hợp lý cho một bài thi sẽ cho đậu những thí sinh thực sự có năng lực và ngược lại đánh rớt những thí sinh không đủ năng lực. Do đó, thiết lập chuẩn đậu có cơ sở lý luận và dựa trên các phương pháp khoa học với chứng cứ rõ ràng là rất quan trọng. Lý tưởng nhất, thiết lập chuẩn đậu nên được thực hiện bởi những chuyên gia có kinh nghiệm dày dặn trong việc xử lý những nhóm học viên với khả năng khác nhau. Kinh nghiệm này sẽ cho phép họ phân biệt không chỉ giữa những người đậu và người rớt, mà quan trọng hơn, họ sẽ có khả năng đặc trưng hóa nhóm những "thí sinh ranh giới" (borderline student). Các thí sinh ở "nhóm ranh giới" thường là những người mà chúng ta không chắc chắn liệu có nên cho đậu hay không vì chúng ta không chắc chắn họ có đủ của kiến thức và kỹ năng trong việc đưa ra những quyết định lâm sàng hiệu quả.<sup>[1]</sup>

Phương pháp Nedelsky thuộc nhóm phương pháp thiết lập chuẩn đậu tuyệt đối, tham chiếu dựa trên tiêu chí (criterion-referenced) và lấy bài thi làm trung tâm (test-centered). Các phương pháp thiết lập chuẩn đậu tuyệt đối và tham chiếu trên tiêu chí, trong đó có sự tham gia của các chuyên gia để xác định mức năng lực tối thiểu (minimally competent candidate - MCC) cần để đậu kỳ thi, là cần thiết để loại bỏ sự tùy tiện và tính chủ quan sẽ thống trị quá trình thiết lập tiêu chuẩn. Ví dụ phổ biến và rõ ràng nhất về điều này là việc đặt ngưỡng điểm ở một con số tròn như 50% hoặc 60% một cách duy ý chí, bất kể độ khó của bài kiểm tra hay năng lực của người thi. Chứng cứ từ các nghiên cứu trong y văn cho thấy việc thực hành lấy mức chuẩn 50% điểm số để phân biệt nhóm thí sinh có năng lực và không có năng lực không có tính biện hộ được (defensible).<sup>[2]</sup> Khắc phục khuyết điểm các phương pháp tương đối, phương pháp Nedelsky đề xuất một cách tiếp cận hệ thống và dựa trên bằng chứng để thiết lập chuẩn đậu nhằm đảm bảo tính giá trị (validity), công bằng (fair), đáng tin cậy (reliability) và tính có tính

biện hộ về pháp lý (legally defensible) trong quá trình lượng giá giúp phản ánh chính xác mức độ năng lực cần thiết cho việc thực hành y học và đáp ứng được các tiêu chuẩn của chuẩn định khu vực và quốc tế (accreditation standards).

*Câu hỏi nghiên cứu:*

Có sự khác biệt về điểm đậu xác định bằng phương pháp dựa trên tiêu chí của Nedelsky so với phương pháp dựa trên quy chuẩn của Cohen trong lượng giá cuối kỳ Học phần Lý thuyết Sản Phụ Khoa dành cho sinh viên Y năm 4 tại Trường Đại học Y khoa Phạm Ngọc Thạch?

*Mục tiêu nghiên cứu:*

1. Tiến hành quy trình thiết lập tiêu chuẩn để xác định điểm đậu theo phương pháp Nedelsky cho lượng giá cuối kỳ Học phần Sản Phụ Khoa năm 4.

2. Xác định tỉ lệ rớt của sinh viên Y khoa năm thứ 4 tham gia kỳ thi kết thúc học phần Lý thuyết Sản Phụ Khoa theo phương pháp Nedelsky.

3. So sánh tỉ lệ rớt của sinh viên giữa 2 phương pháp dựa trên tiêu chí của Nedelsky và phương pháp dựa trên quy chuẩn của Cohen.

## **ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU**

### **Thiết kế nghiên cứu**

Nghiên cứu cắt ngang mô tả (observational, cross-sectional study)

### **Đối tượng nghiên cứu**

Đối tượng nghiên cứu là sinh viên Y khoa năm thứ 4 niên khóa 2019 – 2025 đã hoàn tất học phần Lý thuyết Sản Phụ Khoa năm thứ 4.

### **Phương pháp chọn mẫu**

Phương pháp chọn mẫu là lấy mẫu toàn bộ sinh viên Y khoa năm thứ 4 niên khóa 2019 – 2025 dự thi kỳ thi kết thúc học phần Lý thuyết Sản Phụ Khoa ngày 3/12/2022.

### **Quy trình nghiên cứu gồm hai phần nối tiếp:**

#### **Phần 1. Thiết kế một quy trình gồm các bước áp dụng phương pháp Nedelsky để xác lập chuẩn đậu:**

**Bước 1.** Xác định năng lực (Defining the competency):

Trước khi thiết lập chuẩn đậu, quan trọng phải xác định rõ năng lực cụ thể hoặc lĩnh vực kiến thức đang được đánh giá. Trong nghiên cứu này chúng tôi xác định rõ năng lực được lượng giá là “Kiến thức cơ bản trong lĩnh vực Sản Khoa cho đối tượng sinh viên Y khoa năm thứ 4” với đề thi ngày 3/12/22.

**Bước 2.** Thành lập hội đồng chuyên gia (Establishing an expert panel):

Một hội đồng gồm các chuyên gia giàu kiến thức và kinh nghiệm chuyên môn và được đào tạo về lượng giá cũng như thiết lập chuẩn đậu đồng thời cũng là người giảng dạy có kinh nghiệm về khả năng của thí sinh, đặc biệt nhóm thí sinh có năng lực tối thiểu MCC để có thể đưa ra quyết định nhóm này có thể loại trừ những câu bẫy nào. Đây cũng là điểm khác biệt quan trọng giữa phương pháp Nedelsky và các phương pháp dựa trên tiêu chí khác.

Thành lập hội đồng gồm 3 người đánh giá theo phương pháp Nedelsky là 3 giảng viên thuộc bộ môn Sản Phụ Khoa, trường Đại học Y khoa Phạm Ngọc Thạch, có trình độ chuyên môn chuyên khoa cấp 2, có trình độ về lượng giá trong giáo dục y khoa và nắm vững phương pháp Nedelsky, tốt nghiệp các khóa Đào tạo và phát triển giảng viên nâng cao trong Chương trình hợp tác giáo dục (CEP) với Viện Đại Học Sức Khỏe Texas Tech tại El Paso, Hoa Kỳ.

**Bước 3.** Thảo luận về nhóm học viên MCC (Minimally competent candidates) và nhắc lại về cách thực hiện phương pháp Nedelsky: Mỗi người đánh giá đọc kỹ từng câu trắc nghiệm trong đề thi ngày 3/12/22, đặc biệt các câu bẫy (distractor) và xác định số lượng câu bẫy mà nhóm học viên MCC có khả năng loại trừ được. Với câu hỏi MCQ gồm 4 đáp án, chỉ có 1 đáp án đúng, 3 đáp án còn lại là câu bẫy, vậy có các khả năng:

Số lượng câu bẫy mà nhóm MCC có thể loại trừ được ( $x$ )	$Minimum\ Passing\ Level\ (MPL) = \frac{1}{4 - x}$
0	$\frac{1}{4 - 0} = \frac{1}{4} = 25\%$
1	$\frac{1}{4 - 1} = \frac{1}{3} = 33\%$
2	$\frac{1}{4 - 2} = \frac{1}{2} = 50\%$
3	$\frac{1}{4 - 3} = \frac{1}{1} = 100\%$

**Bước 4:** Các người đánh giá tiến hành đánh giá độc lập, sau đó tổng kết các kết quả theo bảng sau đó gửi kết quả về nhóm nghiên cứu. Nhóm nghiên cứu sẽ tổng hợp và tính điểm đều trung bình của cả 2 người đánh giá và xác định số câu đúng cần để đậu kỳ thi theo phương pháp Nedelsky.

## **Phần 2. Sau khi tổ chức thi, xác định số lượng và tỉ lệ rớt môn dựa trên chuẩn đậu theo phương pháp Nedelsky và so sánh với phương pháp Cohen**

**Bước 1:** Sau khi được Ban chủ nhiệm bộ môn Sản Phụ Khoa chấp thuận đề án, chúng tôi liên hệ Giáo vụ bộ môn để nhận kết quả số câu đúng kỳ thi kết thúc học phần Lý thuyết Sản Phụ Khoa ngày 3/12/2022.

**Bước 2:** Áp dụng điểm đậu theo phương pháp Nedelsky đã thiết lập trong phần 1, từ đó tính tỉ lệ rớt theo phương pháp Nedelsky.

**Bước 3:** Thiết lập điểm đậu theo phương pháp Cohen, từ đó tính tỉ lệ rớt theo phương pháp Cohen.

**Bước 4:** So sánh tỉ lệ rớt giữa phương pháp Nedelsky và phương pháp Cohen dùng phép kiểm định Chi-square và số p.

Các số liệu được phân tích thống kê bởi phần mềm R với thống kê mô tả và thống kê phân tích.

### **Đạo đức trong nghiên cứu**

Chúng tôi tiến hành nghiên cứu này được sự thông qua của Bộ môn Sản Phụ Khoa, Trường Đại học Y khoa Phạm Ngọc Thạch. Thu thập thông tin đầy đủ, trung thực, khách quan

theo mục tiêu nghiên cứu. Mọi thông tin của sinh viên đều được bảo mật, chỉ được sử dụng làm nghiên cứu (không dùng vào mục đích khác).

## KẾT QUẢ NGHIÊN CỨU

**Mục tiêu 1.** Tiến hành quy trình thiết lập tiêu chuẩn để xác định điểm đậu theo phương pháp Nedelsky cho lượng giá cuối kỳ Học phần Sản Phụ Khoa năm 4.

Sau khi tiến hành quy trình ở Phần 1, chúng tôi thu được kết quả độc lập của 3 người đánh giá cho số câu bẫy mà nhóm MCC có thể loại trừ được (x) từ đó tính Ngưỡng đậu tối thiểu (Minimal Passing Level – MLP) cho từng câu hỏi trong 50 câu của đề thi ngày 3/12/22 và cho toàn bộ đề thi (trong Bảng 1).

Bảng 1. Kết quả các đánh giá của rater theo phương pháp Nedelsky

Item	Rater 1		Rater 2		Rater 3	
	x	MPL	x	MPL	x	MPL
1	3	100%	3	100%	1	33%
2	3	100%	3	100%	1	33%
3	2	50%	2	50%	1	33%
4	0	25%	2	50%	3	100%
5	3	100%	2	50%	3	100%
6	3	100%	2	50%	3	100%
7	3	100%	1	33%	2	50%
8	0	25%	0	25%	2	50%
9	0	25%	2	50%	2	50%
10	2	50%	2	50%	3	100%
11	2	50%	2	50%	2	50%
12	2	50%	2	50%	2	50%
13	1	33%	1	33%	2	50%
14	3	100%	3	100%	2	50%
15	1	33%	2	50%	2	50%
16	2	50%	2	50%	2	50%
17	3	100%	3	100%	2	50%
18	3	100%	3	100%	2	50%
19	3	100%	3	100%	2	50%
20	3	100%	3	100%	2	50%
21	3	100%	3	100%	2	50%
22	3	100%	3	100%	2	50%
23	3	100%	3	100%	2	50%
24	1	33%	1	33%	2	50%
25	3	100%	2	50%	3	100%
26	2	50%	2	50%	3	100%
27	2	50%	2	50%	3	100%
28	2	50%	2	50%	1	33%
29	3	100%	3	100%	2	50%
30	1	33%	2	50%	3	100%
31	3	100%	3	100%	2	50%
32	3	100%	3	100%	2	50%
33	1	33%	2	50%	2	50%
34	3	100%	3	100%	3	100%
35	2	50%	3	100%	3	100%
36	3	100%	1	33%	2	50%
37	3	100%	3	100%	2	50%
38	1	33%	1	33%	2	50%
39	3	100%	3	100%	3	100%
40	2	50%	2	50%	2	50%
41	1	33%	2	50%	3	100%
42	2	50%	2	50%	2	50%
43	3	100%	3	100%	2	50%
44	3	100%	3	100%	2	50%
45	2	50%	2	50%	3	100%
46	2	50%	2	50%	2	50%
47	1	33%	1	33%	1	33%
48	2	50%	2	50%	2	50%
49	1	33%	2	50%	2	50%
50	2	50%	2	50%	2	50%
Total		<b>68%</b>		<b>66%</b>		<b>61%</b>

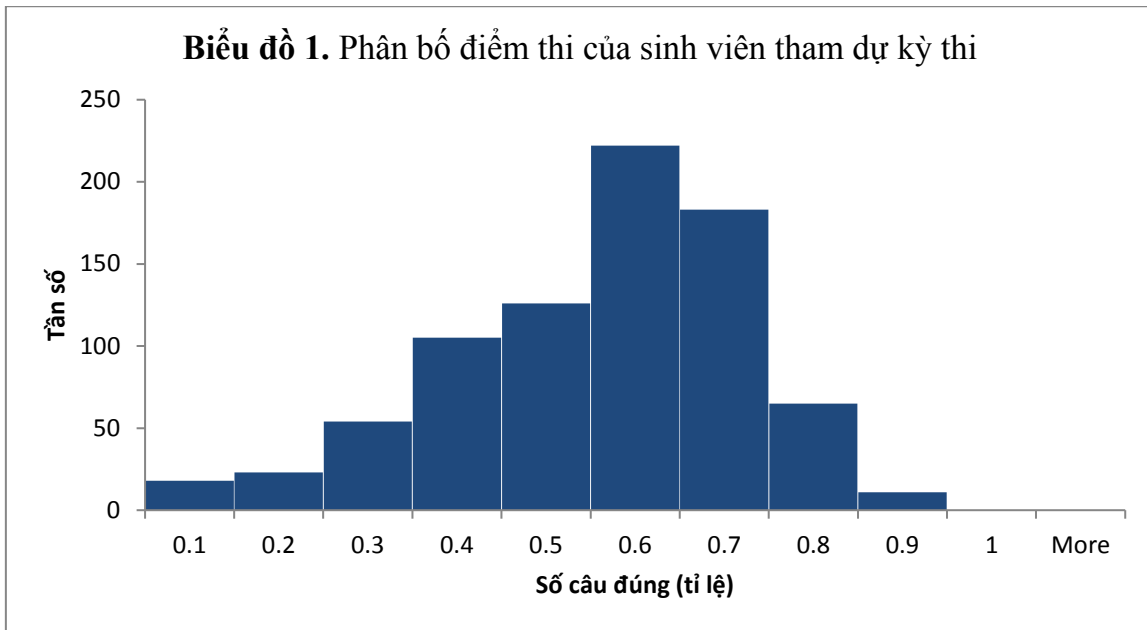
Xác định điểm đậu trung bình cho bài thi theo phương pháp Nedelsky:

$$\text{Điểm đậu} = \frac{\text{Rater 1} + \text{Rater 2} + \text{Rater 3}}{3} = \frac{68\% + 66\% + 61\%}{3} = 65\%$$

Tương ứng với số câu đúng có thể qua môn là 65% x 50 câu = **33 câu đúng**.

**Mục tiêu 2.** Xác định tỉ lệ rớt của sinh viên Y khoa năm thứ 4 tham gia kỳ thi kết thúc học phần Lý thuyết Sản Phụ Khoa theo phương pháp Nedelsky.

Kỳ thi chính thức kết thúc học phần Lý thuyết Sản Phụ Khoa trong năm học 2022 - 2023 dành cho sinh viên Y khoa năm thứ 4 diễn ra ngày 3/12/2022 có 808 học viên dự thi.



*Nhận xét:* Phân bố tỉ lệ số câu đúng của các thí sinh trong kỳ thi không tuân theo luật phân bố chuẩn với tần số cao nhất là tỉ lệ 60%; phần lớn tập trung trên mức 40%. Bách phân vị thứ 95 của tỉ lệ số câu đúng được xác định là 75%.

Với điểm đậu được xác định bởi phương pháp Nedelsky là 65%, tỉ lệ đậu của 808 thí sinh tham gia kì thi là 23% với 183 thí sinh đậu; ngược lại tỉ lệ rớt là 77% tương ứng với 625 thí sinh rớt kỳ thi này.

**Mục tiêu 3.** So sánh tỉ lệ rớt của sinh viên tham gia kỳ thi giữa hai phương pháp Nedelsky và phương pháp Cohen.

Điểm đậu được xác định theo phương pháp Cohen theo công thức:<sup>[3]</sup>

$$\text{Điểm đậu} = C + 0,6*(P - C) = 25\% + 0,6*(75\% - 25\%) = \mathbf{55\%}$$

- C là điểm dự kiến do đoán đúng ngẫu nhiên (theo phần trăm): 25%
- P là điểm của học viên ở bách phân vị thứ 95 (theo phần trăm): 75%

**Bảng 2.** Điểm đậu và tỉ lệ rớt theo phương pháp Cohen và Nedelsky

Số học viên	Phương pháp <b>Cohen</b>		Phương pháp <b>Nedelsky</b>	
	Điểm đậu	Số lượng rớt	Điểm đậu	Số lượng rớt
808	55%	416 (51%)	65%	625 (77%)

*Nhận xét:* Điểm đậu được xác định theo phương pháp Cohen cổ điển là 55% với số câu đúng cần để qua môn là 28/50, theo phương pháp Nedelsky là 65% với số câu đúng cần để qua môn là 33/50, cao hơn so với phương pháp Nedelsky.

**Bảng 3.** So sánh tỉ lệ rớt giữa hai phương pháp Cohen và Nedelsky

Số học viên	Số học viên rớt (Tỉ lệ)		Sai biệt	$\chi^2$	p-value
	Theo <b>Cohen</b>	Theo <b>Nedelsky</b>			
808	416 (51%)	625 (77%)	209 (26%)	117.9276	<0,00001

*Nhận xét:* Có sự khác biệt có ý nghĩa thống kê ( $p < 0,05$ ) giữa tỉ lệ sinh viên rớt theo chuẩn đầu xác định bởi phương pháp Nedelsky và theo phương pháp Cohen, với sự sai biệt tới 26%, chênh lệch lên đến 209 sinh viên.

## BÀN LUẬN

Phương pháp Nedelsky, được đặt tên theo tác giả sáng lập là Tiến sĩ Leo Nedelsky, là một phương pháp định lượng, chặt chẽ được giới thiệu từ năm 1954 nhằm mục đích thiết lập chuẩn đầu cho một bài kiểm tra. Phương pháp Nedelsky được áp dụng trong việc thiết lập chuẩn đầu cho các kỳ thi trắc nghiệm nhiều lựa chọn (MCQ) trong giáo dục y học, lĩnh vực mà việc giảng dạy và lượng giá năng lực được đặt làm trọng tâm. Phương pháp Nedelsky thuộc nhóm phương pháp tham chiếu dựa trên tiêu chí (criterion-referenced) và có những ưu điểm vượt trội hơn so với các phương pháp thiết lập chuẩn đầu tham chiếu trên quy chuẩn (norm-referenced) với điển hình là của tác giả Cohen (Bảng 2).

### Bàn luận về Defensibility giữa phương pháp Nedelsky và Cohen

Defensibility trong bối cảnh thiết lập chuẩn đầu đề cập đến khả năng cung cấp lí lẽ và bào chữa vững chắc cho chuẩn đầu được xác định. Điều này bao gồm việc chứng minh rằng chuẩn đầu được chọn là có tính giá trị, công bằng và biện luận bằng các tiêu chí khách quan và minh bạch.

**Bảng 4.** So sánh phương pháp thiết lập chuẩn đầu Nedelsky và Cohen

Đặc điểm	Phương pháp Nedelsky	Phương pháp Cohen
Xác định năng lực cụ thể	Xác định rõ năng lực hoặc kiến thức cần đánh giá.	Tập trung vào điểm số hoặc độ phân bố điểm của bài kiểm tra.
Hội đồng chuyên gia (Expert panel)	Có sự tham gia của nhóm chuyên gia có kinh nghiệm chuyên môn và giáo dục y học	Thường không cần sử dụng hội đồng chuyên gia, chủ yếu dựa trên sự hiểu biết của các nhà quản lý hoặc giảng viên.
Đánh giá các câu hỏi	Kiểm tra và đánh giá các câu hỏi MCQ theo nhiều tiêu chí.	Chủ yếu xem xét độ khó tổng thể của bài kiểm tra.
Phương pháp standard setting	Criterion-referenced (tuyệt đối)	Norm-referenced (tương đối)
Thông tin về năng lực	Xác định mức độ năng lực tối thiểu cần thiết (MCC) để thực hiện an toàn và hiệu quả trong lĩnh vực cụ thể.	Tập trung vào việc đánh giá đối với điểm số của toàn bộ nhóm thí sinh mà không xác định rõ mức độ năng lực cần thiết.
Sự phức tạp (Complexity)	Thường yêu cầu sự tham gia của một hội đồng chuyên gia và quá trình đánh giá chi tiết các câu hỏi.	Thường đơn giản hơn và không yêu cầu sự tham gia của hội đồng chuyên gia.
Tính giá trị (Validity) và công bằng (Fair)	Cung cấp chuẩn đầu có giá trị và công bằng, phản ánh mức độ năng lực tối thiểu cần thiết cho thực hành y khoa.	Không đảm bảo tính giá trị và công bằng do không phản ánh năng lực.
Defensibility	Có	Không



### **Phương pháp Nedelsky:**

- Defensibility: Phương pháp Nedelsky nhấn mạnh tầm quan trọng mạnh mẽ của một hội đồng chuyên gia đánh giá (panel of raters/assessors) và việc sử dụng các phương pháp standard setting đặc hiệu. Tiếp cận nghiêm ngặt này tăng cường defensibility bằng cách đảm bảo rằng chuẩn đầu được xác định thông qua một quy trình có hệ thống và dựa trên chứng cứ.

- Cơ sở lý luận (Rationale): Phương pháp này yêu cầu việc ghi chép rõ ràng về các bước thực hiện, các tiêu chí được sử dụng và sự đánh giá/phán xét của các chuyên gia trong quá trình standard setting. Các biên bản làm việc này cung cấp một cơ sở mạnh để bào chữa cho chuẩn đầu đã chọn.

- Tính minh bạch (Transparency): Sự tham gia của các chuyên gia và việc sử dụng các phương pháp thiết lập tiêu chuẩn đã được chứng minh góp phần làm tăng tính minh bạch, làm cho việc bào chữa và bảo vệ chuẩn đầu trở nên dễ dàng hơn đối với các bên liên quan.

### **Phương pháp Cohen:**

- Defensibility: Phương pháp Cohen không có tính defensible nếu so với Nedelsky, đặc biệt nếu chuẩn đầu được xác định bằng cách sử dụng các phương pháp tương đối hoặc dựa trên dữ liệu lịch sử. Thiếu quá trình minh bạch và không có sự tham gia của các chuyên gia (SME hoặc rater/assessor) trong việc xác định chuẩn đầu làm cho việc bào chữa cho chuẩn đầu dựa trên phương pháp Cohen trở nên rất khó khăn.

- Cơ sở lý luận (Rationale): Vì phương pháp này dựa vào tiêu chuẩn tương đối phụ thuộc vào kết quả thi của các thí sinh và thay đổi tùy theo năng lực của những nhóm thí sinh khác nhau nên lí lẽ cho xác định chuẩn đầu theo Cohen không minh bạch và không đủ mạnh để bào chữa cho quyết định chọn chuẩn đầu để đánh giá năng lực.

- Tính minh bạch (Transparency): Phương pháp Cohen thiếu tính rõ ràng vì không có sự tham gia của các chuyên gia với các biên bản đánh giá cụ thể cũng như tính tương đối của chuẩn đầu thay đổi theo năng lực của các nhóm thí sinh.

Tóm lại, tính biện hộ (defensibility) của quá trình thiết lập chuẩn đầu thường cao hơn với Phương pháp Nedelsky do tính hệ thống và dựa trên các chuyên gia. Việc sử dụng các phương pháp thiết lập chuẩn đầu dựa trên tiêu chí tuyệt đối và hội đồng chuyên gia làm tăng khả năng bào chữa mạnh cho chuẩn đầu đã chọn. Ngược lại, Phương pháp Cohen, đặc biệt nếu dựa vào các phương pháp tương đối (norm-referenced) hoặc không minh bạch nên không có tính biện hộ.

Ngoài ra, trong phương pháp Nedelsky có thể không có các chuyên gia SME (Subject Matter Expert) về đo lường tâm lý (psychometric) mà dùng các người đánh giá (rater hoặc assessor) đã được huấn luyện về lượng giá, thiết lập chuẩn đầu và cách thực hiện phương pháp Nedelsky. Điều này có ý nghĩa trong các bối cảnh giáo dục y học vẫn chưa có đội ngũ SME, có thể sử dụng được phương pháp thiết lập chuẩn đầu có tính defensible như Nedelsky.

Kết quả nghiên cứu của chúng tôi cho thấy rõ có sự khác biệt đáng kể về điểm đầu giữa phương pháp Nedelsky 65% và phương pháp Cohen 55%. Từ đó tỉ lệ rút được tính theo phương pháp Nedelsky cao hơn theo phương pháp Cohen có ý nghĩa thống kê. Với 209 thí sinh (chiếm

¼ số thí sinh tham gia kỳ thi) được đậu nếu xét theo phương pháp Cohen thì lại trở thành rớt theo phương pháp Nedelsky vì chưa đủ năng lực. Kết quả này đe dọa nghiêm trọng đến quá trình lượng giá dựa trên năng lực, vốn là cốt lõi trong giáo dục y khoa nhằm đào tạo ra các nhân viên y tế đủ chuyên môn để hành nghề. Nói cách khác, thiết lập chuẩn đậu nếu không dựa trên các phương pháp khoa học có tính biện hộ được (defensible) sẽ dẫn đến hậu quả thất bại trong việc đánh rớt các thí sinh không có năng lực – “Failure to Fail”.

#### **Nhược điểm của phương pháp Nedelsky:**

Phương pháp này chỉ áp dụng cho các câu hỏi trắc nghiệm nhiều lựa chọn MCQ, vì nó phụ thuộc vào việc đánh giá xác suất của các lựa chọn. Đây cũng là một sự đơn giản hóa quá mức. Trong bài thi của chúng tôi, các câu hỏi có bốn lựa chọn, chỉ có bốn giá trị có thể có cho điểm Nedelsky là 25%, 33%, 50%, 100%. Điều này trở nên rõ ràng hơn khi xem xét rằng hầu hết các câu hỏi thường có giá trị đúng phần trăm giữa 50% và 100%, và phản ánh một sự thật là không thể ước tính giá trị cụ thể ngoài các mốc trên với phương pháp Nedelsky. Rõ ràng, để trả lời một câu hỏi, cần nhiều khả năng hơn là chỉ loại trừ một hoặc hai distractors. Điều này là một trong những lý do mà phương pháp khác như Angoff thường được ưa chuộng và thay thế phương pháp này.

#### **KẾT LUẬN**

Từ những kết quả và phân tích trên cho thấy phương pháp tính điểm đậu tương đối như phương pháp Cohen không còn phù hợp để xác định chuẩn đậu để phân định người có năng lực và người không có năng lực trong giáo dục y học. Cần ứng dụng các phương pháp xác định chuẩn đậu phù hợp hơn dựa trên các chứng cứ khoa học như phương pháp Nedelsky trong bối cảnh chưa có nhiều các SME tại cơ sở giáo dục, trong thời gian xây dựng hội đồng các SME và nghiên cứu ứng dụng các phương pháp dựa trên tiêu chuẩn như Angoff hay Ebel với sự tham gia của các SME.

#### **Tài liệu tham khảo**

1. Searle J. (2000). Defining competency – the role of standard setting. *Medical Education*, 34:363-366.
2. Nabishah M, Nasri Ab, Salam A, and Ima Nirwana S (2011). Standard Setting for Assessment of Basic Medical Science Modules, *Procedia - Social and Behavioral Sciences*, 18:118-121.
3. Cohen-Schotanus J, van der Vleuten CP (2010). A standard setting method with the best performing students as point of reference: practical and affordable. *Med Teach*, 32(2):154-160.

#### **Acknowledgement**

Chúng tôi xin chân thành cảm ơn Giáo sư Hồ Hội là người Thầy đáng kính đã bình duyệt bản thảo bài báo này (review the manuscript) với những góp ý quý báu giúp hoàn thiện công trình nghiên cứu này.