

## VALIDITY AND RELIABILITY OF STANDARDIZED PATIENT ASSESSMENT FOR COMMUNICATION SKILLS FORMATIVE ASSESSMENT IN SECOND-YEAR MEDICAL STUDENTS

Cao Thi Ngoc Bich\*, Bui Phan Quynh Phuong, Trinh Trung Tien, Nguyen Minh Luan, Tran Tu Trinh, Nguyen Thi Lan Anh, Nguyen Thi Kieu Trang, Nguyen Anh Duy, Chau Thi Kieu Chinh, Nguyen Thi Thanh Thanh

Pham Ngoc Thach University of Medicine - 2 Đường Quang Trung, Hoa Hung ward, Ho Chi Minh city, Vietnam

Received: 11/02/2026

Revised: 01/03/2026; Accepted: 23/04/2026

### ABSTRACT

**Objective:** To determine the validity and reliability of standardized patients in assessing the communication skills formative assessment of second year Viet Duc medical students Pham Ngoc Thach University of Medicine.

**Methods:** A cross-sectional study was conducted on 50 second-year Viet Duc medical students, 5 standardized patients, and 2 lecturers at the Center for Elaboration Competency and Innovation in Clinical Simulation at Pham Ngoc Thach University of Medicine, from June 2025 to June 2026. Standardized patients and lecturers independently assessed students' communication skills using an "abdominal pain" scenario and a checklist with confirmed content validity by six experts (CVI = 0.96, CVR = 0.92). Validity was assessed using Spearman's correlation coefficients and the Mann-Whitney U test. Inter-rater reliability (Inter-RR) and intra-rater reliability (Intra-RR) were evaluated using intraclass correlation coefficient.

**Results:** Standardized patients showed strong correlations with lecturers scores in communication skills ( $r = 0.7-1.0$ ;  $p < 0.01$ ). There were no statistically significant differences in mean scores between standardized patients and lecturers ( $p > 0.05$ ). Inter-rater reliability ranged from moderate to excellent (Inter-RR = 0.57-0.91). Intra-rater reliability of standardized patients after four weeks was good to excellent (Intra-RR = 0.72-0.95).

**Conclusion:** Standardized patients are a valid and reliable assessment tool for evaluating the communication skills formative assessment of second year Viet Duc medical students.

**Keywords:** Standardized patient, communication skills, validity, reliability.

---

\*Corresponding author

Email: bichctn@pnt.edu.vn Phone: (+84) 984 037 426 DOI: 10.52163/yhc.v67i4.4912



# TÍNH GIÁ TRỊ VÀ ĐỘ TIN CẬY CỦA LƯỢNG GIÁ QUÁ TRÌNH CỦA NGƯỜI BỆNH CHUẨN CHO SINH VIÊN Y NĂM HAI VỀ KỸ NĂNG GIAO TIẾP

Cao Thị Ngọc Bích\*, Bùi Phan Quỳnh Phương, Trịnh Trung Tiến, Nguyễn Minh Luân, Trần Tú Trinh, Nguyễn Thị Lan Anh, Nguyễn Thị Kiều Trang, Nguyễn Anh Duy, Châu Thị Kiều Chinh, Nguyễn Thị Thanh Thanh<sup>1</sup>

Trường Đại học Y khoa Phạm Ngọc Thạch - 2 Đường Quang Trung, phường Hòa Hưng, thành phố Hồ Chí Minh, Việt Nam

Ngày nhận bài: 11/02/2026

Ngày chỉnh sửa: 01/03/2026; Ngày duyệt đăng: 23/04/2026

## TÓM TẮT

**Mục tiêu:** Xác định tính giá trị và độ tin cậy của người bệnh chuẩn trong lượng giá quá trình kỹ năng giao tiếp cho SV y Việt Đức năm hai Trường Đại học Y khoa Phạm Ngọc Thạch.

**Đối tượng và phương pháp:** Nghiên cứu cắt ngang mô tả thực hiện trên 50 sinh viên y Việt Đức năm hai, 5 người bệnh chuẩn và 2 giảng viên tại Trung tâm Kỹ năng và Sáng tạo Mô phỏng Lâm sàng, Trường Đại học Y khoa Phạm Ngọc Thạch, từ tháng 6/2025 đến 6/2026. Người bệnh chuẩn và giảng viên đánh giá độc lập kỹ năng giao tiếp của sinh viên qua tình huống “đau bụng” bằng bảng kiểm đã thẩm định giá trị nội dung bởi 6 chuyên gia (CVI = 0,96, CVR = 0,92). Tính giá trị xác định bằng hệ số tương quan Spearman và kiểm định Mann-Whitney U. Độ tin cậy liên giám khảo được đánh giá bằng hệ số tương quan Inter-RR và nội giám khảo là Intra-RR.

**Kết quả:** Điểm đánh giá của người bệnh chuẩn tương quan mạnh với giảng viên trong các kỹ năng giao tiếp ( $r = 0,7-1,0$ ;  $p < 0,01$ ). Không có sự khác biệt có ý nghĩa thống kê về điểm trung bình giữa người bệnh chuẩn và giảng viên ( $p > 0,05$ ). Độ tin cậy liên giám khảo giữa giảng viên và người bệnh chuẩn đạt mức trung bình đến rất tốt (Inter-RR = 0,57-0,91). Độ tin cậy nội giám khảo của người bệnh chuẩn sau 4 tuần đạt mức tốt trở lên (Intra-RR = 0,72-0,95).

**Kết luận:** Người bệnh chuẩn là công cụ đánh giá có tính giá trị và độ tin cậy cao phù hợp cho lượng giá quá trình kỹ năng giao tiếp sinh viên y Việt Đức năm hai.

**Từ khóa:** Người bệnh chuẩn, kỹ năng giao tiếp, tính giá trị, độ tin cậy.

## 1. ĐẶT VẤN ĐỀ

Lượng giá sinh viên là thành phần cốt lõi của quá trình đào tạo y khoa, mục tiêu là đánh giá năng lực và định hướng việc học. Lượng giá kỹ năng giao tiếp (KNGT) gặp nhiều thách thức do tính chất phức tạp và trừu tượng của các kỹ năng mềm. Sử dụng người bệnh là phương pháp đã được chứng minh hiệu quả trên thế giới. Người bệnh chuẩn (NBC) có tác động tích cực trong việc phát triển KNGT và hiệu quả hơn phương pháp đóng vai trong việc nâng cao sự tự tin cho sinh viên (SV) [1-2].

Các nghiên cứu cho thấy NBC có thể đánh giá SV với độ chính xác cao so với giảng viên (GV). Bolstad A.L và cộng sự ghi nhận độ tin cậy NBC và GV dao động từ mức trung bình đến khá tốt [2]. Nghiên cứu tại Mỹ trên SV y khoa năm cuối cũng cho thấy sự nhất quán từ trung bình đến tốt giữa đánh giá của đồng nghiệp, tự đánh giá, NBC và GV đối với KNGT [3]. Tại Việt Nam, việc áp dụng NBC đang dần phổ biến nhưng các nghiên cứu về độ tin cậy của đối tượng này trong vai trò đánh giá SV còn hạn chế. Nghiên cứu gần đây của Bùi Thị Hiền và cộng sự thực hiện tại Trường Đại học VinUni cho thấy sự đồng thuận thấp giữa đánh giá của NBC và đồng nghiệp đối với nhân viên y tế [4]. Điều này đặt ra câu hỏi NBC khi áp dụng cho đánh giá SV y Việt Đức năm

hai có đạt được tính giá trị và độ tin cậy cần thiết để lượng giá quá trình KNGT hay không? Vì vậy, chúng tôi tiến hành nghiên cứu này nhằm xác định tính giá trị và độ tin cậy trong lượng giá quá trình của NBC cho SV y Việt Đức năm hai về KNGT tại Trường Đại học Y khoa Phạm Ngọc Thạch.

## 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

### 2.1. Đối tượng nghiên cứu

Bao gồm 50 SV y Việt Đức năm hai, 5 NBC và 2 GV tại Trung tâm Kỹ năng và Sáng tạo Mô phỏng Lâm sàng, Trường Đại học Y Khoa Phạm Ngọc Thạch.

- Tiêu chuẩn chọn vào:

+ NBC tham gia đầy đủ các buổi huấn luyện, có khả năng diễn xuất kịch bản tốt, có kỹ năng quan sát mọi hành vi của SV, có khả năng nhớ lại các hành vi của SV trong trạm OSCE, hoàn thành chấm điểm chính xác bảng kiểm và có khả năng phản hồi nhận xét cho SV về KNGT.

+ GV có trình độ thạc sĩ trở lên, kinh nghiệm giảng dạy mô phỏng trên 10 năm và có các chứng chỉ giáo dục y khoa quốc tế (CEP, AFDP).

\*Tác giả liên hệ

Email: bichctn@pnt.edu.vn Điện thoại: (+84) 984 037 426 DOI: 10.52163/yhc.v67i4.4912

+ SV năm 2 thuộc Khoa Y Việt Đức đã học lý thuyết về KNGT thầy thuốc - người bệnh và khai thác bệnh sử.

- Tiêu chuẩn loại trừ: các đối tượng không đồng ý tham gia hoặc có xung đột lợi ích.

**2.2. Phương pháp nghiên cứu**

- Thiết kế nghiên cứu: cắt ngang mô tả.

- Thời gian và địa điểm nghiên cứu: từ tháng 6/2025 đến tháng 6/2026 tại Trung tâm Kỹ năng và Sáng tạo Mô phỏng Lâm sàng, Trường Đại học Y Khoa Phạm Ngọc Thạch.

Cỡ mẫu và phương pháp chọn mẫu: cỡ mẫu gồm 5 NBC, 2 GV và 50 SV. Trong cỡ mẫu trên, 5 NBC được chọn ngẫu nhiên trong toàn bộ 14 NBC đang làm việc tại Trung tâm Kỹ năng và Sáng tạo Mô phỏng Lâm sàng; 2 GV được chọn ngẫu nhiên trong 8 GV thỏa mãn tiêu chuẩn chọn mẫu; 50 SV chọn toàn bộ SV y Việt Đức năm hai.

- Công cụ thu thập số liệu: bảng kiểm KNGT tình huống “tiếp cận người bệnh đau bụng” gồm 20 mục, đã được thẩm định tính giá trị nội dung bởi 6 chuyên gia với chỉ số CVI = 0,96, CVR = 0,92 theo quy trình chuẩn hóa của Yusoff M.S.B [5].

- Biến số chính: kết quả đánh giá GV là tiêu chuẩn vàng để so sánh đánh giá của NBC. ICC (hệ số tương quan nội bộ) được sử dụng để đánh giá sự đồng thuận giữa những người đánh giá. Inter-RR (inter-rater reliability) là độ tin cậy liên giám khảo, dùng để đánh giá giữa những người đánh giá khác nhau giữa GV và NBC. Intra-RR (intra-rater reliability) là độ tin cậy lặp lại giữa 2 lần đánh giá của NBC. Độ tin cậy của các hệ số tương quan ICC, Inter-RR và Intra-RR được xác định qua các mức độ: < 0,5 (tin cậy thấp), 0,5-0,75 (tin cậy trung bình), > 0,75-0,9 (tin cậy tốt) và > 0,9 (tin cậy xuất sắc) [6]. Hệ số tương quan Spearman (r) đánh giá tính giá trị giữa điểm số của 2 đối tượng đánh giá khác nhau, với |r| < 0,3 (tương quan yếu), |r| = 0,3-0,5 (trung bình), |r| > 0,5 (tương quan mạnh), với r âm là tương quan nghịch và r dương là tương quan thuận.

- Quy trình nghiên cứu: huấn luyện NBC qua 6 bước, mỗi bước tối thiểu 3 giờ: bước 1 (huấn luyện NBC về kịch bản, bảng kiểm và cách chấm điểm); bước 2 (NBC thực hành đánh giá trên video tình huống mẫu); bước 3 (NBC được chọn thực hành làm đánh giá viên tại trạm OSCE); bước 4 (GV huấn luyện NBC kỹ năng phản hồi); bước 5 (NBC cung cấp phản hồi cho SV tại trạm OSCE, đồng thời GV phỏng vấn SV về cảm nhận đối với phản hồi của NBC); bước 6 (điền bảng kiểm chính xác).

Mỗi SV thực hiện tình huống giao tiếp với NBC trong 6 phút tại trạm OSCE. NBC chấm điểm trực tiếp sau khi kết thúc trạm (lần 1). GV chấm điểm độc lập qua video. Sau 4 tuần, NBC xem lại video và chấm điểm lần 2 để đánh giá độ tin cậy nội giám khảo.

Mỗi GV đánh giá 25 SV, mỗi NBC đánh giá 10 SV.

- Xử lý số liệu: nhập liệu bằng phần mềm Excel, phân tích số liệu bằng phần mềm STATA 17.0. Tính giá trị được xác định bằng hệ số tương quan Spearman và kiểm định Mann-Whitney U so sánh điểm trung bình giữa NBC và GV. Độ tin cậy liên giám khảo và nội giám khảo được đánh giá bằng hệ số tương quan nội bộ.

**2.3. Đạo đức nghiên cứu**

Nghiên cứu đã được thông qua Hội đồng Đạo đức trong nghiên cứu y sinh học, Trường Đại học Y khoa Phạm Ngọc Thạch (số 1374/TĐHYKPNT-HĐĐĐ ngày 18/6/2025). Đối tượng tham gia được cung cấp thông tin và ký bản đồng thuận tham gia. Các dữ liệu thu thập được được bảo mật và chỉ phục vụ cho mục đích nghiên cứu.

**3. KẾT QUẢ NGHIÊN CỨU**

**Bảng 1. Đặc điểm chung của nhóm tham gia nghiên cứu**

Đặc điểm		NBC (n = 5)	GV (n = 2)	SV (n = 50)
Giới tính	Nam	0	1 (50%)	34 (68%)
	Nữ	5 (100%)	1 (50%)	16 (32%)
Tuổi	$\bar{X} \pm SD$	54,0 ± 11,3	37 ± 2,8	20,0 ± 0,8
	Min-max	39-64	35-39	19-23
Trình độ học vấn	Đại học	4 (80,0%)	0	50 (100%)
	Sau đại học	1 (20,0%)	2 (100%)	0
Nghề nghiệp	Giáo viên, giảng viên	1 (20%)	2 (100%)	0
	Diễn viên	2 (40%)	0	0
	Sinh viên	0	0	50 (100%)
	Hưu trí	2 (40%)	0	0

NBC có độ tuổi trung bình là 54,0 ± 11,3; 100% là phụ nữ, có trình độ đại học trở lên, nghề nghiệp là giáo viên, hưu trí (cựu giáo viên) và diễn viên nghiệp dư. GV có độ tuổi trung bình là 37 ± 2,8; nam và nữ bằng nhau. SV có độ tuổi trung bình là 20 ± 0,8 và 68% là nam.

**Bảng 2. Kết quả phân tích tính giá trị của NBC đánh giá SV so với GV**

Loại giao tiếp	NBC	NBC	GV	Spearman		Mann-Whitney
				r	p	p
Thiết lập mối quan hệ	NBC 1	4,8 ± 0,1	4,7 ± 0,2	0,7	0,01	0,61
Giao tiếp không lời		3,7 ± 0,2	3,6 ± 0,2	0,6	0,09	0,89
Giao tiếp chính hỏi bệnh		8,7 ± 0,2	8,6 ± 0,3	1	< 0,001	0,87
Mức độ hài lòng tổng thể		3,4 ± 0,2	3,4 ± 0,2	0,4	0,24	1,00
Thiết lập mối quan hệ	NBC 2	4,9 ± 0,1	4,9 ± 0,1	1	< 0,001	0,32
Giao tiếp không lời		4	4	-	-	-
Giao tiếp chính hỏi bệnh		9,3 ± 0,2	9,2 ± 0,1	0,7	0,01	0,62
Mức độ hài lòng tổng thể		3,7 ± 0,1	3,8 ± 0,1	0,5	0,14	1,00
Thiết lập mối quan hệ	NBC 3	4,7 ± 0,3	4,7 ± 0,3	0,8	0,01	0,63
Giao tiếp không lời		4	4	-	-	-
Giao tiếp chính hỏi bệnh		9,3 ± 0,2	8,7 ± 0,3	1	< 0,001	0,09
Mức độ hài lòng tổng thể		3,7 ± 0,1	3,7 ± 0,2	0,55	0,11	1

Loại giao tiếp	NBC	NBC	GV	Spearman		Mann-Whitney
				r	p	p
Thiết lập mối quan hệ	NBC 4	4,7 ± 0,2	4,7 ± 0,2	1	< 0,001	0,62
Giao tiếp không lời		3,9 ± 0,1	3,9 ± 0,1	1	< 0,001	1
Giao tiếp chính hỏi bệnh		9,2 ± 0,2	9 ± 0,3	1	< 0,001	0,56
Mức độ hài lòng tổng thể		3,9 ± 0,1	3,9 ± 0,1	0,2	0,55	1
Thiết lập mối quan hệ	NBC 5	4,8 ± 0,1	4,8 ± 0,1	0,6	0,06	0,34
Giao tiếp không lời		3,7 ± 0,2	4,6 ± 0,2	1	< 0,001	1
Giao tiếp chính hỏi bệnh		9,6 ± 0,2	9,5 ± 0,2	0,9	< 0,001	0,66
Mức độ hài lòng tổng thể		3,7 ± 0,2	3,7 ± 0,2	0,5	0,13	1

Ghi chú: “-”: Không ước lượng được p-value do dữ liệu không có sự biến thiên giữa hai nhóm (NBC và GV đều cho cùng 1 điểm).

Bảng 2 cho thấy sự tương quan thuận ở mức mạnh (r = 0,7-1,0; p < 0,01) ở hầu hết các mục đánh giá của NBC và GV. Không có sự khác biệt về điểm đánh giá giữa NBC và GV ở tất cả các mục.

**Bảng 3. Kết quả phân tích độ tin cậy liên giám khảo (NBC so sánh với GV)**

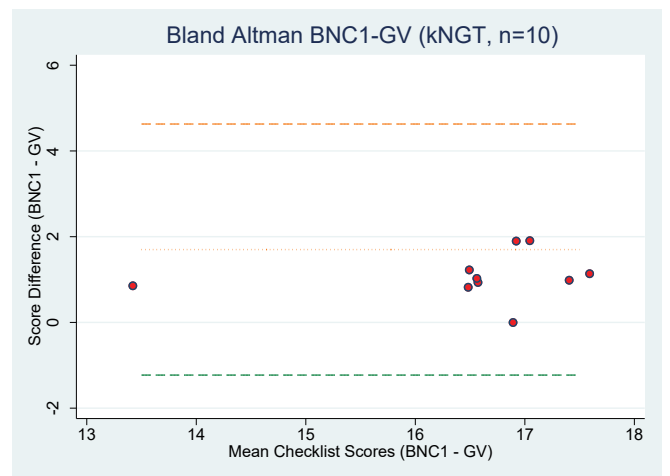
Loại giao tiếp	NBC	Độ tin cậy liên giám khảo	
		Inter-RR (KTC 95%)	p
Thiết lập mối quan hệ	NBC1	0,75 (0,31-0,93)	0,004
Giao tiếp không lời		0,57 (-0,06-0,87)	0,039
Giao tiếp chính hỏi bệnh		0,91 (0,71-0,97)	< 0,001
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC2	0	0,50
Giao tiếp không lời		-	-
Giao tiếp chính hỏi bệnh		0,75 (0,31-0,93)	0,004
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC3	0,85 (0,49-0,96)	0,001
Giao tiếp không lời		-	-
Giao tiếp chính hỏi bệnh		0,59 (-0,085-0,88)	0,003
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC4	0,76 (0,31-0,93)	0,004
Giao tiếp không lời		1	-
Giao tiếp chính hỏi bệnh		0,82 (0,44-0,95)	0,001
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC5	0,57	0,026
Giao tiếp không lời		1	-
Giao tiếp chính hỏi bệnh		0,82 (0,44-0,95)	0,001
Mức độ hài lòng tổng thể		1	-

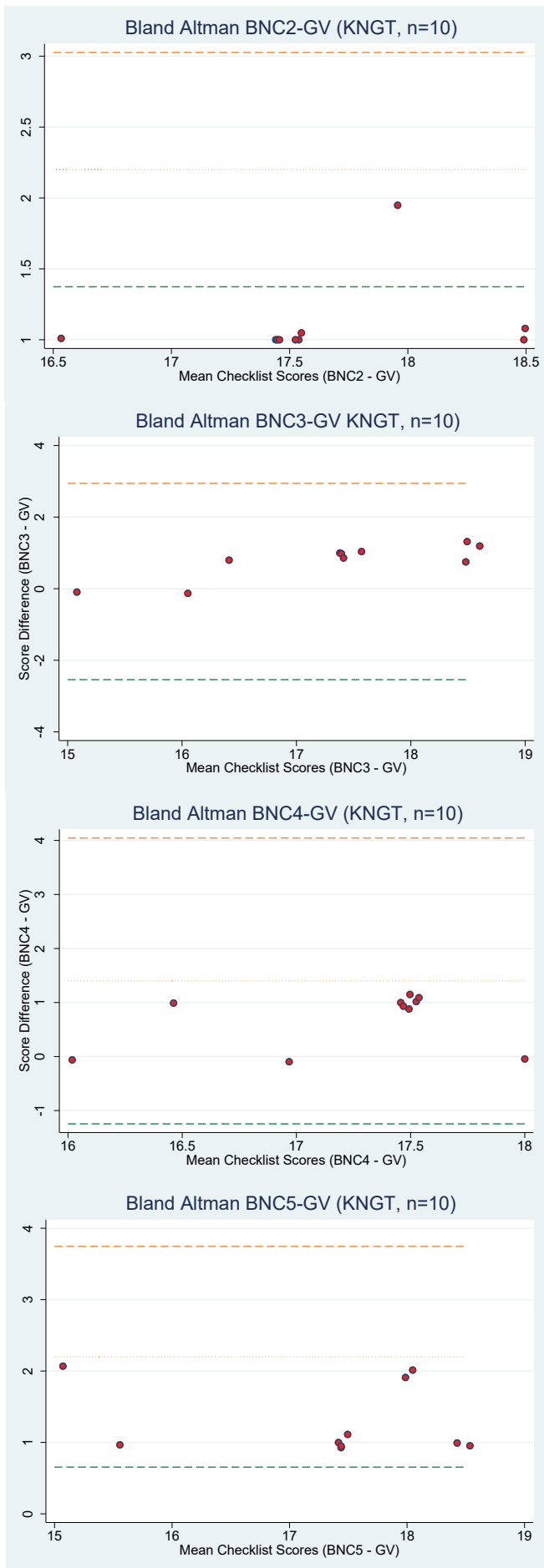
Bảng 3 cho thấy độ tin cậy liên giám khảo đạt mức trung bình đến rất tốt (Inter-RR = 0,57-0,91). Inter-RR rất cao ở KNGT chính hỏi bệnh (0,75-0,91). Tuy nhiên, ở kỹ năng thiết lập mối quan hệ, NBC 2 có độ đồng thuận thấp (Inter-RR = 0).

**Bảng 4. Độ tin cậy nội giám khảo (NBC đánh giá lặp lại sau 4 tuần)**

Loại giao tiếp	NBC	Độ tin cậy nội giám khảo	
		Intra-RR (KTC 95%)	p
Thiết lập mối quan hệ	NBC 1	0,86 (0,47-0,96)	0,004
Giao tiếp không lời		0,72 (-0,13-0,93)	0,039
Giao tiếp chính hỏi bệnh		0,95 (0,83-0,98)	< 0,001
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC 2	0 (-3,02-0,75)	0,50
Giao tiếp không lời		-	-
Giao tiếp chính hỏi bệnh		0,86 (0,47-0,96)	0,004
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC 3	0,91 (0,66-0,98)	0,001
Giao tiếp không lời		-	-
Giao tiếp chính hỏi bệnh		0,74 (-0,18-0,94)	0,003
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC 4	0,86 (0,47-0,96)	0,004
Giao tiếp không lời		1	-
Giao tiếp chính hỏi bệnh		0,89 (0,61-0,97)	0,001
Mức độ hài lòng tổng thể		1	-
Thiết lập mối quan hệ	NBC 5	0,73 (0,03-0,92)	0,026
Giao tiếp không lời		1	-
Giao tiếp chính hỏi bệnh		0,89 (0,61-0,97)	0,001
Mức độ hài lòng tổng thể		1	-

Độ tin cậy nội giám khảo đạt mức tốt ở hầu hết mục đánh giá của NBC (Intra-RR > 0,72, p < 0,001). Tuy nhiên, NBC 2 có độ tin cậy trong thiết lập mối quan hệ thấp với Intra-RR = 0.





Biểu đồ 1. Phân bố điểm đánh giá KNGT của 5 NBC và 2 GV

Các biểu đồ Bland Altman cho thấy mức độ đồng thuận tương đối tốt giữa đánh giá của NBC và GV trong lượng giá KNGT. Phần lớn các điểm nằm trong giới hạn đồng thuận 95%, không có sự sai lệch hệ thống theo mức điểm trung bình.

#### 4. BÀN LUẬN

##### 4.1. Tính giá trị trong lượng giá của NBC cho SV y Việt Đức năm hai về KNGT

Nghiên cứu chúng tôi có tương quan thuận mạnh giữa điểm đánh giá của NBC và GV ( $r = 0,7-1,0$ ) và không có sự khác biệt về điểm đánh giá giữa NBC và GV ở tất cả các mục. Kết quả này tương đồng với Shirazi M và cộng sự, với điểm đánh giá giữa NBC và đánh giá viên đạt tương quan mạnh ( $r = 0,81$ ), NBC có thể đánh giá các KNGT chính xác không thua kém chuyên gia [1]. Tương tự, Zanten V.M và cộng sự cũng ghi nhận mối tương quan cao giữa điểm số của NBC và giám khảo trong các kỳ thi cấp chứng chỉ hành nghề tại Mỹ [7].

##### 4.2. Độ tin cậy trong lượng giá của NBC cho SV y Việt Đức năm hai về KNGT

Độ tin cậy liên giám khảo trong nghiên cứu này đạt mức từ 0,57-0,91. Đáng chú ý, kỹ năng “giao tiếp chính hỏi bệnh” đạt độ tin cậy cao nhất ( $ICC > 0,82$ ). Kết quả này khả quan hơn so với nghiên cứu của Bolstad A.L và cộng sự đã báo cáo hệ số  $ICC = 0,42-0,755$  khi so sánh điểm của NBC và đánh giá viên hai trên đối tượng điều dưỡng quốc tế; sự chênh lệch trong đánh giá thường xuất phát từ góc nhìn khác nhau, nhóm chuyên gia thường tập trung vào việc hoàn thành nhiệm vụ trong khi NBC thường bị ảnh hưởng bởi sự đồng cảm và cảm xúc trong tương tác [2].

Điều này được thể hiện rõ trong nghiên cứu của chúng tôi ở kỹ năng “thiết lập mối quan hệ” có hệ số  $Inter-RR = 0$ , cho thấy sự bất đồng thuận hoàn toàn giữa NBC với GV. Phân tích sâu hơn dựa trên quan điểm của Bolstad A.L và cộng sự [2], có thể thấy rằng các kỹ năng mềm như “tạo thiện cảm” mang tính chủ quan cao. NBC có thể đánh giá dựa trên cảm nhận “được quan tâm” thật sự, trong khi GV đánh giá dựa trên các hành vi quan sát được (như chào hỏi, cười). Kết quả này tương đồng với Bùi Thị Hiền và cộng sự thực hiện tại Trường đại học VinUni cho thấy sự đồng thuận kém giữa đánh giá viên và giữa NBC [4].

Về độ tin cậy nội giám khảo, NBC đánh giá sau 4 tuần cho thấy sự ổn định cao ( $Intra-RR = 0,72-0,95$ ). Sự nhất quán này phù hợp với nhận định của Shirazi M và cộng sự, với NBC được đào tạo bài bản có thể thực hiện vai diễn và chấm điểm một cách kiên định theo thời gian [1]. Sự ổn định này là minh chứng cho hiệu quả của chương trình NBC tại Trung tâm Kỹ năng và Sáng tạo Mô phỏng Lâm sàng của chúng tôi, tương tự các nguyên tắc đảm bảo chất lượng mà Shirazi M và cộng sự đã thực hiện.

Việc khẳng định tính giá trị và độ tin cậy của NBC có ý nghĩa quan trọng trong việc thay đổi phương pháp đánh giá từ truyền thống sang hiện đại. Theo Cahyono B.D và cộng sự, phương pháp giảng dạy mô phỏng với NBC hiệu quả hơn đóng vai trò trong việc nâng cao sự tự tin và KNGT của SV [8]. Việc sử dụng NBC trong lượng giá quá trình giúp SV nhận được phản hồi đa chiều: từ góc độ chuyên

môn của GV và từ góc độ trải nghiệm của người bệnh. Đồng thời cũng hiệu quả hơn vì có thể giảm khối lượng công việc của GV.

Bên cạnh kết quả đạt được, nghiên cứu cũng có một số hạn chế nhất định về cỡ mẫu nhỏ và kịch bản còn hạn chế trong một tình huống mô phỏng “đau bụng”, chưa đa dạng của các tình huống lâm sàng trong thực tế, do đó khả năng áp dụng của nghiên cứu có thể bị hạn chế.

## 5. KẾT LUẬN

NBC là công cụ đánh giá có tính giá trị và độ tin cậy tốt trong lượng giá quá trình KNGT của SV y Việt Đức năm hai Trường Đại học Y khoa Phạm Ngọc Thạch. NBC được đào tạo có thể được sử dụng như một công cụ hợp lệ để lượng giá quá trình KNGT của SV y Việt Đức, giúp lượng giá năng lực SV theo chủ đề cụ thể, đồng thời cũng hiệu quả hơn về chi phí và giảm khối lượng công việc cho GV.

## 6. TÀI LIỆU THAM KHẢO

- [1] Shirazi M, Labaf A, Monjazebi F et al. Assessing medical students' communication skills by the use of standardized patients: emphasizing standardized patients' quality assurance. *Acad Psychiatry*, 2014, 38 (3): 354-360. doi: 10.1007/s40596-014-0066-2.
- [2] Bolstad A.L, Sørli V, Berntsen K et al. Reliability of standardized patients used in a communication study on international nurses in the United States of America. *Nurs Health Sci*, 2012, 14 (1): 67-73. doi: 10.1111/j.1442-2018.2011.00667.x.
- [3] Talwalkar J.S, Murtha T.D, Prozora S, Fortin A.H, Morrison L.J, Ellman M.S. Assessing advanced communication skills via objective structured clinical examination: a comparison of faculty versus self, peer, and standardized patient assessors. *Teaching and Learning in Medicine*, 2020, 32 (3): 294-307. doi: 10.1080/10401334.2019.1704763.
- [4] Bui T.H, Nguyen T.H.H, Tran T.D, Hoang V.L. Communication skills in standardized patient and peer assessment: a cross-sectional study. *Journal of Nursing Science*, 2025, 8 (1): 154-162. doi: 10.54436/jns.2025.01.887.
- [5] Yusoff M.S.B. ABC of content validation and content validity index calculation. *Education in Medicine Journal*, 2019, 11 (2): 49-54. doi: 10.21315/eimj2019.11.2.6.
- [6] Landis J.R, Koch G.G. The measurement of observer agreement for categorical data. *Biometrics*, 1977, 33 (1): 159-174. doi: 10.2307/2529310.
- [7] Zanten V.M, Boulet J.R, McKinley D. Using standardized patients to assess the interpersonal skills of physicians: six years' experience with a high-stakes certification examination. *Health Commun*, 2007, 22 (3): 195-205.
- [8] Cahyono B.D, Zuhroidah I, Sujarwadi M. Effectiveness of standardized patients and role play simulation methods in improving self-confidence and therapeutic communication skills of nursing students. *Nurse and Health Jurnal Keperawatan*, 2020, 9 (1): 9-15.