Published in final edited form as:

J Am Podiatr Med Assoc. 2018 March ; 108(2): 145–150. doi:10.7547/15-175.

Differences in Faculty and Standardized Patient Scores on Professionalism for Second-Year Podiatric Medical Students During a Standardized Simulated Patient Encounter

James M. Mahoney, DPM^{*}, Vassilios Vardaxis, PhD[†], Noreen Anwar, DPM[‡], and Jacob Hagenbucher, DPM[‡]

^{*}Department of Podiatric Medicine, College of Podiatric Medicine and Surgery, Des Moines University, Des Moines, IA.

[†]Department of Research, College of Podiatric Medicine and Surgery, Des Moines University, Des Moines, IA.

[‡]College of Podiatric Medicine and Surgery, Des Moines University, Des Moines, IA. Dr. Anwar is now with St. Vincent's Medical Center, Jacksonville, FL. Dr. Hagenbucher is now with Legacy Health, Portland, OR.

Abstract

Background: This study examined the differences between faculty and trained standardized patient (SP) evaluations on student professionalism during a second-year podiatric medicine standardized simulated patient encounter.

Methods: Forty-nine second-year podiatric medicine students were evaluated for their professionalism behavior. Eleven SPs performed an assessment in real-time, and one faculty member performed a secondary assessment after observing a videotape of the encounter. Five domains were chosen for evaluation from a validated professionalism assessment tool.

Results: Significant differences were identified in the professionalism domains of “build a relationship” ($P = .008$), “gather information” ($P = .001$), and share information ($P = .002$), where the faculty scored the students higher than the SP for 24.5%, 18.9%, and 26.5% of the cases, respectively. In addition, the faculty scores were higher than the SP scores in all of the “gather information” subdomains; however, the difference in scores was significant only in the “question appropriately” ($P = .001$) and “listen and clarify” ($P = .003$) subdomains.

Conclusions: This study showed that professionalism scores for second-year podiatric medical students during a simulated patient encounter varied significantly between faculty and SPs. Further consideration needs to be given to determine the source of these differences.

Corresponding author: James M. Mahoney, DPM, College of Podiatric Medicine and Surgery, Des Moines University, 3200 Grand Ave, Des Moines, IA 50312-4198. (james.mahoney@dmu.edu).

Financial Disclosure: None reported.

Conflict of Interest: None reported.

Publisher's Disclaimer: Copyright of Journal of the American Podiatric Medical Association is the property of American Podiatric Medical Association, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

As public trust in physicians continues to deteriorate,¹ professionalism has become a topic of emphasis in the curriculum of academic health centers. Although there is no clear consensus on its definition, most centers agree that professionalism encompasses the elements of interpersonal skills, empathy, and respect for others.² Ha and Long-necker's³ systematic review found that effective doctor-patient communication regulates the patient's emotional state, facilitates better patient comprehension, better identifies patient needs, and provides for better patient adherence to treatment. Yet, empathetic communication, one of the essential elements of professionalism, continues to decline during the education of medical students and residents, which compromises professionalism training and may lead to lower quality of health care.⁴ Assessments of fourth-year medical students on standardized patient (SP) cases showed that higher scores in overall clinical competency were associated with higher scores in interpersonal and communication skills,⁵ as well as a positive correlation between problems in residency and the frequency of negative comments in the Medical Student Performance Evaluation, which is completed by faculty and a required component of the residency application.⁶ Furthermore, Papadakis et al⁷ found that unprofessional behavior in medical school increased the chance of disciplinary action by medical boards by a factor of three.

Academic health centers recognize the need to teach and model professionalism in their curriculum.¹ However, the persons who perform the actual assessment of professional behavior can vary. Based on their education and experience, physicians who are also faculty in medical academic centers are generally regarded as the gold standard to evaluate professionalism.^{8,9} Others insist that it must be the patient.^{10,11} Standardized patients, or individuals trained to portray specific disease states,¹² have been used in a variety of simulated clinical experiences, both as patient and as evaluator. Each student, therefore, will have the same simulated experience, which allows for uniform assessment. Adequately trained SPs can be satisfactory alternatives to faculty evaluators.^{13,14} As of 2011, 94% of US medical schools involve SPs in clinical training and performance-based assessments.¹⁵ In addition, SPs are involved in the assessment of high-stakes clinical skills board examinations such as the United States Medical Licensing Examination and the Comprehensive Osteopathic Medical Licensing Examination of the United States.¹² Direct faculty observation of students with their patients is the preferred method of assessing professionalism at the College of Podiatric Medicine and Surgery, Des Moines University (Des Moines, Iowa). Our education model first exposes podiatric medical students to patients in the summer after their second year of study. It is vital to offer as many other formal opportunities to teach and assess professionalism before an initial patient encounter. To prepare students for this initial experience, we offer a simulated patient encounter during the spring of their second year. During this event, students complete a history and physical examination on a simulated patient with peripheral vascular disease, document their findings in a progress note, and are provided a formal assessment by both faculty and SPs regarding their performance, including interpersonal communication skills. We compared the assessments of professionalism by trained SPs during this simulated encounter with those of a faculty member who reviewed the videotapes with a validated scoring instrument. It was established previously that SPs and faculty evaluated similarly the history and physical examination components of this encounter.¹⁶ Does the same similarity occur in the

assessment of professionalism? There have been no published articles on this topic in the podiatric medical literature to date.

Methods

A total of 49 professionalism evaluations were used for analysis purposes. Three students were not evaluated by faculty owing to lack of video availability caused by a mechanical delay in beginning the recording. Forty-nine students were evaluated by 11 SPs on the professionalism component of a second-year simulated patient activity and were compared with the assessment scores provided by one faculty member who reviewed the videotapes of all those encounters. The professionalism evaluation was completed by the same SP who portrayed the specific disease state.

According to Ginsburg and McGaghie,¹⁷ the only professional attribute of medical students that can be assessed competently during their initial clinical encounters are their communication skills. The assessment tool we developed and used was modeled after the Kalamazoo Consensus Statement (KCS),¹⁸ which measures seven essential domains of physician-patient communication. The KCS was developed by medical experts representing residencies, medical colleges, and medical education organizations. Schirmer et al¹⁹ evaluated 15 different assessment tools using six different family medicine educators. Interrater reliability for the KCS, as measured by Cronbach alpha, was high at 0.88, and it moderately high (>0.79) for all of the instruments. The present study evaluated five of the seven major professionalism domains from the KCS: “build a relationship,” “open the discussion,” “gather information,” “understand the patient’s perspective,” and “share information.” We expanded the “gather information” section by adding three additional subdomains consisting of seven additional elements: listen and clarify, sequence questions, and sequence physical examination. Each domain/subdomain had a checklist of two or three elements that were observed, with each element assessed as a score of 0 if not observed and 1 if observed. Table 1 shows the components of the assessment tool. Recognizing the limitations of applying mean and standard deviation to the type of data measures having the properties of ordinal scale measurements, the nonparametric statistic Wilcoxon signed rank test was used to assess differences between SP and faculty scoring. Statistical analysis was performed with IBM SPSS for Windows, Version 22.0 (IBM Corp, Armonk, New York).

Results

A total of 49 student professionalism evaluations of the 52-member College of Podiatric Medicine and Surgery class of 2015 were used for analysis purposes. Three students were not evaluated by faculty owing to lack of video availability caused by a mechanical delay in beginning the recording. Each of the 49 students had an evaluation performed by an SP and by the faculty member. The results of the Wilcoxon signed rank test for the five professionalism domains are shown in Table 2. For the purpose of identifying differences in scoring, the mean $\pm$ SD, the z score, and the probability value are presented for the SP and faculty assessments in each professionalism domain (Table 2). Significantly different assessment scores were identified in the professionalism domains of “build a relationship” ($P = .008$), “gather information” ($P = .001$), and “share information” ($P = .002$), where the

faculty scored the students higher than the SPs for 24.5%, 18.9%, and 26.5% of the cases, respectively (Table 3). There were no differences between SP and faculty in the assessment of “open the discussion” and “understand the patient’s perspective.” In the “open the discussion” domain there was 83.7% agreement between SP and faculty scoring, whereas only 38.8% agreement was found for “understand the patient’s perspective” (Table 3). Faculty scores were higher than SP scores in all of the “gather information” subdomains; however, the differences in scores were significant only in the “appropriate questioning” and “listen and clarify” subdomains (Table 4).

Discussion

These findings showed that faculty rated students higher on professionalism than did SPs. Some reasons for the variance in scores between faculty and trained SPs have been described elsewhere. Similar to the present findings, Liew et al²⁰ showed that instructors who did not actively train students in communication skills assigned higher scores than trained SPs and clinical skills instructors. The faculty member in the present study who evaluated the videotapes, although actively involved in daily clinical instruction, is not involved in the curriculum where proper communication skills are taught to students. Perhaps further training of faculty in assessing communication skills is necessary, as Liew et al²⁰ found high agreement between the professionalism assessments of SPs and clinical skills instructors who trained SPs. Zanetti et al¹³ showed that SPs’ rating of students’ professionalism in simulated cases was more severe than that of other raters, which mimics the present findings as well. Furthermore, they concluded that SPs tend not to incorporate a global rating of professionalism into their assessment, which is more common in faculty raters, and rely completely on evaluating each element of the assessment separately and distinctly. Cooper and Mira²¹ showed that communication skills emphasized by teachers did not reflect those skills considered to be important by SPs.

The present results also bring into question the usefulness of the assessment tool. Jah et al²² performed a systematic review and found little evidence of assessment measures that are effective in evaluating attitudes toward professionalism, although several measures are described in the literature. Based on an extensive literature review, the KCS is clearly the most popular instrument that is used and the one that we chose as a blueprint based on its high interrater reliability.¹⁹ However, Schirmer et al¹⁹ rated the KCS as the best assessment tool for use by faculty, and the Rochester Communication Rating Scale²³ was the best for use by SPs. Using the KCS tool for both faculty and trained SPs could have affected the results and accounted for some of the discrepancies in student professionalism grading.

The major limitation of this study is that only one faculty member assessed professionalism during a simulated patient encounter, which does not allow for a statistical analysis of interrater reliability. This is especially important in the area of professionalism, a “soft” skill whose assessment of traits or behaviors is often subjective. Assessing professionalism requires a finer degree of discrimination, unlike assessing the clinical examination skills of medical students, where there are clearly measurable signs of competent performance. However, Jonsson and Svingby²⁴ found in their review of 75 studies that used scoring rubrics in performance assessment, only a couple of studies did not show fair agreement

beyond chance according to the Cohen kappa when evaluating interrater reliability. Despite having only one faculty reviewer for the study, there is a reasonable assumption that the scoring assessment of that individual would be similar to that of additional faculty reviewers and the results should not be entirely ignored.

Differences in scoring in this study could also be attributed to the “halo effect,” which was first coined by Thorndike²⁵ whereby a person makes a biased judgment of another based only on the observation of a single trait, which can either be positive or negative. Hence, exemplary or poor performance in one or two of the domains could have affected the scores of the other domains. Some ways that the halo effect can be reduced is by using observation by multiple reviewers, supplementing formal observation with unobtrusive observation, and educating raters.¹⁷

Checklists to evaluate professionalism, such as the modified KCS checklist, have been shown to be best for inexperienced examiners with sufficient reliability.²⁶ However, such checklists are not without problems. Checklists can be so prescriptive that they can lead a student to merely go through the motions instead of truly interacting with a patient.²⁷ Checklists can also be vague and tend to assess a trait of an individual (“the student is empathetic”) rather than measuring a demonstrated behavior.¹⁷ Van Zanten et al²⁸ suggested that professional behaviors need to be measured using different assessment tools, including checklists, surveys, and self-assessments, to obtain an accurate impression of the student. In addition, professionalism is more accurately measured in a variety of clinical encounters over the course of several years of training. Everyone has a slightly different perception of professional behavior and, therefore, of clinicians, nurses, residents, the students themselves, and their peers should participate in these evaluations. They should occur over multiple encounters and be formative to allow for positive growth as student training increases.²⁹

Conclusions

This study showed that during a standardized performance assessment laboratory for second-year podiatric medical students, a single faculty member rated three of the five domains of the KCS on professionalism higher than did 11 trained SPs and that the difference was statistically significant. At the College of Podiatric Medicine and Surgery, Des Moines University, the opportunities to measure student professionalism before actual patient encounters are limited; therefore, we need to make the assessment reliable and reproducible. Further work will require choosing who is best to assess professionalism—the patient or clinical faculty—and identifying the best tool for assessment of professionalism.

References

1. Blendon RJ, Benson JM, Hero JO: Public trust in physicians: U.S. medicine in international perspective. *N Engl J Med* 371: 1570, 2014. [PubMed: 25337746]
2. Arnold L: Assessing professional behavior: yesterday, today, and tomorrow. *Acad Med* 77: 502, 2002. [PubMed: 12063194]
3. Ha JF, Longnecker N: Doctor-patient communication: a review. *Ochsner J* 10: 38, 2010. [PubMed: 21603354]
4. Neumann M, Edelhauser F, Tauschel D, et al.: Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Acad Med* 86: 996, 2011. [PubMed: 21670661]

5. Colliver JA, Swartz MH, Robbs RS, et al.: Relationship between clinical competence and interpersonal and communication skills in standardized-patient assessment. *Acad Med* 74: 271, 1999. [PubMed: 10099650]
6. Green M, Zick A, Thomas JX: Commentary: accurate medical student performance evaluations and professionalism assessment: "Yes, we can!" *Acad Med* 89: 1105, 2014.
7. Papadakis MS, Teherani A, Banach MA: Disciplinary action by medical boards and prior behavior in medical school. *N Engl J Med* 353: 2673, 2005. [PubMed: 16371633]
8. Ferguson RP: Professionalism: hard to measure but you know it when you see it. *J Community Hosp Intern Med Perspect* 2014, 4: 24226.
9. Bhugra D, Malik A: "Experts and Expertise," in *Professionalism in Mental Healthcare*, p 1, Cambridge University Press, New York, 2011.
10. Woollicroft JO, Howell JD, Patel BP, et al.: Resident- patient interactions: the humanistic qualities of internal medicine residents assessed by patients, attending physicians, program supervisors and nurses. *Acad Med* 69: 216, 1994. [PubMed: 8135980]
11. Boon K, Turner J: Ethical and professional conduct of medical students: review of current assessment measures and controversies. *J Med Ethics* 30: 221, 2004. [PubMed: 15082823]
12. Boulet JR, Smee SM, Dillon GF, et al.: The use of standardized patient assessments for certification and licensure decisions. *Sim Healthcare* 4: 35, 2009.
13. Zanetti M, Keller L, Mazor K, et al.: using standardized patients to assess professionalism: a generalizability study. *Teach Learn Med* 22: 274, 2010. [PubMed: 20936574]
14. Khan AS, Qureshi R, Acemoglu H, et al.: Comparison of assessment scores of candidates for communication skills in an oscE, by examiners, candidates and simulated patients. *creative Educ* 3: 931, 2012.
15. Passiment M, Sacks H, Huang G: Medical simulation in medical education: results of an AAMC survey. Association of American Medical colleges Web site. Available at: <https://www.aamc.org/download/259760/data>. Published 9 2011 Accessed March 22, 2016.
16. Mahoney J, Vardaxis V, Anwar N, et al.: Relationship between faculty and standardized patient assessment scores of podiatric medical students during a standardized performance assessment laboratory. *JAPMA* 106: 116, 2016.
17. Ginsburg S, McGaghie W: "Evaluation and Grading of Students," in *Guidebook for Clerkship Directors*, 3rd Ed, edited by Fincher RME, p 133, Alliance for Clinical Education, Omaha, 2005.
18. Makoul G: Essential elements of communication in medical encounters: the Kalamazoo Consensus Statement. *Acad Med* 76: 390, 2001. [PubMed: 11299158]
19. Schirmer JM, Mauksch L, Lang F, et al.: Assessing communication competence: a review of current tools. *Fam Med* 37: 184, 2005. [PubMed: 15739134]
20. Liew S, Dutta S, Sidhu JK, et al.: Assessors for communication skills: SP's or healthcare professionals. *Med Teach* 35: 626, 2014.
21. Cooper C, Mira M: Who should assess medical students' communication skills: their academic teachers or their patients? *Med Educ* 32: 419, 1998. [PubMed: 9743806]
22. Jah V, Bekker HL, Duffy SRG: A systematic review of studies assessing and facilitating attitudes towards professionalism in medicine. *Med Educ* 41: 822, 2007. [PubMed: 17661891]
23. Epstein RM, Dannefer IF, Nofziger AC, et al.: Comprehensive assessment of professional competence: the Rochester experiment. *Teach Learn Med* 16: 186, 2004. [PubMed: 15276897]
24. Jonsson A, Svingby G: The use of scoring rubrics: reliability, validity and educational consequences. *Educ Res Rev* 2: 130, 2007.
25. Thorndike EL: A constant error in psychological ratings. *J Appl Psychol* 4: 25, 1920.
26. Regehr G, MacRae H, Reznick RK et al.: Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med* 73: 993, 1998. [PubMed: 9759104]
27. Thistlethwaite J, Spencer J: "Assessing Professionalism," in *Professionalism in Medicine*, p 200, Radcliffe Publishing Ltd, oxford, 2008.
28. Van Zanten M, Boulet JR, Norcini J: Using a standardized patient assessment to measure professional attributes. *Med Educ* 39: 20, 2005. [PubMed: 15612897]

29. Ginsburg S, Regehr G, Hatala R, et al.: Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad Med* 75: S6, 2000. [PubMed: 11031159]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.**Professionalism Domains and Assessment Elements**

Domain/Subdomain	Elements^a
Build a relationship	Relates well to patient during history and physical examination Gently examines areas of pain Maintains patient modesty
Open the discussion	Allows patient to complete opening statement Elicits patient's full set of concerns Establishes/maintains a personal connection
Gather information	
Listen and clarify	Uses open-ended and closed-ended questions appropriately Structures, clarifies, and summarizes information Actively listens using nonverbal and verbal techniques
Question appropriately	Uses appropriate questions to clarify chief complaint Avoids ambiguous questions Transitions easily between questions
Sequence questions	Asks questions in appropriate sequence and groups Does not jump between groups of questions
Sequence physical examination	Completes one area of the physical examination before moving to another Does not jump back and forth between examination areas
Understand the patient's perspective	Explores contextual factors (eg, family, culture, sex, age, spirituality) Explores beliefs, concerns, and expectations about health and illness Acknowledges and responds to patient's ideas, feelings, and values
Share information	Uses language that the patient can understand Checks for understanding Encourages questions

^aEach element is scored 0 if not observed, 1 if observed.

Table 2.
Comparison of Faculty and Standardized Patient Scores on All of the Domains of Professionalism

Domain	Score (Mean $\pm$ SD)		Z Score	P Value
	Standardized Patient	Faculty		
Build a relationship	2.71 $\pm$ 0.46	2.94 $\pm$ 0.24	-2.673	.008
Open the discussion	2.87 $\pm$ 0.34	2.96 $\pm$ 0.20	-1.414	.157
Gather information	2.32 $\pm$ 0.25	2.46 $\pm$ 0.13	-3.339	.001
Understand the patient's perspective	2.29 $\pm$ 0.75	2.18 $\pm$ 0.39	-0.985	.325
Share information	2.69 $\pm$ 0.54	2.98 $\pm$ 0.14	-3.116	.002

Table 3.

Percentage Agreement Between Faculty and Standardized Patient (SP) Scores on All of the Domains of Professionalism

Domain	Agreement	Faculty > SP	Faculty < SP
Build a relationship	71.4	24.5	4.0
Open the discussion	83.7	12.2	4.0
Gather information	77.0	18.9	4.0
Understand the patient's perspective	38.8	22.5	38.8
Share information	71.4	26.5	2.0
Average agreement	68.5	20.1	10.6

Table 4.
Comparison of Faculty and Standardized Patient Scores on the Gather Information Subdomains

Gather Information Subdomain	Score (Mean $\pm$ SD)		Z Score	P Value
	Standardized Patient	Faculty		
Question appropriately	2.79 $\pm$ 0.41	3.00 $\pm$ 0.00	-3.317	.001
Sequence questions	1.90 $\pm$ 0.30	1.98 $\pm$ 0.14	-1.633	.102
Listen and clarify	2.79 $\pm$ 0.41	2.98 $\pm$ 0.14	-3.000	.003
Sequence physical examination	1.81 $\pm$ 0.40	1.90 $\pm$ 0.47	-1.147	.251