

## Using 360-degree multi-source feedback to evaluate professionalism in surgery departments: An Iranian perspective

Adel Yazdankhah<sup>1</sup>, Mohammad Tayefeh Norooz<sup>2</sup>, Hadi Ahmadi Amoli<sup>3</sup>, Ali Aminian<sup>4</sup>  
Zhamak Khorgami<sup>5</sup>, Paria Khashayar<sup>6</sup>, Patricia Khashayar<sup>\*7</sup>

Received: 22 January 2015

Accepted: 4 July 2015

Published: 31 October 2015

### Abstract

**Background:** Medical professionalism helps physicians adopt a proper and good healing action for the patients based on their particular circumstance. This study was conducted to assess professionalism in surgical residents, using a 360-degree evaluation technique in several teaching hospitals in Tehran, Iran.

**Methods:** This study was conducted on all the second and third year surgery residents from three university teaching hospitals in Tehran. Multi-source feedback questionnaire contained 10 questions on the residents' professional behavior and was completed by the faculty and staff members (nurses, operation room staff, and medical assistants) as well as other surgery residents, interns and patients to evaluate each resident. Response rates were used to determine feasibility for each of the respondent groups and the mean and standard deviation score for each question was computed to determine the viability of the items. Reliability was assessed using alpha Cronbach coefficient for each respondent group. The correlation between these scores and the residents' final and OSCE grade was also assessed.

**Results:** The internal consistency reliability for 360-degree rating was 0.889. There was no significant difference in the residents' score in different hospitals. While male residents obtained higher total score, there was no significant difference between them. The residents, however, obtained lower scores compared to the staff. The highest score was recorded for question 6, suggesting that the residents treated the patients regardless of their socioeconomic status.

**Conclusion:** This study revealed a strong agreement between the results gathered from different respondents, confirming the reliability of the questionnaire and the respondents' unbiased response. It also revealed that the residents did well in the whole test, showing they were conscientious and learning to become medical professionals.

**Keywords:** Multi-Source Feedback, Professionalism, 360-Degree, Assessment.

**Cite this article as:** Yazdankhah A, Tayefeh Norooz M, Ahmadi Amoli H, Aminian A, Khorgami Zh, Khashayar P, Khashayar P. Using 360-degree multi-source feedback to evaluate professionalism in surgery departments: An Iranian perspective. *Med J Islam Repub Iran* 2015 (31 October). Vol. 29:284.

### Introduction

Medical professionalism helps physicians adopt a right and good healing action for the patients based on their particular cir-

cumstance. It is one of the six competencies proposed by the Accreditation Council for Graduate Medical Education (ACGME) in 1999 that needed to be imparted during res-

<sup>1</sup>. MD, Associate Professor, Surgery Department, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran. yazdankhah@tums.ac.ir

<sup>2</sup>. MD, Assistant Professor, Surgery Department, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran. mtnorooz@gmail.com

<sup>3</sup>. MD, Associate Professor, Surgery Department, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran. ahmadiam@sina.tums.ac.ir

<sup>4</sup>. MD, Assistant Professor, Surgery Department, Imam Khomeini Hospital, Tehran University of Medical Sciences, Tehran, Iran. aaminian@tums.ac.ir

<sup>5</sup>. MD, Assistant Professor, Surgery Department, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran. khorgami@tums.ac.ir

<sup>6</sup>. PhD Candidate, Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran. pkhashayar@yahoo.com

<sup>7</sup>. (Corresponding author) MD, PhD Candidate, Osteoporosis Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran, & Center for Microsystems Technology, Imec and Ghent University, Gent-Zwijnaarde, Belgium. patricia.kh@gmail.com

idency or fellowship training (1).

The professionalism charter defines three fundamental principles: 1) Primacy of patient welfare, which focuses on altruism, trust, and patient interest; 2) Patient autonomy, which stresses on honesty with patients and the need to educate and empower them and their families to make appropriate medical decisions; 3) Social justice, which addresses physicians' societal contract and distributive justice (2).

In recent years, educators have been criticized for not teaching and assessing the core values of medicine that determine communication skills, intra-individual relations and professionalism in clinical practice. These critics stress that medical trainees, particularly the residents, need to learn and practice such behavior during their training period, adding that certain "conscientious behaviors" in medical school predict professionalism later on in practice (3).

Many studies have revealed that certain conceptual frameworks of professionalism need excellence in medical knowledge and clinical skills. In addition, many believe professional behaviors are best assessed by seeking verification from numerous sources, such as peers, senior residents, nursing staff, students, and even patients over time, stressing that these assessments should be transparent, fair and within realistic contexts involving situations likely to challenge professionalism (4).

Multi-source feedback, also known as 360-degree assessments, is a multi-rater multidimensional assessment of residents' professional behaviors and communication skills with high internal consistency reliability (5). Using this technique, a feasible and fair and at the same time a valid and reliable assessment is made based on the feedback gathered from ancillary staff, peers and patients in direct relation with the residents.

As for the surgery department, excellent interviewing skills play an important role in strengthening the bond between the patient and physician (6). Having effective communication skills, intra-individual relations

and professionalism is of great importance in this department as it not only increases patient satisfaction but also it is associated with improved patient compliance and outcome. Therefore, this study was conducted to assess professionalism in surgical residents using a 360-degree evaluation technique in several teaching hospitals in Tehran, Iran.

## Methods

The study took place in three university teaching hospitals in Tehran, all affiliated to Tehran University of Medical Sciences (TUMS). All the 2<sup>nd</sup> and 3<sup>rd</sup> year surgery residents from these hospitals were required to participate in the process. The study was performed independently of other assessment techniques and did not influence passing or rejecting the residency.

The assessment tool used was Multi-source feedback, developed based on the Education Outcomes Service Group (EOS group) of the Arizona Medical Education Consortium and determined by a consensus of three experts in the field of surgery, medical ethics and medical education (7). The adaptation was based on the translation/back-translation methodology. Several meetings were held between the translators, researchers and the expert panel before the final version was approved. The internal consistency ( $\alpha$  cronbach = 0.83), validity and reliability of the questionnaire were tested in a group of surgery residents studying in a hospital not enrolled in the study in two different occasions two weeks apart. During this step, the relative value of each response choice of the questionnaire items was assessed through a visual analogue scale. The form also included an inquiry of the respondents' position and the statement that the opinions would be kept confidential.

Each questionnaire contained 10 questions on the residents' professional behavior (Table 1). For each objective, a 3-likert scale was attached with the grades 'satisfactory', 'not quite satisfactory' and 'unsatisfactory'. The box 'cannot assess this

Table 1. The questionnaire used in this study

- Question 1. Demonstrates responsibility/accountability  
 Question 2. Functions effectively as a member of the team and responds in a timely fashion to requests for help  
 Question 3. Seeks consultation/supervision when appropriate  
 Question 4. Actively seeks feedback and immediately self-corrects  
 Question 5. Shows compassion for patients and their families and communicates sensitively, patiently, and effectively with patients  
 Question 6. Is courteous to patients and families and shows unconditional positive regard for them irrespective of their socioeconomic status  
 Question 7. Solicits comments, feedback, and concerns from other team members, patients, families, and peers  
 Question 8. Handles demanding interpersonal situations in a respectful and effective manner  
 Question 9. Fosters an atmosphere of honesty and mutual respect in daily practice.  
 Question 10. Completes assigned tasks and even more if needed

item' was excluded as the pilot study revealed the option would distort the results.

Every surgery resident was evaluated by the faculty and staff members (nurses, operation room staff, and medical assistants) as well as other surgery residents, interns and patients with whom they worked regularly during the past three months. The entire eleven faculty, 15 nurses (surgery and emergency department), 15 staff (surgery and emergency department and operating room), 8 residents (year 1 and 4), 10 interns and 10 patients were appointed by the authors as a group to answer the surveys. All the investigators were uniformly trained. They were asked to answer the questionnaires anonymously.

#### *Statistical Analysis*

To analyze the data, all questionnaires were returned to the authors and entered into SPSS 16 for analysis. Response rates were used to determine feasibility for each of the respondent groups and the mean and standard deviation score for each question was computed to determine the viability of

the items. Reliability was assessed using Alpha Cronbach coefficient for each respondent group. The correlation between these scores and the residents' final and OSCE grade was also assessed.

#### **Results**

After obtaining approval for the study from the Ethical Board Committee of Tehran University of Medical Sciences, all the 37 2<sup>nd</sup> and 3<sup>rd</sup> year surgery residents studying in TUMS affiliated teaching hospitals in July 2009 were enrolled, producing 1741 assessments. The cooperation rate was 68.8%. Twenty five of the residents were male and the others were female; twelve of them studied in the first grade, 13 in the second, and 12 in the third grade.

The internal consistency reliability for 360-degree rating was 0.889 (Table 2). There was no significant difference between the internal consistency reliability based on the respondents. However, the results were more consistent for female residents. There was no difference in the total score of the residents working in the three

Table 2. Reliability and total score calculated in different groups

		Alpha cronbach	Score (Mean ± SD)	p
All the residents		0.889	7.50 ± 4.20	-
Gender	Male	0.864	7.87 ± 2.61	<0.001
	Female	0.922	7.68 ± 2.97	
Grade	2nd	0.892	7.10 ± 4.40	<0.001
	3rd	0.899	7.20 ± 4.40	
Respondents	Faculty	0.743	8.20 ± 2.20	0.03
	Grade 4 Residents	0.752	8.17 ± 2.33	
	Grade 1 Residents	0.773	8.14 ± 2.81	
	Interns	0.773	7.89 ± 2.97	
	Patients	0.741	7.40 ± 2.48	
	Nurses	0.771	7.82 ± 2.88	
	Department Staff	0.771	6.98 ± 3.21	
	Operation Room Staff	0.767	7.71 ± 2.89	

Table 3. Mean score obtained in each question based on the residents gender

	Total (Mean ± SD)	Female (Mean ± SD)	Male (Mean ± SD)
Question 1	0.77 ± 0.42	0.71 ± 0.45	0.76 ± 0.43
Question 2	0.74 ± 0.44	0.68 ± 0.47	0.72 ± 0.45
Question 3	0.82 ± 0.38	0.78 ± 0.42	0.80 ± 0.40
Question 4	0.74 ± 0.44	0.67 ± 0.47	0.72 ± 0.45
Question 5	0.81 ± 0.39	0.73 ± 0.45	0.80 ± 0.40
Question 6	0.91 ± 0.29	0.82 ± 0.38	0.89 ± 0.31
Question 7	0.81 ± 0.39	0.76 ± 0.43	0.79 ± 0.40
Question 8	0.77 ± 0.42	0.72 ± 0.45	0.76 ± 0.43
Question 9	0.82 ± 0.38	0.75 ± 0.43	0.81 ± 0.39
Question 10	0.68 ± 0.22	0.60 ± 0.49	0.68 ± 0.47

studied hospitals. Residents studying in grade 3 obtained the highest total score (year 1:  $7.3 \pm 3.03$ , year 2:  $7.61 \pm 2.8$ , year 3:  $8.5 \pm 2.14$ ; p-value < 0.001). Although male residents obtained higher total score, there was no significant difference between them ( $7.87 \pm 2.61$  vs.  $7.68 \pm 2.97$ ; p-value: 0.178).

Except for questions 5, 8 and 9, there was a significant difference between the scores obtained from different respondents. The residents attained the lowest score from the department staff and the highest from the faculty. Based on the results, the residents scored highest in question 6 and lowest in question 10 regardless of their gender (Table 3).

The mean global rating of the residents was  $17.5 \pm 1.33$ . There was no correlation between the residents' score in 360 ° and that of the other tests (OSCE:  $13.7 \pm 0.92$  vs. DOPS:  $15.7 \pm 1.45$ ).

## Discussion

Several assessment techniques such as direct observation, simulation-based assessments, Objective Structured Clinical Examinations (OSCEs), global faculty evaluations, In-Training Examination (ITE), mini-clinical evaluation exercise (mini-CEX) scores, 360 degree evaluation, case logs, self-reflection, clinical performance metrics, and portfolios are commonly used to assess different aspects of medical residents' competence in patient care at the end of each semester. Each of them has its strengths and intrinsic flaws (8).

Van der Vleuten described five criteria for determining the usefulness of a particu-

lar assessment method: Reliability (the degree to which the measurement is accurate and reproducible); validity (whether the assessment measures what it claims to measure); impact on future learning and practice, acceptability to learners and faculty; and cost effectiveness (to the trainee, the institution, and society at large) (9).

Direct observation of medical trainees studying at different levels (undergraduate or residency) with actual patients is critical for teaching and assessing clinical and communication skills. A recent report released by the Institute of Medicine showed that improved supervision of trainees can enhance patient safety and quality of clinical education (10). This comes while direct observation occurs infrequently and inadequately (11). Moreover, end-of-rotation global rating is often performed by supervisors who have not directly observed trainees in practice (12).

At the same time, many studies have brought up the idea that colleagues and nursing staff might be well placed to make these judgments as the faculty members are not present to assess all resident interactions, and may miss some of their interactions with both patients and other members of the health care team. Thus, the idea of asking peers to assess professional behavior and performance that are less accessible to conventional assessment techniques such as written and clinical examinations has been increasingly explored in the literature (13,14).

360-degree evaluation incorporates several different areas of inquiry, particularly teamwork, interpersonal skills, communica-

tion skills, management skills, and clinical decision making (15). The technique is frequently used to evaluate residency programs in the US, all foundation programs in the UK and similarly many other countries.

The technique not only provides reliable feedbacks of different groups in contact with the resident but also can be used as a formative evaluation to assess and improve certain aspects of behavior in different environments (16). Several studies, however, have pointed out that the 360-degree feedback can only be effective if certain conditions are met, more importantly having skilled facilitators to encourage reflection (17,18).

Some drawbacks to this type of assessment are as follows: This type of assessment relies on the willingness of the evaluators to be fair; its impact depends on the ability of the end user to assimilate feedback; and there are practical issues of selecting the best questions in each department (19). The results can only be generalized if the test is conducted on large samples (classmates: 8-10, healthcare providers: 8-10, patients>25) (20).

Moreover, in view of the fact that daily events and stimulus differ in a hospital, the response to the questions may vary over days and there is no specific standard to evaluate or interpret the results based on the condition; and this reduces the reliability of the questionnaires. In a study describing the use of this technique in the evaluation of core competencies in cardiothoracic surgery department, a high degree of defensiveness was reported among residents in relation to the evaluation process and feedback from non-physician sources (21). Many studies have also pointed out the influence of rater training on scores, stressing that observers need training to rate learners' performance reliably and discriminate between performance levels (22).

The efficacy of the tool in assessing different aspects of a resident's capability in different specialties has been reported (23, 24). In a systematic review conducted by Al Khalifa et al. MSF was reported to be a

feasible, reliable, and valid method to assess surgical practice (25). They stressed that the tool can well assess non procedural competencies such as communication skills, interpersonal skills, collegiality, humanism, and professionalism, adding that other techniques should be used to assess procedural competence. Stark et al. showed that the instrument improves faculty comfort and self-assessed skill in providing feedback about professionalism (26). Qu et al. also confirmed that the MSF assessment tools are internally valid and reliable for assessing residents' professionalism and interpersonal and communication skills (27).

In another study, the tool was used for formative assessment of the residents as well as their clinical skills and professional behavior in the ward (15). They, however, stressed that gender can alter the results. Based on a study conducted by Lockyer et al. on medical graduates seeking job in Canada, MSF was shown to have a high reliability for self-assessment ( $\alpha_{cronbach}= 0.83$ ) and assessment ( $\alpha_{cronbach}= 0.90$ ). They showed MSF as an effective test for such assessments (10). In another study, MSF was shown as an acceptable tool for evaluating communication skills and professionalism in many specialties. The present study similarly showed a strong agreement between the results gathered from different respondents, confirming the reliability of the questionnaire and the respondents' unbiased response. No significant difference was observed in the residents' score in different hospitals, showing the uniformity of the education provided in different centers. Nevertheless, the residents obtained lower scores from the staff. This was in line with the results of Ogunyemi et al.'s study which similarly found weak correlations between nursing and faculty evaluations and standard medical examination scores (28). Considering the fact that the nursing staff spends more time with the residents, it could be concluded that this group can better and differently assess residents on the competencies

of interpersonal and communication skills and professionalism.

In our study, the highest score was recorded for question 6, suggesting that the residents treated the patients regardless of their socioeconomic status. This suggests there would be no problem in equity distribution in different parts of the country. Nonetheless, the low score achieved in question 10 indicates that the residents consider the jobs refer to them as a task rather than a responsibility which should be changed.

Studying the effect of study level and gender revealed that while studying more years had a positive effect on the score, gender did not influence the results; 360 scores were only correlated with the academic staff's assessment score of the resident and not with that of other assessments. Therefore, it could be concluded that the residents' professional behavior is not related to their knowledge but it considerably affects the academic staff's assessment score. While the present study failed to report any association between the residents' OSCE and DOPS' score and the 360 results, others have revealed that residents' knowledge and clinical skill were correlated with professionalism assessments. Reed et al. showed that residents' clinical skill, as measured by mini-CEX, was the variable most strongly associated with professionalism (29). This comes while many have doubted the accuracy of mini-CEX tests, saying its validity and reliability may have been influenced by the residents' professional attributes during clinical encounters (30).

The results of this study revealed that attention is shifted to professionalism in surgery departments of TUMS, a move that has a great effect on the care provided to the patient and their outcome. Our findings also suggest deficiencies in the studied hospitals and solutions to improve the quality of training as well as patient care in the surgery departments.

There were several limitations to this study. The study was limited to surgery res-

idents studying in three teaching hospitals in Tehran and thus its results may not be generalizable to residents in other specialties or other hospitals. It is possible that the quality of the residents and resources in these hospitals may be different from that of smaller hospitals. Moreover, although our sample size was adequate to detect statistically significant differences between groups, a larger sample size may provide greater stability of the estimates. Furthermore, similar to other studies, there was defensiveness among the respondents in relation to the evaluation process.

### Conclusion

It could be concluded that most of our residents did well in the whole test, showing they were conscientious and learning to become medical professionals. As for the others, it is possible that the residents' busy daily work schedules, limits the time they need to dedicate to adopt professionalism in their practice. Therefore, a follow-up study is needed to determine how the residents used these data and what changes they made as a result of the feedback based on this study.

### Acknowledgements

The authors would like to thank all the faculty, residents and medical students, staff, and patients in these three hospitals, without whom we could not complete this project. We would like to thank Dr. Pouria Khashayar for helping us gathering the questionnaires.

### References

1. Kirk LM. Professionalism in medicine: definitions and considerations for teaching. Proc (Baylor Univ Med Cent) 2007;20(1):13–16.
2. American Board of Internal Medicine Foundation. American College of Physicians–American Society of Internal Medicine Foundation. European Federation of Internal Medicine Medical professionalism in the new millennium: a physician charter. Ann Intern Med. 2002;136(3):243–246.
3. Papadakis MA, Teherani A, Banach MA, Knettler TR, Rattner SL, Stern DT et al. Disciplinary action by medical boards and prior behavior in

- medical school. *N Engl J Med*. 2005;353(25):2673-2682
4. Stern DT, ed. *Measuring Medical Professionalism*. New York, NY: Oxford University Press Inc; 2006:8-12
  5. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. *Med Teach*. 2006;28(7):e185-91.
  6. Oh J, Segal R, Gordon J, Boal J, Jotkowitz A. Retention and use of patient-centered interviewing skills after intensive training. *Acad Med*. 2001;76:647-50
  7. Introduction to the EOSG Manual. <http://azmec.med.arizona.edu/eos.htm>.
  8. Epstein RM. Assessment in Medical Education. *N Engl J Med* 2007; 356:387-396.
  9. Van Der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ* 1996;1:41-67
  10. Ulmer C, ed, Wolman DM, ed, Johns MME, ed. Committee on Optimizing Graduate Medical Trainee (Resident) Hours and Work Schedule to Improve Patient Safety. *Resident Duty Hours: Enhancing Sleep, Supervision and Safety*. Washington, DC: National Academy Press; 2008.
  11. 2008 AAMC Graduation Questionnaire Program Evaluation Survey: All Schools Summary Report Final. [http://www.aamc.org/data/gq/allschoolsreports/2008\\_pe.pdf](http://www.aamc.org/data/gq/allschoolsreports/2008_pe.pdf). Accessed May 14, 2009.
  12. Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356(4):387-396.
  13. Violato C, Marini A, Toews J, Lockyer J, Fidler H. Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. *Acad Med* 1997;72(suppl): 82-4.
  14. Evans R, Elwyn G, Edwards A. Review of instruments for peer assessment of physicians. *BMJ* 2004;328:1240
  15. Hayden SR, Dufel S, Shih R. Definitions and Competencies for Practice-based Learning and Improvement. *Acad Emerg Med*. 2002 Nov; 9(11):1242-8.
  16. Higgins RSD, Bridges J, Burke JM, O'Donnell MA, Cohen NM, Wilkes SB. Implementing the ACGME general competencies in a cardiothoracic surgery residency program using 360-degree feedback. *Ann Thorac Surg* 2004;77: 12-17.
  17. Overeem K, Wollersheim H, Driessen E, Lombarts K, van de Ven G, Grol R, et al. Doctors' perceptions of why 360-degree feedback does (not) work: a qualitative study. *Med Educ* 2009 Sep;43(9):874-82.
  18. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design? *Med Teach* 2006 Nov;28(7):e185-91.
  19. Weigelt JA, Brasel KJ, Bragg D, Simpson D. The 360-degree Evaluation: Increased Work with Little Return? *Curr Surg*. 2004 Nov-Dec; 61(6):616-26;
  20. Ramsey PG, Wenrich MD, Carline JD, Inui TS, Larson EB, LoGerfo JP. Use of peer ratings to evaluate physician performance. *JAMA*. 1993; 269:13:1655-60.
  21. Watling CJ, Lingard L. Toward meaningful evaluation of medical trainees: the influence of participants' perceptions of the process. *Adv in Health Sci Educ* 2012; 17:183-194.
  22. Kogan JR, Holmboe ES, Hauer KE. Tools for Direct Observation and Assessment of Clinical Skills of Medical Trainees. A Systematic Review. *JAMA* 2009;302(12):1316-1326.
  23. Buccieri KM, Rodriguez J, Smith SS, Robinson R, Gallivan SP, Frost JS. Director of Clinical Education Performance Assessment Surveys: A 360-Degree Assessment of the Unique Roles. *Journal of Physical Therapy Education* 2012;26 (3): 13-24.
  24. Whitehouse A, Hassell A, Bullock A, Wood L, Wall D. 360 degree assessment (multisource feedback) of UK trainee doctors: Field testing of team assessment of behaviours (TAB). *Med Teach*. 2007 Mar;29(2-3):171-6.
  25. Al Khalifa Kh, Ansari AA, Violato C, Donnon T. Multisource Feedback to Assess Surgical Practice: A Systematic Review. *J Surg Educ*. 2013 Jul-Aug;70(4):475-86.
  26. Stark R, Korenstein D, Karani R. Impact of a 360-degree Professionalism Assessment on Faculty Comfort and Skills in Feedback Delivery. *J Gen Intern Med*. 2008 Jul;23(7):969-72.
  27. Qu B, Zhao YH, Sun BZ. Assessment of Resident Physicians in Professionalism, Interpersonal and Communication Skills: a Multisource Feedback. *Int J Med Sci* 2012; 9(3): 228-36.
  28. Ogunyemi D, Gonzalez G, Fong A, Alexander C, Finke D, Donnon T, et al. From the eye of the nurses: 360-degree evaluation of residents. *J Contin Educ Health Prof* 2009 Spring;29(2):105-10.
  29. Reed DA, West CP, Mueller PS, Ficalora RD, Engstler GJ, Beckman TJ. Behaviors of Highly Professional Resident Physicians. *JAMA* 2008; 300(11):1326-1333.
  30. Durning SJ, Cation LJ, Markert RJ, Pangaro LN. Assessing the reliability and validity of the mini-clinical evaluation exercise for internal medicine residency training. *Acad Med* 2002; 77(9): 900-904.