DISCUSSION

360-Degree Evaluations of Plastic Surgery Resident Accreditation Council for Graduate Medical Education Competencies: Experience Using a Short Form

Jeffrey E. Janis, M.D.
Dallas, Tex.

This well-written and extremely interesting article by Dr. Pollock and colleagues is an important addition to the literature on the evaluation of resident performance and, as stated by the authors themselves, represents the first article to use Accreditation Council for Graduate Medical Education competency-based ratings to evaluate plastic surgery residents. The authors describe their experience with the plastic surgery residents in their 6-year integrated plastic surgery residency program whereby seven quarterly reviews of individual resident performance were solicited from 40 individuals representing eight separate groups of health care professionals, from attending (board-certified) physicians to residents/interns to nursing staff across different hospital systems. These reviews were two-page anonymous reviews of performance based on the Accreditation Council for Graduate Medical Education core competencies,[1] with a supplementary subcategory on operating/technical skills that was added by the authors. The authors found two defined "clusters" of evaluators that were statistically different in their assessment of resident performance, one being more "positive" (the nurses' cluster) and one being more constructively critical (the surgeon's cluster). The authors conclude that "Based on this clustered arrangement, the resident is able quarterly to enjoy two, independent, formative assessments, potentially over 6 years of integrated training."

Several issues, however, require further discussion and highlighting. The first involves potential bias in the nurses' cluster that may influence the outcomes reported. According to the questionnaire provided in Figure 1, the nurses are asked to evaluate the residents based on the six Accreditation Council for Graduate Medical Education core competencies, which are listed for the evaluator. However, for the uninitiated, these core competencies may not be completely defined and/or understood by the nurses who are not familiar with, or who have not received training in, evaluating specifically by these parameters. Admittedly, when these core competencies were first released by the Accreditation Council for Graduate Medical Education, they were accompanied by documents that required careful review to understand the specific criteria that are encompassed by each competency.[1] Without this associated document, or without specific fundamental training in this regard, there may be some confusion as to how to judge or evaluate the residents, thus leading to bias that may affect the outcome and conclusions. Furthermore, the nurses' cluster rated the residents higher on medical knowledge, practice-based learning and improvement, technical/operative skills, and interpersonal/communication skills. It is difficult to determine based on this study how the nurses are assessing these competencies accurately. For instance, are they witnessing Socratic teaching in the operating room or on ward rounds and using those observations to base their score on medical knowledge? How are the nurses evaluating technical skills? Certainly, the interpersonal/communication skills core competency can be more easily assessed by the nursing cluster, and as might be expected, this was rated more highly (statistically significant) by the nurses' cluster relative to the surgeons' cluster. The authors admit this may be attributable to the intrinsic "nurturing" nature of nurses. In contrast, can we be certain that higher scores on this competency, or even all competencies in general, were not influenced by the fact that residents are more likely to befriend nurses than they are their attending physicians/faculty? This may have led to bias that may have resulted in the statistically significantly higher scores given by the nurses' cluster, even if the bias was not able to be specifically quantified.

From the Division of Plastic Surgery and the Department of Surgery Education, University of Kentucky.
Received for publication February 24, 2008.

DOI: 10.1097/PRS.0b013e31817d66f8

Disclosure: *The author has no financial interest in any of the products, devices, or drugs mentioned in this discussion or the original article.*

The second issue that requires attention is that of the exposure of the evaluators to the resident. During the average resident's day, there are faculty, residents, and nurses with whom he or she interacts. The amount of interaction is not uniform. However, each evaluator is given equal weight in terms of the impact of each evaluator's score relative to the mean score. Simply put, the score for the nurse who may have interacted with the resident twice a week for 10 minutes is given the same credence as the score given by the attending physician who operates and interacts with the residents every day for hours at a time. This also may have led to bias in the results, where some scores may be truly underweighted and others overweighted simply based on the fact that all scores are given equal weight.

The third issue concerns the frequency of the reviews themselves. The authors have adjusted the frequency of the 360-degree evaluations to quarterly to make them less onerous on the evaluators and to prevent "rater burnout." However, there is no indication as to the length of the resident rotations. It is possible that if the rotation length is significantly less than the evaluation interval, the accuracy of the rating may be in question. For instance, if the resident is on a particular rotation one month but the evaluations are distributed every 3 months, the rater may not recall details/specifics of the resident's performance as clearly 2 months later, especially if the amount of exposure or interaction with that particular resident was minimal.

Finally, the authors have chosen to exclude medical student evaluations in the 360-degree reviews based on admission that "none of these groups deemed themselves competent to gauge performance." The literature will suggest, however, that medical student evaluations of the residents are, in fact, a valid means of evaluating resident performance.[2,3] The information gleaned from this cohort is especially valid considering the amount of interaction and exposure between the resident and student. The authors have acknowledged this by stating that "At the request of our residents, students may again be approached regarding their participation." This would be an interesting addition to this study, especially with respect to the number of clusters. It has been well-described that this generation of medical students has a greater sense of entitlement or "what can you do for me" attitude relative to previous generations.[4] They tend to be candid, critical, but fair. Nevertheless, because of their chronologic proximity to the residents they work under, they can also identify with the resident more than the average attending surgeon. In sum, this group seems to have elements of both the nurses' cluster (in terms of commiseration, perhaps) and the surgeons' cluster (in terms of being more critical). Incorporation of this cohort into the 360-degree evaluation would be an intriguing addition to this study design.

This study represents a fantastic leap forward in the evaluation of plastic surgery residents in a 360-degree fashion. These types of evaluations are mandated by the Accreditation Council for Graduate Medical Education, yet no statistical evaluation of this method has been published in the plastic surgery literature until now. Although there are limitations to this study, it does paint a very clear picture that solicited ratings from different cohorts or clusters yield a more comprehensive evaluation of resident performance which, considered in sum, represent a more accurate picture. It cannot be overstressed that these data need to be shared with the resident being evaluated if any substantive change will occur. This information cannot simply be a part of a resident's file in a drawer or cabinet but must be part of a feedback loop to effect real change. The authors should be complimented on their well-written review of their experience and the conclusions that can be drawn from it.

Jeffrey E. Janis, M.D.
Department of Plastic Surgery
University of Texas Southwestern Medical Center
Dallas, Texas 75390-9132
jeffrey.janis@utsouthwestern.edu

REFERENCES

1. ACGME web site. Common program requirements. Available at: http://www.acgme.org/acWebsite/dutyHours/dh_dutyhoursCommonPR07012007.pdf. Accessed March 10, 2008.
2. Brasher, B. A., Chowdhry, S., Hauge, L. S., and Prinz, R. A. Medical students' perceptions of resident teaching: Have duty hours regulations had an impact? *Ann. Surg.* 242: 548, 2005.
3. Haghdoost, A. A., and Shakibi, M. Medical student and academic staff perceptions of role models: An analytical cross-sectional study. *B.M.C. Med. Educ.* 6: 9, 2006.
4. Rohrich, R. J. Training the generation X plastic surgeons: Dispelling the myths. *Plast. Reconstr. Surg.* 108: 1733, 2001.