

An End-of-Year Oral Examination for Internal Medicine Residents: An Assessment Tool for the Clinical Competency Committee

CRISTIN A. MOUNT, MD, FACP
PATRICIA A. SHORT, MD, FACP
GEORGE R. MOUNT, MD
CHRISTINA M. SCHOFIELD, MD, FACP

Abstract

Background Comprehensive evaluations of clinical competency consume a large amount of time and resources. An oral examination is a unique evaluation tool that can augment a global performance assessment by the Clinical Competency Committee (CCC).

Objective We developed an oral examination to aid our CCC in evaluating resident performance.

Methods We reviewed tools used in our internal medicine residency program and other training programs in our institution. A literature search failed to identify reports of a similar evaluation tool used in internal medicine programs. We developed and administered an internal medicine oral examination (IMOE) to our postgraduate year-1 and postgraduate year-2 internal medicine residents annually over a 3-year period. The results were used to enhance our CCC's

discussion of overall resident performance. We estimated the costs in terms of faculty time away from patient care activities.

Results Of the 54 residents, 46 (86%) passed the IMOE on their first attempt. Of the 8 (14%) residents who failed, all but 1 successfully passed after a mentored study period and retest. Less than 0.1 annual full-time equivalent per faculty member was committed by most faculty involved, and the time spent on the IMOE replaced regular resident daily conference activities.

Conclusions The results of the IMOE were added to other assessment tools and used by the CCC for a global assessment of resident performance. An oral examination is feasible in terms of cost and can be easily modified to fit the needs of various competency committees.

Editor's Note: The online version of this article contains a sample case with a scoring checklist and an example of the evaluation form.

Introduction

Since the Accreditation Council for Graduate Medical Education (ACGME) introduced the 6 competencies in

1999, a progressive shift to outcomes-based measurement of resident progress and competence has led residency programs to innovate evaluation tools. The ACGME offers tools for programs to incorporate into their evaluations, including standardized patients, objective structured clinical examinations, simulations and models, videotaped encounters, mini-clinical evaluation exercises (mini-CEXs), and standardized oral examinations.¹

A systematic review of 85 studies by Kogan et al² showed that residency programs had used 55 unique observation tools. Few of these tools had been thoroughly evaluated, although 11 contained multiple elements of validity, and the most studied tool ($n = 20$) was the mini-CEX. Focusing on 391 internal medicine (IM) residency programs, Chaudry et al³ surveyed program directors on the use of 12 evaluation tools to assess each core competency. The vast majority of the respondents used direct observation tools more often than any other tools, and the mini-CEX was favored by 90% of respondents.³ Use of an oral examination was not studied, despite the ACGME's citing an oral examination as the "next best method" for evaluating informed decision making and as a "most desirable tool"

All authors are at the Madigan Army Medical Center. **Cristin A. Mount, MD, FACP**, is Chief, Pulmonary, Critical Care and Sleep Medicine Service; **Patricia A. Short, MD, FACP**, is Program Director, Internal Medicine Residency Program; **George R. Mount, MD**, is Assistant Program Director, Internal Medicine Residency Program; and **Christina M. Schofield, MD, FACP**, is Assistant Program Director, Transitional Year Internship Program.

Funding: The authors report no external funding source for this study.

Conflict of Interest: The authors declare they have no competing interests.

The views expressed are those of the authors and do not reflect the official policy of the Department of the Army, the Department of Defense, or the US government.

Corresponding author: Cristin A. Mount, MD, FACP, Department of Medicine, Madigan Army Medical Center, Building 9040 Fitzsimmons Drive, Tacoma, WA 98431, 253-968.0833, cristin.a.mount.mil@mail.mil

Received October 10, 2013; revisions received January 28, 2014, and April 2, 2014; accepted April 14, 2014.

DOI: <http://dx.doi.org/10.4300/JGME-D-13-00365.1>

for assessing medical knowledge and practice-based learning and improvement.¹

Oral examinations are more prominent in training programs with an oral component to board certification examinations (the American Board of Internal Medicine [ABIM] eliminated its oral examination several decades ago). Oral examinations in these programs are associated with low cost and low-resource utilization compared with such other tools as simulation or standardized patient encounters.⁴

In 2011, our Clinical Competency Committee (CCC) recognized decreased in-training examination scores, less precise global rotation evaluations, and recent poor performance on the ABIM certifying examination. As a result, an end-of-year internal medicine oral examination (IMOE) was developed as an additional evaluation tool. The goal of the examination was to evaluate medical knowledge, critical thinking in evaluation, diagnosis, problem solving and therapeutics, and systems-based practice in areas not easily tested with traditional multiple-choice, board-style questions.

To our knowledge, this is the first report of an oral examination used as an evaluation tool for an IM CCC. We describe the process for obtaining validity evidence to support the assessment tool and how that tool aided in the evaluation of resident performance in the CCC.

Methods

Setting and Participants

Madigan Army Medical Center is the Army's major 240-bed tertiary care center in the Northwest. Our IM residency program trains between 8 and 11 residents each year. Core faculty and the program evaluation committee developed 25 examination cases that were intended to represent bread-and-butter IM practice (TABLE). Past program performance on the IM in-training examination (IM-ITE) was used to guide case topic selection.

The IMOE was administered to postgraduate year (PGY)-1 and PGY-2 residents in the last quarter of each academic year. Examiners were the program and assistant program directors and the core clinical faculty as well as the chief of medical residents and chief of medical residents-select, both of whom were members of the program evaluation committee and CCC at the time of IMOE implementation. To avoid peer bias, the chief of medical residents and chief of medical residents-select were each paired with senior faculty and only participated in examination of the PGY-1 residents. We looked to other residency programs in our institution that use a mock oral certification examination for guidance on testing logistics. A sample case and an example evaluation form are provided as online supplemental material.

What was known

Oral examinations are a useful assessment method for evaluating residents' decision-making skills as well as assessing their medical knowledge and practice-based learning and improvement.

What is new

The authors developed an oral examination for use in an internal medicine residency program that was a reliable tool, with supportive validity evidence, to augment discussions of residents' global performance.

Limitations

Single-site, single-specialty study limit generalizability.

Bottom line

An oral examination is feasible in terms of cost and can easily be modified to fit the needs of various competency committees.

To determine whether the IMOE was cost-effective, we calculated the approximate full-time equivalent commitment per faculty examiner.

Our Institutional Review Board deemed this study exempt, as it did not constitute research.

Methods for Validity Analysis

Fromme et al⁵ described 5 sources of evidence to determine the validity of assessment tools: content, response process, internal structure, relationship to other variables, and consequences.

Content Content evidence describes the process to confirm that an assessment tool tests the learner on what is actually being taught.⁵ Here, we described the process used to ensure that our cases were representative of what residents at each PGY level are expected to know. Program evaluation committee faculty reviewed each case to ensure that the content was appropriate, and we administered an anonymous posttest survey to faculty and resident examinees.

Response Process The response process describes how the residents and examiners respond to the examination and includes familiarity with the examination expectations, procedures, and evaluations. Increasing the number of observations and training of residents and faculty can increase the response process evidence.⁵

A faculty development workshop was conducted each year by the oral examination work group for new faculty examiners and those wishing to refresh their skills. Faculty members received instruction on how to complete the score report and review their assigned cases, and they could ask questions about the cases or the examination process. Before the examination, residents were briefed on the process and schedule. They were given examples of cases with appropriate responses and had the opportunity to ask

TABLE CURRENT CASE CONTENT LIST REPRESENTING BASIC THEMES IN INTERNAL MEDICINE AND MEDICINE SUBSPECIALTIES

Postgraduate Year–1 Case List	Postgraduate Year–2 Case List
Pancreatitis ^a	Hyponatremia ^a
Hyperkalemia ^a	Decompensated congestive heart failure ^a
Angina ^a	Adrenal insufficiency ^b
Chronic obstructive pulmonary disease ^a	Headache ^a
Health care maintenance ^b	Diabetic ketoacidosis ^a
Liver function test abnormality ^b	Atrial fibrillation ^a
Acute kidney insufficiency ^a	Gastrointestinal bleed ^a
Meningitis ^a	Atrial fibrillation ^b
Secondary hypertension ^b	Thrombotic thrombocytopenic purpura ^a
Community-acquired pneumonia ^a	Endocarditis ^a
Depression ^b	Sepsis ^a
	Tumor lysis syndrome ^a
	Stroke ^a
	Lung cancer ^b
	Autoimmune hemolytic anemia ^a

^a Scenarios that begin in an inpatient setting.^b Scenarios that begin in an outpatient setting.

questions; a PowerPoint presentation was also available for individual review. Standardized instructions were read to each group of residents immediately before testing. To increase the number of observations of residents and faculty, faculty members tested different residents over several days, and each resident was observed by 2 faculty members for each case.

Internal Structure Internal validity focuses on the reliability of scores and evaluators.^{5,6} A total of 202 examination cases were administered, and 404 faculty evaluation forms were completed. Interrater reliability was calculated for scores of “pass” or “fail.” Limited opportunities for examiners to retest the same case on multiple occasions prevented the calculation of intrarater reliability at this time.

Relationship to Other Variables We compared resident performance on the IMOE with the IM-ITE. Although the IMOE was designed to offer an alternative evaluation to knowledge assessment, we thought the basic content of the IM-ITE and the oral examination was similar.

Consequences Consequence evidence ensures that groups of residents who score the same on 1 test score similarly in other evaluated settings. We compared results on the IMOE to results of other assessments made during the same academic year, specifically the global rotation evaluations and the IM-ITE. A Pearson correlation coefficient was

calculated using performance on the IMOE grouped into thirds (high pass, pass, fail) and individual IM-ITE percentile scores. If lower performers received additional training and their scores improved, consequence evidence also improved.⁶ All residents who failed the IMOE were assigned a faculty mentor and were given a 6-week study period before retesting on new (to them) cases with new examiners.

Results

From 2011 to 2013, 54 residents were tested with an average of 12 faculty evaluators each year. Of the residents who took the oral examination, 50% (27 of 54) failed at least 1 case regardless of training year, 27% (6 of 22) of PGY-2 residents passed all cases, and 62% (20 of 32) of PGY-1 residents passed all cases. Over the past 3 years, 14% residents (8 of 54) failed the IMOE, and this number was distributed evenly between PGY years. One resident (PGY-1) failed all the cases tested, and 1 resident failed both the PGY-1 and PGY-2 examinations. Of those residents required to retake the IMOE ($n = 6$), only 1 failed the retest. Subsequent CCC outcomes for these residents, taken into context with overall residency performance, have included watch and wait, retake the IMOE after a period of focused study, program-level remediation, and, in 1 case, hospital-level probation.

In an anonymous posttest survey, residents and faculty thought the examination was organized, fair, and an

adequate demonstration of expected medical knowledge and patient care. With respect to evaluation results, interrater reliability was found to be very strong ($\kappa = 0.935$).

Compared with other evaluation tools, performance on the IMOE was found to weakly correlate with performance on the IM-ITE, with a Pearson correlation coefficient of $r = -0.3508$ ($P = .01$). With respect to global rotation evaluations, residents who performed well on the IMOE also generally performed well on global rotation evaluations. Residents who performed poorly on the IMOE often had signals of problems on global evaluations that were more noticeable in hindsight through review at the CCC. A specific correlation coefficient was not calculated for the global rotation evaluations as the structure and output of these evaluations were undergoing constant revision during this time.

Discussion

An oral examination used as an observation tool is a unique evaluative process for IM residency programs. Over our 3-year experience, we have obtained validity evidence for our examination using many of the criteria recommended by Fromme et al.⁵

Our CCC believed the information about resident performance gained from the IMOE was worthwhile and aided in discussions of residents, particularly troubled learners. The IMOE results were used in conjunction with other evaluations and information to initiate remediation or hospital probation as needed. On several occasions, the CCC used performance on the IMOE as a measure of successful completion of program-level remediation. In other cases, residents with marginal performance were more closely monitored during subsequent CCC evaluation periods. Thus, the IMOE served as an early warning system that allowed learning issues to be addressed before program-level remediation became necessary.

In our system, the IMOE was a low-cost success compared with more traditional observation tools such as simulation and standardized patient encounters. And although significant time was spent in the initial development phase, maintenance of the program required few further resources.

To determine whether the IMOE was cost-effective, we calculated the approximate full-time equivalent time commitment per faculty examiner. For the first IMOE, the working group faculty spent an average of 15 hours in preparation for the 6 months before testing. The group devised cases and trained other faculty members. In subsequent years, the preparatory time commitment decreased, though the time spent testing residents remained stable. The 5 faculty members involved in IMOE admin-

istration, case content review, and IMOE testing spent an average of 11 hours working over the month before the testing period and an additional 6 hours during the testing period. The 7 faculty members involved in case content review and IMOE testing spent an average of 1 hour over the month before the testing period and 4 hours during the testing period. Faculty who participated in testing only averaged 3.75 hours.

The IMOE impact on patient care services was minimal as these hours of commitment amounted to less than 0.1 full-time equivalent of work averaged over a week. The examinations were scheduled when clinics were closed for lunch and inpatient morning rounds had been completed. Usual scheduled academic activities were canceled during testing week to allow residents compensatory time for patient care. Conference rooms and office space for the IMOE did not displace patient care activities. There was a negligible cost associated with photocopying the case and evaluation forms.

Further study is needed to increase internal structure validity evidence, paying specific attention to intrarater reliability, generalizability evidence, and ongoing content validation when new cases are written. This will be an important process as the CCC composition undergoes revision to reflect current ACGME guidelines. Additional efforts to correlate the results with IM-ITE, ABIM certifying examination, and existing hospital-wide standardized patient examinations may also boost the validity evidence.

Conclusion

We developed an oral examination for use in an IM residency program that was a reliable tool, with supportive validity evidence, to augment discussions of residents' global performance in the CCC. The format lends itself to adaptation to program or learner needs and is reasonably cost-effective. Oral examinations are useful and should be added to the evaluation toolbox for specialty programs that do not have an oral element to board certification examinations.

References

- 1 Accreditation Council for Graduate Medical Education, ACGME Outcomes Project. Toolbox of Assessment Methods. 2000. <http://dconnect.acgme.org/Outcome/assess/Toolbox.pdf>. Accessed July 16, 2014.
- 2 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.
- 3 Chaudry SI, Holmboe E, Beasley BW. The state of evaluation in internal medicine residency. *J Gen Intern Med*. 2008;23(7):1010–1015.
- 4 Jefferies A, Simmons B, Ng E, Skidmore M. Assessment of multiple physician competencies in postgraduate training: utility of the structured oral examination. *Adv Health Sci Educ*. 2011;16(5):569–577.
- 5 Fromme H, Karani R, Downing S. Direct observation in medical education: a review of the literature and evidence for validity. *Mt Sinai J Med*. 2009;76(4):365–371.
- 6 Sullivan G. A primer on the validity of assessment instruments. *J Grad Med Educ*. 2011;3(2):119–120.