Residents' Ability to Interpret Radiology Images:

Development and Improvement of an Assessment Tool

F. Jacob Seagull, PhD, Janet E. Bailey, MD, Andrew Trout, MD, Richard H. Cohan, MD, Monica L. Lypson, MD, MHPE

Rationale and Objectives: Despite increasing radiology coverage, nonradiology residents continue to preliminarily interpret basic radiologic studies independently, yet their ability to do so accurately is not routinely assessed.

Materials and Methods: An online test of basic radiologic image interpretation was developed through an iterative process. Educational objectives were established, then questions and images were gathered to create an assessment. The test was administered online to first-year interns (postgraduate year [PGY] 1) from 14 different specialties, as well as a sample of third- and fourth-year radiology residents (PGY3/R2 and PGY4/R3).

Results: Over a 2-year period, 368 residents were assessed, including PGY1 ($n = 349$), PGY3/R2 ($n = 14$), and PGY4/R3 ($n = 5$) residents. Overall, the test discriminated effectively between interns (average score = 66%) and advanced residents (R2 = 86%, R3 = 89%; $P < .05$). Item analysis indicated discrimination indices ranging from −0.72 to 48.3 (mean = 3.12, median 0.58) for individual questions, including four questions with negative discrimination indices. After removal of the negatively indexed questions, the overall predictive value of the instrument persisted and discrimination indices increased for all but one of the remaining questions (range 0.027–70.8, mean 5.76, median 0.94).

Conclusions: Validation of an initial iteration of an assessment of basic image-interpretation skills led to revisions that improved the test. The results offer a specific test of radiologic reading skills with validation evidence for residents. More generally, results demonstrate a principled approach to test development.

Key Words: Radiology; image interpretation; assessment; resident education.

Basic radiology skills are critical to almost all practicing physicians. Graduate-level training in radiology has been established as an important skill with varying pedagogical approaches (1). With the recent emphasis on competency-based assessment (2), the need for valid task-based methods of assessment of radiology skills is paramount.

Because first-year residents, in programs other than radiology, must review imaging studies in settings with indirect supervision, it is especially important to assess this skill to establish competence. At times, residents with indirect supervision are asked to interpret radiographs and other imaging studies without input from a radiologist, or provide a preliminary read before receiving a somewhat delayed final radiologist's interpretation. First-year residents not specializing in radiology are often required to independently interpret basic radiologic studies, in most instances while on overnight service, however their ability to do so accurately is not routinely assessed. Residents' image-interpretation errors have been documented in various settings and specialties, although reports of the rate and severity of the errors vary (3–5). Moreover, with the increasing reliance on electronic medical records and computer-based picture archiving and communication systems, clinical radiologic education is occurring on a more limited basis. Largely gone are radiology rounds when entire clinical teams would visit the radiology department for review of the imaging and discussion of the patients on their service.

Although it is likely that interns and residents will be required to interpret imaging on call, imaging interpretation is generally not a skill acquired during undergraduate medical education (UME) (6,7). Medical students are increasingly marginalized in the clinical environment because concerns regarding accurate billing and clinical load have limited student contact with patients in the name of clinical and financial efficiency. In addition, because of increasing time pressures related to patient length of stay, and the increasing number of patient handoffs, it is unclear if sufficient time is

Acad Radiol 2014; 21:909–915

From the Department of Medical Education, University of Michigan Medical School, 221 Victor Vaughan Building, 1111 E. Catherine SPC-2054, Ann Arbor, MI 48109-2054 (F.J.S., M.L.L.); Department of Radiology, University of Michigan Medical School, Ann Arbor, MI (J.E.B., R.H.C.); Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH (A.T.); and Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI (M.L.L.). Received January 20, 2014; accepted March 11, 2014. **Address correspondence to:** F.J.S. e-mail: jseagull@umich.edu

http://dx.doi.org/10.1016/j.acra.2014.03.010

dedicated in UME for radiology education (8,9). Many medical students do take a radiology rotation as a third- or fourth-year elective, but these experiences are variable in length and intensity (10).

Various methods have been used to foster image-interpretation skills and to assess resident ability to interpret imaging, including peer review of image interpretation as a learning tool (11), and comparison between attending and resident interpretation of imaging studies (12). As with other performance-based clinical skills, there is a measurable learning curve for these image-interpretation skills (13).

Evidence in the literature is mixed regarding the competence of residents to interpret imaging studies independently, as there are no widely accepted assessment standards for basic radiology skills, especially for trainees who are not pursuing radiology as their eventual specialty. Radiology residents do acquire a certain level of competence through the course of their training, and there have been efforts reported of providing standard assessments throughout the residency program, with opportunity for remediation when needed (14). In some areas of radiology, there are no significant differences between specialists and on-call residents (15). In other areas, however, discrepancies do exist with the magnitude of interpretation discrepancies between residents and faculty differing depending on the specialty of the resident and the faculty (16,17).

One recommended, accepted method for developing assessment for clinical skills involves assembling stakeholders (subject matter experts and educators) and generating content based on consensus (18). As previously described, such a process was used in the development and implementation of a test of imaging interpretation competency at our institution (19–21).

The purpose of the present study was to evaluate and improve the quality of a radiology image-interpretation test for nonradiology interns. By applying item analysis to the targeted population and a small sample of additional expert populations (advanced radiology residents), we hoped to determine whether our examination could be improved. We also sought to determine whether we could provide a template for how to use validation evidence to evaluate an assessment scale.

MATERIALS AND METHODS

Examination Development

Using an iterative modified Delphi process to gain consensus (21), a group of experts that included clinical radiologists, graduate medical education program directors and PhD-trained researchers in medical education determined common types of radiographs typically ordered and interpreted by interns under indirect supervision. This information was then used by the group to establish educational objectives, generate questions, and compile images to create an examination that assessed trainees' skills similar to those they would be expected to encounter during their first year of training. The result was a 20-question online assessment of radiologic image interpretation to be administered to incoming first-year nonradiology residents (19,22). The institutional review board deemed the development and validation activities described in this study to be exempt from ongoing review as research in an established educational setting involving normal educational practices.

Our examination followed a case-presentation format. Each case consisted of 1) a radiologic image set and 2) a brief query or description of the patient and his or her chief complaint. Images included radiographs and CT images that were deemed to be representative of images that residents were expected to encounter. For example, a posteroanterior chest radiograph was presented with the question, "A 65 year old patient who is short of breath after bronchoscopy." or "Is the tip of the Swan-Ganz catheter placed correctly?" Respondents could select from 22 possible diagnoses or yes/no answers for specific queries. A software package for online test and quiz development and administration (Questionmark, www.questionmark.com) was used to present the text and static images along with a drop-down menu containing possible responses. There were also two multiple-choice questions regarding image radiation exposure risk. See Appendix A for details of examination.

Examination Administration

The developed imaging examination was a component of a comprehensive baseline assessment of competencies (Postgraduate Orientation Assessment [POA]) administered online to all first-year (postgraduate year [PGY] 1) interns at the University of Michigan. Interns in 14 different specialties/residency programs participated in the assessment (Table 1). At the time of assessment, none of these interns had received radiology training as part of their internship. Results from this POA were used as one source of validity evidence.

In an effort to provide additional validity evidence for our imaging examination, it was also administered to radiology residents in their second or third year of radiology training (PGY3/R2 and PGY4/R3). Residents at this level of training are expected to be fully competent to interpret common imaging studies under indirect supervision, with direct supervision available. Thus, for the purpose of assessing the validity of the imaging examination for interns, we compared performances of first-year residents (interns) from various departments with those of residents in their second and third year of radiology training.

Validity Evidence Analysis

Two levels of validity analysis were performed: 1) analysis of the quality of the test as a whole and 2) analysis of the quality of each of the test questions that comprise the test.

Validity evidence at the test level indicates the extent to which the assessment as a whole discriminates between competent and less competent residents. For a test to be useful, it would be expected that junior residents and interns