The Radiation Oncology In-Training Exam: An Appeal for Better Testing

Ann Morris MD
University of Iowa Hospitals & Clinics, Iowa City, IA

Purpose

Resident physician assessment plays a key role in graduate medical education. As of 1992, 20 medical specialties offered an in-training exam for the purposes of formative assessment. In 1984, the American College of Radiology developed an in-training exam (TXIT) for radiation oncology. Consisting of approximately 350 questions, this exam addresses the topics of radiation biology, radiation physics and clinical oncology. Appropriate use of the TXIT has not been determined, yet 33% of radiation oncology residents surveyed in 2004 felt the exam had a moderate to significant impact on their annual review. Given this perception, critical evaluation of the evidence supporting use of the TXIT is needed in order to optimize use of this exam in formative assessment.

Methods

A review was conducted of all literature published on the topic of the American College of Radiology TXIT between 1984 and 2012. All relevant papers were analyzed for evidence of psychometric reliability and validity, using a framework of content validity, construct validity and concurrent/predictive validity.

Results

From 1988 to 2008, four papers were published on the topic of the ACR in-training exam (TXIT). Only one article addressed the issue of reliability, demonstrating a mean item difficulty of 60.7% ± 22.4% with mean item discrimination of 0.19 and an overall reliability coefficient of 0.92. While favorable, these values were generated from a single year of data (2007) and thus may not be generalizable. A review of the 2004-2007 TXIT attempted to address content validity by showing that each exam contains a similar number of questions on each of 12 clinical subsites; no data was provided, however, to demonstrate the clinical relevance of the items themselves. Content analysis of the 2012 TXIT revealed 53 items (17%, n=321) in which all four foils were percentages and 43 items (13%) that focused on a clinical trial; these items may or may not be viewed as directly relevant to clinical practice (Figures 1-3). Regarding construct validity, articles from 1988, 1992 and 2008 showed that TXIT scores tend to improve with years in training and concurrent/predictive validity.

Conclusions

Very little evidence exists regarding the psychometric properties of the ACR TXIT. Reliability data exists for one administration of the test, and average scores show a modest trend towards improvement with increased years in training. Evidence for content, concurrent and predictive validity, however, remains lacking and represents a large gap in the literature, diminishing the exam’s utility. Future research should focus on defining the role of the TXIT as a formative assessment in resident education through rigorous psychometric analysis.

Figures 1-3: Representative questions from the 2011 American College of Radiology In-Training Exam (TXIT). Anecdotal criticisms of the exam have been that the exam seems to focus on recall, requires in-depth knowledge of clinical trials and is not representative of actual clinical practice.

1. What was the 1 year local control rate after administration of individualized stereotactic body radiation therapy for liver metastases with no size limits, according to a multi-institutional phase I study?
   - A) 29%
   - B) 52%
   - C) 71%
   - D) 87%

2. According to the PORTEC-2 (Nout, et al) trial, what is the 5-year risk of pelvic recurrence in a patient who received vaginal brachytherapy for intermediate to high risk endometrial cancer?
   - A) 4.8%
   - B) 3.3%
   - C) 2.4%
   - D) 0.6%

3. What is the tumor stage of a squamous cell carcinoma that arises from the penile urethra and invades the corpus cavernosum?
   - A) T1
   - B) T2
   - C) T3
   - D) T4

Table 1

<table>
<thead>
<tr>
<th>Psychometric Criteria</th>
<th>Application to the TXIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability</td>
<td>Are TXIT scores reproducible?</td>
</tr>
<tr>
<td>Content Validity</td>
<td>How well does the exam represent the medical knowledge required in the practice of radiation oncology? Is that knowledge tested at an appropriate cognitive level, and do the items reflect actual clinical practice?</td>
</tr>
<tr>
<td>Construct Validity</td>
<td>Does the exam produce results that make sense? For example, do residents demonstrate stepwise improvement in their TXIT scores throughout training?</td>
</tr>
<tr>
<td>Concurrent and Predictive Validity</td>
<td>Does the TXIT correlate with other accepted measures of resident success, such as faculty evaluations or the American Board of Radiology examinations?</td>
</tr>
</tbody>
</table>

References