# Analysis of the Radiation Oncology In-Training Examination Content Using a Clinical Care Path Conceptual Framework

Kevin R. Rogacki, MD; Stanley Gutiontov, MD; Chelain R. Goodman, MD, PhD; Elizabeth Jeans, MD, MEd; Yasmin Hasan, MD; Daniel W. Golden, MD, MHPE

### Abstract

**Hypothesis:** The American College of Radiology (ACR) Radiation Oncology In-Training Examination (TXIT) is an assessment administered by radiation oncology training programs to assess resident performance against national benchmarks. Results are currently reported using a disease site conceptual framework. The clinical care path framework was recently proposed as a complementary view of resident education. This study assesses distribution of 2016-2019 TXIT questions using the clinical care path framework with the hypothesis that questions are unequally distributed across the clinical care path framework, leading to underassessment of fundamental clinical skills.

**Methods and Materials:** Using a clinical care path framework, 1,200 questions from the 2016-2019 TXITs were categorized into primary categories and subcategories. The distribution of questions was evaluated.

**Results:** Primary categorization was achieved for 98.7% of questions. Where applicable, subcategorization was achieved for 96.6% of questions. There was substantial inter-rater reliability (primary category κ = 0.78, subcategory κ = 0.79). Distribution of TXIT content by the clinical care path framework was: treatment decision (35%), diagnosis (16%), radiation biology (12%), radiation physics (12%), treatment planning (9%), biostatistics (4%), cancer biology (4%), toxicity and management (4%), brachytherapy (2%), quality assurance (1%), and research methods (1%). Of the 419 questions within the treatment decision primary category, knowledge from randomized clinical trials was most frequently evaluated (43%).

**Conclusions:** TXIT question items are unequally distributed across clinical care path categories, emphasizing treatment decision over other categories such as treatment planning and toxicity and management. Reporting examination data by both clinical care path and disease site conceptual frameworks may improve assessment of clinical competency.

The American of College of Radiology (ACR) Radiation Oncology In-Training Examination (TXIT) is a standardized assessment used to assess radiation oncology trainees' acquisition of knowledge necessary for independent practice. By providing "mean norm-referenced scores at the national, institutional, and individual levels," the TXIT serves as a formative assessment to inform

**Figure 1.** Radiation oncology education clinical care path conceptual framework. From Perez and Brady's Principles and Practice of Radiation Oncology, 7th ed (p. 2243), by D.W. Golden and P. Ingledew, 2018, Lippincott, Williams & Wilkins. Printed with permission.
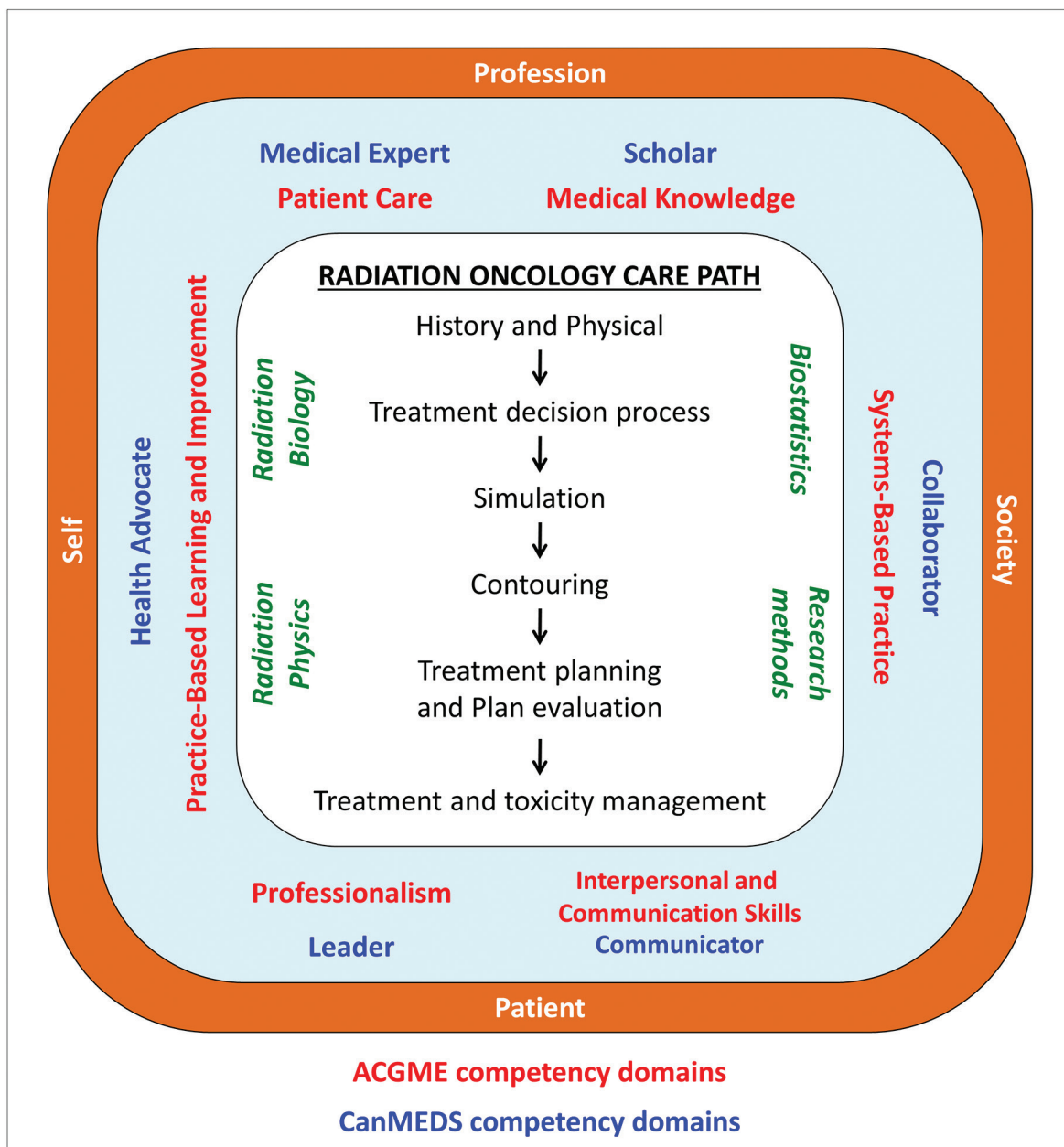


trainees and training programs on content areas that may require additional attention for self-study or formal didactics.[1-4]

The annual TXIT examination consists of 300 single-answer, multiple-choice questions and is sponsored by the ACR Commission on Publications and Lifelong Learning.[3] Examination content is organized into 13 sections according to a disease site conceptual framework (eg, thoracic, breast, lung), or by basic science subject (biology, physics, statistics).

In medical education, a conceptual framework facilitates organization and perception of educational content. A conceptual framework can also "represent [a] way of thinking about a problem" and influence how inter-related topics may be considered.[5] Traditionally, the field of radiation oncology has used a disease site conceptual framework to organize resident education. The use of a disease site framework is reflected in didactic

**Figure 2.** Clinical care path primary categories (A), subcategories (B), and applied sciences categories (C) used to classify TXIT in-training exam questions. *Brachytherapy included as a primary category to emphasize its unique position within the clinical care path.



curricula, clinical rotations, case logs, textbooks, and board examination categories.[6,7] This prevailing conceptual framework also underlies development of assessment tools such as the TXIT examination.[4]

A conceptual framework based on the radiation oncology clinical care path was recently proposed and represents an alternative and complementary lens through which to view radiation oncology education (**Figure 1**).[8] The clinical care path conceptual framework represents the stepwise clinical activities involved in providing care to a patient receiving radiation therapy. These steps begin at the initial consultation and encompass the treatment decision, simulation, contouring, treatment planning, quality assurance, toxicity management, and long-term follow-up. For medical specialties with a technical component, such a framework may better encompass the spectrum of professional activities in which a physician must demonstrate proficiency to be considered competent for independent practice. As a result, the clinical care path framework may provide a more direct link to competency-based medical education.[9-11]

The extent to which the TXIT assesses competency in the clinical activities central to the practice of radiation oncology as defined by the clinical care path conceptual framework is unknown. The purpose of this study was to assess the distribution of questions in the 2016-2019 TXIT examinations using the clinical care path framework. We hypothesized that the TXIT content is unequally distributed across the clinical care path framework, leading to underassessment of fundamental clinical skills.

## Methods and Materials

### Category Definition

The clinical care path framework was adapted to derive clinical care path primary categories and subcategories (**Figure 2**) along with definitions to guide categorization by independent coders (**Table 1**). Subjects outside of the clinical care path but fundamental to the practice of radiation oncology including radiation biology, cancer biology, radiation physics, biostatistics, and research methods were included as separate applied sciences primary categories (**Figure 2**).

### Coding of Question Items

Two independent coders categorized 1,200 questions from the 2016-2019 TXIT examinations based on the content of the question stem, answer choices, and associated rationale. Inter-rater reliability was assessed with Cohen's kappa coefficient test statistic. Items with discordant categorizations were independently reconciled by a third coder. If reconciliation was not achieved, the question was deemed uncategorizable

## Table 1. Definitions Guide for Question Categorization

| PRIMARY CATEGORIES AND SUBCATEGORIES | DEFINITION |
|---|---|
| **Clinical Care Path Categories** | |
| **Diagnosis** | **Questions relating to the diagnosis and integration of clinicopathologic data during initial consultations** |
| ▪History and physical | History and exam findings pertinent to a diagnosis of cancer |
| ▪Epidemiology | Epidemiological facts relating to cancer incidence, prevalence, and population level data |
| ▪Imaging | Imaging modalities or imaging findings |
| ▪Onco-anatomy | Anatomy questions, patterns of spread, and questions about loss or alteration of function due to tumor involvement |
| ▪Pathology | Histological diagnosis and principles |
| ▪Risk factors | Predisposing factors to cancer development |
| ▪Staging | TNM staging, clinical group staging, pathological group staging, and risk groups |
| ▪Workup | Catch-all category for utilization or interpretation of other studies performed during diagnosis (ie, laboratory tests) |
| **Treatment Decision** | **Questions relating to treatment principles or medical knowledge informing treatment decision making** |
| ▪Management | General treatment principles (often derived from consensus guidelines) |
| ▪Nonrandomized data | Phase I, phase II, and observational studies |
| ▪Prognostic/predictive factors | Clinical, pathological and treatment factors that are prognostic for outcome or predictive for treatment effect |
| ▪Randomized control data | Phase III, randomized control studies, including meta-analyses |
| **Treatment Planning** | **Questions relating to radiation treatment planning** |
| ▪Contouring | Tumor volumes, normal volumes, field placement, and field borders |
| ▪Dose constraints | QUANTEC data, dosimetric studies, dose constraints utilized in clinical trials |
| ▪Dosimetry | Radiation planning and dosimetry concepts (ie, measurement of dose) |
| ▪Simulation | Simulation modalities, patient setup, and skills employed during simulation |
| **Quality Assurance** | **Questions pertaining to QA activities performed as part of treatment verification and/or delivery (ie, IGRT)** |
| **Brachytherapy** | **Questions relating to brachytherapy** |
| **Toxicity and Management** | **Questions relating to identification, mitigation, and management of treatment-related toxicities** |
| ▪Acute | Acute toxicity occurring during and just after completion of treatment (fatigue, dermatitis, mucositis, proctitis, cystitis, etc.) |
| ▪Late | Late toxicity occurring because of delayed radiation effects (secondary malignancy, fertility, chronic organ dysfunction, etc.) |
| **Applied Sciences Categories** | |
| **Biostatistics** | **Questions employing statistical principles** |
| **Research Methods** | **Questions pertaining to the research process and trial design** |
| **Cancer Biology** | **Questions about tumor biology, cellular physiology, and molecular mechanisms in the absence of radiation effects (ie, mutations, genomic alterations, cell receptors, downstream effectors, systemic therapy mechanism of action)** |
| **Radiation Physics** | **Questions pertaining to fundamental physics concepts and material commonly taught in radiation physics courses or discussed in radiation physics texts** |
| **Radiation Biology** | **Questions about the biological action of radiation treatment and fundamental radiobiological concepts commonly taught in a radiation biology course or discussed in radiobiology texts** |

*Bold indicates primary category      †Bullets indicate subcategory within a primary category if applicable

Key: TNM = tumor, nodes, metastases; QUANTEC = Quantitative Analyses of Normal Tissue Effects in the Clinic; QA = quality assurance; IGRT = image-guided radiation therapy

## Table 2. Examples of Reconciled or Uncategorized Question Items

| EXAMPLE | QUESTION ITEM | ANSWER AND RATIONALE | CATEGORIZATION BY CODER† | FINAL CATEGORIZATION |
|---|---|---|---|---|
| Primary Category – Reconciled | Which tumor markers can be used as adjunctive studies to assess for disease progression in metastatic breast cancer?<br><br>A. CEA, CA 15-3, CA 19.9<br><br>B. CEA, CA 15-3, CA 27.29<br><br>C. CEA, CA 19.9, CA 27.29<br><br>D. CA 15-3, CA 19.9, CA 27.29 | Key: B. Solution/Rationale: NCCN breast cancer guidelines include optional use of CEA, CA 15-3, CA 27.29 in the assessment of metastatic breast cancer. The ASCO guidelines on biomarkers to guide systemic therapy for metastatic breast cancer state that CEA, CA 15-3, and CA 27.29 may be used to assist in treatment decision-making; however, they should not be used alone for monitoring disease response. CA 19.9 is not used in breast cancer; more often in pancreatic cancer.<br><br>Van Poznak C, Somerfield MR, Bast RC, et al. Use of biomarkers to guide decisions on systemic therapy for women with metastatic breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol*. 2015;33:2695-2704. | 1: Diagnosis - Workup<br><br>2: Treatment Decision-Management<br><br>3: Diagnosis - Workup | Diagnosis - Workup |
| Subcategory – Reconciled | For pleomorphic adenoma of the parotid gland, which feature increases the risk of local recurrence?<br><br>A. Patient age<br><br>B. Tumor size<br><br>C. Male gender<br><br>D. Tumor spillage | Key: D. Solution/Rationale: Surgical excision with a superficial or complete parotidectomy and facial nerve preservation is the recommended treatment for pleomorphic adenoma. Tumor spillage, residual disease, and recurrent disease after initial surgery are risk factors for additional recurrence. With a margin-negative surgical excision, size is not an independent risk factor for recurrence of a pleomorphic adenoma.<br><br>Liu FF, Rotstein L, Davison AJ, et al. Benign parotid adenomas: a review of the Princess Margaret Hospital experience. *Head Neck*. 1995;17(3):177-183. | 1: Retrospective data<br><br>2: Prognostic/predictive factors<br><br>3: Prognostic/predictive factors | Treatment Decision – Prognostic/ predictive factors |
| Primary Category – Uncategorized | What is the MOST common radiation technique for treatment of pterygium?<br><br>A. Electrons with bolus<br><br>B. IMRT<br><br>C. Orthovoltage photons<br><br>D. Beta emitter brachytherapy | Key: D. Solution/Rationale: Contact brachytherapy with strontium-90 is commonly utilized in the postoperative treatment of pterygium. Other methods are less focal.<br><br>Pashtan I, Oh KS, Loeffler JS. Radiation therapy in the management of pituitary adenomas. *Int J Radiat Oncol Biol Phys*. 1994;28(1):101-103. | 1: Radiation Physics<br><br>2: Treatment Decision<br><br>3: Treatment Planning | Uncategorized (primary category) |
| Subcategory – Uncategorized | What is the approximate 5-year local regional failure rate of T4, node positive anal cancer after definitive chemoRT?<br><br>A. 20%<br><br>B. 40%<br><br>C. 60%<br><br>D. 80% | Key: C. Solution/Rationale: According to the secondary analysis of RTOG 98-11 stratified by TN staging, for patients with T4, N1-3 cancer, the rate of 5-year LRF was 60%.<br><br>Gunderson LL, Moughan J, Ajani JA, et al. Anal carcinoma: impact of TN category of disease on survival, disease relapse, and colostomy failure in US Gastrointestinal Intergroup RTOG 98-11 phase 3 trial. *Int J Radiat Oncol Biol Phys*. 2013;87(4):638-645. | 1: Randomized control data<br><br>2: Nonrandomized data<br><br>3: Prognostic factors | Treatment Decision – Uncategorized (subcategory) |

† Coder 1: SG; Coder 2: KRR; Coder 3: CRG
Key: NCCN = National Comprehensive Cancer Network, ASCO = American Society of Clinical Oncology, IMRT = intensity-modulated radiation therapy, RTOG = Radiation Therapy Oncology Group, LRF = locoregional failure

## Table 3. Patterns of Disagreement in TXIT Question Items Requiring Reconciliation

**Disagreement by Subcategory**

| SCENARIO | N (215 TOTAL) | PATTERN OF DISAGREEMENT IN CATEGORIZATION |
|---|---|---|
| Disagreement between all coders | 16 | No pattern |
| Disagreement between initial coders, reconciled by third coder | 199 | Most common disagreements by primary categorization*:<br>▪ diagnosis & treatment decision (34)<br>▪ treatment decision & treatment planning (28)<br>▪ cancer biology & radiation biology (23)<br>▪ radiation physics & quality assurance (12)<br>▪ treatment planning & radiation physics (12)<br>▪ treatment planning & toxicity and management (12)<br>▪ biostatistics & research methods (10) |

**Disagreement by Subcategory**

| SCENARIO | N (128 TOTAL) | PATTERN OF DISAGREEMENT IN CATEGORIZATION |
|---|---|---|
| Disagreement between all coders | 18 | ▪ 16 of 18 disagreements at level of subcategory within treatment decision primary category (ie, management, nonrandomized data, randomized data, retrospective data, prognostic/predictive factors) |
| Disagreement between initial coders, unreconciled by third coder (different primary category) | 25 | ▪ 23 of 25 were categorized by initial coders in primary category of treatment decision<br>▪ 16 of these 23 were labelled by the third coder as diagnosis (6) or treatment planning (10) |
| Disagreement between initial coders, reconciled by third coder | 85 | Most common primary categories for which subcategories were reconciled:<br>▪ treatment decision (53)<br>▪ diagnosis (22)<br>▪ treatment planning (7) |

*n indicates number of questions*
*\*Most frequent pairwise disagreements listed in descending order from most to least common (≥ 10 occurrences)*
*†Number of disagreements by category or category pairs provided in parentheses*

at the primary or subcategorization level. Examples of reconciled and uncategorized questions are provided in **Table 2**. A single coder categorized questions according to disease site and content area. At the time of question categorization, all coders were radiation oncology residents at accredited US residency programs.

Institutional review board review of this study was not obtained as no human subjects were researched.

## Results

### Question Classification

Initial question categorization was achieved with substantial agreement between two independent coders (primary category, κ = 0.78;

subcategory, κ = 0.79). In total, 343 questions (28.6%) required categorization by a third coder. Of the 1,200 question items, 1184 (99%) were successfully categorized by primary category. Of 199 questions for which reconciliation of the primary category was achieved, items were most commonly labeled as treatment decision by one coder and either diagnosis (n = 34, 17%) or treatment planning (n = 28, 14%) by the second coder. Of the 762 question items with a subcategory, 719 (94%) were successfully subcategorized. Of 85 questions for which reconciliation of the subcategory was achieved, most were within the treatment decision (n = 55, 65%), diagnosis (n = 22, 25%), or treatment

planning primary categories (n = 7, 8%). Additional details of reconciled questions are available in **Table 3**.

### TXIT Content by the Clinical Care Path Framework

The distribution of question items from the 2016-2019 TXIT examinations according to the clinical care path framework is reported in **Table 4**. A total of 796 (66%) question items were classified using the clinical care path, with the remaining 404 (34%) categorized as applied sciences. Clinical care path questions assessing treatment decisions were most prevalent, representing approximately 35% of all items (n = 435). These questions most frequently evaluated data

**Table 4. TXIT Question Categorization by a Clinical Care Path Framework for Examination Years 2016-2019**

| PRIMARY CATEGORIES AND SUBCATEGORIES | N (%) |
|---|---|
| *Clinical Care Path Categories* | *796 (66)* |
| **Diagnosis** | **186 (16)** |
| Epidemiology | 18 (2) |
| Imaging | 10 (1) |
| Onco-anatomy | 44 (4) |
| Pathology | 20 (2) |
| Risk factors | 10 (1) |
| Staging | 57 (5) |
| Workup | 10 (1) |
| Uncategorized subcategory | 14 (1) |
| **Treatment Decision** | **419 (35)** |
| Management | 103 (9) |
| Nonrandomized data | 32 (3) |
| Prognostic/predictive factors | 37 (3) |
| Randomized control data | 182 (15) |
| Retrospective data | 15 (1) |
| Uncategorized subcategory | 50 (4) |
| **Treatment Planning** | **112 (9)** |
| Contouring | 37 (3) |
| Dose constraints | 34 (3) |
| Dosimetry | 25 (2) |
| Simulation | 5 (<1) |
| Uncategorized subcategory | 11 (1) |
| **Quality Assurance** | **11 (1)** |
| **Brachytherapy** | **23 (2)** |
| **Toxicity and Management** | **45 (4)** |
| Acute | 9 (1) |
| Late | 34 (3) |
| Uncategorized subcategory | 2 (<1) |
| *Applied Sciences Categories* | *404 (34)* |
| **Biostatistics** | **46 (4)** |
| **Research Methods** | **14 (1)** |
| **Cancer Biology** | **48 (4)** |
| **Radiation Physics** | **140 (12)** |
| **Radiation Biology** | **140 (12)** |
| Uncategorized Primary Category | 16 (1) |
| **TOTAL** | **1200 (100)** |

*Bold items indicate primary categories with subcategories listed beneath if applicable
†Percentages rounded to nearest whole number
Key: TXIT = Radiation Oncology In-Training Exam

derived from randomized clinical trials (n = 182). Questions assessing diagnosis were the second most common (n = 186, 16%) and most frequently assessed tumor staging. Approximately 10% of questions assessed treatment planning (n = 112), of which approximately two-thirds were related to contouring and dose constraints. Questions evaluating treatment toxicity represented approximately 4% of items. Questions assessing brachytherapy included 23 questions within a 4-year testing period (2%).

For the applied sciences questions, radiation and cancer biology represented approximately 16% of all question items (n = 188), followed by radiation physics (n = 140, 12%), and biostatistics/research methods (n = 60, 5%).

**TXIT Content by a Disease Site Framework**

When classifying questions other than those defined as applied sciences according to a disease site framework, disease sites were represented approximately equally, with 6% to 8% of total questions dedicated to most sites (**Table 5**). Breast (n = 95, 8%) and head and neck (n = 93, 8%) were most frequently assessed, followed by lymphoma, pediatrics, genitourinary, and gynecologic disease sites (n = 73 to 83, 7%).

**Discussion**

Content analysis of the TXIT using a clinical care path framework demonstrates an uneven distribution in the number of questions allocated to the different steps of the clinical care path. Specifically, the exam most frequently assesses knowledge used to guide treatment decisions with fewer questions assessing treatment planning skills and management of treatment-related toxicity. This uneven distribution is not apparent when evaluating question content through a disease site framework.

**Table 5. TXIT Question Categorization by a Disease Site Framework for Examination Years 2016-2019**

| DISEASE SITE | TXIT YEAR, N (%) | | | | |
|---|---|---|---|---|---|
| | 2016 | 2017 | 2018 | 2019 | 2016-2019 Average |
| Breast | 24 (8) | 23 (8) | 25 (8) | 23 (8) | 24 (8) |
| Central Nervous System | 14 (5) | 21 (7) | 19 (6) | 16 (5) | 18 (6) |
| Gastrointestinal | 20 (7) | 18 (6) | 21 (7) | 17 (6) | 19 (6) |
| Genitourinary | 22 (7) | 19 (6) | 17 (6) | 22 (7) | 20 (7) |
| Gynecology | 19 (6) | 21 (7) | 19 (6) | 19 (6) | 20 (7) |
| Head and neck | 23 (8) | 22 (7) | 24 (8) | 24 (8) | 23 (8) |
| Lymphoma | 19 (6) | 23 (8) | 20 (7) | 21 (7) | 21 (7) |
| *Other | 2 (1) | 1 (<1) | 0 | 0 | 1 (<1) |
| Palliative | 4 (1) | 7 (2) | 5 (2) | 9 (3) | 6 (2) |
| Pediatrics | 19 (6) | 22 (7) | 21 (7) | 19 (6) | 20 (7) |
| Sarcoma | 3 (1) | 3 (1) | 4 (1) | 4 (1) | 4 (1) |
| Skin | 4 (1) | 1 (<1) | 0 | 1 (<1) | 2 (<1) |
| Thorax | 21 (7) | 16 (5) | 21 (7) | 16 (5) | 19 (6) |
| Statistics | 16 (5) | 15 (5) | 15 (5) | 15 (5) | 15 (5) |
| Biology | 41 (14) | 45 (15) | 42 (14) | 47 (16) | 44 (15) |
| Physics | 49 (16) | 43 (14) | 47 (16) | 47 (16) | 47 (16) |

*Other contains 2 questions about heterotopic ossificans prophylaxis and 1 question about general cardiac dose constraints
†Numbers and percentages rounded to nearest whole number
Key: TXIT = Radiation Oncology In-Training Exam

**Table 6. Strategies to Improve Representation of Clinical Care Path Content on TXIT Examinations**

| SUGGESTION | COMMENT | PERCEIVED EFFORT OF IMPLEMENTATION |
|---|---|---|
| Report scores using a clinical care path framework in addition to scoring reports by a disease site framework | Provides feedback regarding acquisition of clinical competencies represented by the clinical care path framework | Low |
| Rebalance question content for better distribution across the clinical care path | Keep the number of questions fixed, but adjust the question content to increase underrepresented clinical care path content | Intermediate |
| Add questions to increase underrepresented content | • TXIT previously contained as many as 510 question items[1]<br>• Trainees will not welcome a longer testing session | Intermediate |
| Employ case-based questions to facilitate assessment across the clinical care path using a single clinical vignette | • Case-based questions are commonly employed by other medical licensing exams such as the USMLE<br><br>• Case-based questions have previously been included on the in-training exam in very limited capacity[2] | High |
| Restructure the administration of the in-training exam to end-of-rotation subject exams | • Structure subject exams like NBME shelf exams administered during required clinical rotations in medical school<br><br>• Permits more frequent, longitudinal testing of material<br><br>• Allows for cumulatively more questions over the course of a year, as number of questions are not constrained by one testing session | Very High |

*1. Paulino AC, Kurtz E. American College of Radiology In-Training Examination for Residents in Radiation Oncology (2004-2007). Int J Radiat Oncol Biol Phys. 2008;70(3):666-670. doi:10.1016/j.ijrobp.2007.09.049*
*2. Coia LR, Wilson JF, Bresch JP, Diamond JJ. Results of the in-training examination of the American College of Radiology for Residents in Radiation Oncology. Int J Radiat Oncol Biol Phys. 1992;24(5):903-905. doi:10.1016/0360-3016(92)90472-t*
*Key: TXIT = Radiation Oncology In-Training Exam, USMLE = United States Medical Licensing Examination, NBME = National Board of Medical Examiners*

In considering this uneven distribution, it is important to note that the TXIT never intended to serve as a comprehensive trainee assessment. The first chairman of the ACR Committee on Professional Testing established at the outset that "factors of clinical judgment, diagnostic skills and general sophistication in selecting a treatment program for a patient are not assessed in the in-training examination."[1] Moreover, the ACR has emphasized that the TXIT is not to be used as the principle method of assessing performance in residency, predicting success on the American Board of Radiology (ABR) written board examinations, or as a criterion for employment.[5] As a result, relying on the TXIT as a measure of resident competency across all entrustable professional activities in radiation oncology is a task for which the TXIT was not designed.

Underassessment of certain competencies may be inherent to written examinations in medical education, as evidenced by content analysis of the plastic surgery and orthopedic surgery in-training examinations, which showed unequal distributions of in-training exam content relative to the Accreditation Council for Graduate Medical Education (ACGME) Milestones and competencies.[12] In particular, there appears to be a bias for test-makers with regard to the type of questions included on written in-training examinations with respect to available published evidence on which those questions are based. A content analysis of the plastic surgery in-training examination found there were significantly more Level III (decision-making questions) compared to Level I (fact recall) or Level II (interpretation questions). In addition, Level III

questions more frequently justified the correct answer by referencing a journal article, with an overall higher mean number of journal references cited for these questions compared with other question types.[13] One possible explanation for this finding is that decision-making questions may be easier to develop because consensus exists due to the availability of supporting data.

When extrapolating this to the TXIT exam, and our own finding that treatment decision questions are predominant, we hypothesize there may be fewer questions from underrepresented clinical care path categories because there are fewer data on which to base single best answers. In other words, an acceptable range of choices exists. As an example, the preferred method to position a patient for setup during computed tomography (CT) simulation may vary among radiation

oncologists. Multiple set-up positions may be considered correct, so developing a multiple-choice question to assess a trainee's CT simulation knowledge may be challenging. Our analysis of the TXIT exam content is unable to support or refute this hypothesis as an explanation for the bias toward treatment decision questions. To remedy this, exam item writers may benefit from training on how to develop a multiple-choice question assessing knowledge that is not based on journal publications, but that is required to function as a competent radiation oncologist.

Nevertheless, the TXIT continues to serve a singular and influential role in formative assessment of radiation oncology trainees as it is the only assessment tool that allows residency programs to benchmark their trainees against national metrics. Acknowledging the prominence of the TXIT in resident assessment, residency programs should encourage realignment of the exam to better assess underrepresented radiation oncology clinical care path competencies. Although changes to the TXIT as suggested in **Table 6** may accomplish this goal, the effort and resources required to do so may be prohibitive.

Alternatively, instead of retrofitting the TXIT to improve assessment of specific competencies, a better strategy may be to develop new assessment methods that target specific components of the radiation oncology clinical care path in which trainees are underassessed. For example, residents in the US and Canada have identified a general absence of formal instruction and assessment in treatment planning that impedes transition to independent clinical practice.[14-16] To address this curricular deficiency, radiation oncology educators could make a concerted effort to create teaching resources and assessment tools to promote and measure acquisition of treatment planning skills. Potential

advantages of this approach include removing constraints imposed by a multiple-choice format, incorporating performance-based assessment, and using multiple assessment tools to triangulate trainee competency.[17,18]

Finally, it would be of interest to analyze the ABR Clinical Radiation Oncology Qualifying Exam according to the radiation oncology clinical care path framework. If current formative and summative assessments in the US do not assess for clinical competency across the entire clinical care path, then practicing clinicians may be deficient in specific areas. Further inquiry is needed into the development of comprehensive assessment methods to ensure clinical competency across the clinical care path.

This study has several limitations. First, the categorization of questions is inherently subjective despite utilizing multiple coders to avoid incorrect or inconsistent categorization. Among questions in which the initial two coders disagreed there was frequent overlap between the diagnosis, treatment decision, and treatment planning categories. This likely stems from the abundance of treatment decision questions and the inherent overlap of content with adjacent clinical care path primary categories of diagnosis and treatment planning. The small number of quality assurance questions may be due to significant overlap with the radiation physics category, making it difficult to conclude whether this is an underrepresented content area based on our analysis. Furthermore, the level of agreement between initial categorization by two coders is only moderate and not all question items were categorized through our reconciliatory process. These findings suggest there are other conceptualizations of a clinical care path framework that may improve categorization and better facilitate analysis of exam content.

## Conclusions

Radiation oncology ACR TXIT questions are unequally distributed along the radiation oncology clinical care path conceptual framework. The exam contains a higher proportion of questions pertaining to treatment decisions than questions assessing other clinical skills such as treatment planning, toxicity management, and brachytherapy. To the extent that the TXIT reflects national licensing exams and, more broadly, content prioritized in radiation oncology education, deficiencies in education and assessment within specific areas of the radiation oncology clinical care path may manifest as deficiencies in clinical competency among radiation oncology trainees. Acknowledging there is no singular assessment that can holistically measure trainees' preparedness for independent practice, radiation oncologists in training and their future patients would benefit from additional assessment tools to comprehensively assess knowledge and skills fundamental to all aspects of the practice of radiation oncology.

## References

1) Wilson JF, Diamond JJ. Summary results of the ACR (American College of Radiology) experience with an in-training examination for residents in radiation oncology. *Int J Radiat Oncol Biol Phys*. 1988;15(5):1219-1221. doi:10.1016/0360-3016(88)90207-6

2) Coia LR, Wilson JF, Bresch JP, Diamond JJ. Results of the in-training examination of the American College of Radiology for Residents in Radiation Oncology. *Int J Radiat Oncol Biol Phys*. 1992;24(5):903-905. doi:10.1016/0360-3016(92)90472-t

3) DXIT™ and TXIT™ In-Training Exams. Accessed January 20, 2020. https://www.acr.org/Lifelong-Learning-and-CME/Learning-Activities/In-Training-Exams. Published 2020.

4) Hatch SS, Vapiwala N, Rosenthal SA, et al. Radiation Oncology Resident In-Training examination. *Int J Radiat Oncol Biol Phys*. 2015;92(3):532-535. doi:10.1016/j.ijrobp.2015.02.038

5) Bordage G. Conceptual frameworks to illuminate and magnify. *Med Educ*. 2009;43(4):312-319. doi:10.1111/j.1365-2923.2009.03295.x

6) Accreditation Council for Graduate Medical Education (ACGME). ACGME Program Requirements for Graduate Medical Education in Radiation Oncology; 2020. Accessed January 2021. https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/430_RadiationOncology_2020.pdf?ver=2020-02-20-135340-140

7) American Board of Radiology Study Guide for Clinical (General and Radiation Oncology). Accessed July 1, 2020. https://www.theabr.org/radiation-oncology/initial-certification/the-qualifying-exam/studying-for-the-exam/clinical-radiation-oncology. Published 2020.

8) Golden DW, Ingledew P-A. Radiation Oncology Education. In: *Perez and Bradys Principles and Practice of Radiation Oncology*. 7th ed. Lippincott, Williams & Wilkins; 2018:2220-2252.

9) Accreditation Council for Graduate Medical Education (ACGME). Common Program Requirements (Residency). 2020:1-55. Accessed July 1, 2020. https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRResidency2020.pdf

10) The radiation oncology milestone project. *J Grad Med Educ*. 2014;6(1s1):307-316. doi:10.4300/jgme-06-01s1-27

11) Carraccio C, Englander R, Van Melle E, et al. Advancing competency-based medical education: a charter for clinician-educators. *Acad Med*. 2016;91(5):645-649. doi:10.1097/ACM.0000000000001048

12) Ganesh Kumar N, Benvenuti MA, Drolet BC. Aligning in-service training examinations in plastic surgery and orthopaedic surgery with competency-based education. *J Grad Med Educ*. 2017;9(5):650-653. doi:10.4300/JGME-D-17-00116.1

13) Silvestre J, Bilici N, Serletti JM, Chang B. Low levels of evidence on the plastic surgery in-service training exam. *Plast Reconstr Surg*. 2016;137(6):1943-1948. doi:10.1097/PRS.0000000000002164

14) Moideen N, de Metz C, Kalyvas M, Soleas E, Egan R, Dalgarno N. Aligning requirements of training and assessment in radiation treatment planning in the era of competency-based medical education. *Int J Radiat Oncol Biol Phys*. 2020;106(1):32-36. doi:10.1016/j.ijrobp.2019.10.005

15) Best LR, Sengupta A, Murphy RJL, et al. Transition to practice in radiation oncology: mind the gap. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. 2019;138:126-131. doi:10.1016/j.radonc.2019.06.012

16) Wu SY, Sath C, Schuster JM, et al. Targeted needs assessment of treatment planning education for united states radiation oncology residents. *Int J Radiat Oncol Biol Phys*. 2020;106(4):677-682. doi:10.1016/j.ijrobp.2019.11.023

17) Van Der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-214. doi:10.3109/0142159X.2012.652239

18) Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65(9 Suppl):S63-7. doi:10.1097/00001888-199009000-00045