ABR

AMERICAN BOARD OF RADIOLOGY

Item Writers' Guide

Last updated July 2018
Questions or comments? Email: Editing@theabr.org

TABLE OF CONTENTS

Foreword

Thank you for volunteering to write items for the American Board of Radiology (ABR) examinations. We appreciate the gift of your time and expertise.

This guide provides general instructions for choosing content and formatting items for ABR exams. Most of the rules and examples pertain to standard multiple-choice questions with a single best answer. ABR exams also include R-type and drag-and-drop (point-and-click) items and may include fill-in-the-blank and multiple-correct-option items. These types of items follow many of the same principles as standard multiple-choice questions; the differences in format are explained in Appendix C: Advanced Item Types.

This guide is designed for reference as you write. Points of item writing are explained in detail in the body of the manual. As you get comfortable with the process, you may find that all you need is a checklist to guide you from start to finish, such as that provided in Appendix A: Item-Writing Basics.

Getting ready

Start early

One part of successful writing is the review. To be as objective as possible, you should give yourself several days between writing and assessing items before submitting them. This allows you to clear your mind of the process, "forget" what you wrote, and look at the items with fresh eyes. Waiting until just before the deadline to do your writing robs you of this important perspective.

Consider your audience

It is important to consider the candidates as you write. If what we present is not relevant to their experience or to their expected field of work, the item does not serve its purpose. For Initial Certification examinations, items should deal with subjects that are taught in most residency programs. For Maintenance of Certification exams, items should be practice oriented.

Item format

Each item on an ABR exam focuses on one concept to be tested. For each standard multiple-choice question, there is **one best answer** among the choices, and to the prepared candidate it is clearly the right answer. Our objectives are to:

- test knowledge of the subject matter presented;
- test one piece of information per item;
- create items that are clear and concise, so the candidate can answer as many as possible during the time allotted.

Decide on your topic

- **Is it relevant?**
  One thing that makes even a well-written item unusable for testing is if it asks for information that has little to do with the practice of radiology or with case management in general. Remember that we are testing the candidate's ability to practice radiology safely and effectively. Items should not only be relevant, but also be **important** to practice.

- **Is it something the candidate must know to be successful in practice?**
  Candidates must clearly understand certain basic concepts. In addition, perception, interpretation, and deduction skills should be developed in training. All of these skills should be tested.

- **Is it based on controversial opinions, studies, or material?**
  Items that relate to specific studies, regional practices, or the methods of particular doctors should not be used. Likewise, items should not focus on trendy subjects (those that may change in a short period of time) or on practices and theories that are open to debate.

Choose and cite your reference(s)

Decide what sources you will be using. We require at least one complete reference for every item. References need to be credible sources that are widely studied, replicated, and verified. They should **not** be specialized or controversial. Among other things, references are our defense if an item is challenged. Be sure the reference supports the correct answer. At this time, some preapproved websites are acceptable as references. Check with your committee chair.

Decide what single task you want the candidate to perform

An effective item points toward a single option; therefore, the stem should be focused in one direction, making the path to the answer clear. The item should ask the candidate to complete one task and answer one question.

Recall information

Some information should be memorized and easy for a radiologist to bring to mind. The following is an example of a recall item:

What is the most common cause of superior vena cava syndrome?

A. Tuberculosis
B. Malignancy [key]
C. Pericarditis
D. Mediastinal fibrosis

Apply knowledge

The candidate must be able to apply knowledge in specific situations. This type of item builds on recall skills but requires more of the candidate, such as abstracting meaning from data, recognizing implications of clinical findings, identifying abnormalities on radiographs, or comparing possible treatment approaches. The following is an example of an interpretive item:

A guidewire-induced spasm of the popliteal artery occurs during angioplasty of the superficial femoral artery. What is the most appropriate intra-arterial medication for treatment?

A. Alcohol
B. Lidocaine
C. Papaverine
D. Tolazoline
E. Nitroglycerin [key]

Solve a problem (not necessarily through mathematical calculation)

Here we are testing whether the candidate can make judgments and decisions to deduce answers to complex clinical questions or to select a strategy. The following are examples of problem-solving items:

An image of the brain is obtained using a spin-echo MRI pulse sequence with TR = 1000 ms and TE = 100 ms. What anatomical structure has the highest signal intensity?

A. White matter
B. Lens of the eye
C. Cerebrospinal fluid [key]
D. Intracranial fat

Immediately after balloon angioplasty of an atherosclerotic stenosis of the superficial femoral artery, angiography shows acute closure of the vessel. What is the most likely cause?

A. Spasm
B. Thrombosis
C. Dissection [key]
D. Rupture
E. Recoil

Other tasks the candidate may perform

Other items may evaluate a candidate's interpretive and problem-solving skills by asking him or her to do the following:

- Interpret diagnostic images, data, graphs, or tables
- Use pathophysiology to predict findings from certain conditions

- Explain why something has occurred
- Order diagnostic studies
- Determine the cause of a condition or finding
- Design an overall treatment program (including surgery, radiation therapy, and chemotherapy) or quality management program
- Select or optimize a treatment plan
- Calculate a tumor dose
- Determine a prognosis

Design the item so there is only one question to be answered

We recommend that standard multiple-choice items ask for only one answer. Because we do not offer partial credit, items that have multiple-part options do not discriminate as well as those with a single answer. In addition, overlapping options can provide clues for savvy test-takers to guess the correct answer by looking at which parts occur most frequently.

Incorrect:

A 47-year-old woman has mammography, and unilateral axillary adenopathy is detected. She has no known inflammatory or infectious cause. What is the most appropriate BI-RADS assessment and recommended next step?

A. Category 1: Negative; return to screening
B. Category 2: Benign; return to screening
C. Category 3: Probably Benign; follow-up mammography in 6 months
D. Category 4: Suspicious; biopsy [key]

Fix this by splitting the questions into two separate items.

Correct:

A 47-year-old woman has mammography, and unilateral axillary adenopathy is detected. She has no known inflammatory or infectious cause. What is the most appropriate BI-RADS assessment?

A. Category 1: Negative
B. Category 2: Benign
C. Category 3: Probably Benign
D. Category 4: Suspicious [key]

-and-

A 47-year-old woman has mammography, and unilateral axillary adenopathy is detected. She has no known inflammatory or infectious cause. What is the most appropriate next step?

A. Return to screening
B. Ultrasound of the breast
C. Follow-up mammography in 6 months
D. Tissue biopsy [key]

Decide what level of knowledge you want to test

Consider the following:

What city is the capital of Arizona?
A. Albuquerque
B. Dallas
C. Phoenix
D. Sacramento

A person is not required to have knowledge about Western cities to choose the right answer. However, what if the item is written this way?

What city is the capital of Arizona?
A. Tucson
B. Flagstaff
C. Phoenix
D. Yuma

To answer this question, a person must know much more about Arizona cities to make the correct choice. Think about where the stem is going to lead the candidate in his or her "personal database." What level of knowledge do you want to test?

Write the stem

The stem is the part of the item that asks for a response. Stems should be written as complete sentences that end with a question. The stem should present all the information necessary for the candidate to figure out the answer without having to look for clues in the option list.

Stay focused

Suppose a stem is written like this:

In breast MRI, the chemical shift artifact:

or

Which of the following statements about chemical shift artifact is true?

What is being asked about artifacts? Any number of ideas could run through the candidate's mind. The intent is not clear. These stems fail the "cover test." To pass the cover test, an

item must be answerable with the option list covered. In other words, the examinee should be able to predict the type of answer being sought before looking at the options.

Avoid using "Which of the following is true?" construction. Although used in the past, this type of stem is unfocused and usually leads to mixed options. Rewrite the item so it asks a single, focused question that passes the cover test. Often, more than one focused item can be generated from an item with this construction.

This stem is improved from the previous example:

What MRI parameter could be changed to reduce chemical shift artifact?

Now the candidate understands what we are asking. If he or she knows the answer, it's just a matter of finding it in the options.

Be aware of sentence structure and linearity

The most effective way to deliver information is in a linear fashion—one that allows the mind to follow a path, arriving smoothly at the point that an answer must be given. Focusing the stem in this way will also help you focus the options.

The formula most suited to linear thought is:

Background info + Situational info (or equation) + Request for solution

For example,

> A 7-year-old girl presents with a mild fever and swollen lymph nodes. Three days later, she develops a rash that starts on the face and spreads to the neck, chest, and the rest of the body. What is the most likely diagnosis?

It is important that the question is asked at the end of the stem, after all the information needed to answer the question has been provided. In this example of nonlinear structure, the question comes at the beginning:

> What percentage of patients will survive at least 42 months if they are part of a group that exhibits lifetimes that are normally distributed with a mean of 36 months and a standard deviation of 3.0 months?

In this instance, the candidate would already be trying to formulate an answer (or wondering how it can be done) before crucial information is given.

A better, linear way to present the information is:

> A group of patients exhibits lifetimes that are normally distributed with a mean of 36 months and a standard deviation of 3.0 months. What percentage of these patients will survive at least 42 months?

Use common and precise language

We are not testing vocabulary. Try not to use specialized words when common words will be sufficient. Common words are more effective because they ensure that candidates are not laboring to understand the item, and that we are in fact testing knowledge of the subject matter—not language skills.

Likewise, language needs to be medically, scientifically, and technically precise, accurate, and consistent. This may mean refraining from "this is just how we say it" type thinking to increase clarity, reduce ambiguity, and use nomenclature consistently. This is equally important for the medical field in general and for the examinations leading to board certification.

Be clear and concise

Candidates have a limited amount of time to answer questions. Therefore, stems should be clear and concise. Long, involved explanations or histories should be avoided. State only the information needed to answer the question. The following stem is overly detailed:

> A 60-year-old woman presents with rectal bleeding and is found to have an adenocarcinoma of the midsigmoid colon. It is completely resected by low anterior resection, and five out of five lymph nodes are negative. No adjuvant therapy is given. Two years later, she presents with anorexia and a 10-pound weight loss and is found to have liver metastases. What is the most appropriate treatment now?

This stem could be rewritten more simply as:

> A 60-year-old woman presents with an adenocarcinoma of the midsigmoid colon. It is completely resected. Two years later, she develops liver metastases. What is the most appropriate treatment?

We don't want to waste the candidate's time with extraneous reading. A stem that rambles, delivers information disconnectedly, or includes too much information not directly related to the question being asked can be confusing and can draw the candidate away from the task at hand. Although it might be appropriate to include some extraneous information in a stem to test whether the candidate can glean the pertinent points, it is important that the stem maintain a high degree of focus.

Avoid negatives

We are trying to ascertain what the candidate knows, not how easily a test-taker can be confused. It is best to avoid items that require reverse thinking. It complicates the presentation of information. A negative stem presents the following problem:

Background info + Situational info (or equation) + Request for solution, but—*oh, by the way*—give us the opposite of what we've led you to think about.

Avoid writing items that include:

- Which of the following is **FALSE**?
- Which of these is **NOT** an indicator . . . ?
- All of the following are true **EXCEPT**:
- What is the **LEAST** likely . . . ?

This kind of item is easy to write because it only requires three or four facts from references and one fabricated "fact." But asking for a false "answer" is more a test of the candidate's ability to think in reverse than knowledge of the subject. Often, even experts will read this type of stem and mentally proceed to the CORRECT answer. Test-takers tend to choose the first correct answer in the list and often don't remember that they were supposed to be seeking the INCORRECT option.

Negative wording is appropriate for two types of items:

- those that ask what should be excluded from a differential diagnosis;
- those that ask about a practice to be avoided because of the potential for serious side effects.

If a stem must contain a negative term, it should appear in capital letters and boldface type.

Avoid second person

The exam should be testing a candidate's understanding of facts, not asking for their opinions. Write in third person rather than second. Second-person construction is subjective. For example, instead of writing "How would you interpret this image?" write "What is the best interpretation of the image?"

Write the key (the correct answer)

It should offer the most plausible response.

To the prepared candidate, there should be no doubt that the key is the only option that could possibly be selected from the list of choices. The prepared candidate will readily choose it.

It should make sense grammatically.

Whether the key answers a question with a phrase or a full sentence, it should be in proper syntax.

It should be similar in length to the distractors.

Nothing gives away the right answer faster than having one choice that is well thought out and carefully worded and distractors that are choppy, too short, too long, etc. Candidates can also be drawn to distractors that are different from the others. To ensure that options are chosen because they are the best answers, and not because of a clue, try to make them as similar to the distractors as possible.

Write the distractors

Distractors are the other possible answers in the option list. Distractors are perhaps the hardest part of the item to write because they must, on some level, seem like reasonable options.

They should be plausible.

To the unprepared or underprepared candidate, distractors should seem like they could potentially be the right answers. Additionally, the distractors need to be real things. Please do not invent terms to complete the option list.

There may be two to four distractors (three to five total options).

It is better to have fewer plausible distractors than a list of options with more distractors that are obviously incorrect. A distractor that adds nothing (or doesn't fit the list) can negatively skew the statistics for that item.

Three total options can be used, especially when a fourth plausible option is difficult or impossible to provide (e.g., when the question asks the candidate to determine whether a dose or agent increases, decreases, or remains the same under specific circumstances).

They should all be in the same category.

Be sure distractors are rooted in the same material as the correct answer.

Which of the following is associated with osteoarthritis?

A. Patient age < 50 years *[demographic information]*
B. Radiography is preferred to MRI. [key] *[imaging]*
C. Nutritional deficiency *[presentation]*
D. Family history is **not** a risk factor. *[negative + risk factor]*

Mixed options are a common problem in items that have unfocused stems. If you find yourself writing mixed options, go back and look at the structure of your stem. A focused stem would ask a direct question like "What is the most appropriate initial imaging modality?" [key = Radiography]

They should be similar in length to the key.

As with the key, you shouldn't have one distractor that is well thought out and carefully worded, while the others are choppy, too short, too long, more specific, use qualifying phrases, etc. Try to make all options as similar as possible.

Why is Lipowitz metal (Cerrobend) more appropriate than lead for custom blocking?

A. It is much easier to machine than lead.
B. It has a much lower melting point than lead. [key]
C. It is composed primarily of cadmium, a less toxic substance than lead.
D. Custom blocks made of Lipowitz metal (Cerrobend) have a sharper penumbra than custom blocks made of lead.
E. Its linear attenuation coefficient is higher than that of lead.

Option D is longer and worded differently from the others. In this case it is not the key, but its differences could cause candidates to think it is the correct answer. What if it is changed to read as follows?

D. It forms blocks that have a sharper penumbra than custom lead blocks.

Now it more closely matches the other options. Therefore, if it is chosen, it will be because the candidate thinks it is the right answer, not because of a trick. If a distractor cannot be rewritten to have a similar length and structure to the other options, the best solution may be to eliminate it from the list.

They should be the same part of speech as the key.

An option that is a different part of speech from the other options sticks out, especially if it doesn't make sense grammatically as an answer to the question.

They should not create "tricks."

As illustrated above, the way options are worded may "trick" the candidate into selecting an incorrect answer. We are not trying to trick the candidate with language or vague clues. On the contrary, we are trying to make it possible for each candidate to answer as many items as possible in the allotted time, thereby affording the greatest chance of success.

- *Negatives*
  Having an option in the negative is more a test of reverse thinking than of knowledge. The following item, in addition to having an unfocused stem and mixed options, is further weakened by tricky negative language:

  Which of the following statements most accurately describes Turner syndrome?

  A. A webbed neck is **not** found frequently.
  B. The ovaries usually have normal function.
  C. Patients typically do **not** have two complete X chromosomes. [key]
  D. The cause for short stature is clearly understood.
  E. Patients rarely have shield-like chests.

  Option A contains a negative term ("not"), which a candidate could easily miss. Options A, B, and E are examples of reverse truths, two of which are fairly obvious. (A reverse truth is a true statement that is made false simply by changing or inserting a word or two.) Option D is improbable, using the word "clearly." The key also contains a negative term, forcing the candidate to think in the reverse.

- *Jargon, slang, acronyms, etc.*
  Other tricks that can be confusing to the candidate include use of jargon, slang, acronyms, and abbreviations.

  A 74-year-old man is admitted for evaluation of a T4N2b squamous cell cancer of the hypopharynx. The patient has cardiac arrest and a code is called. What is the most appropriate next step?

  A. Bolusing with "roids"
  B. Tubing
  C. Bagging
  D. Scoping

  The issue of fairness arises in this situation because not all candidates may be familiar with these particular terms. Standard medical terminology should be used in all examination items for clarity and fairness.

This also applies to acronyms. With few exceptions, ABR policy is to spell out terms on first use in an item, with the acronym following in parentheses. The acronym can then stand alone in the remainder of the item. This is to avoid confusion of terms and to make information absolutely clear.

They should not contain "clues."

Some candidates have mastered the art of test-taking, including how to figure out what choices are not correct by looking for clues. "Clues" include language or constructions that may help unknowledgeable but test-wise candidates select the correct option. Some giveaways are as follows:

- *Vague and absolute terms*
  Vague terms—might, may, can—are clues to the key because they indicate that almost anything is within the realm of possibility. Absolute terms—always, never—are clues to distractors because there are no exceptions.

  The following item illustrates both points:

  Which of the following is a characteristic of Ewing sarcoma?

  A. It always involves the diaphysis.
  B. It involves the metaphysis more commonly than the diaphysis.
  C. It may involve the epiphysis. [key]
  D. Extraosseous Ewing sarcoma has never been reported.
  E. It almost always involves long tubular bones.

  Notice that the stem is unfocused and that the options are mixed. In addition, for the observant candidate, Options A, D, and E can be eliminated from consideration in this item because the terms "always" and "never" mean there are no exceptions. On the other hand, Option C is almost certain to be the correct response because the term "may" includes all possibilities.

  To improve this item, focus the stem, remove the clues, and make the options homologous:

  Which of the following is usually involved in Ewing sarcoma?

  A. Diaphysis
  B. Metaphysis
  C. Epiphysis [key]

Now the stem asks a clearer question. Although "usually" is vague, moving it to the stem makes it equally applied to all options. The answers are straightforward. There is no ambiguity, and there are no clues.

- *Double/multiple options*
  Double or multiple options are responses that contain two or more pieces of information. As explained under "Design the item so that there is only one question to be answered," these options can present scoring issues. Also, many test-takers understand the thought process of the writer and can figure out the correct answer based on the frequency with which each part of the option occurs in the list. For example, what if the elements of the options appear as follows?

  A. 1 & 2
  B. 1 & 3
  C. 1 & 4
  D. 2 & 5

  The savvy examinee will note that 1 appears 3 times, and 2 appears twice. Therefore, the most likely answer is the one that contains both 1 and 2 (A). It is better to stick with one-element answers whenever possible.

- *All of the above*
  We do not use this construction.
  This type of option is typically not a good measurement tool. If candidates recognize two correct choices, they know the rest of the list probably follows suit and may not even read the other choices. These may seem easy to write—just pull a few facts from the references—unfortunately, candidates know the technique.

- *None of the above*
  The ABR does not use any items in this format.
  There is an argument that "none of the above" responses can be effective. Sometimes there is no definitive answer to the question, and the item is trying to determine that the candidate realizes this. A "none of the above" item may actually entice the candidate to study all of the options more thoroughly. However, we feel that the possible negative effects are more relevant to our situation.

- *Pairs*
  Pairs occur when two options are very similar to each other but different from the rest of the list. A candidate will tend to consider only the two options in the pair and ignore the others. For example:

After a patient has a liver transplant, occlusion of the hepatic artery can cause strictures in what area?

A. Portal vein anastomosis
B. Hepatic vein anastomosis
C. Transplant bile ducts [key]
D. Lymphatic drainage

Candidates will tend to focus on options A and B and figure that one of them is correct. If the key is part of the pair, this is a clue to the correct answer. If the key is not part of the pair, it is a trick.

Option lists should either contain no pairs or two (or three) sets of pairs, so that all options must be considered. For example:

A patient requires a long-term, continuous infusion of medication for pulmonary hypertension. Which of the following venous access devices would be most appropriate?

A. Implanted port in the chest wall with the catheter inserted via the internal jugular vein
B. Implanted port in the upper arm with the catheter inserted via the basilic vein
C. Hickman catheter tunneled from the chest wall to the insertion site in the internal jugular vein [key]
D. Quinton catheter tunneled from the chest wall to the insertion site in the internal jugular vein

Now candidates must look at both sets of pairs. There is no clue, and the odds of guessing are returned to 25%.

Review the item

As previously mentioned, it is best to write items well before the deadline, then set them aside for a while. When writers try to revise their own work right away, they tend to miss areas that actually need correction or clarification; the writer may assume the intended meaning rather than recognizing any errors in the meaning of the actual words. Returning to items later decreases this risk. When you do review the items, consider the following points:

Cover the options and see if you can answer the question

The stem should give sufficient information for the candidate to formulate an answer without having to look at the option list for clues (the "cover test").

Look at the stem structure

Be sure you are delivering the information in the most direct way possible. Be sure the presentation is focused, concise, and linear. Also check to be sure the stem is a complete question, and that it does not contain negative phrasing, if possible.

Check the key

Will the prepared candidate recognize that it is clearly the right answer?
Is it in the same category and similar in concept, structure, and length as the distractors?
Does it properly complete the stem?

Consider the validity of the distractors

Do any of them sound ridiculous?
Are they all real entities?
Are they in the same category and similar to the key in concept, structure, and length?
Do they complete the stem appropriately?
Do they avoid tricks, clues, and convoluted presentation?

Check your reference citation

Each item must have at least one complete reference.

Appendix A: Item-Writing Basics

General

- State the needed information as concisely as possible and at a comfortable language level.
- Focus the stem and options; avoid clues and tricks.
- Make sure the items pass the cover test.
- Ensure all items are clinically relevant, noncontroversial, and up-to-date.
- Add the references.
- Include any necessary images. (It is also permissible to include tables, graphs, or diagrams.)

Stems

- Focus on one concept. To focus an item, think of what the question is asking the candidate to do (e.g., recall information, apply knowledge, solve a problem).
- Present all the information necessary for the candidate to formulate an answer without having to look for clues in the option list (the cover test).
- Use present tense, when possible, and write in a linear fashion. The information should be chronological, and the question to be answered should appear immediately before the list of options.
- Avoid the use of second-person (i.e., "you") construction.
- Make sure the stem contains complete sentences and ends in a complete question.
- Use positive, rather than negative, wording.

Options

- Use plausible options to avoid clues.
- Keep options similar in structure and concept. Test one piece of specific information with each item.
- Make sure the options properly and grammatically complete the stem.
- Write at least four options whenever possible. Multiple-choice items generally have four or five answer options. Only three options are acceptable if a fourth option isn't reasonable or logical.

Appendix B: Ideas to Help You Get Started*

The following questions may help as you start formulating items, as well as "polish" items that are not quite what you want them to be.

Understanding—Concepts

- What is the best definition of this concept?
- What is the meaning of this concept?
- What is synonymous with this concept?
- What is a characteristic of this concept?

Understanding—Principles

- What is the best definition of the principle of …?
- What is the reason for or cause of…?
- What is the relationship between … and…?

Critical Thinking—Predicting Using a Principle

- What would happen if…?
- If (there is an action), what would happen next?
- What is the consequence of an action?
- What is the cause of a result?
- Information given. What is the expected result?
- What distinguishes one concept from another concept?

Critical Thinking—Evaluating Using Facts and Concepts

- Which is the most important, significant, effective…?
- Which is better, higher, lower, farther, nearer, heavier, lighter…?
- Which is most like …?
- What is the difference between … and…?
- What is the similarity between … and…?

Critical Thinking—Evaluating Using a Principle

- Which of the following principles applies to evaluating something?
- What is the most important factor contributing to…?

Problem-Solving—Concepts, Principle, Procedures

- Problem presented. What is the best way to solve this problem?
- Problem presented. What is the solution?

Defining

- What is the main symptom of…?
- What is the most common (cause, complication, symptom, or consequence) of (a procedure, a disorder, or an action)?

Predicting

- What is the most common (cause or symptom) of (a patient problem)?

Evaluating

- Patient illness is diagnosed. What is the most appropriate treatment?
- Patient illness is diagnosed. What is the most appropriate next step?

Applying

- Information [about a patient problem] is presented. What is the most appropriate treatment?

* Modified from *Developing and Validating Multiple-Choice Test Items,* by Thomas M. Haladyna. Reprinted with permission from the publisher, Lawrence Erlbaum Associated (Mahwah, NJ, 1999).

Appendix C: Advanced Item Types

In addition to standard multiple-choice questions (MCQs) with a single best answer, other item types are available for ABR volunteers to write. Two of these—R-type and drag-and-drop (for radiation oncology and medical physics, this is called point-and-click)—have been in use for a while for most disciplines. Two new types—fill-in-the-blank and MCQs with multiple correct options—are currently being pilot tested and may start appearing on exams in a few years. Following is a brief description and example of each type.

R-type

The R-type item provides a way to write multiple stems that all use the same option list. It begins with a lead-in statement that establishes the relationship between the stems and the options. The lead-in should be focused and present a clear task for the candidate, and it always ends with this sentence: "Each option can be used once, more than once, or not at all." The item should not be written so that each option matches with one stem, which would allow the candidate to use "process of elimination" to get the last answer correct. Be sure to write some R-type items that use an option more than once or do not use some of the options.

As in standard MCQs, the option list for an R-type item should be homogeneous; all the options should be in the same category and similar in concept, structure, and length (all single diagnoses, treatments, etc.). The option list should be presented in alphabetical order, unless a different order is more logical.

Stems should follow the same criteria as stems in MCQs: they need to be linear, focused, positively worded, and complete sentences. R-type stems need to pass the cover test (i.e., the candidate should understand what is being asked and be able to think of an answer even if the options are covered). There are no questions in R-type stems. The lead-in establishes the task being presented for each stem. Patient descriptions in R-type stems should be similar in structure and length to the others in the set and present the same type of information. For example, if one patient vignette has imaging, all the stems in the set should have imaging. If one stem provides laboratory results, all the others should as well.

Example:

Lead-in: For each patient with fatigue, select the most likely diagnosis. Each option can be used once, more than once, or not at all.

Options:

A. Acute leukemia
B. Congestive heart failure
C. Depression
D. Epstein-Barr virus infection
E. Folate deficiency
F. Hereditary spherocytosis

G. Hypothyroidism
H. Iron deficiency
I. Lyme disease
J. Microangiopathic hemolytic anemia
K. Miliary tuberculosis
L. Vitamin B_{12} (cyanocobalamin) deficiency

Stem 1: A 15-year-old girl has a 2-week history of fatigue and back pain. She has widespread bruising, pallor, and tenderness over the vertebrae and both femurs. A complete blood count shows a hemoglobin concentration of 7.0 g/dL, a leukocyte count of 2000/mm^3, and a platelet count of 15,000/mm^3.

Key = A

Stem 2: A 19-year-old woman has a 1-week history of fatigue, fever, and a sore throat. She has a temperature of 38.3°C (101°F), cervical lymphadenopathy, and splenomegaly. Initial laboratory studies show a leukocyte count of 5000/mm^3. Serum aspartate aminotransferase (AST, GOT) activity is 200 U/L. Serum bilirubin concentration and serum alkaline phosphatase activity are within normal limits.

Key = D

Drag-and-Drop (Point-and-Click)

The stem for a drag-and-drop item (these are called "point-and-click" for medical physics and radiation oncology exams) should be a complete sentence instructing the candidate to "Drag and drop the label onto [something]." (Or, "Point and click on [something].") The "something" can be a structure, an abnormality, a blood vessel, etc. Only one structure can be labeled per item.

Example:

Drag and drop the label onto the fracture.

The candidate must place the label within an area specified by the item writer.

Fill-in-the-Blank

Fill-in-the-blank items are to be used only with numerical answers. You can provide a range of correct answers to account for rounding, but each number that would be considered correct must be specified. In addition, the required number of significant digits should be indicated at the end of the stem in parentheses. Fill-in-the-blank items are written as complete sentences rather than questions.

Example:

The half-life of ^{67}Ga is _____ hours. (Two significant digits)

A. 78.26

Multiple-Correct-Options

For a multiple-correct-options item, the candidate must select all the correct options to get credit for the item. Indicate in the stem how many answers are correct and add "(Please select # options.)" at the end of the stem. The total number of options must be at least twice as great as the number of correct options.

Example:

Which four of the following are considered benign tumors? (Please select four options.)

A. Adenocarcinoma
B. Adenoma*
C. Fibroma*
D. Glioma
E. Hemangioma*
F. Lipoma*
G. Medulloblastoma
H. Rhabdomyosarcoma

Appendix D: Common Item Edits for ABR Style

- Hyphenate patient ages: *30-year-old man*; NOT "30 year old man"
- Use *boy* or *man* (18 years and older) when used as a noun, NOT "male"
- Use *girl* or *woman* (18 years and older) when used as a noun, NOT "female"
- Use *after*, NOT "following" (when meaning an order of events or period of time)
- Use *more than*, NOT "over" (when meaning to be higher in value or measure)
- Use *because*, NOT "since" (when meaning for reason)
- Use *before*, NOT "prior to"; use *previous*, NOT "prior" when used as an adjective
- Use *unenhanced*, NOT "noncontrast"
- Use *emergency department*, NOT "emergency room" or "ER"
- With diagnosis, use *most likely*, NOT "best"
- With interpretation, use *best*, NOT "most likely"
- Use *most appropriate next step*, NOT "next best step"
- Use *A patient presents with* or *A patient reports*, NOT "A patient complains of"
- A patient is not defined by his/her illness. Use *A patient with diabetes*, NOT "A diabetic" or "A diabetes patient"
- Use *the image* or *the image shown*, NOT "below," "above," "following," etc. when referring to images
- Hyphenate *follow-up* as a noun or adjective, BUT *follow up* as a verb
- Use *history* alone, NOT "past" or "previous" or "prior" history; it's redundant
- Use initial articles (*the*, *a*, *an*, etc.) in complete sentences, BUT eliminate initial articles in options that are phrases or terms
- Use punctuation (periods) in options that are complete sentences, NOT in options that are incomplete sentences
- Use *5 to 10 mg*, NOT "5-10 mg" (with a hyphen) when indicating range
- With eponyms, remove the apostrophe and "s" when appropriate
- Use generic names for drugs (with trade names in parentheses afterwards, if done for all drugs in the item)
- Capitalize the first letter of each option
- If a negative word is in the stem, make the word bold and all caps

Numbers

Use numerals for all numbers to express age, percentage, temperatures, dates, units of measure, time, etc. Use the degree symbol with temperatures, BUT *degrees* with angles.

What vs Which

Use *what* when asking about something for which there are unlimited options: ***What*** *is the most likely diagnosis?* (one option out of every diagnosis that exists). Use *which* when asking about something for which there is a limited or finite series of options presented: *The abnormality appears in* ***which*** *chamber of the heart?*

Appendix E: Editing Example

Original:

49 yo female with chest pain, sp MI 5 years. T1, DWI. The next best step is:

Edited version:

A 49-year-old woman presents with chest pain 5 years after a myocardial infarction. T1- and diffusion-weighted imaging are performed. What is the most appropriate next step?

-or, alternatively-

A 44-year-old woman has a myocardial infarction. Five years later, she presents with chest pain. Based on the T1- and diffusion-weighted images, what is the most appropriate next step?