Writing Good Multiple Choice Questions

Steven Montner
University of Chicago

So Why Bother

- Bad questions diminish the reliability of your test:
  - The failure rate with flawed MCQ’s has been estimated at up to 25% higher than with standard questions.
  - Downing SM. Academic Medicine 2002; 77: S103–S104

"The typical MCQ item author is a medical expert with time constraints and a lack of formal didactic education in state-of-the-art item writing principles."

- The definition of everyone in this room!

So Why Bother

- Bad questions don’t test what you should be testing.
  - Not minutiae
  - Not test-taking savvy
  - Not reading ability
  - Not an “eye test”

So Why Bother?

- SO WHY BOTHER?


  - 43% flawed

So What Makes a Good Multiple Choice Question?
Is It Relevant

» Will the test taker think so?

Constructing Good MC Questions

» 1. Write a good “stem”.

Test recall of important facts!

» “What contrast material should be used for an esophagogram if perforation is suspected?”

» “What is the typical enhancement pattern of hepatocellular carcinoma on MRI?”

Don’t be fancy:

» “What is the most likely diagnosis?”

Writing a Good Stem

» Keep it well focused!

So it’s relevant, now what?
Writing a Good Stem

- The central idea of the question should be in the stem.
  - The stem should be longer than the responses.
  - A lead in statement is often helpful
    - A 65-year-old man has difficulty rising from a seated position and straightening his trunk, but he has no difficulty flexing his leg. Which of the following muscles is most likely to have been injured?

Writing a Good Stem

Image based questions need not be complicated

What is the diagnosis for the images provided?
A. NSIP
B. UIP
C. HP
D. LIP

A Mortal Sin

Some Nicely Focused Stems

- "What segment of the small bowel is most frequently involved by lymphoma?"
- "What mechanism of injury is the most frequent cause of a Jefferson fracture?"

Some unfocused stems

- "Which of the following is true regarding......"
- "It is correct that:"
  - Just becomes 4 unrelated true-false questions.
  - Leaves the answer choices open to all sorts of disparate options.

The Cover Test

- Can I put my hand over all of the options and still answer the question?
From the May 2016 Radiology

- Routine whole-body MR imaging protocols used to detect skeletal involvement in metastatic cancer or multiple myeloma
  - A. Should cover the body from head to toes
  - B. Should include contrast-enhanced sequences
  - C. Should include both anatomic (e.g. T1-weighted, short tau inversion recovery) and functional diffusion-weighted imaging sequences
  - D. Should use the machine-integrated body coil

RSNA Instructions for authors

- The stem:
  1. Should be “focused”: that is, contain the main idea of the question. Use the “cover test”:
     Can you cover the options and still answer the question?

From the May 2016 Radiology

- Routine whole-body MR imaging protocols used to detect skeletal involvement in metastatic cancer or multiple myeloma
  - ?????????????????????????

So what do you think?

Unfocused!

A Random CME Sample from RadioGraphics

- Which of the following statements best describes DBT?
- Which of the following statements most accurately describes retroperitoneal fasciitis?
- Which of the following statements best characterizes use of head CT in geriatric trauma patients?
A Random CME Sample from RadioGraphics

» What is wrong with those stems:
  » They are all simply multiple True/False questions.

Another Example from RadioGraphics

» “Numbness of the left side of the lower lip of a 45 year old man with facial trauma due to a motor vehicle collision is most likely caused by a fractured mandible with displacement of the ____.”

Writing a Good Stem

» Keep it well-focused
  » No extraneous material and no red herrings.

How about just?

» Get rid of the superfluous info:
  "With mandibular fracture, numbness of the lower lip is most likely caused by displacement of what structure."

Just the facts, please

» If you're going to tell me that “A 32 year old woman on oral contraceptives develops hypotension, and chest and abdomen pain”…..
  » Bleeding hepatic adenoma
  » Venous thrombosis and a PE

Writing a Good Stem

» Keep it well-focused.
  » No extraneous info and no red herrings.
  » Stay away from “negative” constructions:
    » “What finding is NOT associated with…”
    » “…all of the following EXCEPT…”
    » “Which of the following is LEAST likely…”
Happens to the Best of us

A question submitted to the Board for the MOC Exam with a negative construct:

Still another in RadioGraphics

In most children aged 12–16 years, ulnar variance is ________.

- Positive
- Negative
- Neutral
- No predominant pattern has been observed.

After the Stem

- Answers and Distractors

A question submitted to the Board for the MOC Exam with a negative construct:

This 38-year-old female underwent laparoscopic cholecystectomy. After four uneventful postoperative days, she developed sudden-onset abdominal pain, nausea, and vomiting. Abdominal CT one week later is shown below. The differential diagnosis should include all of the following EXCEPT:

- A. Severe acute pancreatitis
- B. Pancreatic abscess
- C. Pancreatic necrosis.
- D. Gallstone pancreatitis.

Writing a Good Stem

- Simple wording, both for straightforward questions or clinical vignettes.
- No jargon or unusual abbreviations.
- Make sure the stem and all the choices have matching grammar.

The Answer

Must be absolutely, 100%, incontrovertibly, everyone-agrees-and-no-room-for-discussion correct.
This patient had abdominal pain for 12 hours, fever, and leukocytosis. The next best step for this patient:

A. laparoscopic surgery.
B. CT-directed drainage.
C. intravenous antibiotics.
D. surgery.
E. observation for 48 hours.

The Distractors

- The toughest job of all.
  - How many?
    - 3 is fine
    - That means 4 responses including the key

The Answer

- Must be absolutely, 100%, incontrovertibly, everyone-agrees-and-no-room-for-discussion correct.
- We now insist that a reference be included with each question submitted.

The Distractors

- Make sure that they are plausible

The Answer

- Avoid relative terms: often, frequently, rarely.
- Test takers have wide-ranging opinions on just how frequent is “frequently.”
- And they know that nebulous terms often indicate the right answer

Make sure they are plausible

16 year-old female in a head-on motor vehicle collision was unresponsive on arrival with a severe closed head injury and degloving injury of the scalp.
Make sure they are plausible

What is the most likely cause of the appearance of the small bowel?

A. Hypoperfusion
B. Mesentery Laceration
C. Mural Edema
D. Infection

Keep distractors and key the same length

After viewing Figure 2, injection of the nasobiliary tube two weeks after interval cholecystectomy, what is the most likely diagnosis?

A. Suprapancreatic common bile duct stones
B. Distal common bile duct inflammatory stricture and a few air bubbles.
C. Sclerosing cholangitis complicated by polypoid cholangiocarcinoma
D. Mirizzi syndrome

What's the implausible choice?

What is the most likely cause of the appearance of the small bowel?

A. Hypoperfusion
B. Mesentery Laceration
C. Mural Edema
D. Infection

The Distractors

Make sure they’re plausible.

Keep all of them (and the answer, too) about the same length.

“Avoid Non-homologous Options”

They mean: keep ‘em in the same ballpark.
This 23-year-old woman underwent CT of the abdomen to evaluate right lower quadrant pain. Which clinical question is critical to the differential diagnosis?

A. Does the patient have an elevated white cell count?
B. Does the patient have a history of recent acute pancreatitis?
C. Is the patient on birth control pills?
D. Is the symptom upper rather than lower abdominal pain?

The Distractors

- Make sure they’re plausible.
- Keep all of them (and the answer, too) about the same length.
- Like the stem, keep them focused.
- Keep them simple, simple, simple.

Non-homologous options

- Does the patient have an elevated white cell count? (*Sign*)
- Does the patient have a history of recent acute pancreatitis? (*History*)
- Is the patient on birth control pills? (*History*)
- Is the symptom upper rather than lower abdominal pain? (*Symptom*)

Keep it simple

- 49-year-old woman with acute right upper quadrant pain had a CT scan of the abdomen

Yet one more RadioGraphics

- Which of the following best describes Brenner tumors? (*Unfocused*)

A. They are typically benign (*Prognosis*)
B. They rarely show calcifications (*Findings*)
C. They represent about 10% of ovarian neoplasms (*Prognosis*)
D. They demonstrate little or no enhancement (*Findings*)

Keep it simple

- What is the most likely diagnosis?
  - A. Gallbladder stone with biliary obstruction
  - B. Acute cholecystitis with gallbladder stone
  - C. Acute acalculous cholecystitis
  - D. Gallbladder stone without cholecystitis
So why toss this question?

- It's too confusing...
  - It's too confusing...
  - It's too confusing...
  - It's too confusing...

Watch out for inadvertent clues

- No mutually exclusive options
  - For questions that require a single best answer, options that contradict one another cannot both be correct and therefore mutually exclusive options reduce the number of plausible responses.

It's too confusing.

A. Gallbladder stone with biliary obstruction
B. Acute cholecystitis with gallbladder stone
C. Acute acalculous cholecystitis
D. Gallbladder stone without cholecystitis

Which of the following statements best describes digital breast tomosynthesis acquisition?

- A. Multiple low-dose x-ray projection images are obtained in a 15–30° arc
- B. A single x-ray projection image is obtained while the x-ray tube moves in an arc
- C. A CT image of the breast is obtained
- D. Multiple low-dose x-ray projection images are obtained in 360° arc

Keep it simple

- What is the most likely diagnosis?
  - A. Gallbladder stone with biliary obstruction
  - B. Acute cholecystitis with gallbladder stone
  - C. Acute acalculous cholecystitis
  - D. Gallbladder stone without cholecystitis

Mutually exclusive options

- So can I ever use paired options?
  - Yes, but you need two pairs!
Mutually exclusive options should be paired

- What is the most common site of Crohn’s Disease?
  - A. Proximal esophagus
  - B. Distal esophagus
  - C. Ascending colon
  - D. Descending colon

Quick Review

- Keep it relevant.
- Keep it clear: focus, focus, focus.
- No tricks and no clues.
- Sterling images.

Watch out for inadvertent clues

- No vague “maybe” terms.
- No “always”. No “never”.
- Don’t “highlight” any particular choices.
- Randomize the position of the answer

Real-life examples

Actual questions submitted to the ABR from 2006 through 2012

Special considerations for tests with images

- Images have to be technically superb
- Absolutely classic examples of the diagnosis
- Readily apparent findings
- **NOT AN EYE TEST!**

Submitted to ABR

This 50-year-old female underwent an upper gastrointestinal study to evaluate vague abdominal pain. The differential diagnosis includes all of the following except:

- A. A stone lodged in the distal common bile duct.
- B. ampullary villous adenoma.
- C. choledochal cyst.
- D. annular pancreas.
So, what's the problem?

- A. Negative construction
- B. Non-homologous choices
- C. A red herring clue
- D. Too-complex stem

More to learn from the same question.

- Based on your differential, the patient was next evaluated with endoscopic retrograde cholangiopancreatography. Which of the following statements is true?
  - A. The patient has an increased risk of developing cholangiocarcinoma.
  - B. The patient should undergo pancreatoduodenectomy.
  - C. The patient should undergo percutaneous transhepatic stent placement.
  - D. No action is necessary unless the lesion grows.

Submitted to ABR

This 50-year-old female underwent an upper gastrointestinal study to evaluate vague abdominal pain. The differential diagnosis includes all of the following EXCEPT:

- A. A stone lodged in the distal common bile duct.
- B. Ampullary villous adenoma.
- C. Choleodochal cyst.
- D. Annular pancreas.

So, what's the problem?

- A. Unfocused stem
- B. Disparate length clue
- C. Mutually exclusive pair
- D. Don’t change a thing

But wait! There's more!

- The case came with a second question

Based on your differential, the patient was next evaluated with endoscopic retrograde cholangiopancreatography. Which of the following statements is true?

Failure of the cover test
Based on your differential, the patient was next evaluated with endoscopic retrograde cholangiopancreatography. Which of the following statements is true?

A. The patient has an increased risk of developing cholangiocarcinoma. \( \text{Prognosis} \)
B. The patient should undergo pancreatoduodenectomy. \( \text{Treatment} \)
C. The patient should undergo percutaneous transhepatic stent placement. \( \text{Treatment} \)
D. No action is necessary unless the lesion grows. \( \text{Treatment} \)

**Key = A**

Which type of choledochal cyst does the patient have?

A. Type 1  
B. Type 2  
C. Type 3, choledochocele  
D. Type 4  
E. Type 5, Caroli disease

Non-homologous choices, with the single outlier as the answer

So what’s the pitfall?
Which type of choledochal cyst does the patient have?

A. Type 1  
B. Type 2  
C. Type 3, choledochocele  
D. Type 4  
E. Type 5, Caroli disease

Special Considerations for Audience Response

- "Who is my audience?"
- Simple and short: ONE idea per question
- Target the big stuff (less really is more)
- Groom your images

Which type of choledochal cyst does the patient have?

A. Type 1  
B. Type 2  
C. Type 3, choledochocele  
D. Type 4  
E. Type 5, Caroli disease

Writing True–False Questions

- Incontrovertibly true or false
- Simple language
- ONE idea per question
- Phrase it positively
- Avoid extreme and nebulous modifiers

The Most Important Stuff

1. Relevance
2. Clarity of focus:  
   - stem especially  
   - answers and distractors, too  
3. Keep it positive
4. No “lazy test maker” clues

Reference links

- Writing Good Multiple-Choice Questions: A Brief Guide for Radiologists  
- Prevalence of Flawed MultipleChoice Questions in Continuing Medical Education Activities of Major Radiology Journals  
  http://www.rsna.org/docs/dl/articles/159944/AI.15.11961091.0x--.abstract
- Developing Multiple Choice Questions for the Royal College Certification Examinations, Royal College of Physicians and Surgeons of Canada  
- A Brief Guide to Writing Good Multiple Choice Questions  