**Question 1**

When we talk about error in epidemiology, we are generally talking about the difference between the true underlying population value of interest and our measure (or estimate) of that value.

1. True
2. False

This is true. Error is simply the difference between the truth and our estimate of the truth.

**Question 2**

The two highest-level (most general) categories of error are typically considered to be

1. Bias and confounding
2. Random and systematic
3. Differential and non-differential
4. Ignorable and non-ignorable

When thinking about error in epidemiologic studies, we typically first divide the error into random error (sampling variability) and systematic error (bias).

**Question 3**

We expect less random error to be present in our estimates, on average, as sample size increases?

1. True
2. False

This is true. As an extreme example, let's say we want to know the average height of everyone in the class. If we estimate the average height from a sample of 1 student, there's a pretty good chance that the true average height will be different than that one student's height (error). If we estimate the average height from every student in class, there will be no difference between the truth and our estimate (no error).

**Question 4**

We expect less systematic error to be present in our estimates, on average, as sample size increases?

1. True
2. False

Typically not. When we have *systematic* error, collecting additional data typically just leads to additional biased data. We end up just feeling more confident in our biased estimates.

**Question 5**

We conduct a cohort study to investigate the relationship between diet (vegetarian vs. standard) and heart disease. We find that the relative risk among those with vegetarian diet is 0.9 (10% less risk of heart disease). Later, we find that some of the people in the standard diet group actually began eating vegetarian diets during follow-up. Intuitively, how would you expect this new information to affect our results if vegetarian diet is, in fact, protective against heart disease?

1. We underestimated the true association between vegetarian diet and heart disease.
2. We overestimated the true association between vegetarian diet and heart disease.
3. This new information doesn't affect our results.

We were misclassifying people with a vegetarian diet as having a standard diet. If vegetarian diet is effective, then we would expect these people to have lower risk of heart disease. However, because we misclassified them, we were "taking" this lower risk from the vegetarian group and "giving" it to the standard diet group. Said another way, the misclassification probably lead to the appearance of less risk in the standard diet group than was actually present. Therefore, we expect that our estimate (0.9) underestimates the true association between vegetarian diet and heart disease.

**Question 6**

We conduct a cohort study to investigate the relationship between diet (vegetarian vs. standard) and heart disease. We find that the relative risk among those with vegetarian diet is 0.9 (10% less risk of heart disease). Later, we find that some of the people in the standard diet group actually began eating vegetarian diets during follow-up. After correctly classifying the study participants, we calculated a new relative risk of 0.7. Was our original estimate of the relative risk bias towards or away from the null?

1. Biased towards the null
2. Biased away from the null

0.9 is closer to the null (1.0) than 0.7. Therefore, we say that our original estimate, 0.9, was biased towards the null.

**Question 7**

If we have a true OR of 4 and an observed OR of 2.6, do we say that our observed estimate is biased towards or away from the null?

1. Biased towards the null
2. Biased away from the null

2.6 is closer to the null (1.0) than 4. Therefore, we say that our observed estimate, 2.6, is biased towards the null.

**Question 8**

If we have study results with the following sensitivity and specificity of exposure measurement:

•Sensitivity in classifying exposure among people with outcome = 0.9

•Specificity in classifying exposure among people with outcome = 0.8

•Sensitivity in classifying exposure among people without outcome = 0.9

•Specificity in classifying exposure among people without outcome = 0.8

Would we say that we have misclassification in our exposure?

1. No
2. Yes
3. Unable to tell from sensitivity and specificity alone

Yes. There is some misclassification (in this case, of the exposure) whenever sensitivity or specificity are less than 1.0.

**Question 9**

If we have study results with the following sensitivity and specificity of exposure measurement:

•Sensitivity in classifying exposure among people with outcome = 0.9

•Specificity in classifying exposure among people with outcome = 0.8

•Sensitivity in classifying exposure among people without outcome = 0.9

•Specificity in classifying exposure among people without outcome = 0.8

Would we say that we have differential or non-differential misclassification of our exposure?

1. Differential
2. Non-differential
3. Unable to tell from the information provided

In this case, we have non-differential misclassification of exposure. The sensitivity and specificity of the exposure are not dependent on outcome status. Said another way, there is equal misclassification of the exposure between participants with the outcome and participants without the outcome.

**Question 10**

A \_\_\_\_ study is equivalent to an “unbiased” study—a study that, based on its design, methods, and procedures, will produce (on average) overall results that are close to the truth.

Szklo, Moyses,Nieto, F. Javier. Epidemiology (Kindle Locations 3544-3546). Jones & Bartlett Learning. Kindle Edition.

1. Valid
2. Reliable

A valid study is equivalent to an “unbiased” study—a study that, based on its design, methods, and procedures, will produce (on average) overall results that are close to the truth.

Szklo, Moyses,Nieto, F. Javier. Epidemiology (Kindle Locations 3544-3546). Jones & Bartlett Learning. Kindle Edition.