# Measurement in Health and Disease

Suggested answers: sensitivity and specificity

Question 1

a) *What are the sensitivity and specificity of a test? What do they tell us?* The sensitivity is the proportion of subjects with disease who the test detects. It tells us how good the test is at finding the disease. The specificity is the proportion of subjects without disease who the test correctly identifies. It tells us how good the test is at excluding people who do not have the disease.

b) *How could we present the sensitivity and specificity graphically?* We could draw the receiver operating characteristic curve (ROC curve). This is a plot of the sensitivity versus 1 minus specificity for each possible cut-off, with the points joined up.
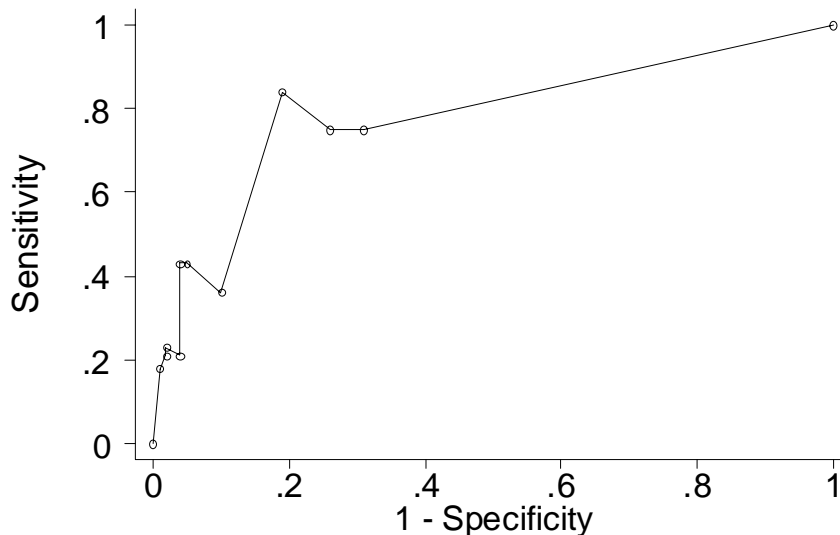
c) *What are positive and negative predictive values and on what do they depend?* The positive predictive value is the proportion of those subjects who are identified as diseased by the diagnostic test, who actually have the disease. It depends on the prevalence of the disease, increasing as prevalence increases. The negative predictive value is the proportion of those subjects who are identified as not having the disease (i.e. as being HIV negative) by the diagnostic test who actually do not have the disease (i.e. really are HIV negative). It also depends on the prevalence of the disease, decreasing as prevalence increases.

d) *Why are the denominators constant in the sensitivity and specificity columns but varying in the positive predictive value column?* The denominator for sensitivity is the number HIV positive, the denominator for specificity is the number HIV negative. These are the same for all the tests. The denominator for positive predictive value is the number positive on the test, which varies from test to test. Similarly, the denominator for negative predictive value is the number negative on the test, which also varies from test to test.

e)   *Why does the sensitivity go up and the specificity go down if we reduce the size of node which we consider positive?* Lowering the cut-off point makes more subjects test positive, while still including all those test positive at the higher cut-off. This increases the numerator of the sensitivity, while the denominator is constant, thus automatically increasing the sensitivity. It therefore also reduces the number test negative and reduces the specificity.

f)   *Which two diagnostic tests give the strongest relationship to HIV and how do they differ?* The largest specificity is for the combination of epitrochlear $\geq$ 0.5 cm and axillary $\geq$ 1 cm and submandibular $\geq$ 1 cm (99%), and the largest sensitivity is for epitrochlear $\geq$ 0.5 cm (84%). The combined test has low sensitivity, only 18%, though it has a high specificity. It does not find many cases, but those found are very likely to be HIV positive (PPV=96%). The epitrochlear $\geq$ 0.5 cm test has higher sensitivity, 84%, and also a fairly high specificity (81%). It would find most of the cases, and those found are fairly likely to be HIV positive (PPV=85%). It is thus the preferred test.

Question 2

In what sense does the reliability of a test increase as the prevalence increases and how is this relevant to testing such a child? As the prevalence of the disease increases the positive predictive value increases and the negative predictive value decreases. Thus a positive test is more likely to indicate the presence of the disease when the prevalence is high. As this child is a member of a high-risk population, a positive test would make the probability of HIV infection being present very high indeed.

Question 3

a)   *How might we describe the agreement between predicted and actual group membership?* Kappa = 0.64 is usually interpreted as 'good' agreement.

b)   *Lowering the criterion for dementia increased sensitivity without lowering specificity. What does this tell us about the relationship between the score and dementia classification?* It tells us that all the members of the non demented group had scores below both of the cut-offs, so there was quite a large area of the scale where there were no non-demented subjects.