Statistical Methods for Diagnostic Tests

Practical Exercise - Solutions

Question 1

(a) The 2 x 2 table summarising the results of the test in the Glasgow population is

```
> with(helico,table(Test,Biopsy))
    Biopsy
Test   0   1
   0 628  36
   1 127 478
```

The *reportROC()* function gives the following statistics:

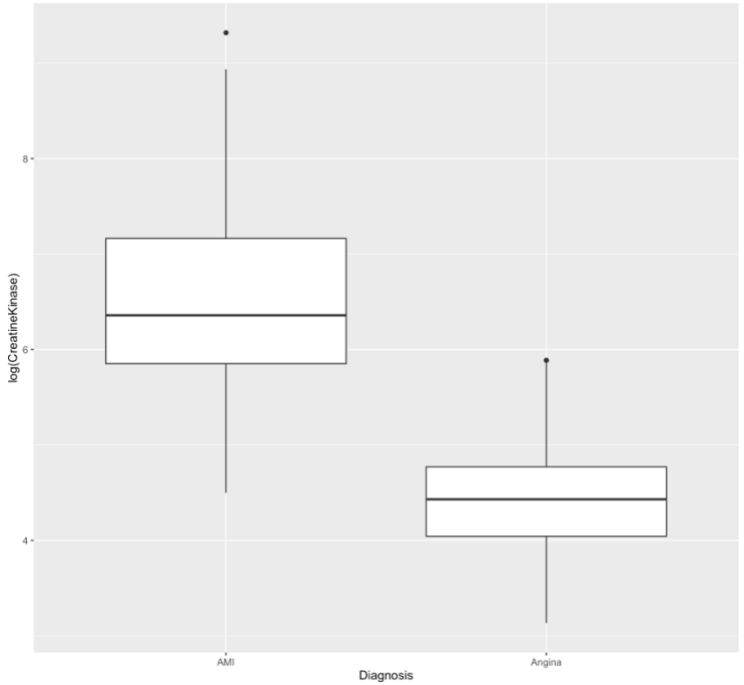
```
> reportROC(gold = helico$Biopsy,predictor.binary = helico$Test,plot=F,positive='1')
   AUC AUC.SE AUC.low AUC.up     P   ACC ACC.low ACC.up   SEN SEN.low SEN.up   SPE SPE.low SPE.up   PLR PLR.low PLR.up   NLR
 0.881  0.012   0.857  0.905 0.000 0.872   0.871  0.872 0.930   0.908  0.952 0.832   0.805  0.858 5.529   4.709  6.490 0.084
 NLR.low NLR.up   PPV PPV.low PPV.up   NPV NPV.low NPV.up
   0.061  0.116 0.790   0.758  0.823 0.946   0.929  0.963
```

This is a high sensitivity test (93%; 95% CI: 90.8%, 95.2%) with reasonable specificity (around 83%, with CI 80.5% to 85.8%). Thus, negative predictive value is also high (95%; 95% CI: 92.9%, 96.3%) and the negative likelihood ratio indicates that a negative test result would reduce ones estimate of the odds of *H. pylori* infection by more than a factor of 10. Thus, this test functions best as a rule-out test – negative results are mostly genuinely negative. Positive results have some confidence attached to them, but may need additional investigation to finalise a diagnostic decision.

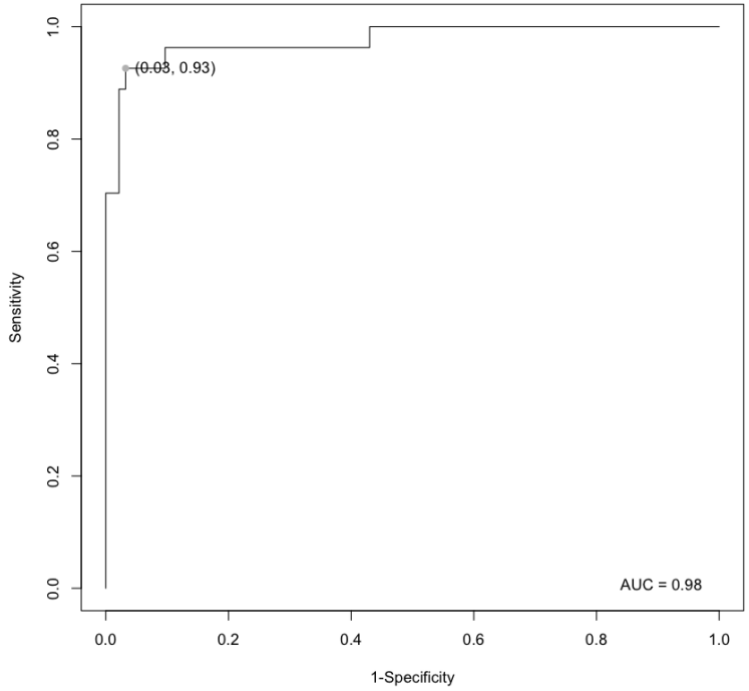
(b) In London, we can recalculate PPV and NPV using the formulae in the lecture. You should carry out the calculations with sensitivity, specificity and prevalence all expressed as decimals, rather than whole number percentages. This gives an updated estimate of PPV = 89.2% and NPV = 88.8%.

Question 2

(a) We have data on 27 AMI cases and 93 angina patients. Looking at a boxplot of CK on the log scale (it is quite skewed) we find reasonably good discrimination between the angina and AMI patients:



This also tells us that higher values of CK are associated with AMI (which we need to know to set up the *reportROC* function correctly). The ROC curve we obtain, therefore, is




```
> reportROC(gold = angina$DiagnosisCat,predictor = angina$CreatineKinase,plot=T,positive='1')
Setting levels: control = 0, case = 1
  Cutoff   AUC AUC.SE AUC.low AUC.up     P   ACC ACC.low ACC.up   SEN SEN.low SEN.up   SPE SPE.low SPE.up    PLR PLR.low
 299.500 0.975  0.013   0.942  1.000 0.000 0.958   0.958  0.959 0.926   0.827  1.025 0.968   0.932  1.004 28.704   9.381
 PLR.up   NLR NLR.low NLR.up   PPV PPV.low PPV.up   NPV NPV.low NPV.up
 87.824 0.077   0.020  0.291 0.893   0.778  1.007 0.978   0.948  1.008
```

We have a high AUC here (0.975; 95% CI: 0.943, 1.000), so the test performs with good discrimination overall. The function suggests an optimal (balanced Sensitivity & specificity) threshold at 299.5 units – lower than that would classify as AMI and above as Angina. Sensitivity and Specificity are both high at this value.

(b) The threshold above is at log(300) = 5.7 on the vertical axis of the boxplot. This value includes some of the lower whisker of the AMI group, so that some AMI cases will incorrectly be assessed as Angina using this threshold (false negatives). Thus, to get better "rule-out" performance, we need to lower the threshold, so that values below it are essentially only indicative of Angina – a log value of around 4.5 might do that, equating to a CK value of about 90. That will increase the sensitivity of the test, moving the threshold up and to the right in the ROC curve.