

Appendix 2 (as supplied by the authors): Sensitivity analyses: blinded versus unblinded reference standard and data with missing variables versus replacement with multiple imputations

When applied to the blinded reference standard, the AUC for the CHD score was 0.79 (95% CI: 0.74-0.84). As measures of agreement of the blinded versus the unblinded reference panel we obtained Kappa=0.62 (95% CI: 0.55, 0.69). Although measures of diagnostic effectiveness were on average reduced with the reference panel blinded to index tests results, they were of the same order of magnitude.

The original model contained 773 complete data sets; in 426 sets one or more variables were missing. We replaced these missing data using the multiple imputations resulting in 5 different complete datasets. In each of these datasets the AUC was calculated, resulting in values ranging from 0.848 to 0.853 with a mean of 0.85. Therefore, the average AUC of the 5 imputed datasets was only slightly lower than the AUC in the actual incomplete dataset.

Internal validation

The ROC was calculated on the base of 200 bootstrap resamples for internal validation in order to reduce overfit bias. The AUCs for those 200 ROC-curves ranged from 0.82 to 0.92 with a mean of 0.869 (95% CI: 0.867; 0.872).