

Supplementary Table S1. The QUADAS tool Item

	Yes	No	Unclear
1. Was the spectrum of patients representative of the patients who will receive the test in practice?	(235)	(-)	(-)
2. Were selection criteria clearly described?	(235)	(-)	(-)
3. Is the reference standard likely to correctly classify the target condition?	(235)	(-)	(-)
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	(198)	(-)	(37)
5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?	(234)	(1)	(-)
6. Did patients receive the same reference standard regardless of the index test result?	(125)	(49)	(61)
7. Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)?	(235)	(-)	(-)
8. Was the execution of the index test described in sufficient detail to permit replication of the test?	(189)	(7)	(39)
9. Was the execution of the reference standard described in sufficient detail to permit its replication?	(158)	(17)	(60)
10. Were the index test results interpreted without knowledge of the results of the reference standard?	(112)	(11)	(112)
11. Were the reference standard results interpreted without knowledge of the results of the index test?	(97)	(5)	(133)
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	(79)	(2)	(154)
13. Were uninterpretable/ intermediate test results reported?	(50)	(175)	(10)
14. Were withdrawals from the study explained?	(229)	(4)	(2)