Supplementary file 1: QUADAS scores for individual studies in each quality domain

Domains:

1. Was the spectrum of patients representative of the patients who will receive the test in practice

2. Were selection criteria clearly described

3. Is the reference standard likely to correctly classify the target condition

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

5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis

6. Did patients receive the same reference standard regardless of the index test result[#]

7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)[#]

8. Was the execution of the index test described in sufficient detail to permit replication of the test

9. Was the execution of the reference standard described in sufficient detail to permit its replication[#]

10. Were the index test results interpreted without knowledge of the results of the reference standard[#]

11. Were the reference standard results interpreted without knowledge of the results of the index test[#]

12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice

13.   Were uninterpretable/ intermediate test results reported

1st author → / Domain ↓	Asahina	Chung	Dooms	Eberhardt 2007	Fielding	Herth 2006	Herth 2002	Kurimoto	Paone	Shirakawa	Yoshikawa	Yamada	Asano	Yang	Huang	Eberhardt 2009
1	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear
2	unclear	unclear	unclear	yes	unclear	yes	unclear	unclear	yes	unclear	unclear	yes	unclear	yes	yes	yes
3	N/A	N/A	N/A	N/A	N/A	N/A	No	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
4	N/A	N/A	N/A	N/A	N/A	N/A	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
5	N/A	N/A	N/A	N/A	N/A	N/A	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
6	N/A	N/A	N/A	N/A	N/A	N/A	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
7	N/A	N/A	N/A	N/A	N/A	N/A	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
8	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9	N/A	N/A	N/A	N/A	N/A	N/A	Yes	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
10	N/A	N/A	N/A	N/A	N/A	N/A	unclear	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
11	N/A	N/A	N/A	N/A	N/A	N/A	unclear	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
12	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear	unclear
13	Yes	Yes	unclear	No	No	Yes	Yes	Yes	unclear	Yes	Yes	No	Yes	No	Yes	Yes
14	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total	3	3	2	3	2	4	8	3	3	3	3	3	3	3	4	4

No reference standard was used in published studies. Instead, a composite end-point of histology by any means OR adequate radiologic surveillance was used to determine final diagnoses for study subjects (as described in table 2)

N/A – not applicable.