### Supplementary Material 2. Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS–2) scale

| QUADAS 2          | Question                                                                                                                                                                                                 |
|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| **Patient selection** | Was a consecutive or random sample of patients enrolled? Was a case-control design avoided? Did the study avoid inappropriate exclusions?                                                                                                      |
| **Risk of bias** | Could the Selection of Patients Have Introduced Bias?                                                                                                                                                      |
| **Applicability** | Are There Concerns That the Included Patients and Setting Do Not Match the Review Question?                                                                                                               |
| **Index test**    | Were the index test results interpreted without knowledge of the results of the reference standard? If a threshold was used, was it pre-specified?                                                            |
| **Risk of bias** | Could the Conduct or Interpretation of the Index Test Have Introduced Bias?                                                                                                                                 |
| **Applicability** | Are There Concerns That the Index Test, Its Conduct, or Its Interpretation Differ From the Review Question?                                                                                               |
| **Reference standard** | Is the reference standard likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of the results of the index test? |
| **Risk of bias** | Could the Reference Standard, Its Conduct, or Its Interpretation Have Introduced Bias?                                                                                                                        |
| **Applicability** | Are There Concerns That the Target Condition as Defined by the Reference Standard Does Not Match the Question?                                                                                             |
| **Flow and timing** | Was there an appropriate interval? between the index test and reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?                                            |
| **Risk of bias** | Could the Patient Flow Have Introduced Bias?                                                                                                                                                              |