eTable 3: QUADAS-2 items and guidelines for scoring

Question	Response	Explanation
D01: Patient selection		
K01: Sampling Was the sampling method adequate?	No = high risk of bias Yes = low risk of bias	Where a sample is used, the designs with consecutive or random samples are least susceptible to producing bias. If the sample selection is based on volunteers or selected participants from a clinic or research institution, a bias is obvious.
K02: Study design Was a case-control or comparable design avoided?	No = high risk of bias Yes = low risk of bias	Design similar to the case-control approach, which could evoke bias, are those designs where the study team deliberately increases or decreases the proportion of study participants with the target disease that can no longer be representative. Some case-control methods can be excluded from the outset if they mix participants from different settings.
K03: Exclusion criteria Are the exclusion criteria described and appropriate?	No = high risk of bias Yes = low risk of bias	The study is automatically classified as unclear if the exclusions are not detailed (depending on the contact with the study authors). Where a detailed description is provided, the study will be considered low-risk if the reviewers consider the exclusion criteria appropriate.
D02: Index test		
K04: Blinding Was the evaluation and interpretation of the biomarker tests performed without knowledge of the clinical CJD diagnosis?	No = high risk of bias Yes = low risk of bias	Terms such as "blinded" or "independent and without knowledge of" are sufficient and comprehensive details of the blinding procedure are not required. The interpretation of the results of the index test could be influenced by the knowledge of the results of the reference standard. If the index test is always evaluated before the reference standard, the evaluator may not be aware of the results of the reference standard. Then this point can be answered with "yes". For specific index tests, the result is objective, and knowledge of the reference standard should not influence the results, e.g. for protein levels in cerebrospinal fluid the quality rating, in this case, is "low risk" even if the blinding has not been implemented.
K05: A priori cutoff Were the biomarker limits pre-specified?	No = high risk of bias Yes = low risk of bias	For scales and biomarkers, there is often a reference point (in units or categories) above which individuals are classified as "test positive". This can be associated with a limit value, a clinical cut-off or a dichotomization point. A study is considered to have a high bias risk once the authors have determined the optimal cut-off post hoc from their study data, as the selection of the threshold to maximize sensitivity and specificity may lead to an over-optimistic estimate of test performance. Some publications may use an alternative method of analysis without applying a limit. These publications should be classified as "not applicable".
D03: Reference standard		
K06: Clinical diagnosis Is the evaluation used for the clinical diagnosis of sCJD acceptable?	No = high risk of bias Yes = low risk of bias	Generally accepted international criteria to support the diagnosis of dementia are explained in ICD-10 and DSM-IV. Criteria for clinical diagnosis of sCJD can also be found in ICD-10[1] and CJD 2010[2,3]. Specific criteria for dementia subtypes include but are not limited to the NINCDSADRDA criteria for Alzheimer's disease; the McKeith criteria for Lewy body dementia; the Lund criteria for frontotemporal dementia; and the NINDS-AIREN criteria for vascular dementia. If the criteria applied are not familiar ("unclear") to the review authors, this point should be assessed as "high bias risk".

K07: Blinding Was the clinical evaluation of CJD performed without knowledge of the biomarkers?	No = high risk of bias Yes = low risk of bias	Terms such as "blinded" or "independent and without knowledge of" are sufficient and comprehensive details of the blinding procedure are not required. The interpretation of the results of the reference standard could be influenced by the knowledge of the results of the index test.
D04: Flow and timing		
K08: Time interval Was there an appropriate interval between biomarker use and clinical evaluation?	No = high risk of bias Yes = low risk of bias	The time interval between index test and reference standard influences the test accuracy. A time variable is therefore used in the evaluation. Moreover, its influence on the test accuracy is investigated. The minimum interval for a follow-up evaluation is set to 1 year. If more than 16% of the participants have received an examination, this aspect is assessed as "no".
K09: Equal treatment Did all patients undergo the same evaluation for dementia, regardless of the biomarkers?	No = high risk of bias Yes = low risk of bias	There may be scenarios in which people tested positive with the index test receive a more detailed examination. Where the assessment of dementia varies between individuals, the study should be rated with a high risk of bias.
K10: Final analysis Were all patients who received a biomarker study included in the final analysis?	No = high risk of bias Yes = low risk of bias	If the number of patients included differs from those shown in the 2x2 contingency table, distortions might occur: If the patients missing due to drop-outs differ systematically from the remaining patients, the estimators of the test performance may differ. If drop-outs were present, they should be quantified. A maximum of 20% has proven to be the maximum proportion to guarantee a low risk of distortion.
K11: Missing values Have missing or non-interpretable biomarker test results been reported?	No = high risk of bias Yes = low risk of bias	For reports of missing or non-interpretable values where there is a significant decrease (arbitrary value of 50% missing data), this should be classified as "no". If such results have not been reported, this should be considered "unclear", and the authors should be contacted.
D05: Applicability		
K12: Representativity Were the patients included representative of the general target population?		The included patients should be consistent with the target population described in the research question of the review. The characterization of the population is done in the form of symptoms, pre-tests, potential disease prevalence, and setting. If there are clear reasons for suspecting an unrepresentative spectrum, this aspect should be considered as little applicable.
K13: Repeatability Were reliable biomarker application data available for the tests to be repeated in independent studies?		Variations in technology, test performance, and interpretation may affect the estimation of accuracy. Also, the background and training/expertise of the assessor should be reported. If the plasma and CSF biomarkers were not consistently applied, this aspect should be assessed as poorly applicable.
K14: Recency Has the clinical diagnosis of sCJD been made in a manner consistent with current clinical practice?		For many reviews, the inclusion criteria and bias risk assessment of a CJD diagnosis will have already been assessed. For some reviews, a judgement on the applicability of the reference standard may not be available. There is a possibility that some form of CJD assessment, although valid, may lead to a diagnosis in a much larger proportion of individuals with the disease than in routine clinical practice. In this case, the aspect should be assessed as poorly applicable.

References:
1. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. 1993.
2. World Health Organization. Global surveillance, diagnosis, and therapy of human transmissible spongiform encephalopathies: report of WHO consultation (WHO/EMC/ZDI/98/9). Geneva; 1998.
3. Zerr I, Kallenberg K, Summers DM, et al. Updated clinical diagnostic criteria for sporadic Creutzfeldt-Jakob disease. Brain. 2009;132:2659–2668.