Received: 8 November 2021 | Accepted: 18 January 2022

DOI: 10.1111/ene.15258

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE

Diagnostic accuracy of cerebrospinal fluid biomarkers for the differential diagnosis of sporadic Creutzfeldt–Jakob disease: a (network) meta-analysis

Nicole Rübsamen[1] | Stephanie Pape[1] | Stefan Konigorski[2,3] | Antonia Zapf[4] | Gerta Rücker[5] | André Karch[1]

[1]Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany

[2]Molecular Epidemiology Research Group, Max Delbrück Center (MDC) for Molecular Medicine in the Helmholtz Association, Berlin, Germany

[3]Digital Health and Machine Learning Research Group, Hasso Plattner Institute for Digital Engineering, Potsdam, Germany

[4]Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[5]Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Freiburg im Breisgau, Germany

Correspondence
Nicole Rübsamen, Institute of Epidemiology and Social Medicine, University of Münster, Albert-Schweitzer-Campus 1, 48149 Münster, Germany.
Email: ruebsame@uni-muenster.de

Funding information
This research received no specific grant from any funding agencies in the public, commercial, or not-for-profit sector.

Abstract

Background: There are no systematic reviews of cerebrospinal fluid and blood biomarkers for sporadic Creutzfeldt–Jakob disease (sCJD) in specialized care settings that compare diagnostic accuracies in a network meta-analysis (NMA).

Methods: We searched Medline, Embase, and Cochrane Library for diagnostic studies of sCJD biomarkers. Studies had to use established diagnostic criteria for sCJD and for diseases in the non-CJD groups, which had to represent a consecutive population of patients suspected as a CJD case, as reference standard. Risk of bias was assessed with QUADAS-2. We conducted individual biomarker meta-analyses with generalized bivariate models. To investigate heterogeneity, we performed subgroup analyses based on QUADAS-2 quality and clinical criteria. For the NMA, we applied a Bayesian beta-binomial ANOVA model. The study protocol was registered at PROSPERO (CRD42019118830).

Results: Of 2976 publications screened, we included 16 studies, which investigated 14–3-3β ($n = 13$), 14–3-3γ ($n = 3$), neurofilament light chain (NfL, $n = 1$), neuron-specific enolase ($n = 1$), p-tau181/t-tau ratio ($n = 2$), RT-QuIC ($n = 7$), S100B ($n = 3$), t-tau ($n = 12$), and t-tau/Aβ42 ratio ($n = 1$). Excluded diagnostic studies had strong limitations in study design. In the NMA, RT-QuIC (0.91; 95% CI [0.83, 0.95]) and NfL (0.93 [0.78, 0.99]) were the most sensitive biomarkers for the diagnosis of definite, probable, and possible sCJD cases. RT-QuIC was the most specific biomarker (0.97 [0.89, 1.00]). Heterogeneity in accuracy estimates was high between studies.

Conclusions: We identified RT-QuIC as the most accurate biomarker, partially confirming currently applied diagnostic criteria. The shortcomings identified in many diagnostic studies for sCJD biomarkers need to be addressed in future studies.

KEYWORDS
diagnostic tests, network meta-analysis, prion diseases, sensitivity, specificity

Nicole Rübsamen and Stephanie Pape contributed equally to the manuscript.

INTRODUCTION

Sporadic Creutzfeldt–Jakob disease (sCJD), the most common human prion disease, accounts for 1–2 incident cases per million and year [1]. The reference standard for definite diagnosis is the post-mortem neuropathological examination of the brain, which is of little benefit to patients during their lifetime. A diagnostic composite reference standard [2] is an alternative: patients are differentiated into non-sCJD patients versus probable and possible sCJD cases based on clinical criteria combined with defined changes in magnetic resonance imaging, a characteristic electroencephalogram, or a positive biomarker test (Table 1).

The proteins 14–3–3 and the PrPSc aggregation assay RT-QuIC (real-time quaking-induced conversion) are currently incorporated in the diagnostic composite reference standard [2]. Other biomarkers, for example, phosphorylated tau (p-tau), total tau (t-tau), neuron-specific enolase (NSE), neurofilament light chain (NfL), and S100B, have been proposed as additions or replacements for the biomarkers in the diagnostic composite reference standard. There is no evidence from systematic reviews and diagnostic meta-analyses that have performed a comparative investigation of biomarkers suitable for the differential diagnosis of sCJD. Previous reviews in the field focused either on a single biomarker [3], or did not formally compare biomarkers in a meta-analysis [4]. We conducted a network meta-analysis based on a systematic review to compare the accuracy of established biomarker tests measured in the blood or cerebrospinal fluid (CSF) to diagnose sCJD in a specialized care setting under real-life conditions.

METHODS

We aimed to synthesize evidence from phase III diagnostic test accuracy studies with recruitment of consecutive patients suspected of sCJD who subsequently received the index test and the reference standard (based on established diagnostic criteria). We systematically searched Medline, Embase, and the Cochrane Library for diagnostic studies that assessed the accuracy of blood or CSF biomarkers to diagnose sCJD (Table S1). S.P. searched all databases on 25 July 2018 (initial search) and 23 September 2020 (update). The search term for Medline was (((((Biomarker) OR biomarkers [MeSH Terms])) OR ((Diagnosis) OR diagnosis [MeSH Terms]))) AND ((((Creutzfeldt Jakob disease) OR CJD)) OR cjd creutzfeldt jakob disease [MeSH Terms]) AND ((CSF OR cerebrospinal) OR (blood or serum or plasma)).

Studies were included if they assessed CSF or blood biomarkers' accuracy for the differentiation of sCJD from other diseases in a specialized care setting (Table S2). Eligible studies had to use established diagnostic criteria of sCJD [2,5] and established diagnostic criteria for diseases in the non-CJD groups (e.g., Alzheimer's disease, other rapid progressive dementia) as reference standard. The non-CJD group had to represent a consecutive population of patients suspected as a CJD case (because of rapidly progressive dementia or rapidly progressive other symptoms of neurodegeneration) and referred to the specialized care centre for a first diagnostic workup. Studies were not eligible if the non-CJD group consisted of healthy individuals or did not focus on patients representing the full spectrum of differential diagnoses as seen in clinical practice. The biomarker tests had to be performed during the patient's first diagnostic workup for her/his symptoms without a predefined diagnosis. Studies were not eligible if blood or CSF samples were selected from biorepositories based on already known diagnoses. The population, the index test, and the reference standard had to be described in sufficient detail to allow replication. If this was not the case, the respective study was not eligible for inclusion because it has been shown that including such studies may overestimate the accuracy of diagnostic tests [6-9].

The studies had to provide sufficient information to construct a diagnostic contingency table. There were no constraints regarding time, language, patient population, type of biomarker, or reference standard. Different restrictions were used for specific biomarkers:

TABLE 1 Diagnostic criteria for sporadic Creutzfeldt–Jakob disease

Definite sCJD	Probable sCJD	Possible sCJD
Progressive neurological syndrome AND (neuropathologically OR immunocytochemically OR biochemically confirmed)	Rapidly progressive cognitive impairment AND 2 clinical manifestations (myoclonus, visual or cerebellar problems, pyramidal or extrapyramidal features, akinetic mutism) AND (typical EEG OR typical MRI brain scan OR positive 14–3–3) OR progressive neurological syndrome AND positive RT-QuIC in CSF or other tissues	Rapidly progressive cognitive impairment AND 2 clinical manifestations (myoclonus, visual or cerebellar problems, pyramidal or extrapyramidal features, akinetic mutism) AND duration <2 years

Abbreviations: CSF, cerebrospinal fluid; EEG, electroencephalogram; MRI, magnetic resonance imaging.

results for 14–3–3 and RT-QuIC were included if there was no risk of incorporation bias [10]i.e., if 14–3–3 or RT-QuIC were not incorporated in the reference standard used). Results for t-tau were included if the cut-off used to classify a test result as positive/negative was within ±10% of the recommended cut-off 1300 pg/mL [11]. For all other biomarkers, the cut-off values used in the study needed to be based on external knowledge or training datasets. We allowed different cut-off values for the same biomarker as long as they fulfilled the abovementioned criteria.

S.P. and A.K. independently reviewed all titles, abstracts, and, if publications were included based on this information, full texts. Discrepancies were resolved in a consensus meeting.

All reference lists of publications included in the final systematic review were searched for additional studies that were missed, and these studies were further included in the review. We did not specifically assess grey literature sources because due to the rarity of sCJD and the clustering of patients in few specialized care settings worldwide, newly obtained findings are generally made available in manuscript form.

S.P. extracted clinical and demographic information as well as details of the assays and cut-offs used. N.R. and S.P. independently extracted the number of true and false positives and negatives (stratified by the level of certainty of sCJD diagnosis). We contacted the corresponding authors if further information was needed.

S.P., A.K., and N.R. performed their risk of bias assessments for the included studies independently using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [12]. QUADAS-2 covers quality evaluations in four different domains (patient selection, index test, reference standard, flow and timing) by using signaling questions related to the research question, and is supplemented by applicability inquiries. A.K. and S.P. first piloted the tool by using two randomly selected publications [11,13]. Satisfactory agreement was reached in all domains so that the QUADAS-2 signaling questions (Table S3) were retained unchanged for the assessment of the included studies.

Data analysis

All analyses were performed with R version 4.0.3 [14]. We calculated Krippendorff's alpha [15] to assess the reliability of the QUADAS-2 assessment between the investigators.

Meta-analyses of individual biomarkers

We tabulated true positives, false negatives, false positives, and true negatives in patients with and without sCJD, stratified by study and level of certainty of sCJD diagnosis (as reported in the individual studies). Based on this, we calculated estimates of sensitivity and specificity and their 95% confidence intervals (CIs). To investigate publication bias, we constructed funnel plots of effective sample size versus the estimated log diagnostic odds ratios and did a regression test of asymmetry [16]. To synthesize data, we implemented the generalized linear mixed model approach by Chu and Cole [17], which uses a bivariate binomial model to jointly analyze pairs of sensitivity and specificity, using the *glmer* function in the R package *lme4* (version 1.1–25) [18]. This approach allows estimation of the correlation between true positive rate (TPR) and false-positive rate (FPR) as well as the between-study standard deviation (SD) for both of them via random effects, which provides information on the heterogeneity of the results [19]. To investigate sources of heterogeneity, we performed subgroup analyses based on QUADAS-2 quality and clinical criteria. We repeated all analyses for different levels of certainty of sCJD diagnosis: (1) definite sCJD cases, (2) definite and probable sCJD cases, and (3) definite, probable, and possible sCJD cases.

Network meta-analysis

To compare the diagnostic accuracy of different biomarkers, we applied a diagnostic network meta-analysis (NMA) approach for evidence synthesis. We used the beta-binomial analysis of variance model for NMA of diagnostic test accuracy data as described by Nyaga et al. [20] to combine direct and indirect evidence simultaneously. This arm-based generalized linear mixed model models sensitivity and specificity as repeated measures jointly through a copula function and assumes that the missing tests/arms are missing at random. The models were fitted in the Bayesian framework with beta(1,1) = U(0,1) as prior distribution on the hyper-parameters using Stan [21] through the R package *rstan* (version 2.21.2) [22]. We repeated the network meta-analyses for different levels of certainty of sCJD diagnosis (as described earlier).

Ethical approval

All data that informed the meta-analysis are available in the public domain [11,23-37]. No approval by an institutional review board or regional review board was needed because there was no use of humans for this meta-analysis. The study protocol was registered with PROSPERO (CRD42019118830).

RESULTS

Systematic review

Our database search retrieved 2976 articles. After reviewing the titles and abstracts, we excluded 2668 articles, mostly because these studies were not diagnostic studies, they duplicated data already included, or the case definition or study population was not appropriate. After a full-text review of the remaining 308 articles, we found that 114 studies were not diagnostic studies and that 32 articles reported duplicated data. Among the remaining 162 diagnostic

studies, we excluded 109 studies because of inappropriate study populations ($n = 63$) or insufficient information on selection and characteristics of the study population ($n = 46$). Thirty-four studies were excluded because of an inappropriate target condition or insufficient information on index test or outcome, leaving 19 articles for risk of bias assessment (Figure 1).

Sufficient inter-rater reliability was achieved in the QUADAS-2 assessment (Table S4). Based on high risks of bias and poor applicability, the studies by Hamlin et al. [13], Rudge et al. [38], and Wang et al. [39] were excluded, leaving 16 studies for further quantitative analyses [11,23-37] (Table 2).

Fourteen of these studies investigated 14–3-3β via Western blot. The results regarding 14–3-3β in the study by Sanchez-Juan et al. [26] were excluded due to high risk of incorporation bias, leaving 13 studies that investigated 14–3-3β. Three studies investigated 14–3-3γ via enzyme-linked immunosorbent assay (ELISA). NfL, NSE, p-tau181/t-tau ratio, S100B, and t-tau/Aβ42 ratio were investigated in one, one, two, three, and one studies, respectively (Table 2). Seven studies investigated RT-QuIC, either with the first-generation assay (PQ-CSF, $n = 1$) or the second-generation assay (IQ-CSF, $n = 6$). Thirteen studies investigated t-tau, but the study by Leitão et al. [24] used the cut-off 1035 pg/mL to classify a test result as positive, so these results were excluded. Studies included in our analyses investigated between one and five biomarkers in the same study population, providing direct comparisons for some of the biomarkers involved.

Meta-analyses of individual biomarkers

The funnel plots did not indicate publication bias for any studied biomarker (Figure S1). Among definite sCJD cases, the range of observed sensitivities was 0.80 to 0.99 for 14–3-3β, 0.88 to 0.96 for 14–3-3γ, 0.82 to 0.96 for RT-QuIC, 0.65 to 0.93 for S100B, and 0.77 to 0.94 for t-tau (Figure 2). IQ-CSF had higher sensitivities for RT-QuIC than PQ-CSF. The range of observed specificities was 0.24 to 0.97 for 14–3-3β, 0.69 to 0.95 for 14–3-3γ, 0.95 to 0.99 for RT-QuIC, 0.90 to 0.93 for S100B, and 0.33 to 0.95 for t-tau (Figure 2). Individual meta-analyses of NfL, NSE, p-tau181/t-tau ratio, and t-tau/Aβ42 ratio were not conducted due to the low number of studies. Heterogeneity was high for all individual meta-analyses so that no pooled estimates were reported (Table S5–Table S9, Figure S2–Figure S6). Ranges of observed sensitivities and specificities were similar when including lower levels of certainty of sCJD diagnosis (Figure S7, Figure S8) except for the specificity of S100B, which was considerably lower when all levels of certainty were included. When including definite sCJD cases only, heterogeneity among studies was highest for the sensitivity of S100B; heterogeneity in specificity was high for all biomarkers but S100B (Figure 2).

There were various sources of heterogeneity for all biomarkers in all scenarios; the most important ones (identified by a decrease of heterogeneity in the respective subgroup analysis) were the definition of the study population (especially concerning the group without CJD), study design, blinding of the reference standard, and clinical characteristics of the patients (Table S5–Table S9, Figure S4–Figure S8).

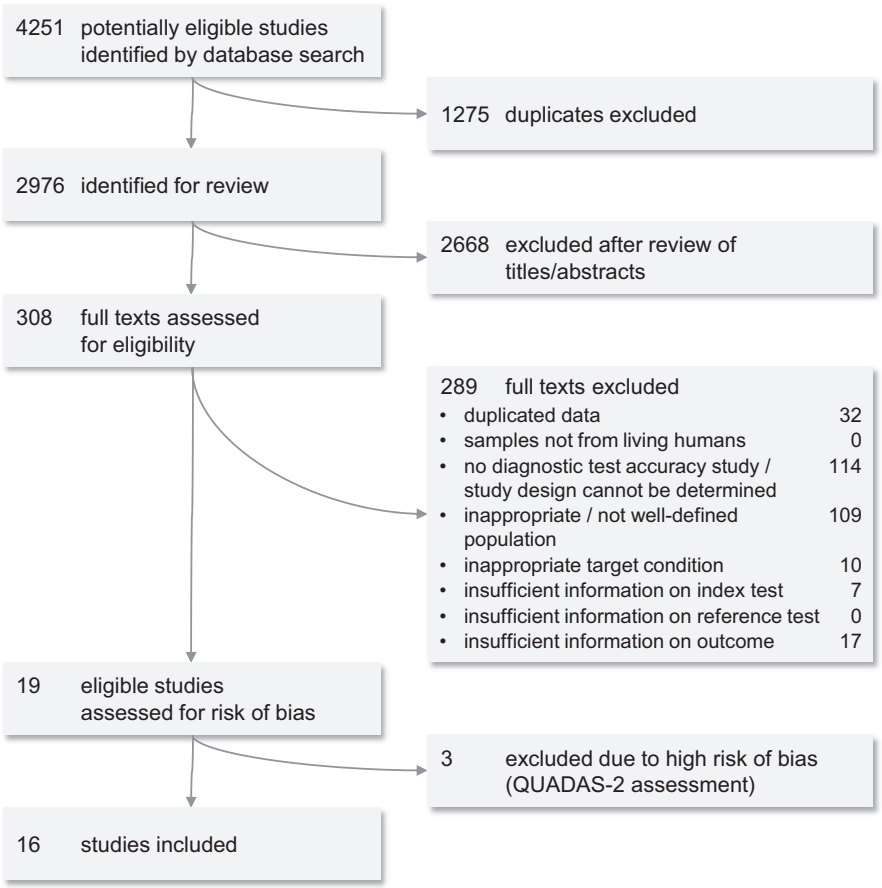

FIGURE 1 Flow chart of study selection

TABLE 2 Results of diagnostic sporadic Creutzfeldt–Jakob disease blood or cerebrospinal fluid biomarker studies that met the inclusion criteria

		All		Definite		Probable		Possible			
		sCJD cases								Non-CJD cases	
Study	Biomarker	TP	FN	TP	FN	TP	FN	TP	FN	TN	FP
Abu-Rumeileh et al. (2019)	14–3-3γ	70	10	70	10	–	–	–	–	75	34
	14–3-3β	64	16	64	16	–	–	–	–	72	37
	NfL	79	1	79	1	–	–	–	–	47	62
	RT-QuIC	60	2	60	2	–	–	–	–	53	0
	t-tau	74	6	74	6	–	–	–	–	86	23
Baldeiras et al. (2009)	14–3-3β	29	1	29	1	–	–	–	–	32	9
	p-tau181/t-tau	28	2	28	2	–	–	–	–	39	2
	S100B	28	2	28	2	–	–	–	–	38	3
	t-tau	27	3	27	3	–	–	–	–	39	2
	t-tau/Aβ42	28	2	28	2	–	–	–	–	39	2
Bizzi et al. (2020)	14–3-3β	295	26	295	26	–	–	–	–	9	30
	RT-QuIC	68	11	68	11	–	–	–	–	9	0
	t-tau	340	51	340	51	–	–	–	–	37	24
Bongianni et al. (2017)	14–3-3β	49	9	26	2	23	7	–	–	10	7
	RT-QuIC	Excluded (samples selected based on negative PQ-CSF findings)									
	t-tau	42	10	23	7	19	3	–	–	3	6
Chohan et al. (2010)	14–3-3β	210	35	210	35	–	–	–	–	127	44
	S100B	158	85	158	85	–	–	–	–	152	17
	t-tau	175	41	175	41	–	–	–	–	115	20
Cuadrado-Corrales et al. (2006)	14–3-3β	155	22	67	8	88	14	–	–	480	15
Fiorini et al. (2020)	14–3-3β	87	15	54	7	33	8	–	–	37	43
	RT-QuIC	98	4	58	3	40	1	–	–	80	0
	t-tau	87	15	56	5	31	10	–	–	56	24
Franceschini et al. (2017)	14–3-3β	110	14	56	7	54	7	–	–	24	18
	RT-QuIC	120	4	61	2	59	2	–	–	42	0
	t-tau	117	7	58	5	59	2	–	–	26	16
Lattanzio et al. (2017)	14–3-3β	259	53	155	31	80	17	24	5	133	79
	RT-QuIC	247	58	148	31	77	20	22	7	162	1
	t-tau	278	34	164	22	90	7	24	5	158	54
Leitão et al. (2016)	14–3-3γ	69	3	69	3	–	–	–	–	69	4
	14–3-3β	72	0	72	0	–	–	–	–	46	27
	p-tau181/t-tau	69	3	69	3	–	–	–	–	64	9
	t-tau	Excluded (cut-off value not within ±10% of 1300 pg/mL)									
Otto et al. (2002)	14–3-3β	94	15	94	15	–	–	–	–	70	15
	t-tau	103	6	103	6	–	–	–	–	75	10
Rhoads et al. (2020)	14–3-3β	Excluded (not only sCJD cases)									
	RT-QuIC	408	31	408	31	–	–	–	–	69	1
	t-tau	Excluded (not only sCJD cases)									

(Continues)

TABLE 2 (Continued)

		All		Definite		Probable		Possible			
		sCJD cases								Non-CJD cases	
Study	**Biomarker**	**TP**	**FN**	**TP**	**FN**	**TP**	**FN**	**TP**	**FN**	**TN**	**FP**
Sanchez-Juan et al. (2006)	14–3-3β	Excluded (high risk of incorporation bias)									
	NSE	379	138	Level of certainty of sCJD diagnosis not reported						12	0
	S100B	483	106							7	3
	t-tau	704	115							13	1
Simon et al. (2020)	14–3-3γ	103	19							452	49
	RT-QuIC	117	5							496	5
	t-tau	112	10							442	59
Van Everbroecket al. (2003)	14–3-3β	52	0	47	0	5	0	–	–	183	15
	t-tau	45	7	–	–	–	–	–	–	193	5
Zerr et al. (1998)	14–3-3β	155	19	62	3	64	5	29	11	155	9

Abbreviations: sCJD, sporadic Creutzfeldt–Jakob disease; FN, false negative test results; FP, false positive test results; PQ-CSF, first-generation RT-QuIC assay; TN, true negative test results; TP, true positive test results.

Network meta-analysis

Since differences in estimates between studies in the individual meta-analyses were consistent across different biomarkers, an evidence-synthesis approach based on the combination of intra-study differences (Figure 3) was used to derive pooled estimates of diagnostic accuracy. In the NMA analysis based on definite sCJD cases only, RT-QuIC was the most specific (0.96 [0.85, 1.00]) and the second most sensitive (0.91 [0.83, 0.96]) biomarker (Figure 4, Table 3). The balanced accuracy ([sensitivity + specificity]/2) was also highest for RT-QuIC (0.93). While NfL had the highest sensitivity in this NMA analysis (0.92 [0.72, 0.99]), its specificity was the lowest among all biomarkers (0.45 [0.15, 0.79]). Ratios involving t-tau (p-tau181/t-tau, t-tau/Aβ42) had higher accuracies than t-tau alone (Table 3); however, sensitivities and specificities were estimated with wider 95% CI because fewer studies investigated these ratios than t-tau alone.

The pooled sensitivities and specificities of most included biomarkers did not change much (±0.01) when including lower levels of certainty of sCJD diagnosis (Table 3). As an exception, specificity increased from 0.75 to 0.79 for 14–3-3γ and from 0.72 to 0.77 for t-tau, while it decreased from 0.84 to 0.82 for S100B. At the same time, the sensitivity increased from 0.74 to 0.77 for S100B. RT-QuIC was the most accurate test in among all levels of certainty of sCJD diagnosis. When excluding RT-QuIC data from Lattanzio et al. [37] who used PQ-CSF, RT-QuIC sensitivity increased to 0.93 (Table S10).

DISCUSSION

In our systematic review, we evaluated for the first time systematically the diagnostic accuracy of CSF and blood biomarkers for the differential diagnosis of sCJD in a specialized care setting. We applied the innovative concept of a diagnostic NMA to compare accuracies between different biomarkers. We found that RT-QuIC had the best-balanced accuracy throughout all scenarios.

Most of the 162 diagnostic accuracy studies retrieved in this systematic review did not meet the highest quality requirements of evidence-based medicine. All diagnostic studies evaluated were subject to methodological limitations, partly due to the rare disease situation, the lack of a uniform reference standard, and various forms of bias. Many studies had to be excluded because inappropriate study populations had been selected: either sCJD patients who already had received their diagnosis or control groups not suspected of sCJD. Even the 16 (10%) studies included in the final analysis had limitations mainly based on the study population's composition and the timing of the definition of the reference standard.

Subgroup analyses indicated that heterogeneity was mainly caused by choice of study population and less by methodological quality or clinical criteria. Specificity estimates showed the highest heterogeneity since most biomarkers used in the sCJD context are unspecific markers of neurodegeneration so that the composition of the non-sCJD group is crucial. One exception was RT-QuIC, which is the only available prion-specific biomarker to date. Although we only included diagnostic studies that reflected the real-world referral setting of a specialized care center, the proportion of other diseases in the non-CJD group might have been heterogeneous based on country-specific referral patterns or local research foci.

Many methodological constraints found in our systematic review were typical for diagnostic accuracy studies dealing with low-prevalence settings. To ensure optimal diagnostic accuracy estimates, Holtman and colleagues provide a general guide for six designs applicable in different low-prevalence situations [40]. Moreover, taking into account the recommendations of the regulatory authorities European Medicines Agency (EMA) [41] and

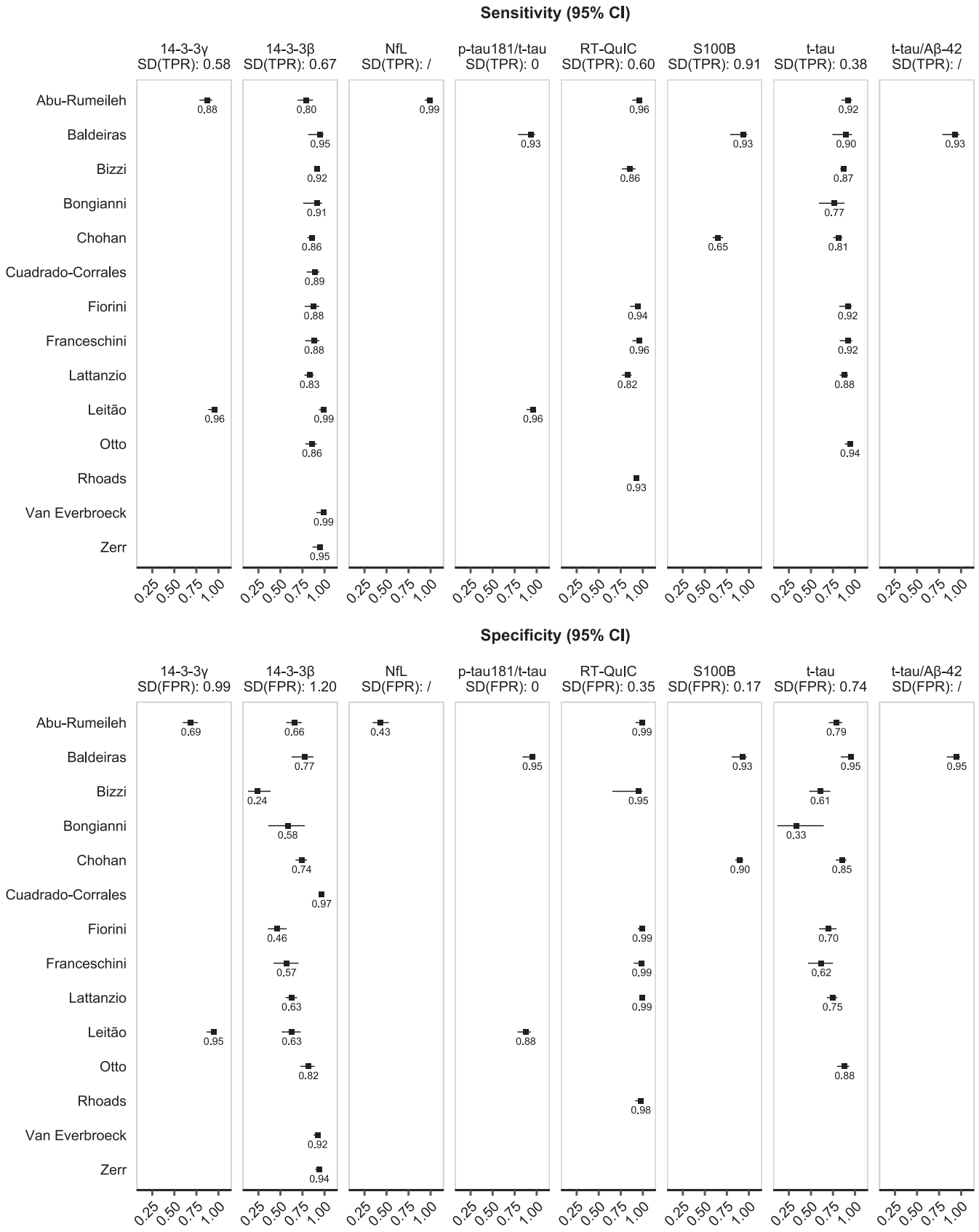

FIGURE 2 Sensitivity and specificity of cerebrospinal fluid biomarkers for the diagnosis of definite sporadic Creutzfeldt–Jakob disease (sCJD). CI, confidence interval; FPR, false positive rate; SD, standard deviation; TPR, true positive rate

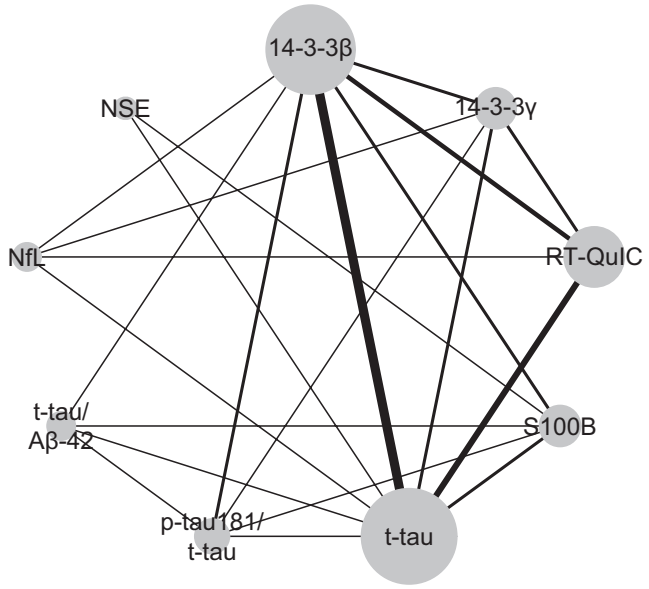

FIGURE 3 Network plot in which the thickness of the nodes is proportional to the number of direct comparisons

US Food and Drug Administration (FDA) [42], the Cochrane Collaboration [43], QUADAS-2 [12], and the STARD checklist [44] can contribute considerably to ensuring high-quality diagnostic accuracy studies. While a perfect diagnostic study would evaluate all study participants against the reference standard of a post-mortem neuropathological evaluation, this cannot be put into practice because of ethical, practical, and self-determination reasons. Selective restriction to autopsied cases only, however, can lead to biased estimates due to partial verification bias [10].

The inclusion/exclusion criteria and signaling questions of the QUADAS-2 tool that we applied do not only represent the prerequisites for conducting valid meta-analyses, but also reflect the patient characteristics and basic medical criteria that are essential for reliable decision-making in a real-world clinical setting. In the absence of this information in diagnostic accuracy studies, such studies provide only limited evidence to support the medical diagnosis of sCJD in everyday clinical practice. If relevant patient data cannot be retrieved, diagnostic studies may be of little use for the reader because the applicability and generalizability of the results remain unclear.

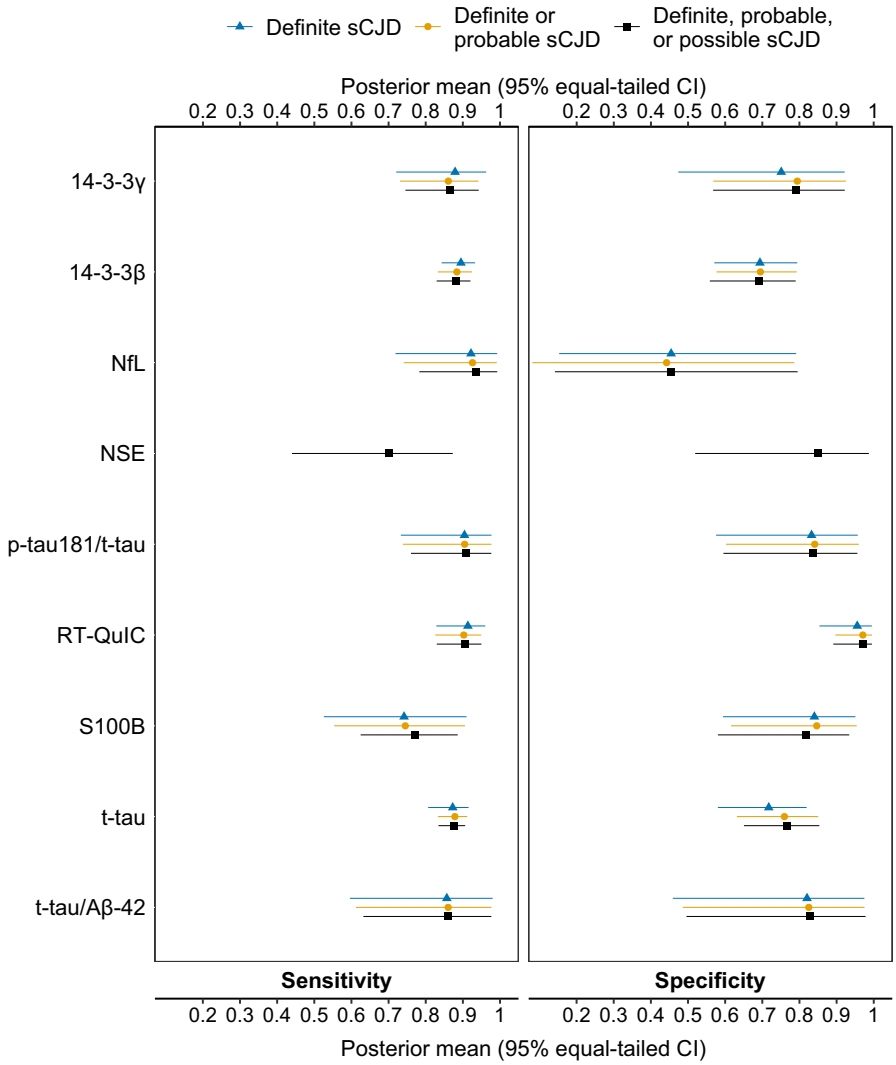

FIGURE 4 Results of network meta-analyses (stratified by level of certainty of sporadic Creutzfeldt–Jakob disease [sCJD] diagnosis). CI, confidence interval [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Results of network meta-analyses (sorted by accuracy)

	Definite, probable, or possible sCJD			Definite or probable sCJD			Definite sCJD		
Biomarker	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy
RT-QuIC	0.91 (0.83, 0.95)	0.97 (0.89, 1.00)	0.94	0.90 (0.83, 0.95)	0.97 (0.90, 1.00)	0.94	0.91 (0.83, 0.96)	0.96 (0.85, 1.00)	0.93
p-tau181/t-tau	0.91 (0.76, 0.98)	0.84 (0.60, 0.96)	0.87	0.90 (0.74, 0.98)	0.84 (0.60, 0.96)	0.87	0.90 (0.73, 0.98)	0.83 (0.58, 0.96)	0.87
t-tau/Aβ42	0.86 (0.63, 0.98)	0.83 (0.50, 0.98)	0.84	0.86 (0.61, 0.98)	0.83 (0.49, 0.98)	0.84	0.86 (0.60, 0.98)	0.82 (0.46, 0.97)	0.84
14–3-3γ	0.87 (0.75, 0.94)	0.79 (0.57, 0.92)	0.83	0.86 (0.73, 0.94)	0.79 (0.57, 0.92)	0.83	0.88 (0.72, 0.96)	0.75 (0.47, 0.92)	0.82
t-tau	0.88 (0.83, 0.91)	0.77 (0.65, 0.85)	0.82	0.88 (0.83, 0.91)	0.76 (0.63, 0.85)	0.82	0.87 (0.81, 0.92)	0.72 (0.58, 0.82)	0.80
S100B	0.77 (0.62, 0.89)	0.82 (0.58, 0.93)	0.80	0.74 (0.55, 0.91)	0.85 (0.62, 0.95)	0.80	0.74 (0.53, 0.91)	0.84 (0.59, 0.95)	0.79
14–3-3β	0.88 (0.83, 0.92)	0.69 (0.56, 0.79)	0.79	0.88 (0.83, 0.92)	0.69 (0.58, 0.79)	0.79	0.89 (0.84, 0.93)	0.69 (0.57, 0.79)	0.79
NSE	0.70 (0.44, 0.87)	0.85 (0.52, 0.99)	0.78						
NfL	0.93 (0.78, 0.99)	0.45 (0.14, 0.80)	0.69	0.93 (0.74, 0.99)	0.44 (0.08, 0.79)	0.68	0.92 (0.72, 0.99)	0.45 (0.15, 0.79)	0.69

Abbreviations: CI, confidence interval; sCJD, sporadic Creutzfeldt–Jakob disease.

By excluding healthy individuals and focusing on patients representing true differential diagnoses such as Alzheimer's disease or other rapid progressive dementia, studies increase their relevance for clinical decision-making. It has to be assessed carefully if certain diagnoses, which are more prevalent than sCJD and can be diagnosed easily using a set of diagnostic criteria for this disease, can be removed from the field of clinically relevant differential diagnoses of sCJD if sCJD diagnosis is already ruled out based on these alternative criteria [45]. In such a case, it may also be reasonable to consider these diagnostic tests directly in the evaluation of potential sCJD patients.

Our NMA results imply that RT-QuIC is overall the most accurate biomarker with both high sensitivity and specificity. RT-QuIC has already been proposed as an addition to 14–3–3 positivity for the composite reference standard [46], and an updated reference standard has been validated by the authors of the German CJD guideline [47]. However, it is currently unclear how to best combine RT-QuIC and 14–3–3 in the setting of the complex composite reference standard. A naïve approach would be to use the easily applicable, but still highly sensitive 14–3–3—or t-tau, which was as sensitive as 14–3–3 in the NMA—as screening test or initial index test in patients with suspected sCJD based on clinical criteria (Table 1). Only patients that tested positively could then be further examined according to the composite reference standard, including RT-QuIC instead of 14–3–3 to confirm sCJD diagnosis. Abu-Rumeileh et al. [23] explored the combinations of two biomarkers (a surrogate marker plus RT-QuIC). Integrating their data in our NMA found that NfL had the highest sensitivity for all certainty levels of sCJD diagnosis, but the lowest specificity, resulting in the poorest balanced accuracy. From our perspective, the issues of biomarker test combinations and screening tests need further validation. High-quality primary studies with paired information on at least two biomarkers for each individual are necessary to develop more sophisticated strategies to combine biomarkers. Moreover, estimates for the expected benefit for the patient are needed to make inferences about the likely impact on patient-important outcomes [48].

Strengths and limitations

Our systematic review provides the most extensive evaluation of biomarkers for the differential diagnosis of sCJD. It is the first one offering head-to-head comparisons. Despite considerable heterogeneity between studies, this comparison was possible since we applied a specific form of diagnostic NMA. We only included studies that mimicked a phase III diagnostic study design, and were performed in the real-life target population. By doing so, we wanted to provide the best available evidence to support clinicians in daily real-world clinical decision-making in the differential diagnosis of sCJD. All the included studies had a moderate risk of bias in at least one of the categories assessed so that residual bias in estimates cannot be ruled out. Of note, the pooled diagnostic accuracy of RT-QuIC in our NMA was mainly based on results from the IQ-CSF assay (six of seven studies), while most European national surveillance centers currently use the PQ-CSF assay.

Compared to CSF biomarkers, serum biomarkers have many advantages (e.g., with respect to invasiveness, patient acceptability, cost and time-effectiveness, and population-level feasibility),

making them beneficial complementary partners of CSF biomarkers or neuroimaging [49]. However, no study using serum biomarkers could be included in our systematic review. Both, comparing biofluid biomarkers to, and combining them with, neuroimaging is a relevant research field, as the recent study of Bizzi et al. [31] showed. We included the biomarker data from this study in our NMA, but decided against an extension of our work to imaging markers, as this would result in more complex requirements about what needs to be reported in the respective studies and about how the studies were performed. Due to the rare disease situation, the same patients might have been included in more than one study within the review. Although no author has explicitly described this, frozen CSF samples from one or more centers were pooled in several studies opening a potential for reuse of already analyzed samples; this could have resulted in an overestimation of the study population's variability. We did not consider that diagnostic accuracy might depend on sCJD subtype; its distribution might have affected individual study estimates. However, we only included studies that enrolled consecutive patients with suspected sCJD or a representative sample of those with suspected sCJD, so that study populations mirror the patient populations to which diagnostic tests are applied in clinical practice.

CONCLUSIONS

Our NMA suggested RT-QuIC as the most powerful biomarker test for the differential diagnosis of sCJD. The indirect comparisons undertaken in the NMA complete the already available head-to-head evidence and confirm that RT-QuIC (potentially in combination with a more easily applicable screening test like 14–3–3 or t-tau) should be implemented in the composite reference standard. Our work also pointed out the methodological limitations of previous diagnostic accuracy studies in the field, and the requirements we consider necessary for the design and conduct of future research projects in this setting. New high-quality studies with appropriate study designs are necessary to provide appropriate diagnostic accuracy data for unanswered research questions such as how to improve diagnostic processes in the pre-symptomatic period.

ACKNOWLEDGMENT

Open access funding was enabled and organized by ProjektDEAL.

CONFLICT OF INTEREST

A.K. published various manuscripts on biomarkers for sporadic Creutzfeldt–Jakob disease. He worked from 2011 to 2013 as research fellow at the German National Reference Centre for Transmissible Spongiform Encephalopathies at University Hospital Göttingen. N.R., S.P., S.K., A.Z., and G.R. have nothing to disclose.

AUTHOR CONTRIBUTIONS

Nicole Rübsamen: Data curation (equal); Formal analysis (lead); Methodology (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (lead). **Stephanie Pape:** Conceptualization (equal); Data curation (lead); Investigation (lead); Writing – original draft (equal); Writing – review & editing (equal). **Stefan Konigorski:** Formal analysis (equal); Writing – review & editing (equal). **Antonia Zapf:** Formal analysis (equal); Methodology (equal); Writing – review & editing (equal). **Gerta Rücker:** Formal analysis (equal); Methodology (equal); Writing – review & editing (equal). **André Karch:** Conceptualization (lead); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Table 2 and in the supplementary material for this article.

ORCID

Nicole Rübsamen https://orcid.org/0000-0003-2198-5577

REFERENCES

1. Heinemann U, Krasnianski A, Meissner B, et al. Creutzfeldt-Jakob disease in Germany: a prospective 12-year surveillance. *Brain*. 2007;130:1350-1359.
2. European Centre for Disease Prevention and Control. EU case definition: sporadic CJD (from 1 January 2017) [online]. 2017. Accessed January 14, 2022. https://www.ecdc.europa.eu/en/infectious-diseases-public-health/variant-creutzfeldt-jakob-disease/eu-case-definition.
3. Muayqil T, Gronseth G, Camicioli R. Evidence-based guideline: diagnostic accuracy of CSF 14-3-3 protein in sporadic Creutzfeldt-Jakob disease. Report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology*. 2012;79:1499-1506.
4. Behaeghe O, Mangelschots E, De Vil B, Cras P. A systematic review comparing the diagnostic value of 14-3-3 protein in the cerebrospinal fluid, RT-QuIC and RT-QuIC on nasal brushing in sporadic Creutzfeldt-Jakob disease. *Acta Neurol Belg*. 2018;118:395-403.
5. Zerr I, Kallenberg K, Summers DM, et al. Updated clinical diagnostic criteria for sporadic Creutzfeldt-Jakob disease. *Brain*. 2009;132:2659-2668.
6. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. 1999;282:1061.
7. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PMM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy. *Ann Intern Med*. 2004;140:189.
8. Rutjes AWS. Evidence of bias and variation in diagnostic accuracy studies. *Can Med Assoc J*. 2006;174:469-476.
9. Ochodo EA, de Haan MC, Reitsma JB, Hooft L, Bossuyt PM, Leeflang MMG. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of "spin". *Radiology*. 2013;267:581-588.
10. Karch A, Koch A, Zapf A, Zerr I, Karch A. Partial verification bias and incorporation bias affected accuracy estimates of diagnostic studies for biomarkers that were part of an existing composite gold standard. *J Clin Epidemiol*. 2016;78:73-82.
11. Otto M, Wiltfang J, Cepek L, et al. Tau protein and 14-3-3 protein in the differential diagnosis of Creutzfeldt-Jakob disease. *Neurology*. 2002;58:192-197.
12. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529-536.
13. Hamlin C, Puoti G, Berri S, et al. A comparison of tau and 14-3-3 protein in the diagnosis of Creutzfeldt-Jakob disease. *Neurology*. 2012;79:547-552.

14. R Core Team. R: *A Language and Environment for Statistical Computing*. R Found. Stat. Comput.; 2019.
15. Zapf A, Castell S, Morawietz L, Karch A. Measuring inter-rater reliability for nominal data – which coefficients and confidence intervals are appropriate? *BMC Med Res Methodol*. 2016;16:93.
16. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. 2005;58:882-893.
17. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol*. 2006;59:1331-1332.
18. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67:1–48.
19. Vogelgesang F, Schlattmann P, Dewey M. The evaluation of bivariate mixed models in meta-analyses of diagnostic accuracy studies with SAS, stata and R. *Methods Inf Med*. 2018;57:111-119.
20. Nyaga VN, Arbyn M, Aerts M. Beta-binomial analysis of variance model for network meta-analysis of diagnostic test accuracy data. *Stat Methods Med Res*. 2018;27:2554-2566.
21. Carpenter B, Gelman A, Hoffman MD, et al. Stan: a probabilistic programming language. *J Stat Softw*. 2017;76.
22. Stan Development Team. RStan: the R interface to Stan [online]. 2019. Accessed January 14, 2022. http://mc-stan.org/
23. Abu-Rumeileh S, Baiardi S, Polischi B, et al. Diagnostic value of surrogate CSF biomarkers for Creutzfeldt-Jakob disease in the era of RT-QuIC. *J Neurol*. 2019;266:3136-3143.
24. Leitão MJ, Baldeiras I, Almeida MR, et al. Sporadic Creutzfeldt-Jakob disease diagnostic accuracy is improved by a new CSF ELISA 14-3-3γ assay. *Neuroscience*. 2016;322:398-407.
25. Rhoads DD, Wrona A, Foutz A, et al. Diagnosis of prion diseases by RT-QuIC results in improved surveillance. *Neurology*. 2020;95:e1017-e1026.
26. Sanchez-Juan P, Green A, Ladogana A, et al. CSF tests in the differential diagnosis of Creutzfeldt-Jakob disease. *Neurology*. 2006;67:637-643.
27. Simon SLR, Peterson A, Phillipson C, et al. Prospective study demonstrates utility of EP-QuIC in Creutzfeldt-Jakob disease diagnoses. *Can J Neurol Sci*. 2021;48:127-129.
28. Van Everbroeck B, Quoilin S, Boons J, Martin JJ, Cras P. A prospective study of CSF markers in 250 patients with possible Creutzfeldt-Jakob disease. *J Neurol Neurosurg Psychiatry*. 2003;74:1210-1214.
29. Zerr I, Bodemer M, Gefeller O, et al. Detection of 14-3-3 protein in the cerebrospinal fluid supports the diagnosis of Creutzfeldt-Jakob disease. *Ann Neurol*. 1998;43:32-40.
30. Baldeiras IE, Ribeiro MH, Pacheco P, et al. Diagnostic value of CSF protein profile in a Portuguese population of sCJD patients. *J Neurol*. 2009;256:1540-1550.
31. Bizzi A, Pascuzzo R, Blevins J, et al. Evaluation of a new criterion for detecting prion disease with diffusion magnetic resonance imaging. *JAMA Neurol*. 2020;77:1141.
32. Bongianni M, Orrù C, Groveman BR, et al. Diagnosis of human prion disease using real-time quaking-induced conversion testing of olfactory mucosa and cerebrospinal fluid samples. *JAMA Neurol*. 2017;74:155.
33. Chohan G, Pennington C, Mackenzie JM, et al. The role of cerebrospinal fluid 14-3-3 and other proteins in the diagnosis of sporadic Creutzfeldt-Jakob disease in the UK: a 10-year review. *J Neurol Neurosurg Psychiatry*. 2010;81:1243-1248.
34. Cuadrado-Corrales N, Jiménez-Huete A, Albo C, et al. Impact of the clinical context on the 14-3-3 test for the diagnosis of sporadic CJD. *BMC Neurol*. 2006;6:25.
35. Fiorini M, Iselle G, Perra D, et al. High diagnostic accuracy of RT-QuIC assay in a prospective study of patients with suspected sCJD. *Int J Mol Sci*. 2020;21:1-10.
36. Franceschini A, Baiardi S, Hughson AG, et al. High diagnostic value of second generation CSF RT-QuIC across the wide spectrum of CJD prions. *Sci Rep*. 2017;7:10655.
37. Lattanzio F, Abu-Rumeileh S, Franceschini A, et al. Prion-specific and surrogate CSF biomarkers in Creutzfeldt-Jakob disease: diagnostic accuracy in relation to molecular subtypes and analysis of neuropathological correlates of p-tau and Aβ42 levels. *Acta Neuropathol*. 2017;133:559-578.
38. Rudge P, Hyare H, Green A, Collinge J, Mead S. Imaging and CSF analyses effectively distinguish CJD from its mimics. *J Neurol Neurosurg Psychiatry*. 2018;89:461-466.
39. Wang LH, Bucelli RC, Patrick E, et al. Role of magnetic resonance imaging, cerebrospinal fluid, and electroencephalogram in diagnosis of sporadic Creutzfeldt-Jakob disease. *J Neurol*. 2013;260:498-506.
40. Holtman GA, Berger MY, Burger H, et al. Development of practical recommendations for diagnostic accuracy studies in low-prevalence situations. *J Clin Epidemiol*. 2019;114:38-48.
41. European Medicines Agency. Clinical evaluation of diagnostic agents. Doc. Ref. CPMP/EWP/1119/98/Rev. 2010.
42. U.S. Food & Drug Administration. Statistical guidance on reporting results from studies evaluating diagnostic tests - guidance for industry and FDA staff. FDA-2020-D-0957 2007.
43. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C eds. *Cochrane Handbook for Systemic Reviews of Diagnostic Test Accuracy Version 1.0.0*. The Cochrane Collaboration; 2009.
44. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
45. Lantero Rodriguez J, Karikari TK, Suárez-Calvet M, et al. Plasma p-tau181 accurately predicts Alzheimer's disease pathology at least 8 years prior to post-mortem and improves the clinical characterisation of cognitive decline. *Acta Neuropathol*. 2020;140:267-278.
46. Zerr I, Budka H, Kallenberg K, Steinhoff BJ, Weber JR, Sturzenegger M. Creutzfeldt-Jakob-Krankheit, S1-Leitlinie, 2018; in: Deutsche Gesellschaft für Neurologie (Hrsg.), Leitlinien für Diagnostik und Therapie in der Neurologie. Accessed January 14, 2022. https://dgn.org/leitlinien/
47. Hermann P, Laux M, Glatzel M, et al. Validation and utilization of amended diagnostic criteria in Creutzfeldt-Jakob disease surveillance. *Neurology*. 2018;91:e331-e338.
48. Gerke O, Høilund-Carlsen PF, Vach W. Analyzing paired diagnostic studies by estimating the expected benefit. *Biometrical J*. 2015;57:395-409.
49. Lewczuk P, Riederer P, O'Bryant SE, et al. Cerebrospinal fluid and blood biomarkers for neurodegenerative dementias: an update of the Consensus of the Task Force on Biological Markers in Psychiatry of the World Federation of Societies of Biological Psychiatry. *World J Biol Psychiatry*. 2018;19:244-328.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Rübsamen N, Pape S, Konigorski S, Zapf A, Rücker G, Karch A. Diagnostic accuracy of cerebrospinal fluid biomarkers for the differential diagnosis of sporadic Creutzfeldt–Jakob disease: a (network) meta-analysis. *Eur J Neurol*. 2022;29:1366–1376. doi:10.1111/ene.15258

MANAGE-PD

Tool for Making Informed Decisions to **Aid Timely Management of Parkinson's Disease**

MANAGE-PD allows you to:

- Identify PD patients inadequately controlled on oral medications
- Determine which patients with PD may be adequately controlled on their current treatment regimen or may require changes to their treatment regimen

Scan the QR code to access to the web

Click here to access to the web

MANAGE-PD is an AbbVie Inc. registered Medical Device. It is a collaborative research and development effort between AbbVie Medical Affairs and Health Economics and Outcomes, the Parkinson's Foundation and an international panel of Movement Disorder Specialists.

PD: Parkinson´s Disease

Parkinson's Foundation

Veeva Code: ABBV-AA-00335-FM Approved: January 2022

abbvie