



Review Article

Diagnostic Accuracy and Feasibility of Brief Cognitive Screening in Adults with Type 2 Diabetes: A Systematic Review

Sharmila Bishowkarma^{1,*} , Gagan Karki² , Ravi Rokaya³ , Bibek Shrestha⁴

¹Department of Endocrinology, The First Affiliated Hospital of Xuzhou Medical University, Xuzhou, China

²Department of Orthopedic Surgery, The First Affiliated Hospital of Xuzhou Medical University, Xuzhou, China

³Department of Cardiology, Chitwan Medical College Teaching Hospital, Bharatpur, Nepal

⁴Department of Ultrasound Radiology, Zhongnan Hospital of Wuhan University, Wuhan, China

Abstract

Background Cognitive impairment and dementia are more common in people with type 2 diabetes yet often go undetected. Brief cognitive screening tests could enable earlier identification in routine care, but their diagnostic accuracy and practicality in this population remain uncertain. We systematically reviewed the diagnostic accuracy and feasibility of brief cognitive screening in this population. **Methods** The review was registered (PROSPERO CRD420261441730) and reported following the PRISMA-DTA guideline. We searched MEDLINE (via PubMed), Embase, CENTRAL and PsycINFO from January 2000 to April 2026, restricted to English-language reports, for studies evaluating a brief cognitive screen against an acceptable reference standard clinical diagnostic criteria or a comprehensive neuropsychological battery or reporting feasibility. Two reviewers independently screened, extracted and appraised studies (QUADAS-2 for accuracy; the Mixed Methods Appraisal Tool for feasibility). Sensitivity and specificity with Wilson 95% confidence intervals were synthesised structurally using paired forest and summary receiver-operating-characteristic plots, and the certainty of evidence was assessed with GRADE. **Results** Six of 557 records were eligible: four diagnostic-accuracy studies (802 adults) and two feasibility studies. All accuracy studies evaluated the Montreal Cognitive Assessment (MoCA) and one also evaluated the Mini-Mental State Examination (MMSE). At the conventional cut-off (<26) the MoCA was highly sensitive but poorly specific (sensitivity 88-94%, specificity 29-51%); lower, diabetes-specific thresholds improved specificity at the expense of sensitivity (for example 49%/91% at 20/21, and 81%/62% at an optimal 24/25). The single MMSE study showed the reverse profile (66%/96%). Risk of bias was appreciable in every study (incorporation bias, spectrum effects, within-sample thresholds), and a meta-analysis was not appropriate. Two studies supported nurse-delivered screening in primary care (MoCA; and the Mini-Cog, mean 118 seconds, 100% completion). Overall certainty of evidence was low to very low. **Conclusions** The MoCA is the only brief cognitive screen with replicated accuracy evidence in type 2 diabetes: sensitive but over-referring at the standard cut-off, with a lower threshold better balancing accuracy. Non-specialist screening appears feasible, but the accuracy evidence is limited and at material risk of bias. Diabetes-specific validation against clinical or neuropsychological reference standards is the priority for future research.

*Correspondence: Sharmila Bishowkarma (sharmilabishowkarma52@gmail.com)

Received: 15 September 2026; **Accepted:** 27 September 2026; **Published:** 29 October 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Keywords

Type 2 Diabetes Mellitus, Cognitive Impairment, Montreal Cognitive Assessment, Diagnostic Test Accuracy, Cognitive Screening

1. Background

Type 2 diabetes affects hundreds of millions of adults worldwide, and its prevalence continues to rise [1]. Alongside its well-recognised vascular complications, type 2 diabetes is now firmly established as a risk factor for cognitive impairment: compared with people without diabetes, those with type 2 diabetes have approximately a 1.5-fold higher risk of dementia and a substantial excess of milder deficits such as mild cognitive impairment (MCI) [2-6]. This matters clinically because cognitive impairment undermines the demanding self-management that diabetes requires taking medications reliably, monitoring glucose, recognising and treating hypoglycaemia, and attending appointments and is associated with poorer glycaemic control, more frequent hypoglycaemia, and higher rates of complications, hospitalisation and death [7-10]. Much of this impairment nonetheless goes unrecognised in routine diabetes care.

Because early identification allows treatment to be simplified and support to be arranged, clinical guidelines including the American Diabetes Association's Standards of Care recommend screening older adults with diabetes for cognitive impairment [11-13], although adherence to guideline-recommended diabetes care varies between routine hospital and specialist settings [14]. Brief cognitive screening instruments such as the Montreal Cognitive Assessment (MoCA) [15], the Mini-Mental State Examination (MMSE) [16] and the Mini-Cog [17] are well suited to this task: they are short, inexpensive, require little training, and can be administered by non-specialists in primary or community care, making them plausible tools for screening at the population level.

The performance of these instruments, however, was largely established in general older or memory-clinic populations, and it may not transfer directly to people with type 2 diabetes. Diabetes-related factors cerebrovascular disease, depression, lower educational attainment, and sensory or motor impairment that can affect individual test items may alter the balance of sensitivity and specificity and the optimal cut-off [18, 19]. The threshold is decisive for any screening test: set too high, it over-refers and overwhelms diagnostic services; set too low, it misses cases. Despite a growing number of studies that apply brief screens in diabetes, it remains unclear which instrument, at what threshold, accurately detects cognitive impairment in this population, and whether such screening is feasible in everyday care.

We therefore undertook a systematic review, reported in accordance with the PRISMA-DTA guideline, of the diagnostic

accuracy and feasibility of brief cognitive screening instruments for detecting cognitive impairment in adults with type 2 diabetes, in order to inform whether and how such screening might be used at the population level.

2. Methods

2.1. Protocol, Registration and Reporting

The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Diagnostic Test Accuracy studies (PRISMA-DTA) [20, 21], and the protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO CRD420261441730; registered 10 July 2026, after the database searches had been run). The review had two objectives: to determine the diagnostic accuracy of brief cognitive screening instruments for detecting cognitive impairment in adults with type 2 diabetes, and to summarise the feasibility of delivering such screening in routine, non-specialist care. Departures from the protocol are noted in the relevant sections below.

2.2. Eligibility Criteria

Studies were eligible if they enrolled adults (aged ≥ 18 years) with type 2 diabetes and evaluated a brief cognitive screening instrument defined as a standardised test that can be administered in a routine or non-specialist setting in approximately fifteen minutes or less, for example the Montreal Cognitive Assessment (MoCA), the Mini-Mental State Examination (MMSE), the Mini-Cog, clock-drawing, the Rowland Universal Dementia Assessment Scale (RUDAS) [22], the General Practitioner Assessment of Cognition (GPCOG) [23], the AD8 [24], or the Addenbrooke's Cognitive Examination. For the diagnostic-accuracy question the index test had to be compared against an acceptable reference standard for cognitive impairment a formal clinical diagnosis using DSM or ICD criteria, the Petersen or NIA-AA criteria [25, 26], an expert consensus panel, or a comprehensive neuropsychological battery and the study had to report, or permit reconstruction of, the numbers of true positives, false positives, false negatives and true negatives at one or more thresholds. The target condition

was cognitive impairment, encompassing mild cognitive impairment (MCI) and dementia. Cross-sectional and cohort diagnostic-accuracy designs were eligible.

For the feasibility question, studies reporting the practicality of brief cognitive screening in type 2 diabetes administration time, completion, acceptability, or deliverability by non-specialists were eligible irrespective of whether a reference standard was applied. Only reports published in English were included.

Studies were excluded when a screening instrument had been applied at a cut-off to estimate the prevalence of, or risk factors for, cognitive impairment without an independent reference standard, so that the screen defined the condition rather than being tested against it; when the reference standard was itself another brief screen (a screen-versus-screen comparison); when the index test fell outside the definition above; and when the report was a conference abstract without a full paper or was not a primary study. Systematic reviews were not eligible as primary studies but were retained for reference-list screening.

2.3. Information Sources and Search Strategy

We searched four electronic databases MEDLINE (via PubMed), Embase, the Cochrane Central Register of Controlled Trials (CENTRAL) and PsycINFO from 1 January 2000 to 30 April 2026, restricted to English-language reports. The search combined two concept blocks, type 2 diabetes AND cognition/brief cognitive screening, using controlled-vocabulary terms (MeSH, Emtree and APA thesaurus descriptors) and free-text terms. The full, reproducible strategy for each database, together with a CINAHL strategy that was developed but not run, is provided in Additional file 1. The database searches were supplemented by screening the reference lists of all included studies and of relevant systematic reviews identified during screening. Records were collated and de-duplicated before screening.

2.4. Study Selection

Two reviewers independently screened titles and abstracts against the eligibility criteria and then assessed the full texts of potentially eligible reports; disagreements were resolved by discussion and, where necessary, by a third reviewer. Reports excluded at the full-text stage are listed with their reasons in Additional file 2, and the flow of records through the review is shown in Figure 1.

2.5. Data Extraction

Data were extracted into a piloted extraction form. For each study we recorded the country and income setting, the design and clinical setting, the number and characteristics of participants (Table 1), the index test and the threshold(s) evaluated, the reference standard and the definition of the target condition, and the numbers of true positives, false positives, false

negatives and true negatives at every reported threshold. Where a 2×2 table was not reported directly it was reconstructed from the reported sensitivity and specificity and the sizes of the diseased and non-diseased groups; where reconstruction was not possible the cell counts were recorded as not reconstructable. Predictive values, likelihood ratios and the area under the curve were taken as reported. Extraction was checked in duplicate and against an independent re-extraction of each study contributing to the synthesis, with discrepancies resolved by returning to the primary source. One study (Alagiakrishnan 2013) reported diagnostic figures that could not be reconciled with its own data no single 2×2 table reproduced the reported sensitivity, specificity and predictive values, and the reported sensitivity conflicted with the study's own table so the internally consistent values were used and the study was treated as supportive only.

2.6. Risk of Bias and Methodological Quality

Risk of bias and applicability were assessed for the diagnostic-accuracy studies with the QUADAS-2 tool [27], across its four risk-of-bias domains (patient selection, index test, reference standard, and flow and timing) and three applicability domains, each rated as low, unclear or high. The two feasibility studies, which do not evaluate accuracy against a reference standard, were appraised with the Mixed Methods Appraisal Tool (MMAT) [28]. Appraisal was undertaken independently by two reviewers, with disagreements resolved by discussion. The protocol also specified QUADAS-C for any formal head-to-head comparative-accuracy synthesis; because no such synthesis was undertaken, each index test was appraised with QUADAS-2 alone. Domain-level judgements with the support for each are provided in Additional file 3 and summarised in Figure 2.

2.7. Measures of Diagnostic Accuracy

The primary measures of accuracy were sensitivity and specificity, calculated from the 2×2 table at each threshold with 95% confidence intervals estimated by the Wilson score method [29]. Positive and negative predictive values and positive and negative likelihood ratios were derived from the same tables, and the area under the receiver-operating-characteristic curve was reported where available. A summary of accuracy at the near-standard threshold for each study is given in Table 2, and the full per-threshold data set for every reported cut-off is provided in Additional file 4.

2.8. Data Synthesis

Diagnostic accuracy was synthesised structurally and narratively. For each instrument we present paired sensitivity and specificity at the near-standard threshold in a forest plot (Figure 3) and the trade-off across all reported thresholds in a summary receiver-operating-characteristic (SROC) plot (Figure 4). A quantitative meta-analysis (a bivariate or hierarchical

SROC model) [30, 31] was pre-specified only where at least four studies evaluated the same instrument at a comparable threshold. This condition was not met no single threshold was shared by four studies — and, in addition, between-study heterogeneity was substantial and attributable to identifiable sources (incorporation bias, spectrum differences, and thresholds selected within the same sample); pooling was therefore judged inappropriate, and accuracy is reported per study with 95% confidence intervals. The subgroup analyses and meta-regression specified in the protocol (by instrument and threshold, reference-standard type, care setting, country income group, and diabetes duration or glycaemic control) were not feasible given the small number of studies, and these factors are considered narratively instead. Feasibility outcomes were

synthesised narratively. Figures were produced in R (version 4.3) with the ggplot2 package [32, 33].

2.9. Certainty of the Evidence

The certainty of the accuracy evidence was assessed with the GRADE approach for diagnostic tests [34, 35], considering risk of bias, indirectness, inconsistency, imprecision and publication bias, and rated as high, moderate, low or very low. The judgements are presented as a summary-of-findings table (Table 3), which expresses the expected consequences of screening per 1,000 people screened at an illustrative prevalence.

3. Results

3.1. Study Selection

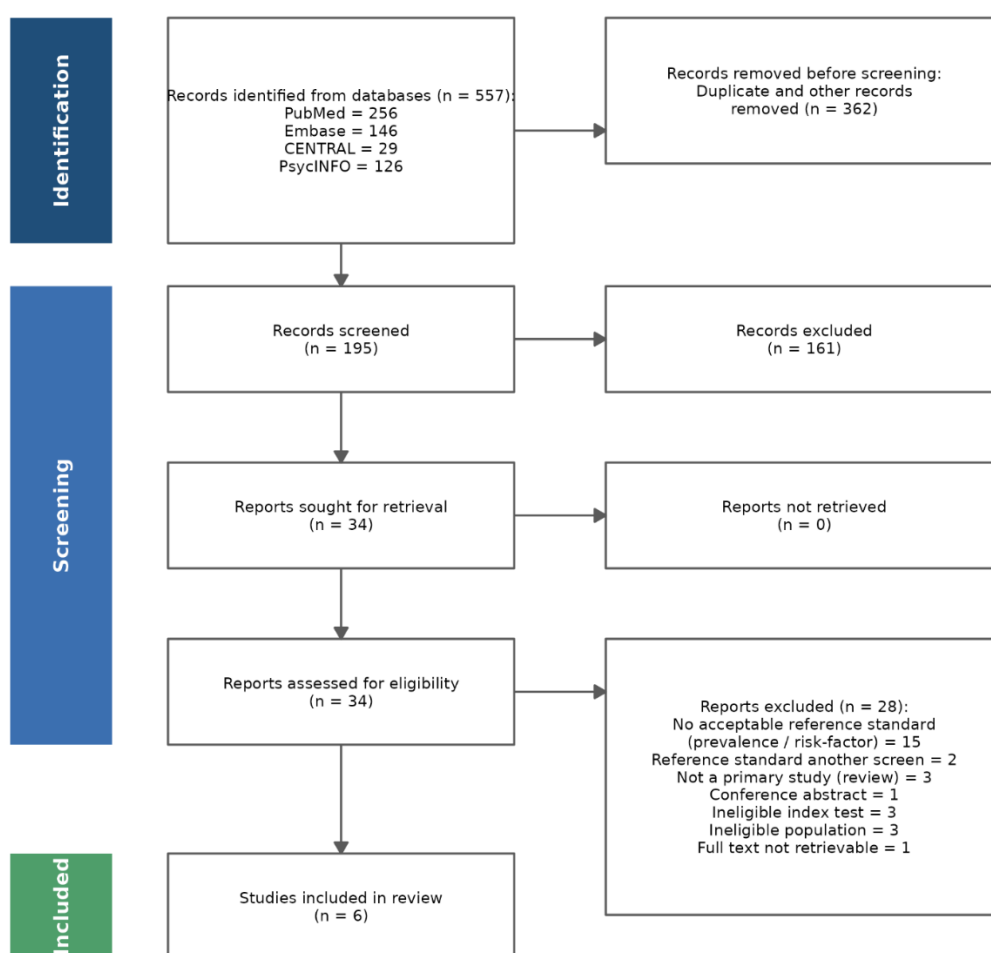


Figure 1. PRISMA 2020 flow diagram of study identification, screening and inclusion.

The database searches identified 557 records (PubMed, n = 256; Embase, n = 146; CENTRAL, n = 29; PsycINFO, n =

126); the full search strategies are provided in Additional file 1. After 362 duplicate and other records were removed before

screening, 195 records were screened on title and abstract, of which 163 were excluded. Thirty-two full-text reports were assessed for eligibility and 26 were excluded most because a cognitive screening instrument had been applied at a cut-off to estimate the prevalence of, or risk factors for, cognitive impairment without an independent reference standard (n = 16); the excluded studies and their reasons are listed in Additional file 2. Six studies met the eligibility criteria: four provided diagnostic-accuracy data (Gupta 2024 [36]; Popović 2025 [37]; Alagiakrishnan 2013 [38]; Wang 2025 [39]) and two provided feasibility data Merzon 2024 [40]; Sinclair 2013 [41]). The study-selection process is summarised in Figure 1.

3.2. Characteristics of Included Studies

The four diagnostic-accuracy studies enrolled 802 adults with type 2 diabetes (range 30-313 participants per study) and

were conducted in India, seven European countries, Canada and China. All four evaluated the Montreal Cognitive Assessment (MoCA), and one (Wang 2025) additionally evaluated the Mini-Mental State Examination (MMSE). The reference standard was a comprehensive neuropsychological battery in two studies (Gupta 2024; Popović 2025) and recognised clinical criteria in the other two (Petersen criteria, Wang 2025; modified Petersen/European-Consortium criteria, Alagiakrishnan 2013). The target condition was mild cognitive impairment (MCI) in all four studies, with a prevalence of 39-59%, and all were conducted in specialist, hospital or secondary-care settings. The two feasibility studies were undertaken in primary care and evaluated nurse-administered screening the MoCA (Merzon 2024; n = 350) and the Mini-Cog (Sinclair 2013; n = 201). Full characteristics of the included studies are presented in Table 1.

Table 1. Characteristics of the included studies.

Study	Country (income)	Design; setting	N	Age, y (mean ± SD)	Female,%	Index test	Reference standard	MCI,%
Gupta 2024	India (LMIC)	Cross-sectional; tertiary OPD	267	65.2 ± 4.2	31.5	MoCA	Neuropsychological battery (ICMR-NCTB) + actuarial criteria	39.0
Popović 2025	7 European countries (HIC)	Cross-sectional; secondary care	313	73 ± 5.4	35	MoCA	Neuropsychological battery (NTB) + modified CDR	59.1
Alagiakrishnan 2013	Canada (HIC)	Cross-sectional pilot	30	59.9 ± 7.1	47	MoCA, SMMSE	Modified Petersen / European-Consortium criteria	50.0
Wang 2025	China (UMIC)	Cross-sectional; hospital	192	68.5 ± 5.5	60.9	MoCA, MMSE	Petersen criteria	49.0
Merzon 2024 ^a	Israel (HIC)	Cohort; primary care	350	73.8 ± 5.8	44.9	MoCA (nurse-administered)	None (feasibility)	19.4 ^b
Sinclair 2013 ^a	UK (HIC)	Cross-sectional; primary care	201	72 (median)	39.3	Mini-Cog (nurse-administered)	MMSE <24 ^c	17.4 ^b

- 1) Feasibility study (secondary objective); contributes to the feasibility synthesis only.
- 2) Proportion screening positive (Merzon, MoCA <19; Sinclair, Mini-Cog <4), not MCI prevalence against a reference standard.
- 3) The MMSE is another brief screen, not an acceptable reference standard, so Sinclair 2013 does not enter the accuracy synthesis.

3.3. Methodological Quality

Risk of bias, assessed with QUADAS-2, was appreciable across all four accuracy studies (Figure 2; domain-level judgements with the support for each are given in Additional file 3). The reference-standard domain was at low risk of bias in Gupta 2024 and Popović 2025, each of which used a comprehensive neuropsychological battery, and unclear in Wang 2025, which applied the Petersen criteria but under-described

the confirmatory neuropsychological testing; in Wang 2025 the MoCA was explicitly not used to establish the diagnosis, thereby avoiding incorporation bias. Concerns for patient selection arose in all studies from spectrum effects, including the exclusion of participants with established dementia and, in Wang 2025, a substantial between-group difference in educational attainment. The index-test domain was rated unclear in three studies because blinding to the reference standard was

not reported and the optimal thresholds were derived within the same sample. Alagiakrishnan 2013 was at high risk of bias across all four domains, principally because the MoCA and standardised MMSE were incorporated within the reference-standard definition, and its reported accuracy statistics were internally inconsistent; it is therefore considered supportive only.

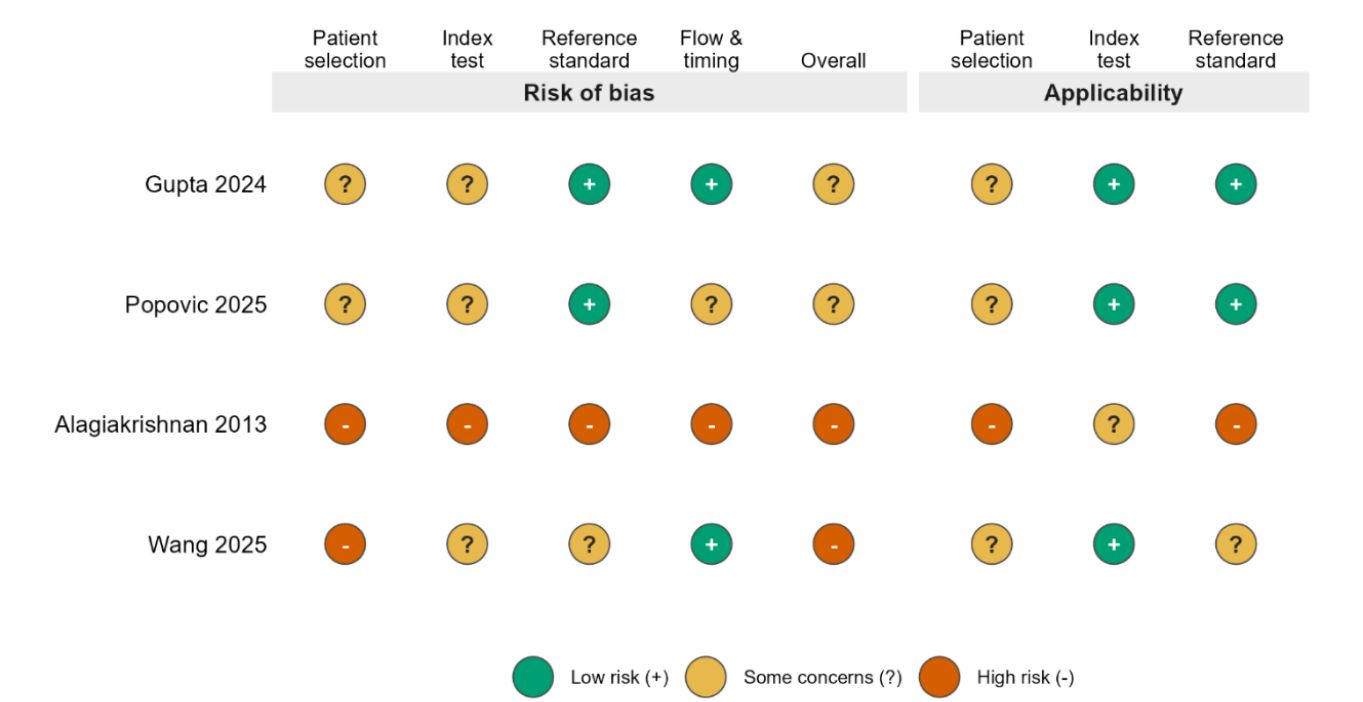


Figure 2. QUADAS-2 risk-of-bias and applicability summary for the four diagnostic-accuracy studies.

3.4. Diagnostic Accuracy

Montreal Cognitive Assessment. At the conventional cut-off (<26), the MoCA was highly sensitive but poorly specific: sensitivity was 94.2% (95% CI 88.0-97.3) with specificity 29.4% (23.0-36.9) in Gupta 2024, and sensitivity 88.1% (82.7-92.0) with specificity 50.8% (42.2-59.3) in Popović 2025. Lowering the threshold increased specificity at the expense of sensitivity: in Gupta 2024 a cut-off of 22/23 yielded 69.2%/71.8% and a cut-off of 20/21 yielded 49.0%/91.4%

with the highest overall accuracy (74.9%), while Popović 2025 identified 24/25 as optimal (81.1%/61.7%; area under the curve [AUC] 0.78). Wang 2025 reported the highest discrimination at 24/25 (95.7% [89.6-98.3]/78.6% [69.5-85.5]; AUC 0.95), albeit in a sample with a wide cognitive and educational spectrum, whereas Alagiakrishnan 2013 reported 46.7%/93.3% at ≤26 (AUC 0.70). The paired sensitivity and specificity estimates with 95% confidence intervals are shown in Figure 3, the threshold trade-off across all cut-offs in the summary ROC plot (Figure 4), and the complete per-threshold data in Additional file 4; a summary is given in Table 2.

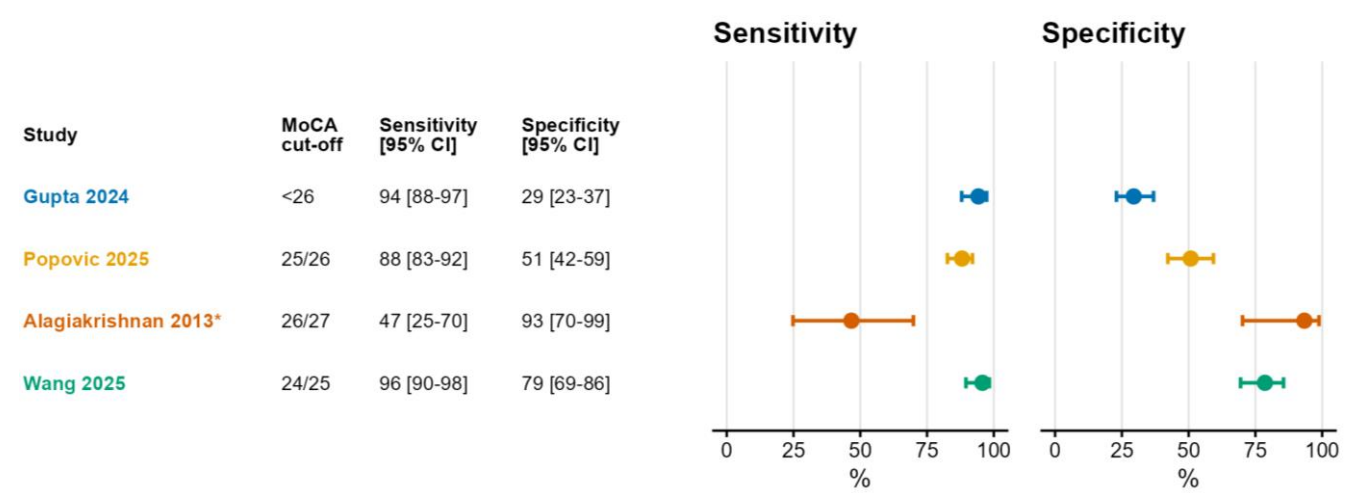


Figure 3. Paired sensitivity and specificity of the MoCA for detecting mild cognitive impairment in adults with type 2 diabetes, each study at its reported near-standard threshold (Wilson 95% confidence intervals). No pooled estimate is shown.

Mini-Mental State Examination. One study evaluated the MMSE against an acceptable reference standard: Wang 2025 reported a cut-off of 26/27 with sensitivity 66.0% (55.9-74.7) and specificity 95.9% (90.0-98.4) (AUC 0.84) a profile of higher specificity and lower sensitivity than the MoCA in the same participants (Table 2).

Table 2. Diagnostic accuracy of the MoCA and MMSE for detecting mild cognitive impairment.

Study	Cut-off	TP	FP	FN	TN	Sensitivity,% (95% CI)	Specificity,% (95% CI)	AUC
Gupta 2024	<26	98	115	6	48	94 (88-97)	29 (23-37)	NR
Gupta 2024	22/23	72	46	32	117	69 (60-77)	72 (64-78)	NR
Gupta 2024	20/21	51	14	53	149	49 (40-59)	91 (86-95)	NR
Popović 2025	25/26	163	63	22	65	88 (83-92)	51 (42-59)	0.78
Popović 2025	24/25	150	49	35	79	81 (75-86)	62 (53-70)	0.78
Alagiakrishnan 2013	26/27 ^a	7	1	8	14	47 (25-70)	93 (70-99)	0.70
Wang 2025 (MoCA)	24/25	90	21	4	77	96 (90-98)	79 (69-86)	0.95
Wang 2025 (MMSE)	26/27	62	4	32	94	66 (56-75)	96 (90-98)	0.84

Each index test was compared with an acceptable reference standard (Table 1); rows show the reported and/or optimal thresholds. Full per-threshold data are in Additional file 4. No pooled (meta-analytic) estimate was calculated. 95% CIs are Wilson intervals.

MoCA and reference standard were not independent (incorporation bias) and the reported sensitivity (67%) is internally inconsistent with the study's own data; the consistent value (47%) is shown. Cut-off 26/27 denotes scores of 26 or below classed as positive.

Because fewer than four studies evaluated any single instrument at a comparable threshold, and because between-study heterogeneity was substantial and attributable to identifiable sources (incorporation bias, spectrum differences and within-sample threshold selection), a bivariate/HSROC meta-analysis was not performed; accuracy is presented per study with 95% confidence intervals (Figures 3 and 4; Additional file 4). The certainty of the evidence for MoCA accuracy, assessed with the GRADE approach for diagnostic tests, was low to very low, downgraded for risk of bias, inconsistency and imprecision (Table 3).

Table 3. GRADE summary of findings.

Test (cut-off)	Studies (participants)	Sensitivity	Specificity	Certainty (GRADE)	Per 1,000 with MCI (n=400): TP (missed)	Per 1,000 without MCI (n=600): FP (cleared)
MoCA, <26 (standard)	2 (580)	88-94%	29-51%	Very low (⊕○○○) ^a	352-377 (23-48 missed)	295-424 (176-305 cleared)
MoCA, 22/23-24/25 (lower) ^b	2 (580)	69-81%	62-72%	Very low (⊕○○○) ^a	277-324 (76-123 missed)	169-230 (370-431 cleared)
MMSE, 26/27	1 (192)	66%	96%	Very low (⊕○○○) ^a	264 (136 missed)	25 (575 cleared)

Population: adults with type 2 diabetes; illustrative MCI prevalence 40%; reference standard: neuropsychological battery or clinical criteria; pooled analysis not performed.

- 1) Downgraded for risk of bias, inconsistency and imprecision. GRADE certainty: ⊕⊕⊕⊕ high, ⊕⊕⊕○ moderate, ⊕⊕○○ low, ⊕○○○ very low.
- 2) Range from Gupta 2024 (22/23) and Popović 2025 (24/25); Wang 2025 at 24/25 (96%/79%) is shown separately (wide spectrum). Alagiakrishnan 2013 is omitted from the ranges (high risk of bias).

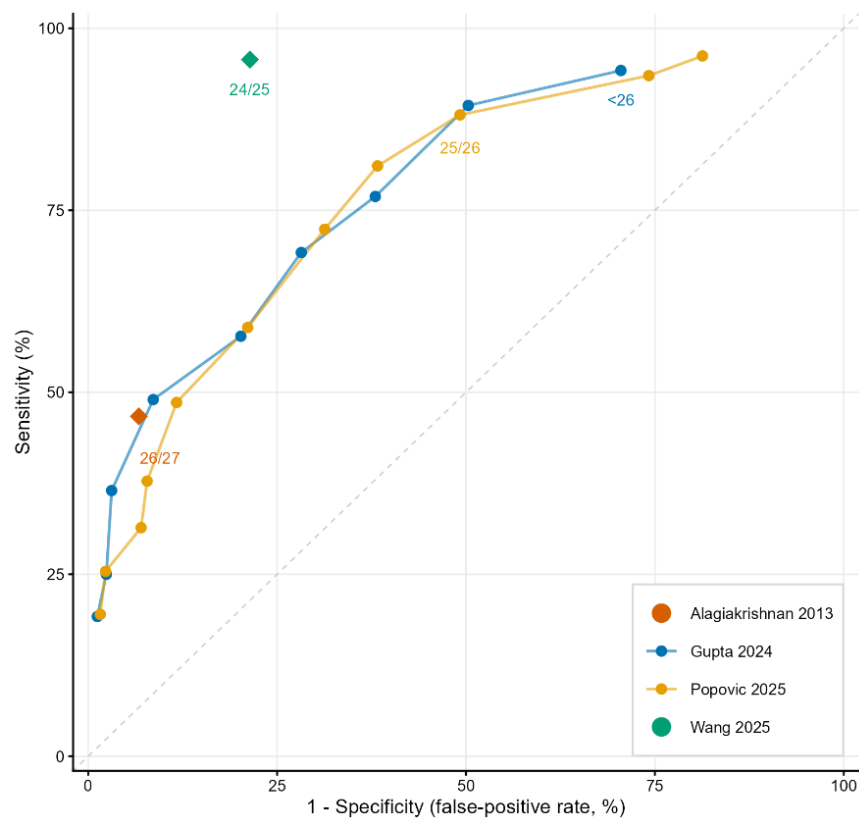


Figure 4. Summary receiver-operating-characteristic (SROC) plot of the MoCA. Connected lines show single studies across their cut-offs (Gupta 2024; Popović 2025); diamonds show single-threshold studies (Wang 2025; Alagiakrishnan 2013).

3.5. Feasibility

Two studies evaluated feasibility, and both supported the

delivery of brief cognitive screening by non-specialists in routine diabetes care. In an Israeli primary-care initiative, the MoCA was administered by trained nurse practitioners and identified previously undiagnosed severe impairment (MoCA

<19) in 19.4% of older adults with diabetes (Merzon 2024). The Mini-Cog was faster, requiring a mean of 118 seconds (SD 45) and never more than five minutes, with 100% completion and 99% agreement between nurse and expert scoring ($\kappa = 0.949$), and was readily integrated into the annual diabetes review (Sinclair 2013); because its accuracy was assessed only against the MMSE, it contributed to the feasibility rather than the accuracy synthesis. Consistent with these findings, the MoCA required approximately ten minutes to administer in the accuracy studies and was completed electronically by 99.3% of a large diabetes cohort.

4. Discussion

This systematic review set out to determine how accurately, and how practically, brief cognitive screening detects cognitive impairment in adults with type 2 diabetes. Despite screening two decades of literature, only six studies were eligible, and the diagnostic-accuracy evidence reduced to four studies all evaluating the Montreal Cognitive Assessment (MoCA), with a single study also reporting the Mini-Mental State Examination (MMSE). Across these studies the MoCA was consistently sensitive but poorly specific at its conventional cut-off (<26), and the choice of threshold, rather than the instrument itself, dominated performance: lowering the cut-off traded sensitivity for specificity in a predictable and reproducible way, with a lower, diabetes-specific threshold (around 21 to 24), giving the best overall balance. The one MMSE study showed the opposite profile specific but insensitive. Two feasibility studies indicated that nurse-delivered screening is quick and acceptable in routine diabetes care. The certainty of this evidence, however, was low to very low.

4.1. Comparison with Existing Evidence

The pattern we observed is consistent with what is known about these instruments in general older populations, where the MoCA is more sensitive than the MMSE at the standard cut-off but at the cost of specificity [42-46], and where lowering the MoCA threshold improves specificity in groups with lower education or greater vascular burden [18, 19, 47]. Our findings suggest that this trade-off is at least as pronounced in type 2 diabetes, a population enriched for cerebrovascular disease, depression [9, 48] and in several of the included cohorts lower educational attainment. The high sensitivity and low specificity of the MoCA at <26 mean that, applied unchanged as a population screen in diabetes, it would refer a large proportion of people without cognitive impairment for further assessment. This over-referral, rather than missed cases, is the principal accuracy concern, and it is the reason a lower, diabetes-specific cut-off has appeal.

4.2. Implications for Practice

Three points follow for clinicians and services. First, brief

cognitive screening is feasible in diabetes care: the MoCA can be delivered by trained nurses in about ten minutes and the Mini-Cog in under two, with near-complete completion so the practical barrier to screening is not administration but what happens next. Second, the standard MoCA cut-off is poorly suited to unselected screening in diabetes; where the MoCA is used, a lower threshold, or a two-stage pathway in which a positive screen is confirmed by fuller assessment [18], would limit unnecessary referrals. Third, any screening programme must be matched to downstream diagnostic and support capacity [49, 50], because the value of earlier detection depends on being able to act on it. These recommendations are necessarily provisional, given the weakness of the underlying evidence.

4.3. Strengths and Limitations

The review has several strengths: it was registered, reported to the PRISMA-DTA standard, and based on a comprehensive multi-database search; screening, extraction and appraisal were undertaken in duplicate; and every accuracy figure was subjected to an independent re-extraction against the primary source, which identified and corrected internally inconsistent data in one study (Alagiakrishnan 2013). Risk of bias was assessed with QUADAS-2 and certainty with GRADE. The main limitations lie in the evidence rather than the methods. Only four accuracy studies were available, all evaluating a single instrument; several were small; and all carried appreciable risk of bias including incorporation bias where the screen formed part of the reference standard, spectrum and education imbalances that inflate apparent accuracy, and “optimal” thresholds selected within the same sample. Only one study reported the MMSE, and no eligible accuracy data were found for other brief or diabetes-specific instruments. Because fewer than four studies shared a common threshold and heterogeneity was substantial and explicable, we did not pool estimates, and a formal assessment of publication bias was likewise not possible. Finally, we restricted inclusion to English-language reports, which may have excluded relevant studies, particularly from regions where much diabetes-and-cognition research is published in other languages.

4.4. Implications for Research

The clear priority is adequately powered, diabetes-specific diagnostic-accuracy studies that pre-specify their thresholds, blind the index test to an independent reference standard (a formal clinical diagnosis or a full neuropsychological battery), and report complete 2×2 data in line with the STARD guideline [51]. External validation of any diabetes-specific MoCA cut-off is needed before it can be recommended. Beyond the MoCA, the accuracy of the Mini-Cog, RUDAS, GPCOG and emerging diabetes-specific tools (for example, the DCAT and SCIP-D) against acceptable reference standards remains essentially untested in this population and is an important gap [52-55].

5. Conclusions

The MoCA is currently the only brief cognitive screen with replicated diagnostic-accuracy evidence in adults with type 2 diabetes. It is sensitive but, at the conventional cut-off, over-refers, and a lower, diabetes-specific threshold appears to balance sensitivity and specificity better. Screening is feasible in

non-specialist diabetes care, but the accuracy evidence base is small, dominated by one instrument, and at material risk of bias, so current practice recommendations can only be provisional. Well-designed, diabetes-specific validation studies with independent reference standards and pre-specified thresholds are needed before brief cognitive screening can be confidently adopted at the population level.

Abbreviations

AD8	Alzheimer's Disease-8 Informant Interview
AUC	Area Under the Receiver Operating Characteristic Curve
CDR	Clinical Dementia Rating
CI	Confidence Interval
GPCOG	General Practitioner Assessment of Cognition
HIC	High-income Country
ICMR-NCTB	Indian Council of Medical Research Neurocognitive Toolbox
LMIC	Low- and Middle-income Country
MCI	Mild Cognitive Impairment
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
NTB	Neuropsychological Test Battery
OPD	Outpatient Department
PRISMA-DTA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Diagnostic Test Accuracy Studies
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies 2
RUDAS	Rowland Universal Dementia Assessment Scale
SMMSE	Standardised Mini-Mental State Examination
SROC	Summary Receiver Operating Characteristic
T2DM	Type 2 Diabetes Mellitus
UMIC	Upper-middle-income Country

Author Contributions

Sharmila Bishowkarma: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Writing – original draft

Gagan Karki: Conceptualization, Methodology, Writing – review & editing

Ravi Rokaya: Formal Analysis, Validation

Bibek Shrestha: Investigation, Visualization

Supplementary Material

The supplementary material can be accessed at <https://doi.org/10.11648/j.ajim.20261405.11>

Data Availability Statement

All data generated or analysed during this study are included in this published article and its additional files.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract.* 2022; 183: 109119. <https://doi.org/10.1016/j.diabres.2021.109119>
- [2] Xue M, Xu W, Ou YN, et al. Diabetes mellitus and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 144 prospective studies. *Ageing Res Rev.* 2019; 55: 100944. <https://doi.org/10.1016/j.arr.2019.100944>
- [3] Chatterjee S, Peters SAE, Woodward M, et al. Type 2 diabetes as a risk factor for dementia in women compared with men: a pooled analysis of 2.3 million people comprising more than 100,000 cases of dementia. *Diabetes Care.* 2016; 39(2): 300-307. <https://doi.org/10.2337/dc15-1588>

- [4] Biessels GJ, Despa F. Cognitive decline and dementia in diabetes mellitus: mechanisms and clinical implications. *Nat Rev Endocrinol*. 2018; 14(10): 591-604. <https://doi.org/10.1038/s41574-018-0048-7>
- [5] Zilliox LA, Chadrasekaran K, Kwan JY, Russell JW. Diabetes and cognitive impairment. *Curr Diab Rep*. 2016; 16(9): 87. <https://doi.org/10.1007/s11892-016-0775-x>
- [6] Dove A, Shang Y, Xu W, et al. The impact of diabetes on cognitive impairment and its progression to dementia. *Alzheimers Dement*. 2021; 17(11): 1769-1778. <https://doi.org/10.1002/alz.12482>
- [7] Munshi MN. Cognitive dysfunction in older adults with diabetes: what a clinician needs to know. *Diabetes Care*. 2017; 40(4): 461-467. <https://doi.org/10.2337/dc16-1229>
- [8] Punthakee Z, Miller ME, Launer LJ, et al. Poor cognitive function and risk of severe hypoglycemia in type 2 diabetes: post hoc epidemiologic analysis of the ACCORD trial. *Diabetes Care*. 2012; 35(4): 787-793. <https://doi.org/10.2337/dc11-1855>
- [9] Srikanth V, Sinclair AJ, Hill-Briggs F, Moran C, Biessels GJ. Type 2 diabetes and cognitive dysfunction-towards effective management of both comorbidities. *Lancet Diabetes Endocrinol*. 2020; 8(6): 535-545. [https://doi.org/10.1016/S2213-8587\(20\)30118-2](https://doi.org/10.1016/S2213-8587(20)30118-2)
- [10] Bashir M, Ijaz S, Farooq U, Kahar A, Quan W. Diabetes educator-led education intervention to improve the medication adherence and quality of life in type 2 diabetes patients: a pre-and post-analysis. *Romanian Journal of Diabetes Nutrition and Metabolic Diseases*. 2025; 32(2): 119-127. Available from: <https://www.rjdnmd.org/index.php/RJDNMD/article/view/1816>
- [11] ElSayed NA, McCoy RG, Aleppo G, et al.; American Diabetes Association. 13. Older adults: Standards of Care in Diabetes-2025. *Diabetes Care*. 2025; 48(Suppl 1): S266-S282. <https://doi.org/10.2337/dc25-S013>
- [12] Biessels GJ, Verhagen C, Janssen J, et al. Cognitive dysfunction in diabetes: how to implement emerging guidelines. *Diabetologia*. 2019; 62(3): 380-392.
- [13] Rovner BW, Casten RJ, Chang AM, et al. Improving glycemic control in African Americans with diabetes and mild cognitive impairment. *J Am Geriatr Soc*. 2020; 68(5): 1015-1022. <https://doi.org/10.1111/jgs.16339>
- [14] Alhussain H, Priyatha V, Bashir MB, Ijaz S, Farooq U, et al. Analysing health professionals' adherence to national guidelines and comparing diabetes care in specialized care centres and hospitals. *American Journal of Internal Medicine*. 2024; 12(1): 1-10. <https://doi.org/10.11648/j.ajim.20241201.11>
- [15] Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005; 53(4): 695-699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- [16] Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975; 12(3): 189-198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- [17] Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The Mini-Cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *Int J Geriatr Psychiatry*. 2000; 15(11): 1021-1027. [https://doi.org/10.1002/1099-1166\(200011\)15:11<1021::aid-gps234>3.0.co;2-6](https://doi.org/10.1002/1099-1166(200011)15:11<1021::aid-gps234>3.0.co;2-6)
- [18] Thomann AE, Berres M, Goettel N, Steiner LA, Monsch AU. Enhanced diagnostic accuracy for neurocognitive disorders: a revised cut-off approach for the Montreal Cognitive Assessment. *Alzheimers Res Ther*. 2020; 12(1): 39. <https://doi.org/10.1186/s13195-020-00603-8>
- [19] Chen KL, Xu Y, Chu AQ, et al. Validation of the Chinese version of the Montreal Cognitive Assessment Basic for screening mild cognitive impairment. *J Am Geriatr Soc*. 2016; 64(12): e285-e290. <https://doi.org/10.1111/jgs.14530>
- [20] McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: the PRISMA-DTA statement. *JAMA*. 2018; 319(4): 388-396. <https://doi.org/10.1001/jama.2017.19163>
- [21] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021; 372: n71. <https://doi.org/10.1136/bmj.n71>
- [22] Storey JE, Rowland JT, Basic D, Conforti DA, Dickson HG. The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale. *Int Psychogeriatr*. 2004; 16(1): 13-31. <https://doi.org/10.1017/s1041610204000043>
- [23] Brodaty H, Pond D, Kemp NM, et al. The GPCOG: a new screening test for dementia designed for general practice. *J Am Geriatr Soc*. 2002; 50(3): 530-534. <https://doi.org/10.1046/j.1532-5415.2002.50122.x>
- [24] Galvin JE, Roe CM, Powlishtta KK, et al. The AD8: a brief informant interview to detect dementia. *Neurology*. 2005; 65(4): 559-564. <https://doi.org/10.1212/01.wnl.0000172958.95282.2a>
- [25] Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004; 256(3): 183-194. <https://doi.org/10.1111/j.1365-2796.2004.01388.x>
- [26] Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups. *Alzheimers Dement*. 2011; 7(3): 270-279. <https://doi.org/10.1016/j.jalz.2011.03.008>
- [27] 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(8): 529-536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- [28] Hong QN, Fabregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inf*. 2018; 34(4): 285-291. <https://doi.org/10.3233/EFI-180221>

- [29] Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc.* 1927; 22(158): 209-212. <https://doi.org/10.1080/01621459.1927.10502953>
- [30] Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol.* 2005; 58(10): 982-990. <https://doi.org/10.1016/j.jclinepi.2005.02.022>
- [31] Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: analysing and presenting results. In: *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*, version 1.0. The Cochrane Collaboration; 2010.
- [32] R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2023.
- [33] Wickham H. *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag; 2016. <https://doi.org/10.1007/978-3-319-24277-4>
- [34] Schunemann HJ, Oxman AD, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ.* 2008; 336(7653): 1106-1110. <https://doi.org/10.1136/bmj.39500.677199.AE>
- [35] Schunemann HJ, Mustafa RA, Brozek J, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *J Clin Epidemiol.* 2020; 122: 129-141. <https://doi.org/10.1016/j.jclinepi.2019.12.020>
- [36] Gupta A, Goyal A, Rajan R, Vishnu VY, Kalaivani M, Tandon N, et al. Validity of Montreal Cognitive Assessment to detect cognitive impairment in individuals with type 2 diabetes. *Diabetes Ther.* 2024; 15(5): 1155-1168. <https://doi.org/10.1007/s13300-024-01549-y>
- [37] Popović N, Lois N, Pérez-Hoyos S, Simó R, Exalto LG; RECOGNISED Consortium. Revisiting the Montreal Cognitive Assessment in a European cohort of elderly living with type 2 diabetes. *J Alzheimers Dis.* 2025; 104(2): 585-594. <https://doi.org/10.1177/13872877251318029>
- [38] Alagiakrishnan K, Zhao N, Mereu L, Senior P, Senthilselvan A. Montreal Cognitive Assessment is superior to Standardized Mini-Mental Status Exam in detecting mild cognitive impairment in the middle-aged and elderly patients with type 2 diabetes mellitus. *Biomed Res Int.* 2013; 2013: 186106. <https://doi.org/10.1155/2013/186106>
- [39] Wang N, Chen X. Diagnostic value of Montreal cognitive assessment combined with olfactory tests for mild cognitive impairment in older adult patients with type 2 diabetes. *Eur Geriatr Med.* 2025; 16: 1639-1646. <https://doi.org/10.1007/s41999-025-01223-x>
- [40] Merzon E, Shpigelman M, Vinker S, Golan Cohen A, Green I, Israel A, et al. Clinical characteristics and healthcare utilisation associated with undiagnosed cognitive impairment in elderly patients with diabetes in a primary care setting: a population-based cohort study. *BMJ Open.* 2024; 14(1): e078996. <https://doi.org/10.1136/bmjopen-2023-078996>
- [41] Sinclair AJ, Gadsby R, Hillson R, Forbes A, Bayer AJ. Brief report: use of the Mini-Cog as a screening tool for cognitive impairment in diabetes in primary care. *Diabetes Res Clin Pract.* 2013; 100(1): e23-e25. <https://doi.org/10.1016/j.diabres.2013.01.001>
- [42] Pinto TCC, Machado L, Bulgacov TM, et al. Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's disease (AD) in the elderly? *Int Psychogeriatr.* 2019; 31(4): 491-504. <https://doi.org/10.1017/S1041610218001370>
- [43] Roalf DR, Moberg PJ, Xie SX, Wolk DA, Moelter ST, Arnold SE. Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimers Dement.* 2013; 9(5): 529-537. <https://doi.org/10.1016/j.jalz.2012.10.001>
- [44] Tsoi KKF, Chan JYC, Hirai HW, Wong SYS, Kwok TCY. Cognitive tests to detect dementia: a systematic review and meta-analysis. *JAMA Intern Med.* 2015; 175(9): 1450-1458. <https://doi.org/10.1001/jamainternmed.2015.2152>
- [45] Davis DHJ, Creavin ST, Yip JLY, Noel-Storr AH, Brayne C, Cullum S. Montreal Cognitive Assessment for the detection of dementia. *Cochrane Database Syst Rev.* 2021; 7: CD010775. <https://doi.org/10.1002/14651858.CD010775.pub3>
- [46] Wang X, Li F, Zhu H, et al. A hierarchical Bayesian latent class model for the diagnostic performance of the Mini-Mental State Examination and Montreal Cognitive Assessment in screening mild cognitive impairment due to Alzheimer's disease. *J Prev Alzheimers Dis.* 2022; 9(4): 589-600. <https://doi.org/10.14283/jpad.2022.70>
- [47] Lee JY, Lee DW, Cho SJ, et al. Brief screening for mild cognitive impairment in elderly outpatient clinic: validation of the Korean version of the Montreal Cognitive Assessment. *J Geriatr Psychiatry Neurol.* 2008; 21(2): 104-110. <https://doi.org/10.1177/0891988708316855>
- [48] Ehtewish H, Arredouani A, El-Agnaf O. Diagnostic, prognostic, and mechanistic biomarkers of diabetes mellitus-associated cognitive decline. *Int J Mol Sci.* 2022; 23(11): 6144. <https://doi.org/10.3390/ijms23116144>
- [49] Janssen J, Koekkoek PS, van Charante EPM, et al. Depressive symptoms and quality of life after screening for cognitive impairment in patients with type 2 diabetes: observations from the Cog-ID cohort study. *BMJ Open.* 2019; 9(4): e024696. <https://doi.org/10.1136/bmjopen-2018-024696>
- [50] Lin JS, O'Connor E, Rossom RC, Perdue LA, Eckstrom E. Screening for cognitive impairment in older adults: a systematic review for the US Preventive Services Task Force. *Ann Intern Med.* 2013; 159(9): 601-612. <https://doi.org/10.7326/0003-4819-159-9-201311050-00730>
- [51] 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. <https://doi.org/10.1136/bmj.h5527>

- [52] Tayob MH, et al. Diagnostic accuracy of the Diabetes Cognitive Assessment Tool (DCAT) for cognitive impairment in type 2 diabetes. *BMC Geriatr.* 2026.
<https://doi.org/10.1186/s12877-026-07498-4>
- [53] Jensen NJ, Zarp J, Vikner M, et al. SCIP-a novel and efficient screening tool for detecting cognitive impairment among older adults with type 2 diabetes (abstract 1594-P). *Diabetes.* 2025; 74(Suppl 1).
<https://doi.org/10.2337/db25-1594-P>
- [54] Zhang J, et al. Cognitive risk stratification score in middle-aged and older adults with type 2 diabetes: a cross-sectional study. *J Clin Endocrinol Metab.* 2025; 110(10): e3313-e3321.
<https://doi.org/10.1210/clinem/dgaf063>
- [55] Valles-Salgado M, Gil-Moreno MJ, Curiel-Cid RE, et al. Comparison of the diagnostic accuracy of five cognitive screening tests for diagnosing mild cognitive impairment in patients consulting for memory loss. *J Clin Med.* 2024; 13(16): 4695.
<https://doi.org/10.3390/jcm13164695>