

## CRITICAL REVIEW

# Screening Tools for Cognitive Impairment in Adults with Substance Use Disorders: A Systematic Review

Katherine Y. Ko<sup>1</sup>, Nicole Ridley<sup>2</sup>, Shayden D. Bryce<sup>3,4</sup> , Kelly Allott<sup>3,4</sup>, Angela Smith<sup>5</sup> and Jody Kamminga<sup>6,\*</sup> 

<sup>1</sup>Department of Psychology, Macquarie University, Sydney, New South Wales, Australia

<sup>2</sup>Drug and Alcohol Services, South Eastern Sydney Local Health District, Sydney, New South Wales, Australia

<sup>3</sup>Centre for Youth Mental Health, The University of Melbourne, Parkville, Victoria, Australia

<sup>4</sup>Orygen, Parkville, Victoria, Australia

<sup>5</sup>Hunter New England Health Libraries, Hunter New England Local Health District, Newcastle, New South Wales, Australia

<sup>6</sup>Drug and Alcohol Clinical Services, Hunter New England Local Health District, Newcastle, New South Wales, Australia

(RECEIVED January 6, 2021; FINAL REVISION May 2, 2021; ACCEPTED July 6, 2021; FIRST PUBLISHED ONLINE August 26, 2021)

## ABSTRACT

**Objectives:** Cognitive impairment is common in individuals with substance use disorders (SUDs), yet no evidence-based guidelines exist regarding the most appropriate screening measure for use in this population. This systematic review aimed to (1) describe different cognitive screening measures used in adults with SUDs, (2) identify substance use populations and contexts these tools are utilised in, (3) review diagnostic accuracy of these screening measures versus an accepted objective reference standard, and (4) evaluate methodology of included studies for risk of bias. **Methods:** Online databases (PsycINFO, MEDLINE, Embase, and CINAHL) were searched for relevant studies according to pre-determined criteria, and risk of bias and applicability was assessed using the Quality Assessment of Diagnostic Accuracy Studies–2 (QUADAS–2). At each review phase, dual screening, extraction, and quality ratings were performed. **Results:** Fourteen studies met inclusion, identifying 10 unique cognitive screening tools. The Montreal Cognitive Assessment (MoCA) was the most common, and two novel screening tools (Brief Evaluation of Alcohol-Related Neuropsychological Impairments [BEARNI] and Brief Executive Function Assessment Tool [BEAT]) were specifically developed for use within SUD populations. Twelve studies reported on classification accuracy and relevant psychometric parameters (e.g., sensitivity and specificity). While several tools yielded *acceptable to outstanding* classification accuracy, there was poor adherence to the Standards for Reporting Diagnostic Accuracy Studies (STARD) across all studies, with high or unclear risk of methodological bias. **Conclusions:** While some screening tools exhibit promise for use within SUD populations, further evaluation with stronger methodological design and reporting is required. Clinical recommendations and future directions for research are discussed.

**Keywords:** Cognition, Screening, Adults, Substance-related disorders, Neuropsychology

Substance use disorder (SUD) is a pathological cluster of cognitive, behavioural, and physiological symptoms stemming from compulsive use of a substance despite adverse effects on psychosocial functioning (American Psychological Association, 2013). Between 30% and 80% of individuals presenting for the treatment of SUDs experience some form of cognitive impairment (CI; Copersino et al., 2009), with neurophysiological changes occurring as a result of intoxication, withdrawal, or chronic use. Substance users frequently

demonstrate deficits in attention, memory, executive functions, and decision-making (Ramey & Regier, 2020), with nature and chronicity of impairment differing between types of substances, quantity, and frequency of use (Bruijnen et al., 2019; Chen, Strain, Crum, & Mojtabai, 2013).

CI may also result from other comorbid factors. Psychiatric illness is highly prevalent in individuals with SUD (Gould, 2010; Kelly & Daley, 2013) and can increase severity of health issues and complicate recovery (Daley & Moss, 2002; Grella, Hser, Joshi, & Rounds-Bryant, 2001). Other common comorbidities that can impact cognition include physical illness, such as vascular conditions (Volkow, Baler, Compton, & Weiss, 2014), respiratory

\*Correspondence and reprint requests to: Jody Kamminga, Drug and Alcohol Clinical Services, Hunter New England Local Health District, Level 3, 670 Hunter Street, Newcastle West, NSW 2304, Australia. E-mail: Jody.Kamminga@health.nsw.gov.au

problems (Owen, Sutter & Albertson, 2014), and liver damage (Brodersen et al., 2014). Between 37% and 66% of individuals with an acquired brain injury (ABI, e.g., due to falls, head trauma, overdose, or chronic substance use; Ridley et al., 2018) misuse alcohol, and 10%–44% use illicit substances (Parry-Jones, Vaughan, & Miles Cox, 2006), suggesting potential cumulative cognitive dysfunction. Developmental disorders (e.g., intellectual or learning disability, attention-deficit hyperactivity disorder) also co-occur with SUDs, with rates varying (between 1% and 45%; Levin, 2007; Lin et al., 2016; Salavert et al., 2018). Many developmental disorders may remain undiagnosed at treatment (Gooden et al., 2020), and it is possible that similar factors that form barriers to diagnosis and support for individuals with developmental disorders (i.e., social and environmental disadvantages; Carroll Chapman & Wu, 2012) are also risk factors for developing SUDs.

Regardless of aetiology, CI in individuals with SUDs can severely impede everyday functioning. The degree of executive dysfunction could determine how effectively individuals with SUDs are able to cope with day-to-day demands (see Verdejo-Garcia, Garcia-Fernandez, & Dom, 2019). Impaired inhibitory control can contribute to behavioural difficulties, including increased impulsivity and poor decision-making (Ramey & Regier, 2020). CI also has impacts on the ability to engage successfully with SUD treatment (Bates, Bowden, & Barry, 2002), with individuals less likely to benefit from the therapeutic process or acquire effective strategies, related to diminished insight and motivation (Aharonovich, Amrhein, Bisaga, Nunes, & Hasin, 2008). CI is also associated with poorer treatment compliance (Copersino et al., 2012). However, individuals with CI can benefit from treatment, particularly when treatment is appropriately matched to cognitive level (Bates, Buckman, & Nguyen, 2013). For instance, cognitive-behavioural approaches are geared to enhance higher-order cognitive processes related to self-awareness and decision-making. These cognitive skills may be fundamentally reduced in individuals with CI (Hagen et al., 2016), who may benefit more from interactional programmes with reduced cognitive demands (Secades-Villa, García-Rodríguez, & Fernández-Hermida, 2015) or increased involvement of social supports (Buckman, Bates, & Cisler, 2007; Buckman, Bates, & Morgenstern, 2008).

Given the impact of CI on SUD treatment, accurate and timely detection of CI is essential. The wide range of factors contributing to CI in individuals presenting for SUD treatment, however, results in heterogeneous cognitive profiles that vary in their clinical presentation, making CI difficult to detect (Horner, Harvey, & Denier, 1999). Consequently, individuals with CI presenting for SUD treatment might not receive the most appropriate support (Baatveit, Torsheim, & Hove, 2018).

While neuropsychological assessment is the gold standard in evaluating cognitive functioning (Roebuck-Spencer et al., 2017), this is not always feasible as it is time-consuming (taking up to several hours) and requires extensive training in

administration and interpretation (Ridley et al., 2018; Roebuck-Spencer et al., 2017). According to the 2014 American Psychological Association Working Group on Screening and Assessment (APA–WGSA) guidelines, cognitive screening tools are used for early identification of individuals at high risk for a specific disorder, are generally brief, monitor treatment progress or change in symptoms over time, can be administered by clinicians with appropriate training, and are not definitively diagnostic (APA–WGSA, 2014). Cognitive screening, therefore, represents a critical preliminary stage of triaging those in need of more comprehensive neuropsychological assessment (Roebuck-Spencer et al., 2017).

Cognitive screening is routinely practiced in primary and tertiary settings in the assessment of adults with mild CI, dementia, ABI, and psychiatric illness (Roebuck-Spencer et al., 2017); however, it is much less established as a standard clinical process for individuals with SUD. Cognitive screening tools should be chosen carefully, with reference to both clinical utility and psychometric properties for the population in which it will be used (Slater & Young, 2013). The measure should have high sensitivity (to correctly detect those with CI) and good specificity (to accurately rule out those without CI; Bunnage, 2017). If specificity is too low, clients may inaccurately be perceived as needing full assessment, with potential for resource overload as well as unnecessary alarm for the individual. Yet, if sensitivity is too low, clients that would benefit from full assessment may not receive this, leading to a potential mismatch in treatment provision.

There exists limited evidence to guide decision-making on which cognitive screening tool is appropriate for use with individuals presenting for the treatment of SUDs. The most widely used cognitive screening tools, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) and the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), have most consistently been used in older adult populations (e.g., for stroke and dementia screening). It remains unclear how well these measures can detect CI in individuals with SUDs, or whether there are more appropriate screening tools for use in this population. Given the impact of CI on SUD treatment outcomes, there is a clear need for evidence-based review and recommendations around the use of cognitive screening tools in adults engaging in SUD treatment.

## Objective

The aims of this systematic review were to: (1) describe types of cognitive screening measures used in adults with SUDs; (2) identify specific substance use populations and settings these tools are utilised in; (3) review diagnostic accuracy of these screening measures in detecting CI versus an accepted objective reference standard; and (4) evaluate methodology of these studies for risk of bias. It was hoped this review would yield clinical recommendations for use of cognitive screening in individuals presenting for the treatment of

SUDs in clinical settings, as well as identify areas for future research in this field.

## METHOD

### Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement guided the reporting standards of this review (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). Results of systematic review registry searches (e.g., PROSPERO) did not reveal similar systematic reviews as of May 2020. The review protocol was registered with PROSPERO on 7 November 2020 (ID: CRD42020185902; [www.crd.york.ac.uk/prospero](http://www.crd.york.ac.uk/prospero)).

A systematic search of the Medline, Embase, PsycINFO (OVID), and CINAHL (EBSCO) databases was undertaken in June 2019 and updated in May 2020. The search strategy was constructed around the aims of the review. Subject headings, keywords, and keyword phrases were developed for each of the search concepts, *substance use disorder*, *cognition*, and *screening or assessment*. Subject headings and keywords for each concept were combined using the Boolean operator 'OR' before the concepts were brought together using the 'AND' operator. The search strategy was developed by a senior research librarian (AS) and was designed in Medline before being translated to the other databases. Searches were limited to adult populations and studies in English. No date limits were applied. Comments, editorials, letters, news, book, and book chapters as well as conference abstracts were excluded (see Supplementary Material 1 for full search strategies). In addition, supplementary searching was undertaken in Google Scholar using a combination of search terms for the first 200 results. The reference lists of all papers that met the final inclusion criteria were screened for any additional relevant citations not captured in the database searches. Screening for articles prepared for publication but not yet in print was performed through a search of above search terms of the Psychology Preprint Archive ([psyarxiv.com](http://psyarxiv.com)); an author known to be developing a measure in this field was also contacted for potential preprints. A comprehensive grey literature search was not conducted. Results were exported to EndNote and duplicates removed. Studies were then exported to Covidence (Veritas Health Innovation; [www.covidence.org](http://www.covidence.org)) for screening.

### Inclusion Criteria

The following inclusion criteria were set: (1) original empirical article, published or prepared for publication (i.e., *under review*); (2) in English; (3) reported on a tool designed to screen for cognitive functioning, or the tool was used for this purpose (in isolation or embedded as part of a broader screening process); (4) examined or provided data that supported inferences regarding criterion validity of the screening tool compared to an accepted standard criterion of cognitive functioning (e.g., neuropsychological assessment);

(5) participants were adults (aged 18 years and above); and (6) participants were reported to have some form of SUD (inclusive of alcohol-use or other specific disorder types). The standards described by the 2014 APA–WGSA statement guided screening tool eligibility, specifying they must: (1) be brief to administer; (2) collect objective test data with an established protocol for scoring; (3) require minimal, non-specialist training; and (4) not used for the purpose of definitive diagnosis. To maximise eligible studies, screening tools were not excluded if they also collected functional or self-report information, or focused on one cognitive domain. No restriction on type of criterion validity analysis was applied (e.g., correlations, regression, receiver operating curve).

### Screening, Extraction, and Risk of Bias

Two authors (JK and KK) independently screened titles and abstracts from the database searches for articles meeting inclusion criteria (Phase 1). JK and KK independently reviewed full-text articles for eligibility using the same review criteria (Phase 2). If there was disagreement between authors on eligibility for an article, a third author (NR) was consulted and decisions were resolved by consensus. Additional studies retrieved from correspondence or Google Scholar searches were included if author consensus (JK, KK, and NR) indicated they met inclusion criteria.

Data extraction from articles meeting criteria was conducted by JK and KK, with a check by author NR. Variables of interest included target population, study design, recruitment method, participant demographics, index test, reference standard, and relevant psychometric data (e.g., sensitivity, specificity, likelihood ratios, and correlation coefficients).

Included studies were also assessed for risk of bias and applicability (JK and KK) using the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS–2) tool according to scoring guidelines presented in Whiting and colleagues (2011). Data extraction, risk of bias, and applicability ratings followed the same dual, independent examination and consensus procedures as described for article screening. NR was not involved in reviewing her respective paper (Ridley et al., 2018); instead, SB was consulted for consensus decisions.

### Risk of bias assessment using QUADAS–2

The QUADAS–2 assessment tool (Whiting et al., 2011) evaluated risk of bias and applicability of diagnostic accuracy studies ([www.quadas.org](http://www.quadas.org)). Guided by the Standards for Reporting Diagnostic Accuracy Studies (STARD) criteria, this scale was developed to allow more transparency in reporting (Cohen et al., 2016; Whiting et al., 2011) and is recommended by the Cochrane Collaboration ([www.cochrane.org](http://www.cochrane.org)) and National Institute for Health and Clinical Excellence (NICE; [www.nice.org.uk](http://www.nice.org.uk)).

The QUADAS-2 rating process has four phases: (1) summarise review question; (2) tailor the tool to the review; (3) construct a flow diagram for the study; and (4) assess risk of bias and applicability concerns. The tool consists of four key domains: (1) patient selection; (2) index test; (3) reference standard; and (4) flow and timing (i.e., flow of participants through the study, timing of administration of index test and reference standard). Each domain is assessed for risk of bias, the first three also for applicability concerns. Signalling questions guided decision-making on high or low risk of bias (Schueler, Schuetz, & Dewey, 2012), rated as 'Yes', 'No', or 'Unclear'. Risk of bias is deemed 'Low' if all relevant signalling questions are rated 'Yes'; however, bias may exist if these questions are rated as 'No' or 'Unclear'. If questions relating to concerns regarding applicability are rated as 'High' or 'Unclear', aspects of the study may not match the review question.

### Statistical Interpretation

Sensitivity, the ability of a screening tool to correctly identify individuals with a condition of interest, and specificity, the ability of a screening tool to detect individuals who do not have a condition, are relative to a reference standard. A receiver operating characteristic (ROC) curve can be plotted by using sensitivity against the false-positive rate (1-specificity) for given cut-off points of a screening test. An area under ROC curve (AUC) value summarises overall test diagnostic accuracy (Trevethan, 2017), ranging from 0 to 1, where 0 and 1 represent perfectly inaccurate or accurate tests, respectively. As a guide, AUC values of 0.5 indicate no meaningful discrimination ability, 0.7 to 0.8 *acceptable*, 0.8 to 0.9 *excellent*, and >0.9 *outstanding* (Mandrekari, 2010).

Sensitivity and specificity are used to obtain predictive power and likelihood ratios. Positive predictive power describes the probability that an individual with a positive test result is a true positive; negative predictive power indicates the probability an individual who obtains a negative test result is a true negative. Both are dependent on the prevalence of the target condition in each population sample (Bunnage, 2017; Trevethan, 2017).

A positive likelihood ratio (LR+) refers to the ratio between the probability of correctly classifying an individual as having CI and the probability of incorrectly classifying an individual who does not have CI (i.e., sensitivity/100-specificity). A LR+ greater than 10 indicates CI is highly likely when a positive test result is yielded. A negative likelihood ratio (LR-) is the ratio between the probability of incorrectly classifying an individual as having CI and the probability of correctly classifying an individual who does not have CI (i.e., 100-sensitivity/specificity). A LR- below 0.1 indicates that impairment is highly unlikely when a negative test result is obtained (McGee, 2002).

For studies that examined correlations between cognitive screening tools and reference tests, coefficients of 0.1, 0.3,

and 0.5 indicated small, medium, and large relationships, respectively (Cohen, 1988).

## RESULTS

### Search Results

Figure 1 presents a flow diagram of search results. Searches of electronic databases of published articles identified 5,135 candidate studies. Removal of 1,672 duplicates resulted in 3,463 studies. An additional study identified from a Google Scholar search and another through correspondence with an author brought the total to 3,465 studies. A search of the Psychology Preprint Archive yielded no additional studies. After screening the titles and abstracts (Phase 1), the number of studies was reduced to 35. Examination of full-text versions of eligible studies identified 14 relevant papers. Full-text articles were excluded for not meeting the following criteria: not appropriate study design (e.g., did not evaluate screening tool against a standard criterion of cognitive functioning); did not meet 2014 APA-WGSA screening criteria; and did not include participants with SUDs. One duplicate was also removed during this phase.

### Study Characteristics

Tables 1-3 present characteristics of the 14 included studies and an overview of the screening measures and reference standards. Studies were conducted in the United States of America ( $n = 3$ ), France ( $n = 3$ ), the Netherlands ( $n = 2$ ), Spain ( $n = 1$ ), Norway ( $n = 1$ ), Brazil ( $n = 1$ ), Belgium ( $n = 1$ ), and Australia ( $n = 2$ ). Across all 14 studies, the average clinical sample size was 88.33 (SD = 116.75); however, the range between studies varied widely from 20 to 501 participants. The majority of studies ( $n = 11$ ; 78.57%) made reference to formal diagnostic criteria (e.g., Diagnostic and Statistical Manual of Mental Disorders) to operationalise SUD. Over a quarter ( $n = 4$ ) included a sample with a single problematic substance (Ewert et al., 2018; Pelletier et al., 2018; Ritz et al., 2015; Wester et al., 2013). All studies included participants with alcohol use disorder, nearly two-thirds also included those with cannabis use ( $n = 9$ ; 64.29%), followed by cocaine use ( $n = 7$ ; 50.00%), then opioids ( $n = 6$ ; 42.86%). Only one study failed to report the specific substance of misuse (To et al., 2015). Over two-thirds ( $n = 8$ ; 57.14%) required a period of abstinence before engaging in cognitive assessment, while three studies did not require any period of abstinence (Berry et al., 2021; Ridley et al., 2018; Rojo-Mota et al., 2017) and one study did not specify abstinence status (Wester et al., 2013). Nearly half of studies ( $n = 6$ ; 46.15%) were in an inpatient setting, two (15.38%) outpatient, four (30.77%) residential rehabilitation, and two (14.29%) did not report the setting.

As per inclusion criteria, each study that was included compared an index test (i.e., the cognitive screening measure) to a reference standard (i.e., neuropsychological assessment).

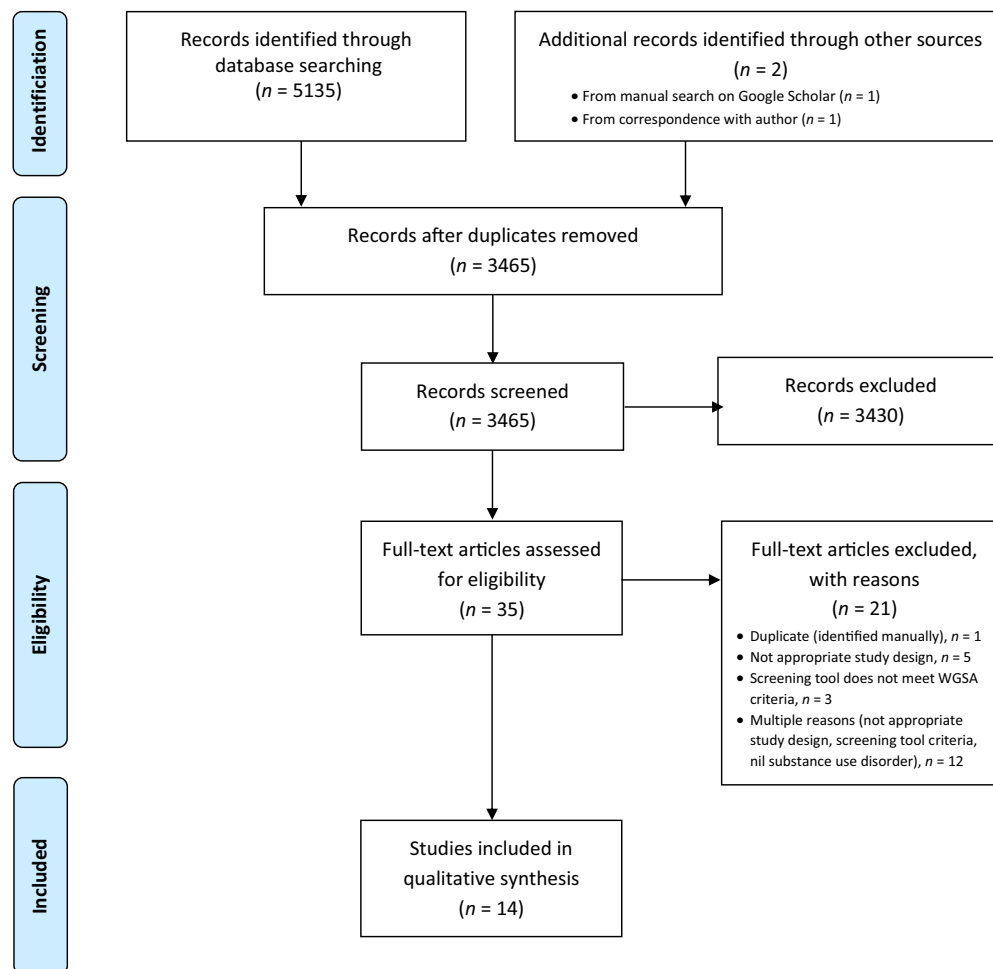


Fig. 1. PRISMA flow diagram.

In over a quarter of studies (28.57%), a particular syndrome of interest (e.g., Korsakoff's syndrome, intellectual disability) influenced test selection for the reference standard. The index tests and reference standards are presented in Tables 2 and 3.

The majority of studies ( $n = 12$ ; 85.57%) examined only one screening tool. Across all 14 studies, 10 different index measures were used, with the most common being the MoCA ( $n = 6$ ; 46.15%). Most ( $n = 8$ ; 80.00%) were developed for use within other populations (e.g., dementia, intellectual disability, brain injury); however, two were developed specifically for use within a SUD population (Brief Evaluation of Alcohol-Related Neuropsychological Impairments [BEARNI] and Brief Executive Function Assessment Tool [BEAT]). Length of administration for cognitive screening measures varied between 10 and 45 minutes. Almost all index measures targeted multiple cognitive domains, including: orientation ( $n = 5$ ; 35.71%), attention ( $n = 6$ ; 42.86%), language ( $n = 6$ ; 42.86%), visual-spatial abilities ( $n = 8$ ; 57.14%), memory ( $n = 7$ ; 50%), motor ability ( $n = 3$ ; 21.43%), and executive functions ( $n = 6$ ; 42.86%). The Frontal Assessment Battery (FAB) focused on one cognitive domain (executive functioning).

Across all 14 studies, reference standards varied in cognitive domains and specific neuropsychological tests. While the reference standard largely encompassed a range of cognitive domains, two studies (14.28%) utilised a reference standard focusing on one cognitive domain (i.e., executive functioning, memory; Berry et al., 2021; Wester et al., 2013). Another study (Cunha et al., 2010) restricted the reference standard to attention, working memory, and executive function. Other variations across studies included time between administration of index and reference standard (ranging from the same day to three months later), and operationalisation of 'impairment' for the reference standard.

## Screening Measure Utility

### Diagnostic accuracy

Table 4 displays psychometric data provided by 12 studies that reported classification accuracy (Berry et al., 2021; Braatveit et al., 2018; Bruijnen et al., 2019; Copersino et al., 2009; Ewert et al., 2018; Fals-Stewart, 1997; Grohman & Fals-Stewart, 2004; Pelletier et al., 2018; Ridley et al., 2018; Ritz et al., 2015; To et al., 2015; Wester et al., 2013). While all

**Table 1.** Study characteristics of articles included for review

First author, year, country	Target population (clinical setting, if applicable)	Exclusion criteria	Study design	Study setting	Recruitment method	Clinical group (n)	Mean age, years (SD; range)	% Male	Education level, mean years (SD; range)
Berry et al., 2021, Australia	Substance use disorder: alcohol, cannabis, methamphetamine, heroin, other stimulants, sedatives/hypnotics/tranquilizers, other opioids	Abstinence: nil Exclusions: unable to provide informed consent, not proficient in English	Cross-sectional	Residential rehabilitation	Consecutive	501	34.7 (11.1; NR)	62	10.7 (2.1; NR)
Braatveit et al., 2018, Norway	Substance use disorder (ICD-10): including alcohol, opioids, cannabis, cocaine, sedatives, hypnotics, other stimulants, hallucinogens, other psychoactive	Abstinence: 6 weeks Exclusions: native language other than Norwegian, WAIS-IV testing within 6 months, under acute influence of substances	Cross-sectional	Inpatient	Convenience	84	33.31 (11.65, 19–64)	67	Mean NR Grade 1–6 4.8%, Grade 7–9 45.2%. Grade 10–12 33.3% Higher 0%
Bruijnen et al., 2019, The Netherlands	Substance use disorder: including alcohol, cannabis, stimulants, and opiates	Abstinence: nil Exclusions: severe cognitive impairment due to neurological injury, acute psychiatric disorder, severe lack of motivation, and insufficient Dutch language	Longitudinal (baseline, 8 weeks)	Outpatient	Convenience, part of larger study (N = 691)	82	44.1 (13.8; NR)	68	Mean NR Low educated 14% Average educated 72% High educated 11%
Copersino et al., 2009, USA	Substance use disorder (DSM-IV): including alcohol, opioids, cocaine, cannabis, benzodiazepine, and amphetamines	Abstinence: 7 days Exclusions: acute intoxication/withdrawal, medical illness or psychiatric condition impeding capacity to consent	Cross-sectional	NR. Recent admission to hospital treatment or residential programme	NR	60	38.3 (13.2; NR)	52	15 (2.4; NR)
Cunha, 2010, Brazil	Cocaine dependence (DSM-IV-TR) with/without alcohol or cannabis use	Abstinence: 2 weeks Exclusions: past/current psychiatric disorders, history of neurological disorders, diagnosis of learning disorder, IQ < 70	Cross-sectional	Inpatient	NR	30	27.17 (7.64; NR)	100	9.93 (2.74; NR)
Ewert et al., 2018, France	Alcohol use disorder (DSM-5)	Abstinence: 1 week Exclusions: severe comorbid neurological or psychiatric diseases (dementia, psychosis, past history of stroke/coma/encephalopathy, cardiac disease, HIV infection)	Cross-sectional	Inpatient	Consecutive	56	Reported by cognitive impairment (CI) group No CI: 49.3 (7.9; NR) CI: 49.6 (9.4; NR)	71	Mean NR No CI: 40% >12 years education CI: 9.7% >12 years education
Fals-Stewart, 1997, USA	Substance abuse or dependence (DSM-III-R): including alcohol, cocaine, heroin, cannabis, other	Abstinence: 2 weeks Exclusions: organic mental disorder, schizophrenia, delusional disorder, other psychotic disorder; physical limitations interfering with testing; significant birth or developmental problems; and psychotropic medications	Cross-sectional	Residential rehabilitation	Consecutive	51	Reported by CI group No CI: 27.9 (4.9; NR) CI: 31.4 (5.2; NR)	75	Mean NR Impaired group: 11.6 (1.1; NR) Intact group: 12.2 (1.4; NR)
Grohman and Fals-Stewart, 2004, USA	Substance use disorder (DSM-IV): including alcohol, cocaine, opiate, cannabis, other	Abstinence: 2 weeks Exclusions: schizophrenia, delusional disorder or other psychotic disorder; on methadone programme; significant head	Cross-sectional	Residential inpatient substance use programme	Consecutive	84	32.0 (6.2; NR)	71	12.9 (1.7; NR)

(Continued)

**Table 1.** (Continued)

First author, year, country	Target population (clinical setting, if applicable)	Exclusion criteria	Study design	Study setting	Recruitment method	Clinical group (n)	Mean age, years (SD; range)	% Male	Education level, mean years (SD; range)
Pelletier et al., 2018, France	Alcohol dependence (DSM-IV)	trauma; medical disorder interfering with testing; and psychotropic medications Abstinence: 7–10 days post withdrawal Exclusions: inability to understand/speak French; severe neurological or psychiatric diseases such as dementia, psychosis, past history of stroke, coma, encephalopathy; and recent consumption of cocaine/cannabis/heroin	Cross-sectional	NR Rehab centre located in hospital	Convenience	90	48.9 (9.6; NR)	74	Mean NR 12 or less: 79% >12 years: 21%
Ridley et al., 2018, Australia	Substance use or dependence (DSM-IV) including alcohol, opioids, and cannabis	Abstinence: not required Exclusions: acute intoxication/withdrawal, limited English, significant sensory deficits	Cross-sectional	Outpatient	Consecutive	CL: 30 CO: 20	CL: 52.3 (10.43; 32–76) CO: NR	CL: 70 CO: NR	CL: 11.07 (2.33; 7–11) CO: NR
Ritz et al., 2015, France	Alcohol dependence (DSM-IV)	Abstinence: undergoing withdrawal treatment Exclusions: lifetime history of neurological pathologies, endocrine or infectious diseases, mental illness, history of other substance misuse, and psychotropic medications	Cross-sectional	Inpatient	NR	73	45.47 (8.85; 26–67)	73	11.18 (1.69; 8–15)
Rojo-Mota et al., 2017, Spain	Substance use disorder (DSM-5), including alcohol, heroin, cocaine, and cannabis	Abstinence: not under effects of substances Exclusions: brain damage, active psychosis, and inability to understand instructions	Cross-sectional	NR	Consecutive	48	39.8 (12.6; 18–67)	79	Mean NR Primary: 25% Secondary: 33% Secondary (advanced): 35.4% Uni: 6.3% NR
To, 2015, Belgium	Substance abuse problem	Abstinence: at least two weeks Exclusions: non-native speakers of Dutch, no WAIS-III last 2 years	Cross-sectional	NR; Recruited from mental health services	Convenience	90	32 (9.80; 18–64)	83	NR
Wester et al., 2013, The Netherlands	DSM-IV-TR criteria for alcohol dependence or alcohol-induced persisting amnesic disorder (Korsakoff's Syndrome)	Abstinence: unclear, but undergoing inpatient treatment Exclusions: NR	Cross-sectional	Inpatient	Convenience	Alcohol dependence (AL): 26 Korsakoff's Syndrome (KS): 20	AL: 54.5 (8.1; NR) KS: 57.6 (8.7; NR)	AL: 76 KS: 75	AL: Mean 7–9 years (range 1–16) KS: Mean 7–8 years (range 6–16) (range 7–16)

Note: Only clinical sample characteristics are presented here. As Ridley et al. (2018) and Wester et al. (2013) combined clinical and control data when calculating accuracy statistics, both groups are presented in this table. DSM-III-R, Diagnostic and Statistical Manual-Third Edition-Revised; DSM-IV, Diagnostic and Statistical Manual-Fourth Edition; DSM-IV-TR, Diagnostic and Statistical Manual-Fourth Edition-Text-Revised; DSM-5, Diagnostic and Statistical Manual-Fifth Edition; ICD-10, International Statistical Classification of Diseases and Related Health Problems-10th Revision; NR, not reported; SD, standard deviation; WAIS-III, Wechsler Adult Intelligence Scale-Third Edition; WAIS-IV, Wechsler Adult Intelligence Scale-Fourth Edition.

**Table 2.** Overview of cognitive screening measures

Screening measure	Domains evaluated	Total score	Standard impairment criteria	Administration time (minutes)	Status in adults with SUDs
Addenbrooke's Cognitive Examination – Revised (ACE-R)	Five subtests assessing orientation, attention, verbal fluency, memory, language, and visuospatial function. The Mini-Mental State Examination (MMSE) is incorporated into this test.	/100 Includes MMSE (/30 points)	<88 for increased sensitivity <82 for increased specificity	12–20	Developed to screen for dementia (Mioshi et al., 2006). Not developed specifically for SUD.
Brief Evaluation of Alcohol-Related Neuropsychological Impairments (BEARNI)	Five subtests assessing visuospatial function, executive functions via alternating verbal fluency, verbal episodic memory, verbal working memory, and ataxia (through a balance task).	/30 Without ataxia /22	Total (including ataxia): <i>Mild impairment:</i> ≤12 years education: ≤19 >12 years education: ≤21 <i>Mod. to sev. impairment:</i> ≤12 years education: ≤16 >12 years education: ≤17 Total (excluding ataxia): <i>Mild impairment:</i> ≤12 years education: ≤15 >12 years education: ≤16 <i>Mod. to sev. impairment:</i> ≤12 years educ: ≤11.5 >12 years educ: ≤12.5	15–20	Specifically developed for use with alcohol use disorder populations. Normative data for adults aged 20 to 66 years (Ritz et al., 2015).
Brief Executive Function Assessment Tool (BEAT)	Contains 20 items evaluating a range of cognitive domains, including self-reported executive functioning, visuoconstruction, learning and memory, attention and working memory, motor sequencing, executive functioning (abstraction, fluency, and delay discounting), reading, and naming.	/60	Intact cognitive functioning: 31–60 <i>Mild impairment:</i> 17–30 <i>Sev. impairment:</i> ≤16	20	Specifically developed to assess cognitive impairment in substance use disorder populations (Berry et al., 2021).
Frontal Assessment Battery (FAB)	Six subtests evaluating executive functioning: conceptualisation, mental flexibility, motor programming, and resistance to interference, inhibitory control, and environmental autonomy.	/18	Not specified. Higher score indicates better performance.	10	Developed to assess the presence and severity of a dysexecutive syndrome (Dubois et al., 2000). Not developed specifically for SUD.
Hayes Ability Screening Index (HASI)	Three subtests evaluating spelling, visuospatial, and visuoconstructional ability, as well as four questions on self-reported learning difficulties.	/96.4	≤85	5–10	Developed to screen for intellectual disability within forensic justice setting (Hayes, 2000). Not developed specifically for individuals with substance use disorder.
Loewenstein Occupational Therapy Cognitive Assessment (LOTCA)	Contains 20 items categorised into 4 subscales: orientation, perception, visuospatial organisation, and thinking operations.	/80	Not specified. Higher score indicates less cognitive impairment.	30–45	Developed to assess basic cognitive performance in brain injured patients (Katz et al., 1989). Not developed specifically for individuals with SUD.

(Continued)



**Table 2.** (Continued)

Screening measure	Domains evaluated	Total score	Standard impairment criteria	Administration time (minutes)	Status in adults with SUDs
Mini-Mental State Examination (MMSE)	Assesses orientation, registration of three words, attention, recall of three words, language, and visuoconstruction.	/30	Score of <21 suggested by authors	5–10	Developed to screen for cognitive impairment in dementia syndromes, delirium, schizophrenia, or affective disorders (Folstein et al., 1975). Not developed specifically for individuals with SUD.
Montreal Cognitive Assessment (MoCA)	Assesses executive functioning, visuospatial abilities, attention, concentration and working memory, language, abstract reasoning, memory, and orientation.	/30	<26 (including +1 point of <13 years education)	10	Developed to screen for mild cognitive impairment and Alzheimer's disease (Nasreddine et al., 2005). Not developed specifically for individuals with SUD.
Neuropsychological Assessment Battery–Screening Module (NAB–SM)	Assesses cognitive functioning across five domains: attention, language, memory, visuospatial, and executive.	Raw scores converted to standard score (M: 100, SD: 15)	<85 (standard score)	35–45	Developed to screen for cognitive impairment across a range of cognitive domains (Stern & White, 2003). Not developed specifically for individuals with SUD.
Neurobehavioural Cognitive Status Examination (NCSE/Cognistat)	Assesses orientation, attention, language (comprehension, repetition, naming), construction, memory, calculations, and executive functions (reasoning, judgement).	10 independent domain scores	Impairment defined by cut-off scores in each domain	10–30	Designed to assess cognition briefly within five major cognitive ability areas (Kiernan et al., 1987). Not developed specifically for individuals with SUD.

M, mean; SD, standard deviation; SUD, substance use disorder.

**Table 3.** Overview of screening measures and reference standards for included studies

First author (year)	Target population	Time between index and reference	Index test(s)	Index impairment criteria	Reference	Reference impairment criteria
Berry et al. (2021)	Substance use disorder	Same day	BEAT	<i>Intact cognitive functioning:</i> 31–60 <i>Mild impairment:</i> 17–30 <i>Sev. Impairment:</i> ≤16	<i>Executive Functions:</i> Alpha Span Task (Total Score), FPT (Unique Designs), Stroop (Interference Trial)	Mean z-score across three tasks: ≤−1.5 SD below the mean was classified as impaired
Braatveit et al. (2018)	Intellectual disability in those with SUD	Same day (screen prior to NPA)	HASI, Norwegian version	N/A: range of cut-scores	<i>Global Intellectual Ability:</i> WAIS-IV	WAIS-IV FSIQ < 70; Vineland II score <70; Self-reported learning difficulties
Bruijnen et al. (2019)	Substance use disorder	Baseline MoCA at intake Repeat MoCA and NPA administered same day 6–8 weeks later (screen prior to NPA)	MoCA, Dutch version 7.1 (baseline), Dutch version 7.2 (follow-up)	N/A: range of scores	<i>Orientation:</i> CST-14; <i>Processing Speed:</i> SCWT word reading, colour naming; <i>D-KEFS TMT</i> motor speed; <i>Attention:</i> DSF and DSB; <i>Visuospatial Abilities:</i> WAIS-IV-NL Block Design, ROCF copy; <i>Memory:</i> RAVLT total correct and delayed recall, ROCF Immediate Recall; <i>Abstract Reasoning:</i> WAIS-IV-NL Similarities, DART; <i>Executive Functions:</i> D-KEFS TMT (letter sequencing, number sequencing, and letter-number switching), SCWT interference	At least 2 of 7 domain z-scores ≤−1.00
Copersino et al. (2009)	Substance dependent	Same day (counterbalanced)	MoCA, English version	N/A: range of scores	NAB-Screening Module (NAB-SM): <i>Attention:</i> Orientation, Digits Forward, Digits Backward, Numbers and Letters; <i>Language:</i> Picture Naming, Auditory Comprehension <i>Spatial:</i> Visual Discrimination and Design Construction; <i>Memory:</i> Story Learning (immediate and delayed recall), Shape Learning <i>Executive Functions:</i> Mazes, Word Generation	NAB-SM Total Screening Index score of <85
Cunha et al. (2010)	Substance dependent	Same day (unclear order)	FAB, Portuguese version	N/A: range of scores	<i>Attention and Working Memory:</i> WAIS-IV DSF, DSB (Portuguese version) <i>Executive:</i> SCWT (Portuguese version), WCST	N/A
Ewert et al. (2018)	Alcohol use disorder	Screen conducted day prior to NPA	MoCA, French version 7.1	N/A: range of scores	<i>Attention and Working Memory:</i> DSF, DSB, DSS; <i>Visuospatial Skills:</i> ROCF copy; <i>Episodic Memory:</i> CVLT learning, short delay free recall, short delay cued recall, delayed free recall, delayed cued recall, recognition; <i>Executive Functions:</i> TMT B, Stroop Interference-Naming, Fluency (Letter and Animal)	1+ score within a domain either <1.65 SD or <5th percentile or 2+ for executive domain
Fals-Stewart (1997)	Substance abuse or dependence	Same day (counterbalanced)	NCSE	Classified as impaired if they scored in the impaired range on any of the 10 subscales.	<i>Premorbid IQ:</i> NART Neuropsychological Screening Battery; <i>Psychomotor Speed:</i> SDMT <i>Sequencing Efficiency:</i> TMT A and B; <i>Visual Attention:</i> Numerical attention test; <i>Verbal Learning:</i> WMS Story; <i>Non-Verbal Learning:</i> ROCF; <i>Delayed Recall:</i> WMS Story, ROCF; <i>Visuoconstruction:</i> ROCF; <i>Language:</i> MAE (fluency, auditory sequencing, naming, repetition, speech articulation)	For each subtest score, <1 SD = mildly impaired = 1 score; <5th percentile = severely impaired = 2 points; a cumulative score of 6 points or more indicated overall impairment
Grohman and Fals-Stewart (2004)	Substance use disorder	One day apart (counterbalanced)	NAB-SM	Index score of 84 or less	<i>Premorbid IQ:</i> NART Neuropsychological Screening Battery; <i>Motor Speed:</i> SDMT <i>Sequencing Efficiency:</i> TMT A and B; <i>Visual Attention:</i> Numerical attention Test; <i>Verbal Learning:</i> WMS Learning Scores; <i>Non-Verbal Learning:</i> ROCF learning; <i>Delayed Recall:</i> WMS Memory, ROCF memory; <i>Visuoconstructional Skills:</i> ROCF copy; <i>Language:</i> MAE (fluency, auditory sequencing, confrontation naming, repetition)	As above
Pelletier et al. (2018)	Alcohol use disorder	One day apart (random order)	MoCA, French version 7.1	25 or less, no educational adjustment	French versions of validated tests <i>Attention and Working Memory:</i> Digit Span Forward, Backward; <i>Visuospatial skills:</i> ROCF copy; <i>Executive Functions:</i> TMT A and B, SCWT, fluency (letter, animals) <i>Verbal Memory:</i> FCSRT, cued recall trial 3, and delayed cued recall trial	1+ subtest <1.65 SD or 5th percentile within any domain (2+ for executive domain)

(Continued)

Table 3. (Continued)

First author (year)	Target population	Time between index and reference	Index test(s)	Index impairment criteria	Reference	Reference impairment criteria
Ridley et al. (2018)	Alcohol use disorder	Same day (screens counterbalanced order with NPA)	MoCA (English version 7.1) ACE-R (Form A, Australian version) MMSE (in ACE-R)	N/A: range of cut-off scores	<i>Attention:</i> RBANS Digit Span, Coding; <i>Language:</i> RBANS Naming, Semantic Fluency; <i>Visuospatial:</i> RBANS Line Orientation, Figure Copy; <i>Memory:</i> RBANS List-Learning, Story-Learning, List Recall, List Recognition, Story Recall, Figure Recall; <i>Executive Functions:</i> Letter Fluency, Digits Backwards, TMT B, Hayling and Brixton Tests	<1.5 SD on domain index scores, 2+ subtests impairment in executive domain
Ritz et al. (2015)	Alcohol use disorder	NPA completed within 2 days following screening measure	BEARNI MMSE	BEARNI Total: <i>Mild impairment:</i> ≤12 years ed: ≤19 >12 years ed: ≤21 <i>Mod.-sev. impairment:</i> ≤12 years ed: ≤16 >12 years ed: ≤17 BEARNI Cognition: <i>Mild impairment:</i> ≤12 years ed: ≤15 >12 years ed: ≤16 <i>Mod.-sev. impairment:</i> ≤12 years ed: ≤11.5 >12 years ed: ≤12.5 MMSE Total: <24	<i>Working Memory:</i> DSF, DSB, LNS; <i>Visuospatial:</i> ROCF; <i>Verbal Memory:</i> CVLT (French); <i>Executive Functions:</i> TMT, RFFT, SCWT, fluency (letter, animal); <i>Ataxia:</i> Walk-a-line Ataxia Battery	NR
Rojo-Mota et al. (2017)	Substance use disorder	All within one week (order unclear)	LOTCA	N/A, correlations only	<i>Orientation:</i> BCSE; <i>Processing Speed:</i> WAIS-IV Coding, Symbol Search; <i>Working Memory:</i> Digit Span, Symbol Span; WMS-IV <i>Verbal Memory:</i> Logical Memory; <i>Visual Memory:</i> ROCF; <i>Executive Functions:</i> FDT; <i>Motor Ability:</i> ACLS-5	N/A
To (2015)	Substance abuse problems	Same day	HASI (Dutch version)	N/A: Range of cut-off scores	<i>Global Intellectual Ability:</i> WAIS-III (Dutch version) RBMT-3 (Dutch version)	<70 FSIQ
Wester et al. (2013)	Alcohol-related cognitive impairment, Korsakoff Syndrome	MoCA-D at intake, RBMT-3 (Dutch version) up to 3 months later	MoCA-D	N/A: Range of cut-off scores	<i>Verbal memory:</i> Names, Stories; <i>Visual memory:</i> Face and picture recognition; <i>Spatial memory:</i> immediate and delayed recall of a route; <i>Prospective Memory:</i> remembering appointments, personal belongings, shopping items; <i>Spatial and temporal orientation; New learning:</i> immediate and delayed recall of a novel complex puzzle	Raw scores changes to standard scores, converted to GMI score; mean of 100, SD of 15 Grouped into: Sev. impairment: GMI of <70; Mild deficits: 70–84; Unimpaired: GMI of ≥85

ACE-R, Addenbrooke's Cognitive Examination-Revised; ACLS-5, Allen Cognitive Level Screen-Fifth Edition; BCSE, Brief Cognitive Status Examination; BEARNI, Brief Evaluation of Alcohol-Related Neuropsychological Impairments; BEAT, Brief Executive Function Assessment Tool; CI, cognitive impairment; CST-14, Cognitive Screening Test-14; CVLT, California Verbal Learning Test; CVLT-II, California Verbal Learning Test-Second Edition; DART, Dutch Adult Reading Test; D-KEFS TMT, Delis-Kaplan Executive Function System Trail Making Test; DSB, Digit Span Backward; DSF, Digit Span Forward; DSS, Digit Span Sequencing; FAB, Frontal Assessment Battery; FCSRT, Free and Cued Selective Reminding Test; FPT, Five-Point Test; FSIQ, Full Scale Intelligence Quotient; GMI, General Memory Index; HASI, Hayes Ability Screening Index; LNS, letter-number sequencing; LOTCA, Loewenstein Occupational Therapy Cognitive Assessment; MAE, Multilingual Aphasia Exam; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NAB-SM, Neuropsychological Assessment Battery-Screening Module; NART, National Adult Reading Test; NCSE, Neurobehavioural Cognitive Status Examination; NPA, neuropsychological assessment; NR, not reported; NSB, Neuropsychological Screening Battery; RAVLT, Rey Auditory Verbal Learning Test; RBANS, Repeatable Battery for the Assessment of Neuropsychological Status; RBMT-3, Rivermead Behavioural Memory Test-Third Edition; RFFT, Ruff Figural Fluency Test; ROCF, Rey-Osterrieth Complex Figure; SCWT, Stroop and Colour Word Test; SD, standard deviation; SDMT, Symbol Digit Modalities Test; SUD, substance use disorder; TMT, Trail Making Test; TMT A, Trail Making Test-Part A; TMT B, Trail Making Test-Part B; TMT B-A, Trail Making Test Part A minus Part B; Vineland II, Vineland Adaptive Behaviour Scales-Second Edition; WAIS-R, Wechsler Adult Intelligence Scale-Revised; WAIS-III, Wechsler Adult Intelligence Scale-Third Edition; WAIS-IV, Wechsler Adult Intelligence Scale-Fourth Edition; WAIS-IV-NL, Wechsler Adult Intelligence Scale-Fourth Edition (Dutch version); WCST, Wisconsin Card Sorting Test; WMS, Wechsler Memory Scale; WMS-IV, Wechsler Memory Scale-Fourth Edition.

of these studies provided sensitivity and specificity data for at least one defined cut-point, two did not provide overall AUC data (Fals-Stewart, 1997; Pelletier et al., 2018). Although several studies included control groups in the overall study, two included control groups in diagnostic accuracy data (Ridley et al., 2018; Wester et al., 2013).

Of the studies that provided overall AUC data, the MoCA ranged from *acceptable* to *outstanding* in the detection of CI (range: 0.75–0.96). The single study examining the Addenbrooke's Cognitive Examination–Revised (ACE–R; Ridley et al., 2018) revealed *excellent* discriminative ability (AUC = 0.85), as did the Hayes Ability Screening Index (HASI; AUC = 0.94–0.95) across two studies (Baatveit et al., 2018; To et al., 2015), and the Neuropsychological Assessment Battery–Screening Module (NAB–SM; AUC = 0.88; Fals-Stewart, 1997). The MMSE by comparison had lower accuracy that ranged from *no discrimination ability* to *acceptable* (AUC = 0.21–0.79) across two studies (Ridley et al., 2018; Ritz et al., 2015). For index measures specifically developed for SUD populations (BEAT and BEARNI), at least *excellent* accuracy in detecting CI was reported (Berry et al., 2021; Ritz et al., 2015).

Half of studies that reported diagnostic accuracy data ( $n = 6$ ) included Positive Predictive Power and Negative Predictive Power; only two studies reported positive and negative likelihood ratios (Berry et al., 2021; Copersino et al., 2009). Berry and colleagues (2021) was the sole study that reported classification accuracy as well as predictive power and likelihood ratios. Using a LR+ of  $\geq 10$  and LR– of  $\leq 0.1$  as threshold (McGee, 2002), a score of  $\leq 14$  on the BEAT (Berry et al., 2021) was needed to rule in likely impairment, while a score of  $\geq 31$  was needed to likely rule out impairment. Copersino et al. (2009) indicated that a cut-score of  $\leq 22$  on the MoCA had strong evidence for ruling in CI and a score of  $\geq 27$  had strong evidence.

### Correlational data

More than half of the 14 studies ( $n = 8$ ) examined correlations between index tests and neuropsychological measures, with most ( $n = 6$ ) presenting this in addition to classification accuracy data. Two studies (Cunha et al., 2010; Rojo-Mota et al., 2017) provided correlational data only. As outlined in Table 5, various correlations were performed, including: index total summary scores as compared to reference standard total summary score; index domain score as compared to corresponding reference standard domain scores; and total index score versus specific test outcome on the reference standard. One study (Ridley et al., 2018) combined control and clinical groups to produce correlational data, while all others used clinical samples only. The largest correlations reported were between total summary scores derived from a reference standard (i.e., the Repeatable Battery for the Assessment of Neuropsychological Status; RBANS) and total scores on the MoCA ( $r = 0.75$ ) and the ACE–R ( $r = 0.81$ ; Ridley et al., 2018). Similarly, large correlations were observed between

HASI total score and Wechsler Adult Intelligence Scale–Third Edition Full Scale IQ ( $r = 0.69$ – $0.70$ ; To et al., 2015). For the MoCA, medium to large correlations between MoCA total score and neuropsychological tasks assessing executive function ( $r = 0.32$ – $0.46$ ), learning and memory ( $r = 0.45$ – $0.56$ ), working memory ( $r = 0.27$ – $0.29$ ), and fluency ( $r = 0.37$ – $0.53$ ) were found (Ewert et al., 2018). MoCA domain scores also yielded medium to large correlations to corresponding neuropsychological domains (Bruijnen et al., 2019; Ridley et al., 2018), with the exception of the domains of language and attention/orientation items (Ridley et al., 2018). For the ACE–R, Ridley and colleagues (2018) reported medium to large correlations between reference domains and relevant ACE–R subscales, with the exception of language. For the FAB, Cunha and colleagues (2010) reported medium to large correlations ( $r = 0.41$ – $0.57$ ) between FAB tasks and performance on specific executive function measures; however, not all correlations were significant ( $r = 0.03$ – $0.27$ ). For the Loewenstein Occupational Therapy Cognitive Assessment (LOTCA), medium to large correlations between index and reference tasks were found for measures of visual-perceptual abilities and psychomotor abilities (Rojo-Mota et al., 2017). However, reference measures assessing executive functioning and memory did not correspond consistently with LOTCA subscales. For the BEAT, there were medium to large correlations between overall score and performance on reference tasks measuring executive functioning ( $r = 0.43$ – $0.57$ ; Berry et al., 2021).

### Quality

Ratings of risk of bias, applicability, and quality of study design are presented in Table 6. Almost all studies were judged to have a high or unclear risk of bias in terms of patient selection (QUADAS–2 Domain 1; 92.86%) due to convenience or unclear sampling methods (57.14%; Baatveit et al., 2018; Bruijnen et al., 2019; Cunha et al., 2010; Copersino et al., 2009; Pelletier et al., 2018; Ritz et al., 2015; To et al., 2015; Wester et al., 2013); inappropriate or unclear exclusions (57.14%; Cunha et al., 2010; Ewert et al., 2018; Fals-Stewart, 1997; Grohman & Fals-Stewart, 2004; Pelletier et al., 2018; Ritz et al., 2015; Rojo-Mota et al., 2017; Wester et al., 2013); inclusion of a control group in analyses (14.29%; Ridley et al., 2018; Ritz et al., 2015); or a combination of these factors. Nearly half of the included studies ( $n = 6$ ; 42.86%) had the main aim of validating the index measure within a SUD population (Berry et al., 2021; Baatveit et al., 2018; Copersino et al., 2009; Ritz et al., 2015; Rojo-Mota et al., 2017; To et al., 2015). Half of the studies had high or unclear risk of bias for the index test (QUADAS–2 Domain 2; 50.00%), while a greater number had a high or unclear risk of bias for the reference standard (QUADAS–2 Domain 3; 71.43%). The high risk was largely due to the non-blinding of index or reference test results; only four studies (Ewert et al., 2018; Pelletier et al., 2018; Rojo-Mota et al., 2017; Wester et al., 2013) used different

**Table 4.** Psychometric data retrieved from primary classification accuracy studies

First author (year)	Impairment Rate (reference standard)	Comparison parameters for impairment	Sensitivity	Specificity	Area Under curve	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Berry et al. (2021)	7.8%	Optimal cut-off score for BEAT of $\leq 24$ against executive tasks	87.2%	71.4%	0.86	20.5%	98.5%	3.04	0.18
Braatveit et al. (2018)	NR	HASI at standard cut-off score of $\leq 85$ against diagnosis of intellectual disability (WAIS-IV $\leq 70$ )	100%	65.4%	0.94	NR	NR	NR	NR
		HASI performance at adjusted cut-off score of $\leq 80.7$ against diagnosis of intellectual disability (including WAIS-IV $\leq 70$ )	100%	80.8%	NR	NR	NR	NR	NR
Bruijnen et al. (2019)	54.9% (sample)	MoCA at baseline (Version 7.1) with cut-off score of 24 against NPA	55.6%	62.2%	0.68	64.1%	53.5%	NR	NR
		MoCA at follow-up (Version 7.2) with optimal cut-off score of 25 against NPA	66.7%	73.0%	0.745	75.0%	64.3%	NR	NR
Copersino et al. (2009)	NR	Optimal cut-off score for MoCA of 25, against NAB-SM Total Screening Index impairment	83.3%	72.9%	0.86	NR	NR	3.08	0.23
Ewert et al. (2018)	NR	MoCA score of $\geq 25$ uncorrected for education, against overall impairment on NPA	68%	96%	0.83	NR	NR	NR	NR
		MoCA score of $\geq 25$ uncorrected for education, against one cognitive domain impaired on NPA	62%	96%	0.82	NR	NR	NR	NR
		MoCA score of $\geq 25$ uncorrected for education, against > one cognitive domain impaired on NPA	73%	96%	0.92	NR	NR	NR	NR
		MoCA score of $\geq 26$ uncorrected for education, against overall impairment on NPA	84%	80%	0.83	NR	NR	NR	NR
		MoCA score of $\geq 26$ uncorrected for education, against one cognitive domain impaired on NPA	75%	80%	0.82	NR	NR	NR	NR
		MoCA score of $\geq 26$ uncorrected for education, against > one cognitive domain impaired on NPA	93%	80%	0.92	NR	NR	NR	NR
		MoCA score of $\geq 26$ corrected for education, against overall impairment on NPA	71%	88%	0.84	NR	NR	NR	NR
		MoCA score of $\geq 26$ corrected for education, against one cognitive domain impaired on NPA	62%	88%	0.80	NR	NR	NR	NR
		MoCA score of $\geq 26$ corrected for education, against > one cognitive domain impaired on NPA	80%	88%	0.90	NR	NR	NR	NR
		MoCA score of $\geq 27$ corrected for education, against overall impairment on NPA	84%	76%	0.84	NR	NR	NR	NR
		MoCA score of $\geq 27$ corrected for education, against one cognitive domain impaired on NPA	75%	76%	0.80	NR	NR	NR	NR
MoCA score of $\geq 27$ corrected for education, against > one cognitive domain impaired on NPA	93%	76%	0.90	NR	NR	NR	NR		

(Continued)

**Table 4.** (Continued)

First author (year)	Impairment Rate (reference standard)	Comparison parameters for impairment	Sensitivity	Specificity	Area Under curve	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Fals-Stewart (1997)	43% (sample)	NCSE impairment (in at least one domain) against impairment on NSA (summary score $\geq 6$ )	36%	86%	NR	67%	64%	NR	NR
Grohman and Fals-Stewart (2004)	NR	NAB-SM at established cut-off $\geq 6$ compared with the NSB	81%	92%	0.88	87%	88%	NR	NR
Pelletier et al. (2018)	NR	MoCA at traditional cut-off <26 (uncorrected for education), against impairment based on NPA	79%	65%	NR	NR	NR	NR	NR
		BEARNI total (including ataxia), against impairment based on NPA	100%	2%	NR	NR	NR	NR	NR
		BEARNI cognition (excluding ataxia) against impairment based on NPA	98%	13%	NR	NR	NR	NR	NR
Ridley et al. (2018)	Two values: 67% and 33% (general)	MoCA at traditional cut-off score of <26, against impairment based on NPA	70%	83%	0.84	67%: 89% 33%: 67%	67%: 58% 33%: 85%	NR	NR
		MoCA at optimal cut-off score of <27, against impairment based on NPA	80%	73%	0.84	67%: 86% 33%: 59%	67%: 64% 33%: 41%	NR	NR
		ACE-R at traditional cut-off score of <88, against impairment based on NPA	55%	90%	0.85	67%: 92% 33%: 73%	67%: 50% 33%: 80%	NR	NR
		ACE-R at optimal cut-off score of <93, against impairment based on NPA	90	73%	0.85	67%: 87% 33%: 62%	67%: 78% 33%: 93%	NR	NR
		MMSE at traditional cut-off score of <27, against impairment based on NPA	15%	100%	0.79	67%:100% 33%:100%	67%: 37% 33%: 70%	NR	NR
		MMSE at optimal cut-off score of <29, against impairment based on NPA	60%	93%	0.79	67%: 94% 33%: 81%	67%: 53% 33%: 82%	NR	NR
Ritz et al. (2015)	Overall: Mild: 86% Mod. to Sev.: 53% Without ataxia: Mild: 84% Mod. to Sev.: 45%	BEARNI total optimal score (education corrected), against mild impairment on NPA	98%	50%	0.92	93%	83%	NR	NR
		BEARNI Total optimal score (education corrected) mod. to sev. impairment on NPA	95%	71%	0.84	79%	92%	NR	NR
		BEARNI cognitive optimal score (education corrected) against mild impairment on NPA	93%	50%	0.86	90%	60%	NR	NR
		BEARNI cognitive optimal score (education corrected) against mod. to sev. impairment NPA	82%	85%	0.84	82%	85%	NR	NR
		MMSE optimal score (education adjusted), against mild impairment on the NPA	10%	90%	0.21	86%	14%	NR	NR
		MMSE optimal score (education adjusted), against mod. to sev. impairment on NPA	8%	88%	0.24	43%	46%	NR	NR
To et al. (2015)	NR	HASI at cut-off score of 84.95, against WAIS-III at IQ 70	91%	80%	0.95	NR	NR	NR	NR
		HASI at cut-off score of 85.55, against WAIS-III at IQ 70	91%	79%	0.95	NR	NR	NR	NR
		HASI at cut-off score of 86.40, against WAIS-III at IQ 70	100%	91%	0.95	NR	NR	NR	NR

(Continued)

Table 4. (Continued)

First author (year)	Impairment Rate (reference standard)	Comparison parameters for impairment	Sensitivity	Specificity	Area Under curve	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Wester et al. (2013)	NR	HASI at cut-off score of 87.05, against WAIS-III at IQ 70 MoCA-D optimal score of $\leq 23$ , against severe memory disorder (GMI < 70) MoCA-D traditional cut-score of $\leq 25$ , against severe memory disorder (GMI < 70) MoCA-D optimal score of $\leq 24$ , against mild memory disorder (GMI 70–84) MoCA-D traditional cut-score of $\leq 25$ , against mild memory disorder MoCA-D optimal score of $\leq 20$ , to classify mild versus severe memory disorder	100% 91% 78% 88% 78% 81%	79% 88% 92% 71% 76% 69%	0.95 0.96 0.96 0.82 0.82 0.75	NR NR NR NR NR NR	NR NR NR NR NR NR	NR NR NR NR NR NR	NR NR NR NR NR NR

ACE-R, Addenbrooke's Cognitive Examination-Revised; BEARNI, Brief Evaluation of Alcohol-Related Neuropsychological Impairments; BEAT, Brief Executive Function Assessment Tool; FAB, Frontal Assessment Battery; GMI, General Memory Index; HASI, Hayes Ability Screening Index; IQ, intellectual quotient; LOTCA, Loewenstein Occupational Therapy Cognitive Assessment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; MoCA-D, Montreal Cognitive Assessment (Dutch version); NAB-SM, Neuropsychological Assessment Battery-Screening Module; NCSE, Neurobehavioural Cognitive Status Examination; NPA, neuropsychological assessment; NR, not reported; NSB, Neuropsychological Screening Battery; WAIS-III, Wechsler Adult Intelligence Scale-Third Edition; WAIS-IV, Wechsler Adult Intelligence Scale-Fourth Edition.

individuals to administer the index and reference measures, while in another seven the order of administration or prior knowledge of test results by the tester was unclear (Berry et al., 2021; Braatveit et al., 2018; Bruijnen et al., 2019; Cunha et al., 2010; Fals-Stewart, 1997; Grohman & Fals-Stewart, 2004; Ritz et al., 2015). For participant flow and timing (QUADAS-2 Domain 4), nearly half of included studies had either high or unclear risk of bias (42.86%) due to inappropriate or unclear time interval between the administration of the index test and the reference standard ( $n = 2$ ), participants not all receiving the reference standard ( $n = 3$ ), or not accounting for missing participant data, either clinically or statistically. Overall, quality rating results based on QUADAS-2 indicated generally poor adherence to STARD criteria across the included studies. Two studies (Berry et al., 2021; Bruijnen et al., 2019) were not rated as high risk across any domains; however, unclear ratings meant that risk was unable to be fully assessed.

By contrast, the majority of studies met applicability criteria for the current review. The exception was in relation to the reference standard (QUADAS-2 Domain 3) across two studies (Braatveit et al., 2018; To et al., 2015). These studies sought to evaluate presence of intellectual disability within SUD populations, and consequently, it is possible the reference standards could fail to detect CI specifically related to SUD, or CI more broadly in this population.

## DISCUSSION

This review had four primary aims: (1) describe types of cognitive screening measures used in adults with SUDs; (2) identify specific SUD population and settings in which these tools are utilised; (3) review diagnostic accuracy of these screening measures in relation to detecting CI versus an accepted objective reference standard; and (4) evaluate methodology of these studies to determine risk of bias.

### Aims 1 and 2: Overview of Cognitive Screening Measures in Adults with SUDs

The current review, to the authors' knowledge, is the first systematic overview of cognitive screening tools utilised in adult SUD populations as compared to a reference standard (e.g., neuropsychological assessment). Overall, 14 studies were identified which broadly aligned with definitions of cognitive screening as proposed by Roebuck-Spencer and colleagues (2017). In summary, more than three-quarters of studies made reference to specific SUD diagnostic criteria. The majority included individuals with poly-substance use, while the remaining specifically examined alcohol use. Alcohol was the most common substance used by participants, with a period of abstinence also frequently endorsed, and nearly half of studies were conducted in an inpatient setting. However, a number of studies failed to report key study details including study setting, specific substance of misuse, and abstinence status. Sample sizes were

**Table 5.** Summary of studies examining correlations between cognitive screening tools and comprehensive cognitive assessment

First author (year)	Comparisons	Summary of main findings*
Berry et al. (2021)	Correlations between BEAT and executive tasks (reference battery)	BEAT total score significantly associated with Stroop Interference Trial ( $r = 0.43$ ), FPT Unique Designs ( $r = 0.47$ ), and Alpha Score ( $r = 0.52$ ).
Braatveit et al. (2018)	Pearson correlations of HASI total score as compared to WAIS-IV FSIQ	HASI total score was significantly associated with the WAIS-IV FSIQ ( $r = 0.70$ ). Data were used in support of convergent validity.
Bruijnen et al. (2019)	Pearson correlations of MoCA domain scores compared to domains scores on neuropsychological assessment	At <i>baseline</i> , correlations between MoCA and NPA domains were significant for executive functioning ( $r = 0.24$ ), abstract reasoning ( $r = 0.30$ ), and memory ( $r = 0.42$ ). Correlations not reported for other domains. At <i>follow-up</i> , all MoCA domain scores were significantly correlated with the corresponding NPA domain: executive functioning: ( $r = 0.33$ ); visuospatial abilities ( $r = 0.24$ ); attention ( $r = 0.40$ ); abstract reasoning ( $r = 0.54$ ); memory ( $r = 0.46$ ); and orientation ( $r = 0.23$ ).
Cunha et al. (2010)	Spearman correlations of FAB total and domain scores with executive function tasks on NPA	Significant negative correlation between FAB score total and time to complete interference task on SCWT ( $r_s = -0.41$ ). Significant negative correlation between FAB mental flexibility subtest and number of perseverative errors on WCST ( $r_s = 0.44$ ). Significant positive correlation between FAB mental flexibility and performance on digits forward ( $r_s = 0.57$ ). Significant positive correlation between performance on FAB motor programming subtest and failures to maintain set on WCST ( $r_s = 0.41$ ).
Ewert et al. (2018)	Zero-order Pearson's correlations of total MoCA score with NPA tasks	Significant negative correlation between MoCA score and completion time ( $r = -0.46$ ) and errors on the TMT ( $r = -0.32$ ), as well as completion time on the Word-Colour part of the Stroop task ( $r = -0.37$ ) and related interference measure ( $r = -0.32$ ). MoCA score correlated positively with CVLT immediate ( $r = 0.55$ ), long delay free ( $r = 0.56$ ), long delay cued ( $r = 0.45$ ), DSF ( $r = 0.27$ ), and DSS ( $r = 0.29$ ), ROCF (F(3,53) = 2.76, $p = 0.05$ ; multiple R = 0.02) and Letter ( $r = 0.53$ ) and Animal Fluency ( $r = 0.37$ ).
Ridley et al. (2018)	Correlations (Pearson or Spearman) of MoCA and ACE-R domains with NPA total scores and domains across whole sample (clinical and control groups combined)	RBANS total score strongly correlated with MoCA ( $r = 0.75$ ) and the ACE-R ( $r = 0.81$ ), moderately correlated with the MMSE ( $r = 0.38$ ). MoCA memory domain was significantly associated with NPA domains of delayed memory ( $r = 0.61$ ), executive ( $r = 0.51$ ), language ( $r = 0.45$ ), and immediate memory ( $r = 0.37$ ). MoCA visuospatial domain was significantly associated NPA domains of visuospatial ( $r = 0.59$ ), executive ( $r = 0.46$ ), and immediate memory ( $r = 0.37$ ). MoCA language domain significantly associated with NPA domains of executive ( $r = 0.55$ ) and attention ( $r = 0.53$ ). MoCA executive domain significantly associated with NPA of delayed memory ( $r = 0.49$ ), executive ( $r = 0.47$ ), attention ( $r = 0.37$ ), immediate memory ( $r = 0.42$ ), and visuospatial ( $r = 0.35$ ). MoCA attention/orientation was not significantly associated with any neuropsychological domain. ACE-R memory domain significantly associated with NPA domains of delayed memory ( $r = 0.67$ ), executive ( $r = 0.50$ ), and immediate memory ( $r = 0.48$ ). ACE-R visuospatial domain significantly associated with NPA executive

(Continued)



Table 5. (Continued)

First author (year)	Comparisons	Summary of main findings*
Rojo-Mota et al. (2017)	Spearman's correlations between the LOTCA scale scores and scale scores obtained with the NPA	<p>(<math>r = 0.64</math>), visuospatial (<math>r = 0.59</math>), delayed memory (<math>r = 0.48</math>), immediate memory (<math>r = 0.42</math>), and attention (<math>r = 0.35</math>). ACE-R language significantly associated with immediate memory (<math>r = 0.56</math>), executive (<math>r = 0.41</math>), and visuospatial (<math>r = 0.37</math>). ACE-R attention/orientation significantly associated with NPA attention (<math>r = 0.39</math>) and executive (<math>r = 0.37</math>). ACE-R fluency domain significantly associated with NPA language (<math>r = 0.73</math>), delayed memory (<math>r = 0.68</math>), immediate memory (<math>r = 0.61</math>), executive (<math>r = 0.60</math>), visuospatial (<math>r = 0.44</math>), and attention (<math>r = 0.39</math>).</p> <p>LOTCA orientation subscale significantly associated with the BCSE (<math>r_s = 0.26</math>), ROCF Copy (<math>r_s = 0.33</math>), Digit Span (<math>r_s = 0.33</math>), Symbol Span (<math>r_s = 0.28</math>), and Logical Memory I (<math>r_s = 0.27</math>). LOTCA Visuo-perceptual subscale significantly associated with ACLS-5 (<math>r_s = 0.34</math>), BCSE (<math>r_s = 0.55</math>), ROCF Copy (<math>r_s = 0.47</math>) and Immediate Recall (<math>r_s = 0.44</math>), WAIS-IV Coding (<math>r_s = .51</math>), Symbol Search (<math>r_s = .46</math>), PSI (<math>r_s = 0.54</math>), Digit Span (<math>r_s = 0.68</math>), WMS-IV Symbol Span (<math>r_s = 0.50</math>), Logical Memory II (<math>r_s = 0.27</math>), FDT Reading (<math>r_s = 0.41</math>), FDT Counting (<math>r_s = 0.25</math>), and FDT Choosing (<math>r_s = 0.35</math>). LOTCA Psychomotor abilities significantly correlated with ACLS-5 (<math>r_s = 0.53</math>), BCSE (<math>r_s = 0.54</math>), ROCF Copy (<math>r_s = 0.55</math>) and Immediate Recall (<math>r_s = 0.56</math>), WAIS-IV Coding (<math>r_s = 0.59</math>), Symbol Search (<math>r_s = 0.64</math>), PSI (<math>r_s = 0.68</math>), Digit Span (<math>r_s = 0.59</math>), WMS-IV Symbol Span (<math>r_s = 0.52</math>), Logical Memory II (<math>r_s = 0.37</math>), FDT Reading (<math>r_s = 0.33</math>), Counting (<math>r_s = 0.31</math>), and Switching (<math>r_s = 0.25</math>). LOTCA thinking operations significantly associated with ACLS-5 (<math>r_s = 0.48</math>), BCSE (<math>r_s = 0.51</math>), ROCF Copy (<math>r_s = 0.43</math>) and Immediate Recall (<math>r_s = 0.41</math>), WAIS-IV Coding (<math>r_s = .59</math>), Symbol Search (<math>r_s = 0.50</math>), PSI (<math>r_s = 0.60</math>), Digit Span (<math>r_s = 0.51</math>), WMS-IV Symbol Span (<math>r_s = 0.54</math>), Logical Memory I (<math>r_s = 0.30</math>), and Logical Memory II (<math>r_s = 0.34</math>). LOTCA execution time significantly associated with ACLS-5 (<math>r_s = -0.44</math>), BCSE (<math>r_s = -0.58</math>), ROCF Copy (<math>r_s = -0.32</math>) and Immediate Recall (<math>r_s = -0.31</math>), WAIS-IV Symbol Search (<math>r_s = -0.37</math>), Digit Span (<math>r_s = -0.41</math>), WMS-IV Symbol Span (<math>r_s = -0.31</math>), Logical Memory II (<math>r_s = -0.30</math>), and FDT Switching (<math>r_s = -0.29</math>).</p>
To et al. (2015)	Pearson correlations between the HASI total and subscores, and WAIS-III FSIQ and subtest scores	<p>Significant correlation between the total HASI score and WAIS-III FSIQ score (<math>r = 0.69</math>). Significant correlations between the correlations between the HASI subtests and the WAIS-III subtests were all significant: background information (<math>r = 0.58</math>), spelling (<math>r = 0.50</math>), puzzle (<math>r = 0.46</math>), and clock drawing (<math>r = 0.45</math>). The HASI also correlated significantly with the verbal subscale of the WAIS-III (<math>r = 0.70</math>) and the performance subscale of the WAIS-III (<math>r = 0.63</math>).</p>

\*Included if at least significant at the .05 level or higher.

ACE-R, Addenbrooke's Cognitive Examination-Revised; ACLS-5, Allen Cognitive Level Screen-Fifth Edition; BCSE, Brief Cognitive Status Examination; BEAT, Brief Executive Function Assessment Tool; DSF, Digit Span Forward; DSS, Digit Span Sequencing; FAB, Frontal Assessment Battery; FDT, Five-Digit Test; FPT, Five-Point Test; FSIQ, Full Scale Intelligence Quotient; HASI, Hayes Ability Screening Index; LOTCA, Loewenstein Occupational Therapy Cognitive Assessment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; NAB-SM, Neuropsychological Assessment Battery-Screening Module; NPA, neuropsychological assessment; PSI, Processing Speed Index; RBANS, Repeatable Battery for the Assessment of Neuropsychological Status; ROCF, Rey-Osterrieth Complex Figure; TMT, Trail Making Test; SCWT, Stroop and Colour Word Test; WAIS-III, Wechsler Adult Intelligence Scale-Third Edition; WCST, Wisconsin Card Sorting Test; WMS-IV, Wechsler Memory Scale-Fourth Edition.

**Table 6.** QUADAS–2 ratings for studies included in the review

First author, year	Domain 1: Patient selection		Domain 2: Index test(s)		Domain 3: Reference standard		Domain 4: Flow and timing
	Risk of bias	Applicability	Risk of bias	Applicability	Risk of bias	Applicability	Risk of bias
Berry et al., 2021	LR	LR	?	LR	?	LR	LR
Braatveit et al., 2018	HR	LR	LR	LR	?	HR	HR
Bruijnen et al., 2019	?	LR	LR	LR	?	LR	?
Copersino et al., 2009	HR	LR	HR	LR	HR	LR	LR
Cunha et al., 2010	HR	LR	?	LR	?	LR	LR
Ewert et al., 2018	HR	LR	LR	LR	LR	LR	LR
Fals-Stewart, 1997	HR	LR	?	LR	?	LR	LR
Grohan and Fals-Stewart 2004	HR	LR	?	LR	?	LR	LR
Pelletier et al., 2018	HR	LR	LR	LR	LR	LR	?
Ridley et al., 2018	HR	LR	LR	LR	HR	LR	LR
Ritz et al., 2015	HR	LR	HR	LR	?	LR	HR
Rojo-Mota et al., 2017	HR	LR	LR	LR	LR	LR	HR
To et al., 2015	HR	LR	HR	LR	HR	HR	LR
Wester et al., 2013	HR	LR	LR	LR	LR	LR	HR

QUADAS–2, Quality Assessment of Diagnostic Accuracy Studies – Version 2; HR, high risk; LR, low risk; ? = unclear risk.

generally small, with the exception of Berry and colleagues (2021), who commendably captured a large sample.

Across the 14 included studies, 10 different cognitive screening tools were evaluated. The majority of studies examined only one measure, and all were paper and pen tasks. The range of cognitive domains and tasks assessed by the screening tool and length of administration varied widely across studies. Most studies sought to validate existing cognitive screening tools in this new population, primarily the MoCA, with two seeking to validate new measures for specific use in SUD populations (BEAT and BEARNI). In over a quarter of studies, a particular syndrome of interest was evaluated within a SUD sample, with all other studies seeking to identify CI more generally.

### Aim 3: Classification Accuracy

Out of the 14 retrieved studies, 12 reported on classification accuracy data. All reported sensitivity and specificity that varied widely (range: 36%–100%, 2%–100%, respectively). A further two used correlational data to assess associations between screening tools and reference measures. Of the 10 studies that provided overall AUC data, the six that related to the MoCA indicated an *acceptable* to *outstanding* ability to detect CI versus respective reference standards. By contrast, the MMSE, another well-known cognitive screening measure, yielded much lower accuracy that ranged from *no discrimination* to *acceptable* across two studies.

Although the HASI exhibited *excellent* discriminant ability (Braatveit et al., 2018; To et al., 2015), this measure was designed to screen specifically for intellectual disability and the reference standards in both instances were restricted to an intellectual disability diagnosis or intellectual cognitive

domains. Therefore, classification accuracy of the HASI for detecting CI more broadly in individuals with SUDs remains unknown.

Berry and colleagues (2021) reported the most extensive psychometric data on the BEAT, a measure specifically developed for use in those with SUDs. The BEAT demonstrated excellent discriminative ability and good sensitivity and specificity at the recommended cut-score of  $\leq 24$ . However, low prevalence of CI in this group (7.8%) contributed to a low positive predictive power (Abu-Akel, Bousman, Skafidas, & Pantelis, 2018). The authors suggest this was due to using a conservative reference standard (i.e., impairment across three tests); however, it is also possible that this relates to sample characteristics (residential participants only), as well as choice of the reference battery (executive tasks only). Given that the prevalence rate in this study is much lower than that in other treatment groups and the SUD literature more broadly (e.g., Copersino et al., 2009), further evaluation of the psychometric properties of the BEAT in broader SUD populations is necessary.

Another cognitive screening tool specifically designed for SUDs, the BEARNI, exhibited good discriminative abilities in the original study but a follow-up analysis demonstrated that despite good sensitivity, the specificity of the tool was poor (Pelletier et al., 2018). To this end and given superior MoCA data, Pelletier and colleagues (2018) suggest that the MoCA may be a more suitable choice for routine use.

Less than half of studies provided predictive value data for their screening tools. These data take into account the prevalence of CI in the population of interest. Including this information is crucial as in the event of a mismatch between the prevalence of CI in the screening sample and in individuals attending SUD treatment, classification data may

misrepresent a tool's screening ability (Slater & Young, 2013). Of the studies that provided predictive data, the rate of CI across samples varied widely (8%–86%), which is unsurprising given the broad range of study settings and samples. Clinicians must be mindful of the expected prevalence of CI in their particular cohort, as screening tools examined in groups with different prevalence rates may differentially capture CI in their cohort (Molinari, 2015).

Two studies relied on correlational data to examine screening measures in individuals with SUD. Medium to large correlations were reported between index and reference measures on the FAB (Cunha et al., 2010) and LOTCA (Rojo-Mota et al., 2017). However, not all domains corresponded, and no further classification data were provided to support these associations, resulting in only indirect inferences regarding criterion validity. As such, further evaluation of these screening measures is required. There were more consistent medium to large correlations between MoCA scores (total and domain) and reference measures, providing initial evidence for construct validity. The ACE-R, BEAT, and HASI also had promising correlational data between index and reference measures; however, these findings need to be replicated.

#### Aim 4: Risk of Bias

This review represents the first time that the methodological quality of studies evaluating cognitive screening tools in SUDs have been systematically reviewed. The studies included in this review were assessed overall to have high or unclear risk of bias according to the QUADAS-2, suggesting significant methodological limitations or insufficient reporting. Patient selection represented the most problematic domain with high or unclear risk of bias occurring in 85.71% of studies. This was primarily due to the use of convenience sampling methods and overly stringent exclusion criteria. This included exclusion of individuals who had a history of psychiatric illness, had head injuries, or were on opioid treatment programmes. As individuals with SUDs present clinically with a range of comorbid and co-occurring conditions (see Gooden et al., 2020), excluding participants based on these factors has the potential to significantly limit translation of study findings to applied clinical settings.

An additional source of high or unclear risk of bias related to whether index measures were interpreted without the knowledge of the reference standard (i.e., blind ratings), or vice versa, an issue in almost three-quarters of studies. Administration of both measures by the same individual is problematic, as it can inadvertently lead to bias in administration and scoring of measures. This is a clear area in need of improvement in future studies. There were also issues with flow and timing, with nearly half of studies classified as having high or unclear risk of bias. This was primarily due to inappropriate or unclear time intervals between the administration of the index test and reference standard. In a SUD population, administration of both the index and reference

standard at a close time interval is critical given the potential of ongoing cognitive recovery with continued abstinence (Stavro, Pelletier, & Potvin, 2013). Other issues with flow and timing included not all participants receiving the reference standard, or not accounting for missing participant data, creating an unknown amount of potential bias. It is possible that participants who are more cognitively impaired may drop out more readily due to increased complexity of tasks. It is essential that all aspects of participant flow are tracked, including those invited, those that accept, and those that complete the study, in order to establish whether the population sample reflects the population at large.

The selection of the reference standard was also a source of discrepancy across studies. Given the aim of the study was to include *all* studies examining CI in general amongst individuals with SUD, three studies focused on specific sub-populations (intellectual disability, Korsakoff syndrome; Braatveit et al., 2018; Wester et al., 2013; To et al., 2015). For this reason, the reference standard used for classification was specific and focused, raising concerns on the applicability of identifying CI more broadly in individuals with SUDs.

Another concern regarding reference standards was inconsistency in impairment ratings. Some studies operationalised impairment as ranging from  $-1$  to  $-2$  standard deviations below normative data (the equivalent to a difference of 15th percentile versus 2nd percentile). Other studies required impairment on a certain number of tasks, by domain, or an average overall score. What constitutes psychometric impairment is a source of debate in the neuropsychological community (Schoenberg & Rum, 2017), and this inconsistency is reflected in the studies included. Interestingly, Godefroy et al. (2014) analysed various methods of classifying impairment and suggested that a cut-off of the 5th percentile is most appropriate. A consensus on how to characterise impairment is needed if analyses of screening measures are to be accurately compared.

Two studies (Berry et al., 2021; Wester et al., 2013) presented a unique methodological challenge that was not captured via the QUADAS-2. In both studies the reference standard focused on a limited area of cognitive functioning (executive functioning and memory, respectively), while the index measures evaluated cognitive function more broadly. In both instances, despite the apparent differences in screening and reference tasks, screening measures were able to detect individuals impaired on the reference standard at an acceptable level ( $AUC \geq 75$ ). This suggests that the BEAT and MoCA were able to adequately detect executive functioning and memory impairments, respectively, despite assessing a broader range of skills. It remains unclear how accurately these screening measures detect impairment in other domains. Use of a comprehensive reference standard that includes a range of cognitive domains would allow evaluation of both how well these screens pick up on a range of impairments other than the domain of interest and establish further construct validity.

Overall, findings suggest that clear adherence to STARD criteria was generally poor in studies evaluating the utility of

cognitive screening tools in adults with SUDs (Bossuyt et al., 2015; Bossuyt et al., 2003; Cohen et al., 2016). The current results reflect the methodological limitations of the studies in this area to date; not one study was classified as low risk across all domains. This primarily relates to insufficient reporting of study procedure and data as in many cases domains were rated unclear, resulting in risk being unable to be determined. This is a novel and important finding. Not only does it provide direction for further research, but it highlights caution that should be taken in interpretation of the findings to date, given that studies with risk of bias are likely to over- or under-estimate diagnostic accuracy.

## Limitations

The current study intentionally set broad limits on search strategy and inclusion criteria to capture as many studies as possible. A key limitation of applying the QUADAS-2 to evaluate risk of bias in captured studies includes dealing with studies that omit key methodological details versus those which report on these more comprehensively and which may be disproportionately penalised as a result. For example, a number of studies did not report how many individuals were approached versus declined to participate, and the reasons for refusal. This could have potentially introduced bias in the recruitment and exclusion of participants in some studies; however, with omission of this information, it cannot be evaluated.

Another limitation is that a comprehensive grey literature search was not incorporated in the search strategy. This was partially mitigated by searching the Psychology Preprint Archive, as well as through inclusion of a publication that was under review and subsequently published (Berry et al., 2021). A more comprehensive grey literature search would have enriched the evidence base for this review by including findings from a broader range of sources, as well as reduce potential publication bias (Paez, 2017).

## Research Recommendations

While in a research context a homogeneous sample may be useful for interpreting study findings, this does not reflect the rich diversity and heterogeneity of individuals presenting for SUD treatment in health settings. To maximise ecological validity and translation of study findings to applied clinical practice, future studies seeking to validate cognitive screening tools in adults with SUDs should seek to adhere to STARD criteria (Bossuyt et al., 2015; Bossuyt et al., 2003), while also maximising ecological validity. This includes keeping exclusion criteria to a minimum, such as instances where reduced capacity to provide fully informed consent is evident, or there are acute medical, psychiatric, or other risks that warrant priority treatment. This ensures the study sample is more representative of individuals presenting for the treatment of SUD and captures the potential wide range of co-occurring conditions that may impact

cognition in a real-world setting. Studies should also clearly report on sample characteristics (including study setting, substances of misuse, abstinence status) to assist the clinician in understanding whether the cognitive screening tool may be useful for a particular client group. Additionally, including comprehensive classification accuracy statistics supports a greater understanding of the utility of the particular cognitive screening measure. Future studies should: use blinding to ensure objectivity of results; not integrate control with clinical data (as this can conflate results); and provide a clear description of study flow; accounting for any missing data. Future systematic reviews should also adhere to the recently updated PRISMA guidelines (Page et al., 2020), which were not available at the time of the current review. Given the wide number of cognitive screening tools available and critiqued in this review, some of which show good promise in use in a drug and alcohol clinical setting (e.g., the MoCA), comprehensively validating existing tools should be an initial first step.

## Clinical Recommendations

CI is frequent in individuals with substance use and, given the negative impact on treatment outcomes, should be screened for on treatment entry and periodically over treatment engagement (Bates, Buckman, & Nguyen, 2013). A cognitive screen used in this context must also be able to capture a range of cognitive disorders, given that the heterogeneity of contributing factors (e.g., ABI, learning or intellectual disability) forms part of a typical client's presentation. Of the tools reviewed, the MoCA was the most studied and demonstrated the most consistently adequate diagnostic classification accuracy in this population. Practically, the MoCA is brief to administer, evaluates a broad range of cognitive domains, and is used widely across healthcare settings, which may aid in effective communication of health information. It also has alternate versions available, which facilitate re-testing over short periods of time (e.g., at the beginning and end of a detoxification stay, and then entry into outpatient clinics). While the MMSE is another well-known cognitive screening measure, it failed to demonstrate adequate classification accuracy in SUD populations and therefore cannot be recommended for clinical use in this setting. Other measures (e.g., ACE-R and BEAT) have initial promise for use in this area but require further empirical validation before they can be recommended for wide clinical use.

Regardless of the specific screening tool, a key consideration is the risk of over-reliance on a cut-off score alone to determine presence or absence of CI. A *cognitive screen is not diagnostic* and can only flag clients who may be in need of further assessment, and *understanding the context in which the screen was administered is essential* in an applied clinical setting. For example, capturing whether a client who performed below the cut-off score on a particular cognitive screen was acutely intoxicated, distracted, or in an uncharacteristically high state of distress is not an uncommon

dilemma, and are all relevant details when deciding on clinical recommendations. Incorporating this information can reduce excessive use of resources but is also heavily reliant on clinical acumen. Given the complexities of clinical presentation and subtlety in interpreting cognitive screening results in adults with SUDs, it is recommended that a clinician skilled in the cognitive field provides oversight of cognitive screening processes within this setting, with clinical neuropsychologists ideally situated to perform such a role. Additionally, the development of clear clinical pathways within a service that provide further assessment and comprehensive support for clients identified as at risk of CI is also essential.

## Conclusions

The current review provides the first systematic overview of cognitive screening tools utilised in adult SUD populations as compared to an accepted standard criterion of cognitive functioning (e.g., neuropsychological assessment). A strength of this review was the broad inclusion criteria; that is, that CI from a range of aetiologies and client populations was included. The MoCA was the most frequently reported cognitive screening tool, and it exhibited promising classification accuracy data across studies, suggesting the potential for use within this population. However, based on our search strategy and inclusion criteria, studies examining cognitive screening in adults with SUDs were significantly limited in number and quality. Other tools indicate initial promising classification data (ACE-R and BEAT); however, these also require further validation. Overall, given the very small number of studies, diverse sample characteristics, wide range of index measures and reference standards used, as well as high risk of bias identified amongst studies, further evaluation with stronger and more transparent methodological design and reporting is required. Without a clearly established tool to screen for CI, there is the potential that CI may not be systematically evaluated within SUD services, and this may negatively impact individual treatment outcomes. Crucially, future studies need to focus on adhering to the STARD criteria (Bossuyt et al., 2015; Bossuyt et al., 2003), while also seeking to maximise ecological validity. Positively, there remains promise in the measures utilised to date such as the MoCA, and it is pleasing to see the development of measures specifically designed for use in the SUD population (e.g., Berry et al., 2021; Ritz et al., 2015). Future studies should continue to build on the learnings of existing studies to ensure evidence-based measures are available to direct treatment and care of individuals who have comorbid substance use and cognitive needs.

## SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S135561772100103X>

## ACKNOWLEDGEMENTS

The authors thank Hunter New England and South East Sydney Local Health District Drug and Alcohol Services for their support in producing this review.

## AUTHOR CONTRIBUTIONS

JK and NR contributed equally to conception, design, interpretation, and writing of the manuscript. KK was involved in analysis, interpretation, and writing of the manuscript. AS designed and conducted the systematic search strategy. SB and KA contributed to design and revisions of the manuscript. All authors contributed to final manuscript review.

## FINANCIAL SUPPORT

KA is supported by a Career Development Fellowship from the National Health and Medical Research Council (NHMRC; APP1141207). KK, NR, SB, AS, and JK have no financial support to disclose.

## CONFLICTS OF INTEREST

The authors have nothing to disclose.

## ETHICAL STANDARDS

This study did not involve human subject participation.

## REFERENCES

- Abu-Akel, A., Bousman, C., Skafidas, E., & Pantelis, C. (2018). Mind the prevalence rate: overestimating the clinical utility of psychiatric diagnostic classifiers. *Psychological Medicine*, *48*, 1225–1227. doi: [10.1017/S0033291718000673](https://doi.org/10.1017/S0033291718000673)
- Aharonovich, E., Amrhein, P.C., Bisaga, A., Nunes, E.V., & Hasin, D.S. (2008). Cognition, commitment language, and behavioral change among cocaine-dependent patients. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, *22*, 557–562. doi: [10.1037/a0012971](https://doi.org/10.1037/a0012971)
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.
- American Psychological Association Practice Organization. (2014). Distinguishing between screening and assessment for mental and behavioral health problems: a statement from an American Psychological Association Practice Organization work group on screening and psychological assessment. Retrieved from: <https://www.apaservices.org/practice/reimbursement/billing/assessment-screening>
- Bates, M.E., Bowden, S.C., & Barry, D. (2002). Neurocognitive impairment associated with alcohol use disorders: implications for treatment. *Experimental and Clinical Psychopharmacology*, *10*, 193–212. doi: [10.1037//1064-1297.10.3.193](https://doi.org/10.1037//1064-1297.10.3.193)
- Bates, M.E., Buckman, J.F., & Nguyen, T.T. (2013). A role for cognitive rehabilitation in increasing the effectiveness of treatment for alcohol use disorders. *Neuropsychology Review*, *23*, 27–47. doi: [10.1007/s11065-013-9228-3](https://doi.org/10.1007/s11065-013-9228-3)

- Berry, J., Shores, E.A., Nardo, T., Sedwell, A., Lunn, J., Marceau, E., . . . Batchelor, J. (2021). The brief executive-function assessment tool: a new cognitive impairment screening tool for alcohol and other drug services. *Applied Neuropsychology: Adult*. doi: [10.1080/23279095.2021.1895791](https://doi.org/10.1080/23279095.2021.1895791)
- Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L., . . . Cohen, J.F. (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ: British Medical Journal*, *351*, h5527. doi: [10.1136/bmj.h5527](https://doi.org/10.1136/bmj.h5527)
- Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L.M., . . . Lijmer, J.G. (2003). The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Annals of Internal Medicine*, *138*, W1–W12.
- Braatveit, K.J., Torsheim, T., & Hove, O. (2018). Screening for intellectual disabilities: a validation of the Hayes Ability Screening Index for in-patients with substance use disorder. *Nordic Journal of Psychiatry*, *72*, 387–392. doi: [10.1080/08039488.2018.1468923](https://doi.org/10.1080/08039488.2018.1468923)
- Bruijnen, C.J., Jansen, M., Dijkstra, B.A., Walvoort, S.J., Lugtmeijer, S., Markus, W., . . . & Kessels, R.P. (2019). The Montreal Cognitive Assessment (MoCA) as a cognitive screen in addiction health care: a validation study for clinical practice. *Journal of Substance Use*, *24*, 47–54. doi: [10.1080/14659891.2018.1497102](https://doi.org/10.1080/14659891.2018.1497102)
- Brodersen, C., Koen, E., Ponte, A., Sánchez, S., Segal, E., Chiappella, A., . . . & Lemberg, A. (2014). Cognitive function in patients with alcoholic and nonalcoholic chronic liver disease. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *26*, 241–248. doi: [10.1176/appi.neuropsych.12040091](https://doi.org/10.1176/appi.neuropsych.12040091)
- Buckman, J.F., Bates, M.E., & Cisler, R.A. (2007). Social networks and their influence on drinking behaviors: differences related to cognitive impairment in clients receiving alcoholism treatment. *Journal of Studies on Alcohol and Drugs*, *68*, 738–747. doi: [10.15288/jsad.2007.68.738](https://doi.org/10.15288/jsad.2007.68.738)
- Buckman, J.F., Bates, M.E., & Morgenstern, J. (2008). Social support and cognitive impairment in clients receiving treatment for alcohol- and drug-use disorders: a replication study. *Journal of Studies on Alcohol and Drugs*, *69*, 738–746. doi: [10.15288/jsad.2008.69.738](https://doi.org/10.15288/jsad.2008.69.738)
- Bunnage, M. (2017). How do I know when a diagnostic test works? In S. Bowden (Ed.), *Neuropsychological assessment in the age of evidence-based practice: Diagnostic and treatment evaluation* (pp. 223–238). New York, NY: Oxford University Press.
- Carroll Chapman, S.L. & Wu, L.T. (2012). Substance abuse among individuals with intellectual disabilities. *Research in Developmental Disabilities*, *33*, 1147–1156. doi: [10.1016/j.ridd.2012.02.009](https://doi.org/10.1016/j.ridd.2012.02.009)
- Chen, L.Y., Strain, E.C., Crum, R.M., & Mojtabai, R. (2013). Gender differences in substance abuse treatment and barriers to care among persons with substance use disorders with and without comorbid major depression. *Journal of Addiction Medicine*, *7*, 325–334. doi: [10.1097/ADM.0b013e31829b7afe](https://doi.org/10.1097/ADM.0b013e31829b7afe)
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J.F., Korevaar, D.A., Altman, D.G., Bruns, D.E., Gatsonis, C.A., Hooft, L., . . . Bossuyt, P.M. (2016). STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*, *6*, e012799. doi: [10.1136/bmjopen-2016-012799](https://doi.org/10.1136/bmjopen-2016-012799)
- Copersino, M.L., Fals-Stewart, W., Fitzmaurice, G., Schretlen, D.J., Sokoloff, J., & Weiss, R.D. (2009). Rapid cognitive screening of patients with substance use disorders. *Experimental and Clinical Psychopharmacology*, *17*, 337–344. doi: [10.1037/a0017260](https://doi.org/10.1037/a0017260)
- Copersino, M.L., Schretlen, D.J., Fitzmaurice, G.M., Lukas, S.E., Faberman, J., Sokoloff, J., & Weiss, R.D. (2012). Effects of cognitive impairment on substance abuse treatment attendance: predictive validation of a brief cognitive screening measure. *The American Journal of Drug and Alcohol Abuse*, *38*, 246–250. doi: [10.3109/00952990.2012.670866](https://doi.org/10.3109/00952990.2012.670866)
- Covidence Systematic Review Software. *Veritas Health Innovation*. Melbourne, Australia. Available at [www.covidence.org](http://www.covidence.org)
- Cunha, P.J., Nicastrì, S., de Andrade, A.G., & Bolla, K.I. (2010). The frontal assessment battery (FAB) reveals neurocognitive dysfunction in substance-dependent individuals in distinct executive domains: abstract reasoning, motor programming, and cognitive flexibility. *Addictive Behaviors*, *35*, 875–881. doi: [10.1016/j.addbeh.2010.05.005](https://doi.org/10.1016/j.addbeh.2010.05.005)
- Daley, D.C. & Moss, H.B. (2009). *Dual disorders: Counseling clients with Chemical Dependency and Mental Illness*. Center City: Simon and Schuster.
- Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB: a Frontal assessment battery at bedside. *Neurology*, *55*, 1621–1626. doi: [10.1212/wnl.55.11.1621](https://doi.org/10.1212/wnl.55.11.1621)
- Ewert, V., Pelletier, S., Alarcon, R., Nalpas, B., Donnadiou-Rigole, H., Trouillet, R., & Perney, P. (2018). Determination of MoCA cutoff score in patients with alcohol use disorders. *Alcoholism: Clinical and Experimental Research*, *42*, 403–412. doi: [10.1111/acer.13547](https://doi.org/10.1111/acer.13547)
- Fals-Stewart, W. (1997). Detection of neuropsychological impairment among substance-abusing patients: accuracy of the Neurobehavioral Cognitive Status Examination. *Experimental and Clinical Psychopharmacology*, *5*, 269.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.
- Godefroy, O., Gibbons, L., Diouf, M., Nyenhuis, D., Roussel, M., Black, S., Bugnicourt, J.M., & GREFEX Study Group. (2014). Validation of an integrated method for determining cognitive ability: implications for routine assessments and clinical trials. *Cortex*, *54*, 51–62. doi: [10.1016/j.cortex.2014.01.016](https://doi.org/10.1016/j.cortex.2014.01.016)
- Gooden, J.R., Cox, C.A., Petersen, V., Curtis, A., Manning, V., & Lubman, D.I. (2020). Characterisation of presentations to a community-based specialist addiction neuropsychology service: cognitive profiles, diagnoses and comorbidities. *Drug and Alcohol Review*. doi: [10.1111/dar.13135](https://doi.org/10.1111/dar.13135)
- Gould, T.J. (2010). Addiction and cognition. *Addiction Science & Clinical Practice*, *5*, 4–14.
- Grella, C.E., Hser, Y.I., Joshi, V., & Rounds-Bryant, J. (2001). Drug treatment outcomes for adolescents with comorbid mental and substance use disorders. *The Journal of Nervous and Mental Disease*, *189*, 384–392.
- Grohman, K. & Fals-Stewart, W. (2004). The detection of cognitive impairment among substance-abusing patients: the accuracy of the Neuropsychological Assessment Battery–screening module. *Experimental and Clinical Psychopharmacology*, *12*, 200. doi: [10.1037/1064-1297.12.3.200](https://doi.org/10.1037/1064-1297.12.3.200)
- Hagen, E., Erga, A.H., Hagen, K.P., Nesvåg, S.M., McKay, J.R., Lundervold, A.J., & Walderhaug, E. (2016). Assessment of executive function in patients with substance use disorder: a comparison of inventory- and performance-based assessment. *Journal of Substance Abuse Treatment*, *66*, 1–8. doi: [10.1016/j.jsat.2016.02.010](https://doi.org/10.1016/j.jsat.2016.02.010)

- Hayes, S.C. (2000). *Hayes Ability Screening Index (HASI) Manual*. Sydney: Faculty of Medicine, University of Sydney.
- Horner, M.D., Harvey, R.T., & Denier, C.A. (1999). Self-report and objective measures of cognitive deficit in patients entering substance abuse treatment. *Psychiatry Research*, *86*, 155–161.
- Katz, N., Itzkovich, M., Averbuch, S., & Elazar, B. (1989). Loewenstein occupational therapy cognitive assessment (LOTCA) battery for brain-injured patients: reliability and validity. *American Journal of Occupational Therapy*, *43*, 184–192.
- Kelly, T.M. & Daley, D.C. (2013). Integrated treatment of substance use and psychiatric disorders. *Social Work in Public Health*, *28*, 388–406. doi: [10.1080/19371918.2013.774673](https://doi.org/10.1080/19371918.2013.774673)
- Kiernan, R.J., Mueller, J., Langston, J.W., & Van Dyke, C. (1987). The Neurobehavioral Cognitive Status Examination: a brief but quantitative approach to cognitive assessment. *Annals of Internal Medicine*, *107*(4), 481–485. doi: [10.7326/0003-4819-107-4-481](https://doi.org/10.7326/0003-4819-107-4-481)
- Levin, F.R. (2007). Diagnosing attention-deficit/hyperactivity disorder in patients with substance use disorders. *Journal of Clinical Psychiatry*, *68*, 9–14.
- Lin, E., Balogh, R., McGarry, C., Selick, A., Dobranowski, K., Wilton, A.S., & Lunsby, Y. (2016). Substance-related and addictive disorders among adults with intellectual and developmental disabilities (IDD): an Ontario population cohort study. *BMJ Open*, *6*, e011638. doi: [10.1136/bmjopen-2016-011638](https://doi.org/10.1136/bmjopen-2016-011638)
- Mandrekar, J.N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, *5*, 1315–1316. doi: [10.1097/jto.0b013e3181ec173d](https://doi.org/10.1097/jto.0b013e3181ec173d)
- McGee, S. (2002). Simplifying likelihood ratios. *Journal of General Internal Medicine*, *17*, 646–649. doi: [10.1046/j.1525-1497.2002.10750.x](https://doi.org/10.1046/j.1525-1497.2002.10750.x)
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J.R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, *21*, 1078–1085. doi: [10.1002/gps.1610](https://doi.org/10.1002/gps.1610)
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine*, *6*, e1000097. doi: [10.1371/journal.pmed1000097](https://doi.org/10.1371/journal.pmed1000097)
- Molinario, A.M. (2015). Diagnostic tests: how to estimate the positive predictive value. *Neuro-Oncology Practice*, *2*, 162–166. doi: [10.1093/nop/npv030](https://doi.org/10.1093/nop/npv030)
- Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*, 695–699. doi: [10.1111/j.1532-5415.2005.53221.x](https://doi.org/10.1111/j.1532-5415.2005.53221.x)
- Owen, K.P., Sutter, M.E. & Albertson, T.E. (2014). Marijuana: respiratory tract effects. *Clinical Reviews in Allergy & Immunology*, *46*, 65–81. doi: [10.1007/s12016-013-8374-y](https://doi.org/10.1007/s12016-013-8374-y)
- Paez, A. (2017). Gray literature: an important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*, 233–240. doi: [10.1111/jebm.12265](https://doi.org/10.1111/jebm.12265)
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D. ... & Shamseer, L. (2020). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *MetaXiv*. <https://osf.io/preprints/metaarxiv/v7gm2/>
- Parry-Jones, B.L., Vaughan, F.L., & Miles Cox, W. (2006). Traumatic brain injury and substance misuse: a systematic review of prevalence and outcomes research (1994–2004). *Neuropsychological Rehabilitation*, *16*, 537–560. doi: [10.1080/09602010500231875](https://doi.org/10.1080/09602010500231875)
- Pelletier, S., Alarcon, R., Ewert, V., Forest, M., Nalpas, B., & Perney, P. (2018). Comparison of the MoCA and BEARNI tests for detection of cognitive impairment in in-patients with alcohol use disorders. *Drug and Alcohol Dependence*, *187*, 249–253. doi: [10.1016/j.drugalcdep.2018.02.026](https://doi.org/10.1016/j.drugalcdep.2018.02.026)
- Ramey, T. & Regier, P. (2019). Cognitive impairment in substance use disorders. *CNS Spectrums*, *24*, 102–113. doi: [10.1017/S1092852918001426](https://doi.org/10.1017/S1092852918001426)
- Ridley, N., Batchelor, J., Draper, B., Demirkol, A., Lintzeris, N., & Withall, A. (2018). Cognitive screening in substance users: diagnostic accuracies of the mini-mental state examination, Addenbrooke's Cognitive Examination-Revised, and Montreal Cognitive Assessment. *Journal of Clinical and Experimental Neuropsychology*, *40*, 107–122. doi: [10.1080/13803395.2017.1316970](https://doi.org/10.1080/13803395.2017.1316970)
- Ritz, L., Lannuzel, C., Boudehent, C., Vabret, F., Bordas, N., Segobin, S., ... & Beaunieux, H. (2015). Validation of a brief screening tool for alcohol-related neuropsychological impairments. *Alcoholism: Clinical and Experimental Research*, *39*, 2249–2260. doi: [10.1111/acer.12888](https://doi.org/10.1111/acer.12888)
- Roebuck-Spencer, T.M., Glen, T., Puente, A.E., Denney, R.L., Ruff, R.M., Hostetter, G., & Bianchini, K.J. (2017). Cognitive screening tests versus comprehensive neuropsychological test batteries: a National Academy of Neuropsychology Education Paper. *Archives of Clinical Neuropsychology*, *32*, 491–498. doi: [10.1093/arclin/acx021](https://doi.org/10.1093/arclin/acx021)
- Rojo-Mota, G., Pedrero-Pérez, E.J., Ruiz-Sánchez de León, J.M., León-Frade, I., Aldea-Poyo, P., Alonso-Rodríguez, M., ... & Morales-Alonso, S. (2017). Loewenstein occupational therapy cognitive assessment to evaluate people with addictions. *Occupational Therapy International*, 1–7. doi: [10.1155/2017/2750328](https://doi.org/10.1155/2017/2750328)
- Salavert, J., Clarabuch, A., Fernández-Gómez, M.J., Barrau, V., Giráldez, M.P., & Borràs, J. (2018). Substance use disorders in patients with intellectual disability admitted to psychiatric hospitalisation. *Journal of Intellectual Disability Research: JIDR*, *62*, 923–930. doi: [10.1111/jir.12514](https://doi.org/10.1111/jir.12514)
- Schoenberg, M.R. & Rum, R.S. (2017). Towards reporting standards for neuropsychological study results: a proposal to minimize communication errors with standardized qualitative descriptors for normalized test scores. *Clinical Neurology and Neurosurgery*, *162*, 72–79. doi: [10.1016/j.clineuro.2017.07.010](https://doi.org/10.1016/j.clineuro.2017.07.010)
- Schueler, S., Schuetz, G.M., & Dewey, M. (2012). The revised QUADAS-2 tool. *Annals of Internal Medicine*, *156*, 323–324. doi: [10.7326/0003-4819-156-4-201202210-00018](https://doi.org/10.7326/0003-4819-156-4-201202210-00018)
- Secades-Villa, R., García-Rodríguez, O., & Fernández-Hermida, J.R. (2015). Contingency management for substance use disorders in Spain: implications for research and practice. *Preventive Medicine*, *80*, 82–88. doi: [10.1016/j.ypmed.2015.07.001](https://doi.org/10.1016/j.ypmed.2015.07.001)
- Slater, H. & Young, J. (2013). A review of brief cognitive assessment tests. *Reviews in Clinical Gerontology*, *23*, 164–176. doi: [10.1017/s0959259813000038](https://doi.org/10.1017/s0959259813000038)
- Stavro, K., Pelletier, J., & Potvin, S. (2013). Widespread and sustained cognitive deficits in alcoholism: a meta-analysis. *Addiction Biology*, *18*, 203–213. doi: [10.1111/j.1369-1600.2011.00418.x](https://doi.org/10.1111/j.1369-1600.2011.00418.x)
- Stern, R.A., & White, T. (2003). *NAB, Neuropsychological Assessment Battery: Administration, scoring, and interpretation manual*. Lutz, FL: Psychological Assessment Resources.

- To, W.T., Vanheule, S., Vanderplasschen, W., Audenaert, K., & Vandeveldel, S. (2015). Screening for intellectual disability in persons with a substance abuse problem: exploring the validity of the Hayes Ability Screening Index in a Dutch-speaking sample. *Research in Developmental Disabilities, 36*, 498–504. doi: [10.1016/j.ridd.2014.10.046](https://doi.org/10.1016/j.ridd.2014.10.046)
- Trevethan, R. (2017). Sensitivity, specificity, and predictive values: foundations, pliabilitys, and pitfalls in research and practice. *Frontiers in Public Health, 5*. doi: [10.3389/fpubh.2017.00307](https://doi.org/10.3389/fpubh.2017.00307)
- Verdejo-Garcia, A., Garcia-Fernandez, G., & Dom, G. (2019). Cognition and addiction. *Dialogues in Clinical Neuroscience, 21*, 281–290. doi: [10.31887/DCNS.2019.21.3/gdom](https://doi.org/10.31887/DCNS.2019.21.3/gdom)
- Volkow, N.D., Baler, R.D., Compton, W.M., & Weiss, S.R. (2014). Adverse health effects of marijuana use. *The New England Journal of Medicine, 370*, 2219–2227. doi: [10.1056/NEJMra1402309](https://doi.org/10.1056/NEJMra1402309)
- Wester, A.J., Westhoff, J., Kessels, R.P., & Egger, J.I. (2013). The Montreal Cognitive Assessment (MoCA) as a measure of severity of amnesia in patients with alcohol-related cognitive impairments and Korsakoff syndrome. *Clinical Neuropsychiatry, 10*, 134–141.
- Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B., ... QUADAS–2 Group. (2011). QUADAS–2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine, 155*, 529–536. doi: [10.7326/0003-4819-155-8-201110180-00009](https://doi.org/10.7326/0003-4819-155-8-201110180-00009)