<The AHRQ-generated report cover goes on this page, color or B&W>

<Title page for draft reports>

Comparative Effectiveness Review

Number xx

Diagnosis of Attention-Deficit/Hyperactivity Disorder in Adults: A Systematic Review

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. xxx-xx-xxxx

Prepared by:
<Name> Evidence-based Practice Center
<City, State>

Investigators:
First and Last Names, X.X.
First and Last Names, X.X.

AHRQ Publication No. xx-EHCxxx
<Month Year>

This report is based on research conducted by the XXXXX Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-20XX-XXXXX). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

A representative from AHRQ served as a Contracting Officer's Representative and reviewed the contract deliverables for adherence to contract requirements and quality. AHRQ did not directly participate in the literature search, determination of study eligibility criteria, data analysis, interpretation of data, or preparation or drafting of this report.

Suggested citation: <Authors>. <Topic in Title Caps>. <Report Series Name in Title Caps No.> <#>. (Prepared by the <EPC Name> Evidence-based Practice Center under Contract No. <##>.) AHRQ Publication No. XX-EHCXXX-EF. Rockville, MD: Agency for Healthcare Research and

Quality. <Month Year>. Posted final reports are located on the Effective Health Care Program search page.
<doi>.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of healthcare in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new healthcare technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see https://effectivehealthcare.ahrq.gov/about/epc/evidence-synthesis

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the healthcare system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the website (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

If you have comments on this systematic review, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

<Name>
Director
Agency for Healthcare Research and Quality

<Name>.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Craig A. Umscheid, M.D., M.S.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

<Name>
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: XXX.

Key Informants

In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who provided input to this report follows:

<Name>
<Place>
<City>, <ST>

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who provided input to this report follows:

<Name>
<Place>
<City>, <ST>

<Name>
<Place>
<City>, <ST>

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers. AHRQ may also seek comments from other Federal agencies when appropriate.

Peer Reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential nonfinancial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

The list of Peer Reviewers follows:

<Name>
<Place>
<City>, <ST>

<Name>
<Place>
<City>, <ST>

Diagnosis of Attention-Deficit/Hyperactivity Disorder in Adults: A Systematic Review

Abstract

Objectives. Different diagnostic tests to diagnose attention-deficit/hyperactivity disorder (ADHD) in adults have been suggested. The review synthesizes the results of evaluations for available diagnostic tools.

Review methods. Following a detailed published protocol and informed by a technical expert panel, we reviewed the evidence for diagnostic tools. We searched multiple sources to identify evaluations of tools used for the diagnosis of ADHD in adults. Registration CRD42025638106.

Results. We identified 114 diagnostic studies. Studies evaluated self report questionnaires, peer review questionnaires, neuropsychological tests, neuroimaging, electroencephalogram, diverse biomarkers, clinician tools, combinations of modalities, and tools to identify feigning ADHD.

Conclusions. XXX

Contents

Tables

Figures

Appendixes

Executive Summary

1. Introduction

1.1 Background

Attention-deficit/hyperactivity disorder (ADHD) is characterized by persistent symptoms in the domains of inattention, hyperactivity, and impulsivity.[1] Clinically significant symptoms, especially inattention, persist into adulthood in most individuals.[1-5] The lifetime prevalence of ADHD is approximately 5.3%,[6] although epidemiological studies that have not required a childhood onset have suggested that its prevalence in adults may be as high as 6.7%.[7-10] Although many adults with ADHD adopt lifestyles that help compensate for their symptoms, they often need to exert excess energy to overcome impairments. They are often distressed by their inability to realize their full potential and by persistent symptoms of restlessness, erratic moods, and poor self-esteem.[11, 12] Impaired productivity because of poor time management, procrastination, and distractibility can limit work productivity and lower overall quality of life.[11] The mainstay of treatments for ADHD in adults are U.S. Food and Drug Administration (FDA)-approved medications, especially stimulant medications. More than 2% of US adults misuse stimulant medications, mostly to aid concentration and alertness,[13] but also as a form of substance abuse. A wide range of other treatments have been developed and may be effective, including psychotherapy, psychosocial interventions, cognitive training, neurofeedback, neuromodulation, and nutritional supplements. Even when ADHD treatments prove helpful, however, long-term adherence is often poor.[14-17] Whether treatments that improve the core symptoms also improve any problems associated with ADHD is unclear, and the long-term effectiveness in adults is generally unknown.

ADHD is most often first diagnosed in elementary or middle school age years or, less commonly, in high school or college when experiencing difficulty meeting academic demands. It can also be first diagnosed in adulthood, when impairments in attention, organization, and impulsivity produce recurrent problems with occupational, social, or family functioning. Adult diagnosis is often difficult because the outward manifestations most readily evident to others, hyperactivity and impulsivity, often improve during adolescence and no longer meet diagnostic criteria.[18] The symptoms of inattention (e.g., easy distractibility, poor organization, being "spacey," avoiding and trouble completing tasks that require sustained attention, losing things, forgetfulness) are more subtle and may not reach the level of obvious functional impairment until adulthood, within an occupational setting or a marriage.

The diagnosis of ADHD in adults, as in childhood, is complicated by the overlap of symptoms with other disorders.[19, 20] Attention and concentration, for example, can be impaired in persons who have depression, bipolar disorder, anxiety, psychosis, post-traumatic stress disorder, or substance abuse. Hyperactivity can be confused with anxiety-related behaviors and the excessive movements of tic and obsessive-compulsive disorders. Impulsivity is often prominent in bipolar and substance use disorders. The accurate diagnosis of adult ADHD is further complicated by a substantial and growing number of individuals, especially college students[21] and highly driven working professionals, who seek stimulant medications to aid cognitive performance. Stimulants have long been known to improve sustained attention and reduce distractibility in healthy individuals who do not have ADHD,[22-26] prompting some success-oriented individuals to feign symptoms to their doctors or on neuropsychological test assessments. Claims of exceptional diagnostic performance of these tools, the differing measures of performance, and the differing performance characteristics of different versions of a given tool,[27] are controversial and often confusing to clinicians, patients, and other stakeholders. In

addition, whether the performance of diagnostic tools varies with the characteristics of the ADHD participants or comparator sample is unknown.[28] These diagnostic challenges can complicate the accurate and reliable diagnosis of adult ADHD even for experienced mental health clinicians.

Despite established diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), diagnosing ADHD in adults remains challenging due to the frequent absence of hyperactivity and impulsivity symptoms, the subtlety of inattention symptoms, the inaccuracy of recall in adults for their retrospective assessments of ADHD symptoms in childhood (required to meet DSM-5 diagnostic criteria), and the common overlap with other mental health conditions.[18, 19, 29] The DSM-5 diagnostic criteria, developed primarily for children, may not be suitable for adult diagnosis, and the requirement of symptoms beginning before age 12 has been debated.[30-34] Furthermore, the accurate diagnosis of adult ADHD is complicated by the large number of individuals,[21-26] including healthy college students,[35, 36] who seek stimulant medications to improve cognitive performance, The absence of a true "gold-standard" diagnosis, the variability in performance of diagnostic tools among clinicians and settings, and the lack of clear practice guidelines further add to diagnostic complexity.[37-40]

The diagnosis of ADHD in adults is often made by primary care providers and nurse practitioners[41] rather than specialists, and the dispensing of ADHD medications has been increasing steadily, highlighting the need for effective diagnostic tools and guidelines.[33] The existing standards and guidelines for diagnosing ADHD in adults are limited, and the use of diagnostic tools and assessments varies widely in practice.[42-44] No guidelines for the diagnosis of adults with ADHD have thus far been developed in the United States, though one is in development.[45] The accuracy of diagnosis directly affects the management and treatment of ADHD, as well as the prevention of medication misuse. The diagnostic accuracy of tools and assessments used in adult ADHD diagnosis is unclear, however, and their performance may vary depending on the characteristics of the ADHD participants and comparator samples.[27, 28]

1.2 Purpose and Scope

This systematic review aims to provide a comprehensive and unbiased assessment of the diagnostic tools used to diagnose ADHD in adults to inform patients, clinicians, and policy makers. Initiated by the FDA, this Agency for Healthcare Research and Quality (AHRQ) report will document the available evidence. We will explore the effects of the setting and participant characteristics that may influence the diagnostic performance A contextual question is which tools are frequently being used in current clinical practice.

2. Methods

The systematic review followed a systematic review protocol that outlines the methods in detail.[46] The methodology followed the EPC Methods Guide.[47] The review is registered as Registration CRD42025638106. The project was supported by a multidisciplinary technical expert panel (TEP). The panel was designed to provide different perspectives of a broad group of stakeholders to ensure the evidence report on diagnosis of ADHD is relevant to a large audience. The panel includes experts specifically in adult ADHD and considered the needs of affected patients as well as family members.

2.2 Key Questions

The systematic review was guided by the following key questions:

- Key Question 1. What is the comparative diagnostic accuracy, unintended consequences and impact of tools that can be used in the primary care practice setting or by specialists to diagnose ADHD among adults?
    - Key Question 1a: How does the comparative diagnostic accuracy of these tools vary by clinical setting, including primary care or specialty clinic, or patient characteristics, including age, sex, cultural background, and risk factors associated with ADHD?

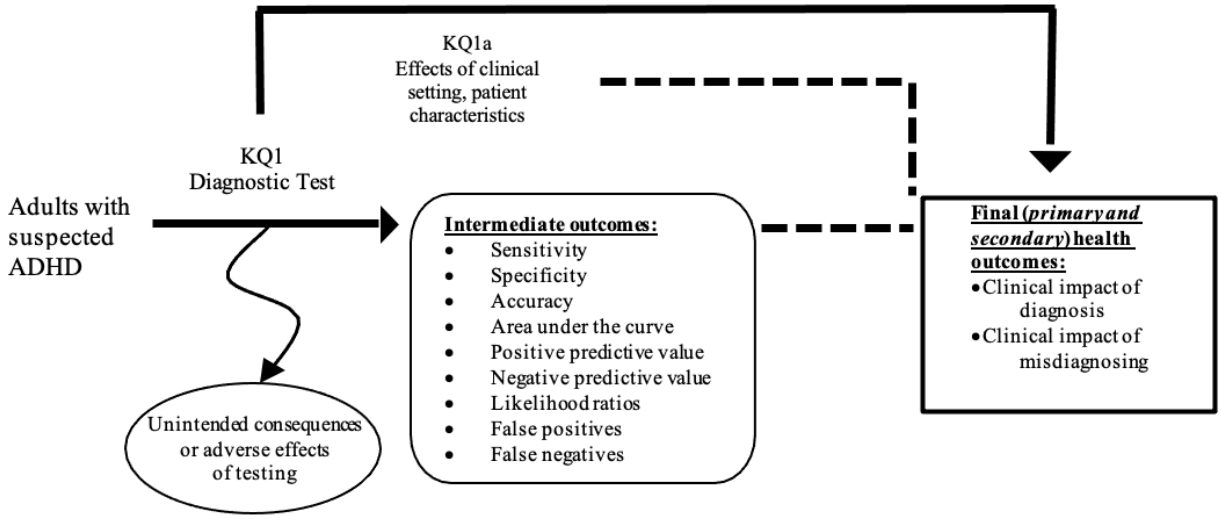
In addition, a contextual question provided additional information:

- Contextual Question. How frequently are the various tools to diagnose ADHD in adults currently being used?

2.3 Logic Model

Figure 1 illustrates the scope of the review.

Figure 1. Logic Model for Diagnosis of ADHD in Adults



Notes: ADHD attention deficit hyperactivity disorder, KQ key question

The model shows the population of interest (adults with suspected ADHD) and depicts the key question (diagnostic test performance) and sub-question (effect modifiers). The model also shows outcomes, ranging from side effects of the testing modality (e.g., relevant for invasive

tests), to intermediate outcomes such as the diagnostic accuracy established in the study, to final outcomes such as the impact of a diagnosing or misdiagnosing ADHD.

2.4 Search Strategy

The literature search for studies on diagnosing ADHD in adults used a combination of known tests to diagnose ADHD and general search terms for diagnostic accuracy studies to identify novel tools. We searched PubMed (biomedical literature), EMBASE (pharmacology emphasis), and PsycINFO (psychological research) without search date restriction and restricted to English language. The search strategy was peer reviewed within the EPC program. We used existing reviews for reference-mining; these were identified through the same databases plus searching the Cochrane Database of Systematic Reviews, Campbell Collaboration, and PROSPERO. We also searched the ECRI repository, G-I-N, and ClinicalKey for published guidelines and used these for reference-mining cited literature. All searches will be updated during the public comment period.

In addition, we leveraged technical experts to ensure that relevant research studies have been identified. We provided a list of included studies, together with all associated publications, and a list of excluded studies to facilitate this process. A Supplemental Evidence And Data for Systematic Reviews (SEADs) portal was available from January 10 to February 4 for this review. Additional data and publications suggested to us from any source, including peer and public review, will be screened applying the outlined eligibility criteria.

2.5 Inclusion/Exclusion Criteria

The eligibility criteria for this review are shown in Table 1.

Table 1. Eligibility Criteria

	Inclusion Criteria	Exclusion Criteria
Population	Adults 18 years and older with symptoms of ADHD and without the diagnosis of ADHD	Individuals 17 years of age or younger unless findings are reported separately for older participants
Intervention	Any ADHD diagnostic tool used for the diagnosis of ADHD in adults	Studies not reporting on diagnostic performance; non-English language questionnaires and interview guides
Comparator	Confirmation of diagnosis by a specialist (reference standard), such as a psychologist, psychiatrist or other healthcare provider using a well validated and reliable process of confirming a clinical diagnosis of ADHD	Comparison to diagnosis with another diagnostic instrument
Outcome	Diagnostic accuracy (e.g., sensitivity, specificity, accuracy, area under the curve, positive predictive value, negative predictive value, likelihood ratios, false positives, false negatives); unintended consequences and impact associated with diagnosing ADHD	Provider opinion of tests, cost without performance measure
Timing	Diagnostic follow-up must be completed before treatment is initiated	Any other timing
Setting	Primary or specialty care settings, including telehealth	Settings where diagnosis is for nonclinical or not research purposes
Study Design	Diagnostic accuracy studies	Editorials, nonsystematic reviews, letters, case series, case reports, pre-post studies. Systematic reviews were not eligible for inclusion but were retained for reference mining

Included ADHD tests were not limited to a set of pre-specified tools; instead, the review documents all tools that have been evaluated in the scientific literature and for which diagnostic accuracy evidence exists. There are no publication date restrictions. Studies with data exclusively published in non-English language publications were excluded to ensure transparency. We obtained all published reports providing data on a study (a study is defined by the included participants), including trial records and multiple publications, and consolidated the information into one study record.

2.5.1 Screening Process

We used an online database designed for systematic reviews to screen the literature search output. The team designed detailed citation and full text screening forms to ensure a transparent, consistent, and unambiguous approach. All citations were screened by two independent literature reviewers. Citations found to be potentially relevant by at least one reviewer were obtained as full text. All citations were also screened by a DistillerSR software machine learning algorithm trained by the human reviewers to ensure that no relevant citation was missed. Any citations identified as potentially relevant by the algorithm that were not selected for full text publication review were rescreened for relevance by an independent literature reviewer.

Full text screening applied the detailed eligibility criteria. Training ensured a shared understanding of all inclusion and exclusion criteria. Full text publications were screened by two independent reviewers to reduce errors and bias, and any discrepancy was resolved through discussion in the review team. The screening decisions and reasons for exclusion of publications were tracked in the online database and citation management software. These citations were shared with the technical expert panel and were documented with the review to ensure that the literature flow was transparent and objective.

2.6 Data Extraction

The data abstraction captured detailed information about eligible studies. One literature reviewer abstracted the data and an experienced methodologist checked the data for accuracy and completeness. We designed a detailed data extraction form in the software DistillerSR.

The data abstraction documented the targeted population, abstracted reported characteristics for all participants (participants with ADHD and those without). We documented the clinical setting, abstracted the method of establishing the reference standard (a clinical ADHD diagnosis), and abstracted diagnostic tool characteristics (format, name of the tool, employed cut offs, use of a training and validation set). We collected data for a diagnostic meta-analysis where possible (i.e., number of false positives, number of false negatives) along with the summary diagnosis accuracy measures reported by the authors such as sensitivity, specificity, area under the curve, positive predictive value. We differentiated between the diagnostic accuracy to diagnose ADHD and the diagnostic accuracy to detect faking ADHD. For all studies reporting multiple results, we selected the best accuracy performance model (either based on the authors' opinion, accuracy data, or trying to maximize sensitivity and specificity simultaneously).

2.7 Risk of Bias Assessment

The critical appraisal for individual studies applied criteria consistent with QUADAS 2.[48] QUADAS-2 evaluates four domains: *patient selection, index test* characteristics, *reference standard* quality, as well as *flow and timing:*

- Patient selection: The domain addresses whether the selection of patients could have introduced bias, taking into account whether the study enrolled a consecutive or random sample, whether the data are not based on a retrospective case-control design, and whether the study avoided inappropriate or problematic exclusions from the patient pool.
- Index test: The domain evaluates whether the conduct or interpretation of the test could have introduced bias, taking into account whether the results of the test were interpreted without knowledge of the results of the reference standard and whether any thresholds or cut-offs were pre-specified (e.g., instead of determined during the study to maximize diagnostic performance).
- Reference standard: The domain evaluates whether the reference standard, its conduct, or its interpretation may have introduced bias, taking into account the quality of the reference standard in correctly classifying the condition and whether the reference standard test results were interpreted without knowledge of the results of the index test.
- Flow and timing: The last domain evaluates whether the conduct of the study may have introduced bias. The assessment takes into account whether the interval between the test and the reference standard was appropriate, whether all patients received the reference standard and whether they received the same reference standard, and whether all patients were included in the analysis.

For each domain, we assessed the potential risk of bias in the study in order to identify high risk of bias and low risk of bias studies. One literature reviewer assessed risk of bias, and a methodologist reviewed individual studies and rating across studies to ensure accuracy and consistency of ratings. As outlined in the applicability section, we also evaluated for each study and appraisal domain whether there were concerns regarding the applicability of the study results to the review question. This encompassed whether the patients included in the studies matched the review question; whether the test, its conduct, or interpretation differed from the review question; or whether the target condition as defined by the reference standard fully matched the review question. The information was incorporated into the strength of evidence assessment.

2.8 Data Synthesis and Analysis

We answered the key question with the available evidence. We broadly differentiated diagnostic tools as

- Self reports
- Peer reports
- Neuropsychological tests
- Neuroimaging
- EEG
- Biomarker
- Observational data
- Clinician tools
- Combination predictions using more than one modality
- Tests to detect feigning of ADHD

We documented comparative effect results where studies compared the performance of more than one tool. In addition, we documented the range of results reported in studies within each tool category (e.g., self reports). We documented the diagnostic accuracy results for all outcomes as reported by the authors in the individual studies. Sensitivity estimates were documented

together with specificity estimates given that the estimates are not independent. A detailed evidence table displays key characteristics, the reference standard, psychometric properties and diagnostic accuracy outcomes for all included studies. In addition, we abstracted the number of true positives, true negatives, false positives, and false negatives where clearly reported for use in a diagnostic meta-analysis. All studies were considered for the narrative synthesis accompanying the summary of findings table.

We documented the results for available diagnostic tools across studies in a comprehensive summary of findings table documenting all assessed outcomes related to the diagnosis, reliability, and impact of the tool. Key outcomes for the summary of findings table were determined with the help of the TEP:

- Misdiagnosis (risk of missed condition that can appear as ADHD)
- Sensitivity
- Specificity
- Administration and scoring time
- Inter-rater reliability
- Costs
- Diagnostic concordance of primary care provider with specialist

The synthesis took study limitations and the risk of bias of individual studies contributing to estimates into account. In particular, we determined whether summary estimates corresponded to data reported in low risk of bias studies or were primarily based on high risk of studies.

To address the sub-question, we reported on subgroup results for different clinical settings (differentiating general and specialty care settings), patient characteristics (differentiating sex, age, cultural background, and comorbidity groups), and ADHD presentation (differentiating predominantly inattentive, hyperactive-impulsive, combined). We assessed whether these variables can explain heterogeneity identified in results across studies.

To address the contextual question we documented the frequency of identified research for each individual tool. In addition, we summarized data sources that reported on the frequency of tool use in clinical practice with emphasis on the U.S. healthcare setting in the discussion.

2.9 Grading the Strength of the Body of Evidence

We applied the EPC strength of evidence criteria to evaluate the body of evidence. In determining the quality of the body of evidence, the following domains were evaluated:

- Study limitations: The extent to which studies reporting on a particular outcome were likely to be protected from bias. The aggregate risk of bias across individual studies reporting an outcome was considered; graded as low, medium, or high level of study limitations.
- Inconsistency: The extent to which studies reported the same direction and/or magnitude of effects for a particular outcome; graded as consistent, inconsistent, or unknown (in the case of a single study or the absence of studies).
- Indirectness: Determines whether the test and the comparator were directly (i.e., within studies) or indirectly (e.g., across studies) compared. The domain was graded as direct or indirect.
- Imprecision: Describes the level of certainty of the estimate of effect for a particular outcome, where a precise estimate is one that allows a clinically useful conclusion. The domain was graded as precise or imprecise.

- Reporting bias: Publication bias, selective outcome reporting, and selective analysis reporting are types of reporting bias. Reporting bias is difficult to assess as systematic identification of unpublished evidence is challenging.

A final strength of evidence grade for each evidence statement was assigned by evaluating and weighing the combined results of the above domains. We differentiated an overall grade of high, moderate, low, or insufficient according to a four-level scale:

- High: We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable (i.e., another study would not change the conclusions).
- Moderate: We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- Low: We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
- Insufficient: We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available, or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

Summary tables included reasons for downgrading or upgrading the strength of evidence. The strength of evidence assessment documented uncertainty and communicated our confidence in the evidence statements that can be drawn from the literature.

2.10 Assessing Applicability

Results are based on the international literature and applicability ratings provided assessments regarding the generalizability of samples, settings, and tool results for U.S. clinical practice. For each study, we assessed the population included in the study to identify studies with narrow eligibility criteria (e.g., looking for a specific subgroup of ADHD participants only), studies that excluded participants with comorbidities, or studies that had more complex participants than typically seen in the community (e.g., dually diagnosed participants). We assessed whether studies described tools not used as recommended or commonly used in practice, the presence of highly trained test team or set up (e.g., analysis via complex machine learning models), or assessors that were not qualified for the assessment. We assessed whether the reference standard was ambiguous, different from standard clinical practice, or insufficiently described.

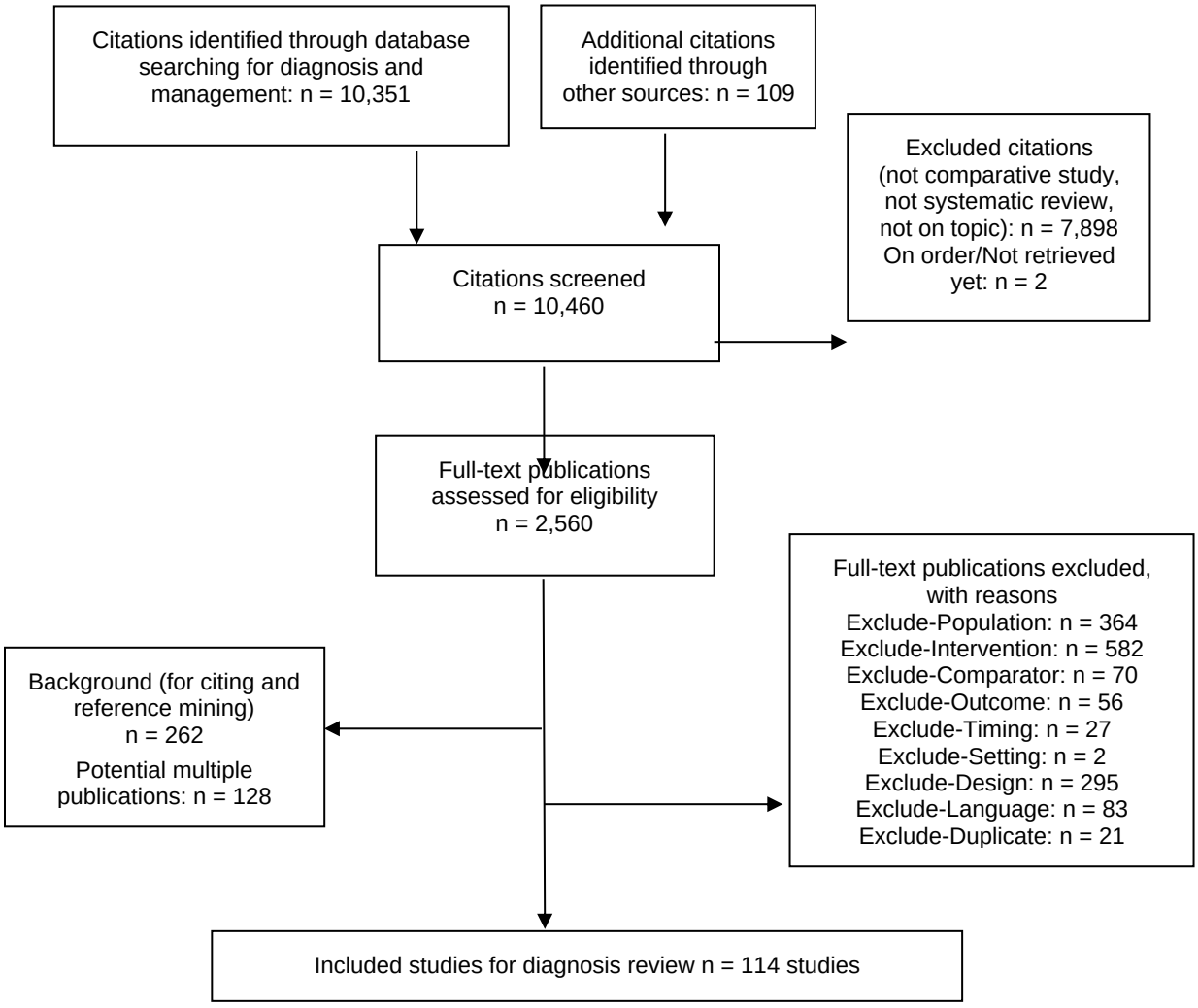
3. Results

The chapter is organized by the literature search results, the comparative diagnostic accuracy, results for individual tests, reporting on the diagnostic accuracy, unintended consequences, and information on the impact associated testing.

3.1 Results of Literature Search

The flow diagram documents the literature flow of the systematic review.

Figure 2. Study Flow Diagram

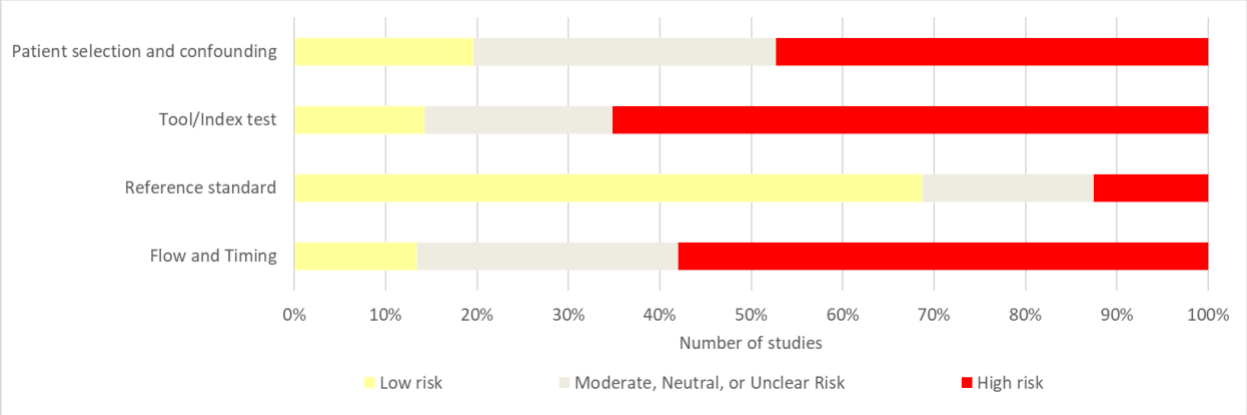


We identified 114 studies meeting inclusion criteria. The earliest identified study was published in 1998. Studies evaluated tools in Brazil, Canada, China, Denmark, Germany, Greece, India, Ireland, Israel, Korea, the Netherlands, Norway, Sweden, Switzerland, Turkey, UK, USA, or combined evaluations in multiple countries. Sample sizes varied widely, from a dozen participants to large samples with over a thousand participants, respectively.[50-52] Studies included participants diagnosed with ADHD and compared to different non-ADHD samples. These included neurotypical adults not diagnosed with ADHD, adults from a clinical sample

evaluated or diagnosed for another clinical condition, and/or adults feigning ADHD. Half of the included studies (51%) included a neurotypical group of adults that did not meet criteria for ADHD and in some cases were also selected specifically because they also never had a childhood diagnosis of ADHD. Many studies (40%) compared to a clinical sample of participants that was being evaluated for another clinical condition. In addition, two studies each compared to participants with autism,[53, 54] participants with conduct disorder or anger dysregulation,[55, 56] or participants with depression,[57, 58] respectively. A quarter (23%) of the identified studies included participants identified or specifically trained to pretend to have ADHD. Studies varied in whether they included an additional group (e.g., a neurotypical or clinical sample), but some studies included only participants feigning ADHD, which were compared to participants with a diagnosis of ADHD.[59-62]

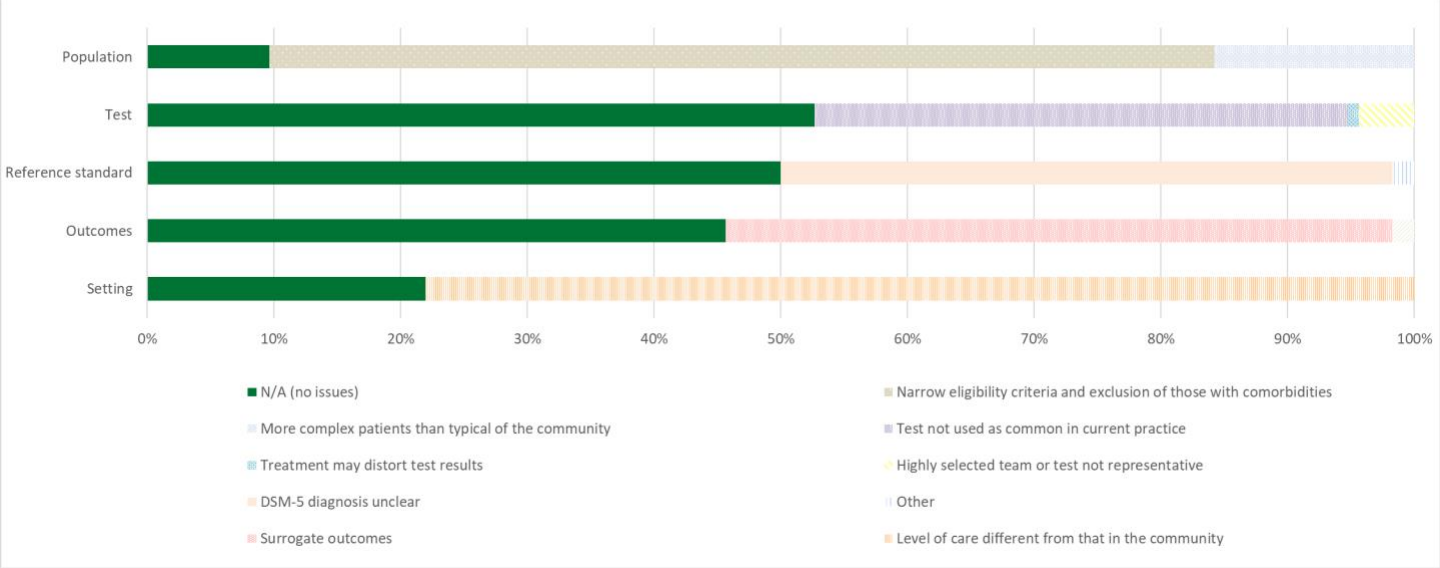
The risk of bias across studies is shown in figure 3.

Figure 3. Risk of Bias



The applicability assessment is summarized in figure 4.

Figure 4. Applicability



Identified studies reported on self-report questionnaires, peer review tools, neuropsychological tests, neuroimaging, electroencephalogram (EEG), diverse biomarkers, clinician tools, combinations of modalities, and tools to identify feigning ADHD. Studies reported on the success of identifying ADHD, success in identifying feigning and exaggerating of ADHD symptoms, or both.

3.2 Results of Key Question 1 comparative diagnostic accuracy, unintended consequences, and impact of tools that can be used in the primary care practice setting or by specialists to diagnose ADHD among adults?

We identified a number of studies that included multiple tools used alone or in combination. However, not all studies reported diagnostic performance for all tools and combinations, and only selected studies allowed direct comparisons.

The 11 studies with head-to-head comparisons between modalities compared primarily self report questionnaires with other modalities: compared to parent rating,{#685} peer report,{#5349}{#10626} a combination of self and other rating,{#949}{#5349}{#10626} neuropsychological tests,{#5349}{#870} a combination of self report and EEG;{#10732} and a clinician tool.{#4185}{#5349} Three studies compared neuropsychological test results to combinations of input;{#10710}{#711}{#325} one compared a neuropsychological test and one compared EEG data under Go-NoGo task conditions with task performance indicators{#73} Table 2 documents the results for the comparative studies.

Table 2. Comparative Studies

Study ID Participants	Self-report	Peer rating	Combined prediction	Neuropsychological tests	EEG	Clinician interview
Biederman, 2017{#73} N = 60 Specialty care				Go-NoGo task errors, participants were seated in a dimly lit room at a distance of 70 cm from a 17-inch CRT screen; Go stimuli were white letters appearing in equal proportions, the NoGo stimulus was a white x symbol, stimuli were presented on the center of a black background computer screen for 150 ms and were located between 2 vertical white lines, 10 trial practice block, analyzed reaction time, error rates (commission and misses) **Misdiagnosis**: **Sensitivity**: **Specificity**: **Admin time**: 12 minutes across all tests. **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A	Event-related potential data to analyze brain activity patterns during Go/NoGo task, Go condition **Misdiagnosis**: 5% **Sensitivity:** 86% NoGo condition: 76 **Specificity**: 95% NoGo condition: 91% **Admin time**: 12 minutes across all tests. **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A	
Dvorsky, 2016{#10626} N = 86 College	BAARS-IV (Barkley Adult ADHD Rating Scale-IV) for self-reported assessment of ADHD symptoms on a 4-point scale (0 = never or rarely to 3 = very often), cut off > 3 symptoms presence **Misdiagnosis**: N/A **Sensitivity**: 89% **Specificity**: 30% **Admin time**: N/A **Rater reliability**: BAARS-IV self report vs BAARS-IV parent ratings Parent ratings compared against student self-reports Current inattention ICC 0.43, current hyperactivity ICC **Costs**: N/A	BAARS-IV (Barkley Adult ADHD Rating Scale-IV) used for parent ratings, 4-point scale (0 = never or rarely to 3 = very often), total sum score equal or larger than 25 **Misdiagnosis**: N/A **Sensitivity**: 60 **Specificity**: 77 **Admin time**: N/A **Rater reliability**: Parent ratings compared against student self-reports Current inattention ICC 0.43, current hyperactivity ICC	Combination prediction model with BAARS parent and self rating of current and childhood ADHD diagnosis **Misdiagnosis**: N/A **Sensitivity**: **Specificity**: **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A			

	Concordance: N/A	0.31, current impulsivity ICC 0.32, retrospective children inattention ICC 0.42, retrospective childhood hyperactivity/impulsivity ICC 0.37 **Costs**: N/A **Concordance**: N/A				
Groom, 2016{#325} N = 57 College	CAARS-E (Conners Adult ADHD Rating Scale-subscale E) **Misdiagnosis**: N/A **Sensitivity**: **Specificity**: **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A		Integration of CAARS-E (Conners Adult ADHD Rating Scale - ADHD Index), the AQ10 (Autism Quotient - 10), and the QbTest (computerized Continuous Performance Test with motion tracking) to differentiate ADHD from Autism Spectrum Disorder **Misdiagnosis**: N/A **Sensitivity**: 94% **Specificity**: 84% **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A	QbTest is a computerized continuous performance test with infra-red motion tracking system, designed to assess attention, impulsivity, and activity levels; participants respond to stimuli on a screen while their movements are tracked, and scores are calculated based on attention accuracy, reaction time, and movement data **Misdiagnosis**: N/A **Sensitivity**: 84% **Specificity**: 80% **Admin time**: Approximately 20 minutes. **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A		
Kingston, 2013{#5349} N = 120 Specialty care	ASRS-v1.1 Part A, a scale based on nosological criteria and pertain to frequency, rather than severity, of ADHD symptoms; Part A comprises 6 screening questions and is considered to be the most predictive of symptoms consistent with ADHD **Misdiagnosis**: 16% **Sensitivity**: 76% ASRS-v1.1 (Part B): .66 (.53–.78); Brown ADD Scale: .84 (.71–.92); CAARS-Self ADHD Index: .63 (.49–.75); WURS: .82 (.69–.91) **Specificity**: 84% ASRS-v1.1 (Part B): .93 (.82–.98); Brown ADD Scale: .73 (.59–.83);		Integration of ASRS-v1, CAARS-Self and CAARS-Observer, Brown ADD scale, and WURS in a discriminant function **Misdiagnosis**: 18% **Sensitivity**: 91% **Specificity**: 82% **Admin time**: N/A **Rater reliability**: **Costs**: N/A **Concordance**: N/A	IVA + Plus FSRCQ (Integrated Visual and Auditory Continuous Performance Test Full Scale Response Control Quotient), a computerized continuous performance test utilizing visual and auditory stimuli to assess response control; constant and sustained attention is required, as participants respond or inhibit their response to 500 counterbalanced trials; FSRCQ measures impulsivity and commission errors,		CAARS-O ADHD Index (Observer), a 66-item measure that contains 9 empirically-derived scales related to adult ADHD symptoms completed by psychiatrist **Misdiagnosis**: 25% **Sensitivity**: 76% **Specificity**: 75% **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A

	CAARS-Self ADHD Index: .91 (.79–.97); WURS: .69 (.54–.80) **Admin time**: N/A **Rater reliability**: rater agreement between self-report measures (ASRS-v1.1, CAARS-Self, WURS, and Brown ADD Scale) and observer-rated measures (CAARS-Observer) r 0.51 **Costs**: N/A **Concordance**: N/A			normative quotient scores have a mean of 100 and a standard deviation of 15 **Misdiagnosis**: 26% **Sensitivity**: 30% IVA + Plus (FSAQ): .39 (.29–.54) **Specificity**: 74% IVA + Plus (FSAQ): .69 (.53–.82) **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A		
Kumar, 2011{#4185} N = 110 Specialty care	CAARS-S:SV (Conners' Adult ADHD Rating Scales: Screening Version), 30-item self-report tool that screens for ADHD symptoms in adults, using a 4-point rating scale to assess the frequency of symptoms based on DSM-IV criteria, cut off point wasT score>70 **Misdiagnosis**: 31% **Sensitivity**: 83% **Specificity**: 69% **Admin time**: N/A **Rater reliability**: Correlation self report CAARS-S:SV and MINI r 0.58 **Costs**: N/A **Concordance**: N/A					MINI (International Neuropsychiatric Interview), a short, structured diagnostic interview designed to assess a range of different mental health disorders **Misdiagnosis**: 48% **Sensitivity**: 83% **Specificity**: 52% **Admin time**: 10-25 minutes **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A
Nikolas, 2019{#10710} N = 246 Specialty care			Combination of self/informant symptom ratings (BAARS-IV), family history, and reactiontime variability from TOVA (Test of Variables of Attention) **Misdiagnosis**: N/A **Sensitivity**: **Specificity**: **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A	TOVA ommission errors, cutoff <95 **Misdiagnosis**: 15% **Sensitivity**: 50% **Specificity**: 85% **Admin time**: N/A **Rater reliability**: **Costs**: N/A **Concordance**: N/A		
Palmer, 2023{#685} N = 71	CAARS-S (Conners Adult ADHD Rating Scales Self Report) ADHD Index assessed ADHD symptoms with a cutoff of	CAARS-P (Conners Adult ADHD Rating Scales Peer Report) ADHD Index cutoff >56;				

Community	≥56; adminstered together with the SDQ (Strengths and Difficulties Questionnaire), cutoff of ≥9 **Misdiagnosis**: N/A **Sensitivity**: 57% SDQ>9: 28 **Specificity**: 81% SDQ>9: 100 **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A	administered together with ABC (Aberrant Behavior Checklist) Hyperactivity/Non-compliance subscale (a cutoff of ≥3) is a parent-reported tool designed to measure hyperactive and **Misdiagnosis**: N/A **Sensitivity**: 94% ABC scale: 91% **Specificity**: 57% ABC scale: 42% **Admin time**: N/A **Rater reliability**: **Costs**: N/A **Concordance**: N/A				
Pettersson, 2018{#711} N = 108 Specialty care			Model with DIVA report, QbTest cardinal variable Acticity, QbTest cardinal variable Inattention, and CpT II Commission errors, combining neuropsychological tests, DIVA clinician report, and self report ASRS Screener **Misdiagnosis**: 17% **Sensitivity**: 90% **Specificity**: 83% **Admin time**: N/A **Rater reliability**: N/A **Costs**: **Concordance**: N/A	Model with CPT II Commission errors, QbTest cardinal variable Inattention, and QbTest cardinal variable Activity **Misdiagnosis**: 33% **Sensitivity**: 80% **Specificity**: 67% **Admin time**: 20 minutes **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A		
Robeva, 2004{#10732} N = 12 College	WURS (Wender Utah Rating Scale), a 61-item retrospective questionnaire witha cutoff score of 30 on the short form with higher cutoff values **Misdiagnosis**: N/A **Sensitivity**: **Specificity**: **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A		Bayesian probability model integrated three diagnostic tools (WURS, ConsistencyIndex (EEG), Alpha Blockade Index (EEG) **Misdiagnosis**: N/A **Sensitivity**: **Specificity**: **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A			

			Concordance: N/A			
Solanto, 2004{#870} N = 93 Specialty care	BADDS (Brown Attention-Deficit Disorder Scale), assesses executive and adaptive functioning across five clusters (Activation, Attention, Effort, Affect, and Memory), cutoff 50; adminstered together with CAARS (Conners Adult ADHD Rating Scale), cutoff ≥65 for inattention, hyperactivity-impulsivity, and total ADHD scores **Misdiagnosis**: 67% **Sensitivity**: 92% **Specificity**: 33% **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A			C-CPT (Conners Continuous Performance Test), a 14-minute computerized task where participants respond to non-target stimuli **Misdiagnosis**: 14% **Sensitivity**: 47% **Specificity**: 86% **Admin time**: 15 minutes **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A		
Van Voorhees, 2011{#949} N = 349 Specialty care	CAARS:S (Conners' Adult ADHD Rating Scales, Self Rating, Long Version), 66-items rated on a 4-point scale (0 to 3) to assess ADHD symptoms **Misdiagnosis**: 39% **Sensitivity**: 65% **Specificity**: 61% **Admin time**: N/A **Rater reliability**: Self-reports (CAARS-S) and observer reports (CAARS-O including ratings from friends, parents, and spouses) Ranged from r 0.24 ("distractible") through r 0.46 ("on the go/driven by a motor") **Costs**: N/A **Concordance**: N/A		CAARS-LV combining self-report CAARS:S and observer-report CAARS-O **Misdiagnosis**: 17% **Sensitivity**: 43% **Specificity**: 83% **Admin time**: N/A **Rater reliability**: N/A **Costs**: N/A **Concordance**: N/A			

Several studies compared multiple tests, but not all reported coefficients for every test. All available comparative results across test modalities (e.g., combinations of variables vs neuropsychological test results alone) are documented in the summary of findings table.

Combination

Seven studies reported on a combination of input from different modalities.{#325}{#949}{#5100}{#5349}{#10626}{#10710}{#10732} Several included self and informant symptom ratings, and some also used demographic variables, neuropsychological assessment results, or EEG data. Studies varied in their complexity of the combination; one study, supported by machine learning, used 93 variables.{#5100} The Appendix Table A1 documents results for all studies that evaluated a combination. The table shows the specific combinations used to correctly diagnose ADHD. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

Self Reports

Forty-one studies reported at least one self report measure evaluated for its performance in diagnosing ADHD. The studies reported on numerous self report measures: ADHD Rating Scale, AHA, ALS-SF, APQ, ARS, ASRS, ASSET-BS, BAARS-IV, BADDS, CAARS, CBS, EarlyDetect Questionnaire, IPDE-SQ, PAI, PDI-4, SR-WRAADDS, and WURS (see evidence table in the appendix for more information). The Appendix Table A2 documents results for all studies that evaluated a written self report for the diagnosis of ADHD. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

Peer Ratings

Two studies evaluated peer reports.{#685}{#10626} One of the studies asked parents to rate their young adults with autism using the CAARS-P and the ABC (Aberrant Behavior Checklist).{#685} The other study included parent ratings of undergraduates using the BAARS-IV (Barkley Adult ADHD Rating Scale-IV).{#10526} The Appendix Table A3 documents results for the small number of studies that evaluated a written peer report for the diagnosis of ADHD. In both cases, the peer report was a parent rating of young adults. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

Neuropsychological Assessment

Twenty-six studies reported on the performance of neuropsychological assessment to diagnose ADHD. Studies evaluated test batteries such as AQT, BQSS, C-CPT, DII, IVA, MOXO-dCPT, QbTest, SCWT, SNST, or the performance of individual tasks such as the Go-No-Go task or WAIS-IV Processing Speed Index. The Appendix Table A4 documents results for all neuropsychological test evaluated to diagnose ADHD. The index test description shows the evaluated test selection that showed the best performance, together with all administered tests. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

Neuroimaging

Five studies evaluated neuroimaging.{#28}{#134}{#812}{#870}{#1011} Studies used Brain Perfusion SPECT (single-photon emission computed tomography),{#28} 3-D SPECT,{#812}

structural MRI and diffusion tensor imaging,{#134} and resting state functional MRI.{#970}{#1011} The Appendix Table A5 documents results of studies that evaluated neuroimaging for the diagnosis of ADHD. The table provides details on the final selection model where reported. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

EEG

Eleven studies evaluated EEG (electroencephalogram) data.{#73}{#334}{#436}{#448}{#454}{#640}{#642}{#722}{#729}{#840}{#425} Studies tested very different conditions, ranging from analyzing resting state EEG,{#448}{#722} event-related potentials during neuropsychological tasks,{#73}{#640} to EEG recording during transcranial magnetic stimulation.{#334} The Appendix Table A6 documents results for the studies that evaluated the use of EEG data for the diagnosis of ADHD. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

Biomarker

Five studies evaluated biomarkers other than EEG or neuroimaging-based.{#327}{#417}{#832}{#5003}{#5685} Studies evaluated genetic marker,{#327} eye tracking,{#417} blood oxidative status,{#832} physiological data from a wearable device,{#5003} or used MFNU (Motor Function Neurological Assessment){#5685} to diagnose ADHD. The Appendix Table A7 documents results for the small number of studies that evaluated the use of biomarkers other than neuroimaging or EEG for the diagnosis of ADHD. Results for key outcomes are synthesized in the Summary of Findings table (Table 3).

Clinician Tool

Three studies reported on a clinician interview or questionnaire that was assessed for congruence with an external reference standard.{#684}{#4185}{#5349} Studies evaluated the MINI (Mini-International Neuropsychiatric Interview),{#4185} MINI-Plus,{#684} and the CAARS-O (Conners' Adult ADHD Rating Scales){#5349 in clinical samples. The reference standards were assessments from trained clinicians,{#684} the recorded chart diagnosis,{#4185} or clinical consensus meeting with multiple clinicians.{#5349}

Key Question 1a: How does the comparative diagnostic accuracy of tools vary by clinical setting or patient characteristics?

Because raw data for diagnostic accuracy were often not reported, we were not able to detect effect modifiers in meta-regressions by adding variables to the meta-analytic model. Results are based on subgroups as reported by the authors and analyses conducted within the original studies.

Clinical setting: Half of the identified studies were conducted in specialty care (n=55). The next most frequent setting was college (n=37). Very few studies were conducted in primary care (n=2). In addition, none of the identified studies analyzed the effect of the setting on the diagnostic process. Hence the question which tests should be used in primary care is difficult to answer. However, several studies addressed the effect of the reference standard and comparator sample, i.e., study characteristics. In addition, several studies addressed the effect of

comorbidities. Although primarily a patient characteristic, participants evaluated for other clinical conditions was more typical of a specialty care clinical setting.

Reference standard and comparator sample: Three studies addressed the effect of the method of establishing a clinical ADHD diagnosis, but all addressed different aspects. One study comparing self reports and neuropsychological tests highlighted that diagnostic accuracy measures were high when comparing ADHD-diagnosed participants to the general population but were less effective when distinguishing ADHD from other psychiatric conditions, with overlapping scores noted for anxiety and depression.{#870} Similarly, a self report study reported a high false-positive rate in patients with depression.{#212} Another self-report study reported that comorbidities such as anxiety and depression were associated with elevated scores on scales which may overlap with ADHD symptoms and potentially contribute to misclassification and highlighted the importance of considering comorbid conditions during assessment.{#495} One study evaluating neuromuscular assessment reported that they found several patients with subthreshold ADHD in a clinical sample suggesting possible diagnostic overlap and the need for further evaluation.{#5685}

ADHD presentation: Four studies addressed diagnosis in different presentations of ADHD with some conflicting results. One-self report study reported that diagnostic accuracy did not significantly vary across ADHD presentations/subtypes (inattentive, hyperactive-impulsive, and combined) but noted that combined type ADHD was the most frequently identified subtype, which could influence overall sensitivity and specificity estimates.{#857} Similarly, another study reported that sensitivity and specificity were consistent across ADHD presentation types (inattentive, hyperactive, and combined), but noted that misdiagnosis rates were slightly higher for the inattentive subtype in self-reports compared to clinician diagnoses, and inter-rater reliability between self-report and clinical ratings was moderate, with higher concordance for combined presentation.{#5176} Another self-report study highlighted that sensitivity for the inattentive subtype was 100 percent on the Inattentive Symptoms subscale, with specificity at 25%.{#490} One self-report study pointed out that inattention symptoms were more predictive of ADHD persistence into adulthood than hyperactivity-impulsivity symptoms; and individuals with the combined-type ADHD in childhood were more likely to retain a diagnosis in adulthood, whereas hyperactive-impulsive presentations were more likely to remit.{#444}

Participant age: Several studies reported on the effect of the age of the participants or specifically on the age at diagnosis, but studies focused on different aspects. A self report study reported that executive functioning impairments were more predictive of ADHD persistence in older adults, while hyperactivity-impulsivity symptoms were more prevalent in younger adults, suggesting age-related shifts in symptom expression and diagnostic criteria applicability; sensitivity and specificity of ADHD diagnoses were higher in younger adults (18–30 years) compared to older adults (31–44 years), likely due to better recall of childhood symptoms and reduced cognitive decline in memory-based reporting.{#444} A self report and neuropsychological test study reported that age was inversely correlated with scores on scales for attention and effort, suggesting that older participants exhibited fewer ADHD-related symptoms, potentially reflecting developmental improvements in executive functioning.{#870} One study did not comment on differential effects on the diagnosis, but suggested that a potential biomarker, oxidative stress, may increase with the duration of the disease.{#832} Another study found that age represented as independent variable in a multiple regression did not significantly influence parameters measured by the QbTest.{#111} A further study reported that ADHD diagnosis based on CAARS-S or MINI were not correlated with age.{#4185}

Participant sex: Several studies reported on the effect of the sex of the participants on the diagnostic performance, but studies reported conflicting results. One EEG study reported lower sensitivity in females compared to males.{#425} A neuroimaging study noted that classification performance was higher in the male-only subgroup compared to the mixed-gender subgroup, suggesting that male ADHD patients may have more significant neuroanatomical deviations from controls.{#134} A self report study did not find lower sensitivity but lower specificity in females versus males.{#960} One study concluded that sex did not influence parameters of the neuropsychological test.{#111} A self report study did not detect differences in sensitivity and specificity between sexes.{#1916} A study reporting on a self report and a clinician interview noted that ADHD diagnosis based on the tests were not correlated with sex.{#4185}

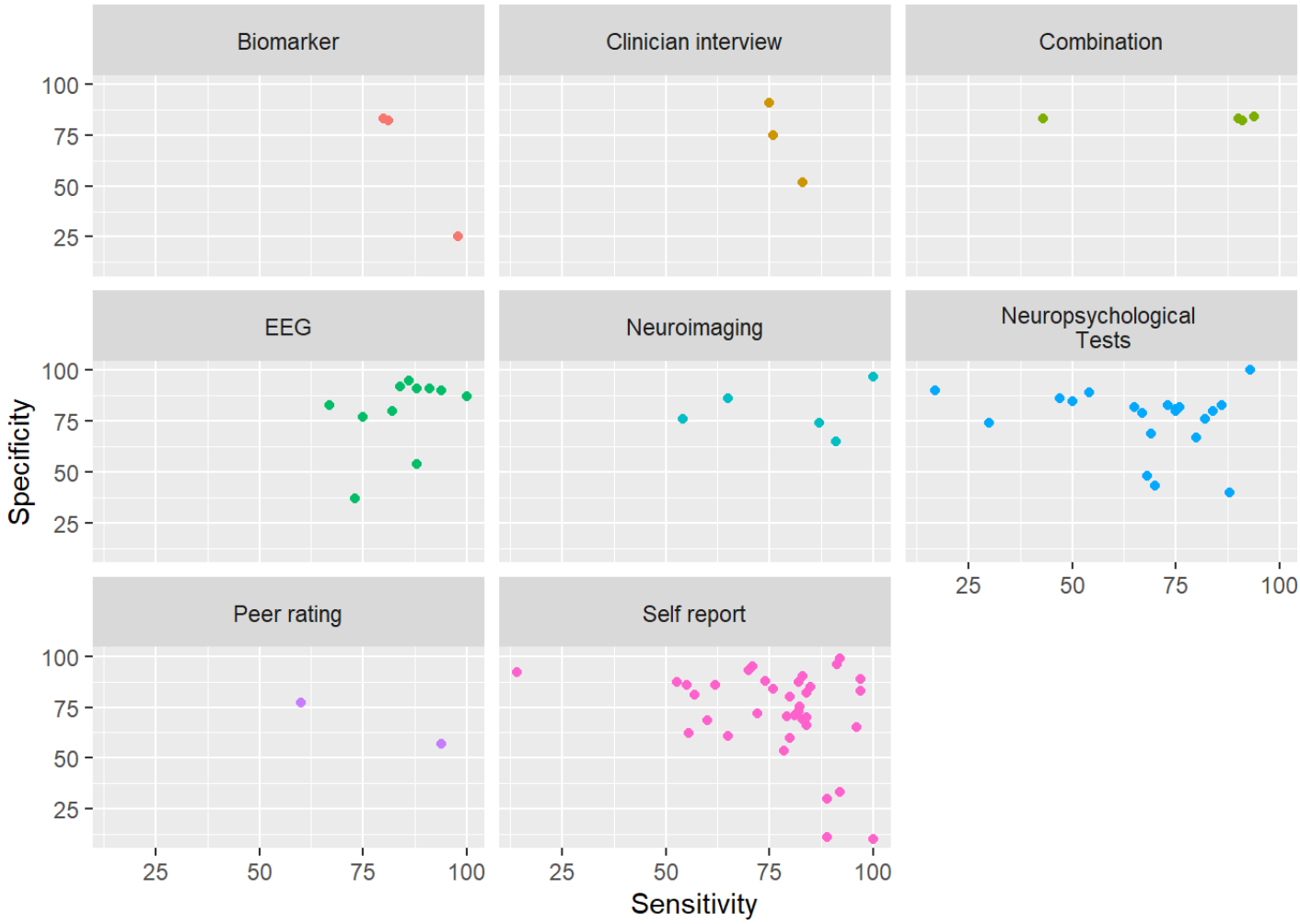
Participant ethnicity: None of the studies stratified diagnostic performance by race or ethnicity.

Comorbidities: Multiple studies reported on the effect of comorbidities in participants with ADHD on diagnostic performance, but results and conclusions differed. One college study reported that comorbidities contributed to challenges in specificity but not sensitivity and that functional impairment was higher in participants with comorbid conditions.{#490} Similarly, a study in addiction centers reported on variability in specificity values across subgroups while sensitivity remained similar.{#942} Another study reported reduced specificity in participants with overlapping symptoms of borderline personality disorder and bipolar disorder in neuropsychiatric clinics.{#220} One study reported lower sensitivity and higher specificity in participants with comorbidity in a mental health center.{#536} Two studies in outpatient centers reported that diagnostic performance was unaffected by comorbidities.{#111}{#495} Some studies pointed out the high prevalence of comorbid conditions such as depression and anxiety.{#857}{#4485}{#444}{#323} One study suggested that participants with ADHD and depression reported higher levels of anxiety.{#212}

Key Question 1 Summary of Findings

Despite the large number of studies, many did not report on the exact number of true positives, true negatives etc. The most common metrics were the author reported sensitivity. Given that sensitivity and specificity are not independent of each other, we plotted both for all reported tests in Figure 5.

Figure 5. Sensitivity and Specificity of Tests



The figure visualizes the much larger evidence base for self reports compared to all other modalities. The figure also illustrates the wide variability reported in the individual studies for the same modality. In addition, the visualization shows that tests were sometimes able to maximize sensitivity or specificity, but not both. Finally, with few exceptions, the evaluated tests were limited in their success of detecting a clinical diagnosis of ADHD.

The summary of findings table (Table 3) provides a synthesis of the results for the key outcomes. Direct comparisons between test modalities are shown first, followed by the test performance for individual tests, and the summary of the subquestion. The summary of findings table shows results for the key outcomes for which at least one study with data was identified. The clinical misdiagnosis results were limited to studies reporting on clinical samples and/or studies comparing to another clinical condition such as anxiety.

To broadly characterize the magnitude of effects, we described false positive rate as moderate when the majority of studies reported values below 20 and as substantial for rates close to 50 percent of incorrectly classified. We described sensitivity and specificity as fair for values below 80, as acceptable above 80 but below 90 percent, as good if most values fell between 90 and 95 percent, and as excellent for values above 95 percent. We describe administration time under 30 minutes as short administration. For rater agreement we described kappa below 0.8 and correlations below 0.40 as limited.

Results of individual studies for all abstracted outcomes are shown in the evidence tables in the appendix.

Table 3. Summary of Findings Table Comparative Performance, Performance of Combinations, and Performance of Individual Tools against a Reference Standard

Key question Direct comparisons Individual Test performance Outcomes	Contributing studies	Results	Reasons for downgrading	GRADE
KQ1 Comparative clinical misdiagnosis Combinations vs self report	2 studies{#5349} {#949}	Conflicting results: 1 study reported a 16% misdiagnosis rate for the ASRS compared to 18% for a combination of variables,{#5349} while 1 study reported a misdiagnosis rate of 39% for the CAARS-S compared to 17% for a combination of self and observer reports{#949}	Inconsistency (conflicting results)	Insufficient for comparative statements
KQ1 Comparative clinical misdiagnosis Combinations vs neuropsychological tests	1 study{#711}	Favors combination: 1 study reported a misdiagnosis rate of 17% for a combination form multiple sources compared to 33% for a neuropsychological test{#711}	Inconsistency (no replication)	Insufficient for comparative statements
KQ1 Comparative clinical misdiagnosis Self report vs clinician tools	2 studies{#5349} {#4185}	Favors self reports: 1 study reported a 16% misdiagnosis rate for the ASRS compared to 25% for clinician rating tool;{#5349} another study reported a 31% misdiagnosis rate for the CAARS-S compared to 48% for the MINI {#4185}	Study limitation (studies assessed different tests)	Low for favoring self report over clinician tool
KQ1 Comparative clinical misdiagnosis Self report vs neuropsychological tests	2 studies{#5349} {#870}	Conflicting results: 1 study reported a 16% misdiagnosis rate for the ASRS compared to 26% for a CPT;{#5349} 1 study reported a 47% misdiagnosis rate for the BADDS compared to 14% for the C-CPT{#4185}	Inconsistency (conflicting results)	Insufficient for comparative statements
KQ1 Comparative sensitivity Combination vs self report	2 studies{#5349} {#949}	Conflicting results: 1 study reported a sensitivity of a combination of 91% (corresponding specificity 82%) vs 76% for the ASRS (corresponding specificity 84%),{#5349} 1 study reported a sensitivity of a combination of 43% (corresponding specificity 83%) vs 65% for the CAARS-S (corresponding specificity 61%);{#949}	Inconsistency (conflicting results)	Insufficient for comparative statements
KQ1 Comparative sensitivity Combination vs neuropsychological tests	3 studies{#325}{#5349}{#711}	Favors combination: Estimates ranged for a combination from 94% (corresponding specificity 84%) vs QbTest 84% (corresponding specificity 80%){#325} to 90% for a combination (corresponding specificity 83%) vs a CPT test with 80% (corresponding specificity 67%){#711}	Study limitation (compared different tests and combinations)	Low for favoring combinations over neuropsychological tests
KQ1 Comparative sensitivity Self vs parent report	2 studies{#10626} {#685}	Conflicting results: 1 study reported a sensitivity of 89% for the BAARS (corresponding specificity 30%) vs BAARS parent rating 60% (corresponding specificity 77%);{#10626} 1 study reported a sensitivity of 57% for CAARS-S (corresponding specificity 81%) vs 94% for a parent rating (corresponding specificity 57%){#685}	Inconsistency (conflicting results)	Insufficient for comparative statements

KQ1 Comparative sensitivity Self report vs neuropsychological tests	2 studies{#5349} {#870}	Favors self report: 1 study reported a sensitivity of 92% for the BADDS (corresponding specificity 33%) vs 47% for the C-CPT (corresponding specificity 86%);{#870} 1 study reported 76% for the ASRS (corresponding specificity 84%) vs 30% for a CPT (corresponding specificity 74%){#5349}	Study limitation (compared different tests and combinations)	Low for favoring self reports over neuropsychological tests
KQ1 Comparative sensitivity Self report vs clinician tool	2 studies{#5349} {#4185}	No difference: 1 study reported a sensitivity of 83% for CAARS-S (corresponding specificity 69%) vs 83% for MINI (corresponding specificity 52%);{#4185} 1 study reported 76% for the ASRS (corresponding specificity 84%) vs CAARS-O sensitivity of 76% (corresponding specificity 75%){#5349}	Study limitation (compared different tests and combinations)	Low for no difference between self reports and clinician tools
KQ1 Comparative specificity Combination vs self report	2 studies{#5349} {#949}	Conflicting results: 1 study reported a specificity of 83% for a combination (corresponding sensitivity 43%) vs 61% for the CAARS-S (corresponding sensitivity 65%);{#949} 1 study reported a specificity of 82% for a combination (corresponding sensitivity 91%) vs 84% for the ASRS (corresponding sensitivity 76%){#5349}	Inconsistency (conflicting results)	Insufficient for comparative statements
KQ1 Comparative specificity Combination vs neuropsychological tests	3 studies{#325}{#5349}{#711}	Favors combination: Estimates ranged from 84% for a combination (corresponding sensitivity 94%) vs 80% for the QbTest (corresponding sensitivity 84%){#325} to 83% for a combination (corresponding sensitivity 90% vs 67% for CPT (corresponding sensitivity 80%){#711}	Study limitation (compared different tests and combinations)	Low for favoring combination over neuropsychological tests
KQ1 Comparative specificity Self vs parent report	2 studies{#10626} {#685}	Conflicting results: 1 study reported a specificity of 81% for the CAARS-S (corresponding sensitivity 57%) vs 57% for the CAARS-P (corresponding sensitivity 94%);{#685} 1 study reported a specificity of 30% for the BAARS self report (corresponding sensitivity 89%) vs 77% for the BAARS parent report (corresponding sensitivity 60%){#10626}	Inconsistency (conflicting results)	Insufficient for comparative statements
KQ1 Comparative specificity Self report vs neuropsychological tests	2 studies{#5349} {#870}	Conflicting results: 1 study reported a specificity of 84% for the ASRS (corresponding sensitivity 76%) vs 74% for a CPT (corresponding sensitivity 30%);{#5349} 1 study reported a specificity of 33% for the BADDS (corresponding sensitivity 92%) vs 86% for the C-CPT (corresponding sensitivity 47%){#870}	Inconsistency (conflicting results)	Insufficient for comparative statements
KQ1 Comparative specificity Self report vs clinician tool	2 studies{#5349} {#4185}	Favors self report: 1 study reported specificity of 84% for the ASRS (corresponding sensitivity 76%) vs 75% for the CAARS-O (corresponding sensitivity 76%);{#5349} 1 study reported a specificity of 69% for the CAARS-S (corresponding sensitivity 83%) vs 52% for the MINI (corresponding sensitivity 83%){#4185}	Study limitation (compared different tests and combinations)	Low for favoring self reports over clinician tools
KQ1 Comparative administration and scoring time	0 studies	N/A	N/A	N/A
KQ1 Comparative inter-rater reliability	0 studies	N/A	N/A	N/A
KQ1 Comparative costs	0 studies	N/A	N/A	N/A
KQ1	0 studies	N/A	N/A	N/A

Comparative diagnostic concordance of primary care provider with specialist				
KQ1 Combination Combining self and informant symptom ratings, demographic variables, neuropsychological assessments and/or EEG data to diagnose ADHD Clinical misdiagnosis	3 studies{#325}{#711}{#5349}	Reported false positive rate ranged from 16% in a study combining self ratings and QBTest data to distinguish ADHD from Asperger's syndrome{#325} to 18% in a study combining multiple self reports and an observer report to distinguish from aggression{#5349}	Study limitation (can not be replicated based on reported detail)	Low for moderate false positive rate
KQ1 Combination Combining self and informant symptom ratings, demographic variables, neuropsychological assessments and/or EEG data to diagnose ADHD Sensitivity	3 studies{#325}{#5349}{#949}	Sensitivity ranged from 94% (corresponding specificity 84%){#325} to 43% (corresponding specificity 83%){#949}	Imprecision	Low for good sensitivity
KQ1 Combination Combining self and informant symptom ratings, demographic variables, neuropsychological assessments and/or EEG data to diagnose ADHD Specificity	3 studies{#325}{#5349}{#949}	Specificity ranged from 84 (corresponding specificity 94%){#325} to 82 (corresponding specificity 91%){#5349}	Imprecision	Low for acceptable specificity
KQ1 Self report Clinical misdiagnosis	22 studies{#571}{#757}{#323}{#495}{#5349}{#555}{#857}{#685}{#1016}{#536}{#390}{#55}{#762}{#4185}{#942}{#172}{#611}{#603}{#870}{#45}{#490}{#16}	Reported false negative rates ranged from 12% differentiating from depression or generalized anxiety using the WURS{#757} to 90% in students with academic or psychological difficulties using the CAARS-S{#490}	Imprecision (values ranged widely)	Low for substantial false positive rate

KQ1 Self report Sensitivity	38 studies{#351} {#323}{#495} {#960}{#685} {#212}{#299} {#949}{#863} {#444}{#667} {#757}{#5349} {#603}{#762} {#1016}{#611} {#55}{#390} {#5176}{#536} {#4185}{#1017} {#525}{#857}{#942} {#10752}{#10626} {#45}{#940} {#10633}{#870} {#16}{#172} {#555}{#571} {#490}{#141}	Sensitivity ranged from 100% (CAARS-S corresponding specificity 10%{#490} or ASRS-v1.1 with specificity not reported{#141}) to 14% (CAARS-S, corresponding specificity 92%){#351}	Imprecision (values ranged widely)	Low for good sensitivity with corresponding limited specificity
KQ1 Self report Specificity	36 studies{#490} {#45}{#10626} {#870}{#603} {#611}{#949} {#960}{#172} {#942}{#212} {#4185}{#525} {762}{#55}{#390} (#536}{#1016} {#685}{#857} {#1017}{#863} {#555}{#5349} {#10752}{#495} {#299}{#5176} {#323}{#757}{#571} {#351}{#444}{#667} {#940}{#10633}	Specificity ranged from 99% (CBS corresponding specificity 90%){#10633} to 10% (CAARS-S, corresponding specificity 100%){#490}	Imprecision	Low for good specificity with corresponding limited sensitivity
KQ1 Self report Administration and scoring time	1 study{#525}	1 study explicitly stated that the newly developed ADHD rating scale took about 15 minutes to complete{#525}	Inconsistency (no replication)	Low for short administration and scoring time
KQ1 Self report Rater agreement	8 studies{#949} {#1016}	1 study reported on kappa and found 0.006 agreement between WURS-brief vs DIVA	Inconsistency (reporting on different	Moderate for limited rater agreement

	{#10626}{#571} {#5349} {#4185}{#45} {#555}	rating;{#45} 1 study reported 89% agreement between self and informant report;{#555} 1 study reported an ICC of 0.43 for self vs parent BAARS-IV ratings;{#10626} 6 studies reporting Pearson self-observer correlations reported ranges from r 0.24 for a CAARS subscale{#949} to r 0.58 for CAARS-S:SV vs MINI report{#4185}	measures and questionnaires)	
KQ1 Peer report Sensitivity	2 studies{#685} {#10626}	Sensitivity ranged from 94% (CAARS:P, corresponding specificity 57%){#685} to 60% (BAARS-IV, corresponding specificity 77%){#10626}	Inconsistency (reporting on different questionnaires), imprecision (range from excellent to poor)	Insufficient
KQ1 Peer report Sensitivity	2 studies{#685} {#10626}	Specificity ranged from 77% (BAARS-IV, corresponding sensitivity 60){#10626} to 57% (CAARS:P, corresponding sensitivity 60%){#685}	Inconsistency (reporting on different questionnaires), imprecision (range from excellent to poor)	Insufficient
KQ1 Peer report Rater agreement	1 study{#10626}	1 study reported ICCs ranging from 0.43 to0.31 for BAARS-IV subscales{#10626}	Inconsistency (no replication), study limitation (subscales only)	Insufficient
KQ1 Neuropsychological tests Clinical misdiagnosis	9 Studies{#447} {#870} {#10710}{#325} {#4544}{#5349} {#711}{#111}{#8}	Reported false negative rates ranged from 11% in a study using Stroop test variables for participants referred for neuropsychological evaluation{#447} to 60% in a model based on QbTest with motion tracking variables{#869}	Inconsistency (studies used different combinations of variables), Study limitation (unclear if conditions can be replicated	Low for substantial false positive rate
KQ1 Neuropsychological tests Sensitivity	21 studies{#872} {#5349}{#870} {#10710}{#447} {#638}{#4544} {#111}{#3913} {#8}{#4474} {#152}{#818} {#4209}{#711} {#234}{#325} {#220}{#869} {#987}{#659}	Reported sensitivity ranged from 93% (corresponding specificity 100%) integrating AQT variables{#659} to 17% for an individual subtest of the C-CPT (corresponding specificity 90%){#872}	Imprecision (wide range of results)	Low for good sensitivity

KQ1 Neuropsychological tests Specificity	21 studies{#872}{#5349}{#870}{#10710}{#447}{#638}{#4544}{#111}{#3913}{#8}{#4474}{#152}{#818}{#4209}{#711}{#234}{#325}{#220}{#869}{#987}{#659}	Reported specificity ranged from 100% (corresponding sensitivity 93%) integrating AQT variables{#659} to 40% for a model integrating QbTest Plus variables (corresponding sensitivity 88%){#869}	Imprecision (wide range of results)	Low for good specificity
KQ1 Neuropsychological tests Administration and scoring time	15 studies{#638}{#937}{#325}{#283}{#8}{#220}{#234}{#711}{#111}{#869}{#479}{#4209}(#3913}{#870}{#152}	There was some variation, but 7 studies estimated the duration of the test to be 20 minutes{#325}{#8} {#220}{#234}{#711}{#111{#869}	Study limitation (scoring / data interpretation not mentioned)	Low for short administration time
KQ1 Rater agreement	1 study{#818}	1 study reported that the six summary scores had good to excellent reliability{#818}	Inconsistency (no replication, very specific test)	Low for good rater agreement
KQ1 Neuroimaging Clinical misdiagnosis	1 study{#812}	1 study reported a 24% false positive rate in a sample with various psychiatric and neuropsychiatric disorders using 3D thresholded SPECT{#812}	Inconsistency (no replication, very specific test)	Low for moderate clinical false positive rate
KQ1 Neuroimaging Sensitivity	5 studies{#28}{#1011}{#970}{#134}{#812}	Performance ranged from 100% (SPECT, corresponding specificity 97%){#28} to 54% in a clinical sample (SPECT, corresponding specificity 76%)	Imprecision (wide range of values)	Low for acceptable sensitivity
KQ1 Neuroimaging Specificity	5 studies{#28}{#1011}{#970}{#134}{#812}	Performance ranged from 97% (SPECT, corresponding sensitivity 100%){#28} to 65% (functional MRI, corresponding sensitivity 91%){#1011}	Imprecision (wide range of values)	Low for acceptable sensitivity
KQ1 Neuroimaging Administration and scoring time	1 study{#812}	1 study reported a procedure duration of 15-20 mins	Inconsistency (no replication, very specific test)	Low for moderate duration
KQ1 Neuroimaging Rater agreement	1 study{#28}	1 study reported kappa 0.79 for agreement in visual interpretation of scans	Inconsistency (no replication, very specific task)	Insufficient
KQ1 EEG Clinical misdiagnosis	1 study{#425}	1 study reported a false positive rate of 3.8% using auditory brainstem response profiling test{#425}	Inconsistency (no replication,	Insufficient

			very specific test)	
KQ1 EEG Sensitivity	10 studies{436} {#729}{#640}{#334} {#73}{#840} {#454}{#642} (#448}{#722}	Reported sensitivity ranged from 100% (machine learning assisted, corresponding specificity 87%) {#436} to 67% (resting state EEG, corresponding specificity 83%){#722}	Imprecision (values ranged widely)	Low for good sensitivity
KQ1 EEG Specificity	10 studies{436} {#729}{#640}{#334} {#73}{#840} {#454}{#642} (#448}{#722}	Reported sensitivity ranged from 95% (event-related potential, corresponding specificity 86%){#73} to 37% (resting state EEG, corresponding specificity 73%){#448}	Imprecision (values ranged widely)	Low for good sensitivity
KQ1 EEG Administration and scoring time	6 studies{#454} {#448}{#642} {#722}{#436}{#73}	Reported session duration ranged from 6 minutes{#448} to 26 minutes{#642}	Imprecision (substantial variation)	Low for moderate duration
KQ1 Biomarkers Clinical misdiagnosis	2 studies{#417} {#5685}	One study reported a false positive rate of 17% in an eye tracker study in sample of participants with conduct disorder,{#417} another study reported a false positive rate of 75% in a MFNU study in a psychiatric outpatient clinic (some participants had subthreshold ADHD){#5685}	Imprecision (value ranged widely), Inconsistency (very different biomarkers, only 1 study each)	Insufficient
KQ1 KQ1 Biomarkers Sensitivity	4 studies{#5685} {#327}{#5003} {#417}	Performance ranged from 98% (MFNU, corresponding specificity 25%){#5685} to 80% (eye tracker, corresponding specificity 83%){#417}	Imprecision (values varied)	Low for acceptable sensitivity
KQ1 Biomarker Specificity	4 studies{#5685} {#327}{#5003} {#417}	Performance ranged from 83% (eye tracker, corresponding sensitivity 80%) to 25% (MFNU, corresponding sensitivity 98%)	Imprecision (values ranged widely)	Low for acceptable specificity
KQ1 Biomarker Administration and scoring time	1 study{#417}	1 study reported that the eye tracking task took about 15 minutes	Inconsistency (no replication, very specific task)	Insufficient
KQ1 Clinician tools Clinical misdiagnosis	3 studies{#684} {#4185}{#5349}	Reported false positive rates ranged from 9% for the MINI-Plus in an addiction treatment center{#684} to 48% in an inpatient psychiatric hospital unit for the CAARS-O{#4185}	Inconsistency (different tools), Imprecision (values ranged widely), Study limitation (likely dependent on patient population)	Insufficient
KQ1 Clinician tool Sensitivity	3 studies{#684} {#4185}{#5349}	Performance ranged from 83% (corresponding specificity 52%) to 75% (corresponding specificity 91%){#684}	Imprecision (values varied), Inconsistency (different tools), Study	Insufficient

			limitation (likely dependent on patient population)	
KQ1 Clinician tool Specificity	3 studies{#684}{#4185}{#5349}	Performance ranged from 91% (corresponding sensitivity 75%){#684} to 52% (corresponding sensitivity 83%){#4185}	Imprecision (values varied), Inconsistency (different tools), Study limitation (likely dependent on patient population)	Insufficient
KQ1 All tests Costs	0 studies	N/A	N/A	Insufficient
KQ1 All tests Concordance primary care and specialty	0 studies	N/A	N/A	Insufficient
KQ1a Effect of clinical setting All outcomes	N/A	N/A	Inconsistency (lack of primary care studies)	Insufficient
KQ1a Effect of comparator sample Clinical misdiagnosis	N/A	4 studies noted that tests were less effective when distinguishing ADHD from other clinical conditions (rather than the general population) due to overlapping symptoms{#870}{#212}{#495}{#5685}	Study limitation (not all tests addressed)	Low for higher risk of clinical misdiagnosis in clinical samples
KQ1a Effect of ADHD presentation Sensitivity	N/A	Conflicting results across 4 studies{#857}{#5176}{#490}{#444}	Inconsistency (conflicting results)	Insufficient
KQ1a Effect of age of participants and age at diagnosis Sensitivity and specificity	N/A	1 study reported that sensitivity and specificity were higher in younger adults (18-44) compared to older adults (31-44){#444}	Inconsistency (no replication)	Insufficient
KQ1a Effect of participant sex Sensitivity and specificity	N/A	Conflicting results across 5 studies{#425}{#134}{#960}{#111}{#1916}{#4185}	Inconsistency (conflicting results)	Insufficient
KQ1a Effect of comorbidities Sensitivity	N/A	Conflicting results: while 2 studies reported no effect of comorbidities on sensitivity,{#490}{#942} 1 study reported lower sensitivity,{#546} and 2 studies reported that diagnostic performance was unaffected by comorbidities{#111}{#495}	Inconsistency (conflicting results)	Insufficient
KQ1a Effect of comorbidities Specificity	N/A	3 studies reported challenges for specificity,{#490}{#942}{#220} 1 study reported higher specificity in participants with comorbidities,{#546} 2 studies reported that diagnostic performance was unaffected by comorbidities;{#111}{#495} 4studies pointed out the high prevalence of comorbid conditions	Study limitation (not all tests addressed)	Low for lower specificity in participants with comorbidities

		such as depression and anxiety.{#857}{#4485}{#444}{#323}		

Notes: KQ key question, N/A not applicable or not available

We identified numerous studies reporting on the performance of tests for detecting feigning ADHD. Studies used subjective and objective tests. Results are shown in the evidence table in the appendix (Appendix Table A9).

The included studies did not report on the impact for participants of being correctly or incorrectly diagnosed. Studies reported only on the performance of the tests but not the effect a diagnosis (or a misdiagnosis) had on participants or similar.

None of the included studies reported on unintended consequences, adverse events, adverse effects, or side effects of the diagnostic tools, including blood-based biomarker, EEG, neuroimaging, and neuropsychological test studies.

4. Discussion

4.1 Findings in Relation to Decisional Dilemmas

- Contextual Question. How frequently are the various tools to diagnose ADHD in adults currently being used?

KQ: What is the comparative diagnostic accuracy, unintended consequences and impact of tools that can be used in the primary care practice setting or by specialists to diagnose ADHD among adults?

XXX

4.2 Implications

XXX

4.3 Strengths and Limitations

Systematic review

Evidence base

4.4 Next Steps

XXX

5. References

1. Weiss G, Hechtman LT. Hyperactive Children Grown Up, Second Edition: ADHD in Children, Adolescents, and Adults: Guilford Publications; 1993.

2. Mannuzza S, Klein RG, Bessler A, et al. Adult psychiatric status of hyperactive boys grown up. Am J Psychiatry. 1998 Apr;155(4):493-8. doi: 10.1176/ajp.155.4.493. PMID: 9545994.

3. Biederman J, Mick E, Faraone SV. Age-dependent decline of symptoms of attention deficit hyperactivity disorder: impact of remission definition and symptom type. Am J Psychiatry. 2000 May;157(5):816-8. doi: 10.1176/appi.ajp.157.5.816. PMID: 10784477.

4. Wilens TE, Biederman J, Spencer TJ. Attention deficit/hyperactivity disorder across the lifespan. Annu Rev Med. 2002;53:113-31. doi: 10.1146/annurev.med.53.082901.103945. PMID: 11818466.

5. Mannuzza S, Klein RG, Bessler A, et al. Adult outcome of hyperactive boys. Educational achievement, occupational rank, and psychiatric status. Arch Gen Psychiatry. 1993 Jul;50(7):565-76. doi: 10.1001/archpsyc.1993.01820190067007. PMID: 8317950.

6. Center for Disease Control and Prevention. Data and Statistics About ADHD. https://www.cdcgov/ncbddd/adhd/datahtml. Accessed on June 10, 2024.

7. Caci HM, Morin AJ, Tran A. Prevalence and correlates of attention deficit hyperactivity disorder in adults from a French community sample. J Nerv Ment Dis. 2014 Apr;202(4):324-32. doi: 10.1097/nmd.0000000000000126. PMID: 24647218.

8. Das D, Cherbuin N, Butterworth P, et al. A population-based study of attention deficit/hyperactivity disorder symptoms and associated impairment in middle-aged adults. PLoS One. 2012;7(2):e31500. doi: 10.1371/journal.pone.0031500. PMID: 22347487.

9. Estevez N, Eich-Hochli D, Dey M, et al. Prevalence of and associated factors for adult attention deficit hyperactivity disorder in young Swiss men. PLoS One. 2014;9(2):e89298. doi: 10.1371/journal.pone.0089298. PMID: 24586672.

10. Moulin F, Chollet A, Ramos-Quiroga JA, et al. Prevalence and Psychosocial Correlates of ADHD Symptoms in Young Adulthood: A French Population-Based Study. J Atten Disord. 2018 Jan;22(2):167-81. doi: 10.1177/1087054717706758. PMID: 28490216.

11. Asherson P, Akehurst R, Kooij JJ, et al. Under diagnosis of adult ADHD: cultural influences and societal burden. J Atten Disord. 2012 Jul;16(5 Suppl):20s-38s. doi: 10.1177/1087054711435360. PMID: 22377849.

12. Cook J, Knight E, Hume I, et al. The self-esteem of adults diagnosed with attention-deficit/hyperactivity disorder (ADHD): a systematic review of the literature. Atten Defic Hyperact Disord. 2014 Dec;6(4):249-68. doi: 10.1007/s12402-014-0133-2. PMID: 24668198.

13. Compton WM, Han B, Blanco C, et al. Prevalence and Correlates of Prescription Stimulant Use, Misuse, Use Disorders, and Motivations for Misuse Among Adults in the United States. Am J Psychiatry. 2018 Aug 1;175(8):741-55. doi: 10.1176/appi.ajp.2018.17091048. PMID: 29656665.

14. Caisley H, Müller U. Adherence to medication in adults with attention deficit hyperactivity disorder and pro re nata dosing of psychostimulants: a systematic review. Eur Psychiatry. 2012 Jul;27(5):343-9. doi: 10.1016/j.eurpsy.2012.01.002. PMID: 22521805.

15. Castells X, Blanco-Silvente L, Cunill R. Amphetamines for attention deficit hyperactivity disorder (ADHD) in adults. Cochrane Database Syst Rev. 2018 Aug 9;8(8):Cd007813. doi: 10.1002/14651858.CD007813.pub3. PMID: 30091808.

16. Cunill R, Castells X, Tobias A, et al. Atomoxetine for attention deficit hyperactivity disorder in the adulthood: a meta-analysis and meta-regression. Pharmacoepidemiol Drug Saf. 2013 Sep;22(9):961-9. doi: 10.1002/pds.3473. PMID: 23813665.

17. Cunill R, Castells X, Tobias A, et al. Efficacy, safety and variability in pharmacotherapy for adults with attention deficit hyperactivity disorder: a meta-analysis and meta-regression in over 9000 patients. Psychopharmacology (Berl). 2016 Jan;233(2):187-97. doi: 10.1007/s00213-015-4099-3. PMID: 26446868.

18. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. Psychol Med. 2006 Feb;36(2):159-65. doi: 10.1017/S003329170500471X. PMID: 16420712.

19. Kooij SJ, Bejerot S, Blackwell A, et al. European consensus statement on diagnosis and treatment of adult ADHD: The European Network Adult ADHD. BMC Psychiatry. 2010 Sep 3;10:67. doi: 10.1186/1471-244X-10-67. PMID: 20815868.

20. Sibley MH, Swanson JM, Arnold LE, et al. Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity. J Child Psychol Psychiatry. 2017 Jun;58(6):655-62. doi: 10.1111/jcpp.12620. PMID: 27642116.

21. Benson K, Flory K, Humphreys KL, et al. Misuse of Stimulant Medication Among College Students: A Comprehensive Review and Meta-analysis. Clinical Child and Family Psychology Review. 2015 2015/03/01;18(1):50-76. doi: 10.1007/s10567-014-0177-z.

22. Agay N, Yechiam E, Carmel Z, et al. Methylphenidate enhances cognitive performance in adults with poor baseline capacities regardless of attention-deficit/hyperactivity disorder diagnosis. J Clin Psychopharmacol. 2014 Apr;34(2):261-5. doi: 10.1097/jcp.0000000000000076. PMID: 24525641.

23. Hester R, Nandam LS, O'Connell RG, et al. Neurochemical enhancement of conscious error awareness. J Neurosci. 2012 Feb 22;32(8):2619-27. doi: 10.1523/jneurosci.4052-11.2012. PMID: 22357846.

24. Rapoport JL, Buchsbaum MS, Weingartner H, et al. Dextroamphetamine. Its cognitive and behavioral effects in normal and hyperactive boys and normal men. Arch Gen Psychiatry. 1980 Aug;37(8):933-43. doi: 10.1001/archpsyc.1980.01780210091010. PMID: 7406657.

25. Rapoport JL, Buchsbaum MS, Zahn TP, et al. Dextroamphetamine: cognitive and behavioral effects in normal prepubertal boys. Science. 1978 Feb 3;199(4328):560-3. doi: 10.1126/science.341313. PMID: 341313.

26. Turner DC, Robbins TW, Clark L, et al. Cognitive enhancing effects of modafinil in healthy volunteers. Psychopharmacology

(Berl). 2003 Jan;165(3):260-9. doi: 10.1007/s00213-002-1250-8. PMID: 12417966.

27. Taylor A, Deb S, Unwin G. Scales for the identification of adults with attention deficit hyperactivity disorder (ADHD): a systematic review. Res Dev Disabil. 2011 May-Jun;32(3):924-38. doi: 10.1016/j.ridd.2010.12.036. PMID: 21316190.

28. Peterson BS, Trampush J, Brown M, et al. Tools for the Diagnosis of ADHD in Children and Adolescents: A Systematic Review. Pediatrics. 2024 Apr 1;153(4). doi: 10.1542/peds.2024-065854. PMID: 38523599.

29. Sibley MH, Swanson JM, Arnold LE, et al. Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity. J Child Psychol Psychiatry. 2017 Jun;58(6):655-62. doi: 10.1111/jcpp.12620. PMID: 27642116.

30. Chandra S, Biederman J, Faraone SV. Assessing the Validity of the Age at Onset Criterion for Diagnosing ADHD in DSM-5. J Atten Disord. 2021 Jan;25(2):143-53. doi: 10.1177/1087054716629717. PMID: 26922806.

31. Caye A, Rocha TB, Anselmi L, et al. Attention-Deficit/Hyperactivity Disorder Trajectories From Childhood to Young Adulthood: Evidence From a Birth Cohort Supporting a Late-Onset Syndrome. JAMA Psychiatry. 2016 Jul 1;73(7):705-12. doi: 10.1001/jamapsychiatry.2016.0383. PMID: 27192050.

32. Agnew-Blais JC, Polanczyk GV, Danese A, et al. Evaluation of the Persistence, Remission, and Emergence of Attention-Deficit/Hyperactivity Disorder in Young Adulthood. JAMA Psychiatry. 2016 Jul 1;73(7):713-20. doi: 10.1001/jamapsychiatry.2016.0465. PMID: 27192174.

33. Asherson P, Buitelaar J, Faraone SV, et al. ADHD Management in Adolescents Transitioning to Adulthood: Challenges and Opportunities. Postgraduate Medicine. 2016;128(8):774-83.

34. Moffitt TE, Houts R, Asherson P, et al. Is Adult ADHD a Childhood-Onset Neurodevelopmental Disorder? Evidence From a Four-Decade Longitudinal Cohort Study. Am J Psychiatry. 2015 Oct;172(10):967-77. doi: 10.1176/appi.ajp.2015.14101266. PMID: 25998281.

35. Plumber N, Majeed M, Ziff S, et al. Stimulant Usage by Medical Students for Cognitive Enhancement: A Systematic Review. Cureus. 2021 May 22;13(5):e15163. doi: 10.7759/cureus.15163. PMID: 34178492.

36. Sharif S, Guirguis A, Fergus S, et al. The Use and Impact of Cognitive Enhancers among University Students: A Systematic Review. Brain Sci. 2021 Mar 10;11(3). doi: 10.3390/brainsci11030355. PMID: 33802176.

37. Epstein JN, Kollins SH. Psychometric properties of an adult ADHD diagnostic interview. J Atten Disord. 2006 Feb;9(3):504-14. doi: 10.1177/1087054705283575. PMID: 16481667.

38. Marshall P, Hoelzle J, Nikolas M. Diagnosing Attention-Deficit/Hyperactivity Disorder (ADHD) in young adults: A qualitative review of the utility of assessment measures and recommendations for improving the diagnostic process. Clin Neuropsychol. 2021 Jan;35(1):165-98. doi: 10.1080/13854046.2019.1696409. PMID: 31791193.

39. Gorlin EI, Dalrymple K, Chelminski I, et al. Reliability and validity of a semi-structured DSM-based diagnostic interview module for the assessment of Attention

Deficit Hyperactivity Disorder in adult psychiatric outpatients. Psychiatry Res. 2016 Aug 30;242:46-53. doi: 10.1016/j.psychres.2016.05.020. PMID: 27259136.

40. Wilens TE, Adler LA, Adams J, et al. Misuse and diversion of stimulants prescribed for ADHD: a systematic review of the literature. J Am Acad Child Adolesc Psychiatry. 2008 Jan;47(1):21-31. doi: 10.1097/chi.0b013e31815a56f1. PMID: 18174822.

41. Staley BS, Robinson LR, Claussen AH, et al. Attention-Deficit/Hyperactivity Disorder Diagnosis, Treatment, and Telehealth Use in Adults - National Center for Health Statistics Rapid Surveys System, United States, October-November 2023. MMWR Morb Mortal Wkly Rep. 2024 Oct 10;73(40):890-5. doi: 10.15585/mmwr.mm7340a1. PMID: 39388378.

42. Adler LA, Faraone SV, Spencer TJ, et al. The reliability and validity of self- and investigator ratings of ADHD in adults. J Atten Disord. 2008 May;11(6):711-9. doi: 10.1177/1087054707308503. PMID: 18025250.

43. Pagán AF, Huizar YP, Schmidt AT. Conner's Continuous Performance Test and Adult ADHD: A Systematic Literature Review. J Atten Disord. 2023 Feb;27(3):231-49. doi: 10.1177/10870547221142455. PMID: 36495125.

44. Varela JL, Magnante AT, Miskey HM, et al. A systematic review of the utility of continuous performance tests among adults with ADHD. Clin Neuropsychol. 2024 Feb 29:1-62. doi: 10.1080/13854046.2024.2315740. PMID: 38424025.

45. APSARD. U.S. Based Guidelines for Adults with ADHD. n.d. https://apsard.org/us-guidelines-for-adults-with-adhd/. Accessed on November 10, 2024.

46. Diagnosis of Attention-Deficit/Hyperactivity Disorder in Adults: A Systematic Review. Rockville, MD: Agency for Healthcare Research and Quality; February 2025. https://effectivehealthcare.ahrq.gov/products/hyperactivity-disorder/protocol. Accessed on March 7, 2025.

47. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD Effective Health Care Program, Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/products/collections/cer-methods-guide. Accessed on October 7, 2024.

48. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529-36. doi: 10.7326/0003-4819-155-8-201110180-00009. PMID: 22007046.

49. Erhardt D, Epstein JN, Conners CK, et al. Self-ratings of ADHD symptoms in adults: II. Reliability, validity, and diagnostic sensitivity. Journal of Attention Disorders. 1999;3(3):153-8. doi: 10.1177/108705479900300304.

50. Amen DG, Henderson TA, Newberg A. SPECT Functional Neuroimaging Distinguishes Adult Attention Deficit Hyperactivity Disorder From Healthy Controls in Big Data Imaging Cohorts. Front Psychiatry. 2021;12:725788. doi: 10.3389/fpsyt.2021.725788. PMID: 34899414.

51. Palma-Álvarez RF, Barta C, Carpentier PJ, et al. Validity of the ADHD module of the Mini International Neuropsychiatric Interview PLUS for screening of adult ADHD in treatment seeking substance use

disorder patients: ADHD screening with MINI-Plus. Span J Psychiatry Ment Health. 2023 Jan-Mar;16(1):11-5. doi: 10.1016/j.rpsm.2020.04.013. PMID: 32561156.

52. van de Glind G, van den Brink W, Koeter MW, et al. Validity of the Adult ADHD Self-Report Scale (ASRS) as a screener for adult ADHD in treatment seeking substance use disorder patients. Drug Alcohol Depend. 2013 Oct 1;132(3):587-96. doi: 10.1016/j.drugalcdep.2013.04.010. PMID: 23660242.

53. Groom MJ, Young Z, Hall CL, et al. The incremental validity of a computerised assessment added to clinical rating scales to differentiate adult ADHD from autism spectrum disorder. Psychiatry Res. 2016 Sep 30;243:168-73. doi: 10.1016/j.psychres.2016.06.042. PMID: 27400220.

54. Palmer M, Fang Z, Hollocks MJ, et al. Screening for Attention Deficit Hyperactivity Disorder in Young Autistic Adults: The Diagnostic Accuracy of Three Commonly Used Questionnaires. J Autism Dev Disord. 2023 Oct 28. doi: 10.1007/s10803-023-06146-9. PMID: 37898580.

55. Jiménez EC, Avella-Garcia C, Kustow J, et al. Eye Vergence Responses During an Attention Task in Adults With ADHD and Clinical Controls. J Atten Disord. 2021 Jul;25(9):1302-10. doi: 10.1177/1087054719897806. PMID: 31959011.

56. Kingston DA, Ahmed AG, Gray J, et al. The assessment and diagnosis of attention deficit hyperactivity disorder in adult forensic psychiatric outpatients. Journal of Psychopathology and Behavioral Assessment. 2013;35(3):293-300. doi: 10.1007/s10862-013-9346-5.

57. Reimherr FW, Marchant BK, Gift TE, et al. Psychometric data and versions of the Wender Utah Rating Scale including the WURS-25 & WURS-45. Data Brief. 2021 Aug;37:107232. doi: 10.1016/j.dib.2021.107232. PMID: 34235235.

58. Nikolas MA, Marshall P, Hoelzle JB. The role of neurocognitive tests in the assessment of adult attention-deficit/hyperactivity disorder. Psychological assessment. 2019;31(5):685.

59. Abramson DA, White DJ, Rhoads T, et al. Cross-validating the Dot Counting Test Among an Adult ADHD Clinical Sample and Analyzing the Effect of ADHD Subtype and Comorbid Psychopathology. Assessment. 2023 Mar;30(2):264-73. doi: 10.1177/10731911211050895. PMID: 34643101.

60. Potts HE, Lewandowski LJ, Lovett BJ. Identifying Feigned ADHD in College Students: Comparing the Multidimensional ADHD Rating Scale to Established Validity Measures. J Atten Disord. 2022 Oct;26(12):1622-30. doi: 10.1177/10870547221092095. PMID: 35466735.

61. Robinson A, Reed C, Davis K, et al. Settling the Score: Can CPT-3 Embedded Validity Indicators Distinguish Between Credible and Non-Credible Responders Referred for ADHD and/or SLD? J Atten Disord. 2023 Jan;27(1):80-8. doi: 10.1177/10870547221121781. PMID: 36113024.

62. Rogers R, Velsor SF, Donnelly JW, 2nd, et al. Embedded WAIS-IV Detection Strategies and Feigned Cognitive Impairment: An Investigation of Malingered ADHD. Assessment. 2021 Jan;28(1):44-56. doi: 10.1177/1073191120927788. PMID: 32495690.

Abbreviations and Acronyms

ACRONYM	Definition
ACRONYM	Definition
ACRONYM	Definition
ACRONYM	Definition
ACRONYM	Definition
ACRONYM	Definition

Appendix A. Search Strategy

Date: October 14, 2024

PubMed

"Attention Deficit Disorder with Hyperactivity"[Mesh] OR "attention deficit hyperactivity disorder"[tiab] OR "ADHD"[tiab] OR "attention deficit disorder"[tiab])
AND
Adult[MESH] OR Aged[MESH] OR Middle Aged[MESH] OR Young Adult[MESH] OR Adult[Title/Abstract] OR Adults[Title/Abstract]
AND
"Attention Deficit and Disruptive Behavior Disorders/diagnosis"[Majr] OR mass screening[mesh] OR questionnaires[mesh] OR Interviews as Topic[Mesh] OR Psychometrics[Mesh] OR Psychiatric Status Rating Scales[Mesh] OR diagnosis[mesh:noexp] OR "Diagnostic Techniques and Procedures"[Mesh] OR "Referral and Consultation"[Mesh] OR questionnaire[tiab] OR questionnaires[tiab] OR screening[tiab] OR screen[tiab] OR scale[tiab] OR instrument[tiab] OR instruments[tiab] OR interview[tiab] OR interviews[tiab] OR diagnosis[tiab] OR diagnostic[tiab] OR diagnosed[tiab] OR Measure [tiab] OR test[tiab] OR tests[tiab] OR testing[tiab] OR "Attention Deficit Disorder with Hyperactivity/diagnostic imaging"[Majr] OR (("Adaptive Behavior Assessment System"[Title/Abstract] OR "ABAS-3"[Title/Abstract] OR "Advanced Clinical Solutions"[Title/Abstract] OR "Word Choice Test"[Title/Abstract] OR "Test of Premorbid Functioning"[Title/Abstract] OR "Social Cognition"[Title/Abstract] OR "Beck Anxiety Inventory"[Title/Abstract] OR "BAI"[Title/Abstract] OR "Beck Depression Inventory"[Title/Abstract] OR "BDI-2"[Title/Abstract] OR "Behavioral Assessment System for Children"[Title/Abstract] OR "Self-Report of Personality"[Title/Abstract] OR "BASC-3 SRP Adolescent"[Title/Abstract] OR "Behavioral Assessment System for Children"[Title/Abstract] OR "Parent Rating Scales"[Title/Abstract] OR "BASC-3 PRS Adolescent"[Title/Abstract] OR "BASC-3 SRP College"[Title/Abstract] OR "Teacher Rating Scales"[Title/Abstract] OR "BASC-3 TRS Adolescent"[Title/Abstract] OR "Brown Executive Function/Attention Scales"[Title/Abstract] OR "Brown EF/A Self"[Title/Abstract] OR "California Verbal Learning Test"[Title/Abstract] OR "CVLT-3"[Title/Abstract] OR "Standard Form California Verbal" "CVLT-3 Brief"[Title/Abstract] OR "California Verbal Learning Test"[Title/Abstract] OR "CVLT-C"[Title/Abstract] OR "Childhood Autism Rating Scale"[Title/Abstract] OR "CARS-2"[Title/Abstract] OR "Childhood Autism Rating Scale"[Title/Abstract] OR "High-Functioning Version"[Title/Abstract] OR "CARS-2 HF"[Title/Abstract] OR "Clinical Evaluation of Language Fundamentals"[Title/Abstract] OR "CELF-5"[Title/Abstract] OR "Comprehensive Executive Function Inventory"[Title/Abstract] OR "CEFI Adult Observer"[Title/Abstract] OR "Comprehensive Executive Function Inventory"[Title/Abstract] OR "CEFI Adult Self-Report"[Title/Abstract] OR "Conners' Adult ADHD Diagnostic Interview for DSM-IV"[Title/Abstract] OR "CAADID Part 1"[Title/Abstract] OR "CAADID Part 2"[Title/Abstract] OR "CAARS–O:L"[Title/Abstract] OR "CAARS–S:L"[Title/Abstract] OR "CAARS-2 Observer"[Title/Abstract] OR "Conners' Adult ADHD Rating Scales"[Title/Abstract] OR "CAARS-2 Self-Report"[Title/Abstract] OR "Delis-Kaplan Executive Function System"[Title/Abstract] OR "D-KEFS"[Title/Abstract] OR "Dot Counting Test"[Title/Abstract] OR "Grooved Pegboard Test Kaufman Test of Educational Achievement"[Title/Abstract] OR "KTEA-3"[Title/Abstract] OR "Neuropsychological Assessment

Battery"[Title/Abstract] OR "Attention, Language, Memory, Spatial, and Executive Functions Modules"[Title/Abstract] OR "NIH Executive Abilities–Measures and Instruments for Neurobehavioral Evaluation and Re-search"[Title/Abstract] OR "NIH EXAMINER"[Title/Abstract] OR "Personality Assessment Inventory"[Title/Abstract] OR "PROMIS Sleep Assessments Pediatric Parent Proxy"[Title/Abstract] OR "Repeatable Battery for the Assessment of Neuropsychological Status"[Title/Abstract] OR "RBANS"[Title/Abstract] OR "Rey-Osterrieth Complex"[Title/Abstract] OR "Wechsler Abbreviated Scale of Intelligence"[Title/Abstract] OR "WASI-2"[Title/Abstract] OR "Wechsler Adult Intelligence Scale"[Title/Abstract] OR "WAIS-4"[Title/Abstract] OR "WAIS-IV"[Title/Abstract] OR "Wechsler Individual Achievement Test"[Title/Abstract] OR "WIAT-4"[Title/Abstract] OR "Wechsler Intelligence Scale "[Title/Abstract] OR "Wechsler Memory Scale"[Title/Abstract] OR "WMS-4"[Title/Abstract] OR "Wide Range Achievement Test"[Title/Abstract] OR "WRAT-5"[Title/Abstract] OR "Adult ADHD Rating Scale"[Title/Abstract] OR "ADHD-RS"[Title/Abstract] OR "Brown ADD scales"[Title/Abstract] OR "Continuous Performance Tests"[Title/Abstract] OR "Conners CPT"[Title/Abstract] OR "QB Test"[Title/Abstract] OR "TOVA"[Title/Abstract] OR "Wender Utah Adult ADHD Scale"[Title/Abstract]))
AND
"Sensitivity and Specificity"[Mesh] OR "Diagnostic Errors"[Mesh] OR sensitivity[tiab] OR specificity[tiab] OR (accura*[tiab] AND (diagnos*[tiab] OR classif*[tiab])) OR "ROC curve"[tiab] OR "positive predictive value"[tiab] OR "negative predictive value"[tiab] OR "false positive"[tiab] OR "false negative"[tiab] OR "likelihood ratio"[tiab]
NOT
Editorial[ptyp] OR Letter[pt] OR Case Reports[pt] OR Comment[pt] address[pt] OR "autobiography"[pt] OR "bibliography"[pt] OR "biography"[pt] OR "case report"[tw] OR "case reports"[tw] OR "case series"[tw] OR "comment on"[All Fields] OR congress[pt] OR "dictionary"[pt] OR "directory"[pt] OR "festschrift"[pt] OR "historical article"[pt] OR lecture[pt] OR "legal case"[pt] OR "legislation"[pt] OR "news"[pt] OR "newspaper article"[pt] OR "patient education handout"[pt] OR "periodical index"[pt]
NOT
"animals"[mesh] NOT "humans"[mesh])

EMBASE

((((('adaptive behavior assessment system':ti OR 'abas-3':ti OR 'advanced clinical solutions':ti OR 'word choice test':ti OR 'test of premorbid functioning':ti OR 'social cognition':ti OR 'beck anxiety inventory':ti OR 'bai':ti OR 'beck depression inventory':ti OR 'bdi-2':ti OR 'self-report of personality':ti OR 'basc-3 srp adolescent':ti OR 'behavioral assessment system for children':ti OR 'parent rating scales':ti OR 'basc-3 prs adolescent':ti OR 'basc-3 srp college':ti OR 'teacher rating scales':ti OR 'basc-3 trs adolescent':ti OR 'brown executive function/attention scales':ti OR 'brown ef/a self':ti OR 'california verbal learning test':ti OR 'cvlt-3':ti OR 'standard form california verbal':ti) AND 'cvlt-3 brief':ti OR 'california verbal learning test':ti OR 'cvlt-c':ti OR 'cars-2':ti OR 'childhood autism rating scale':ti OR 'high-functioning version':ti OR 'cars-2 hf':ti OR 'clinical evaluation of language fundamentals':ti OR 'celf-5':ti OR 'cefi adult observer':ti OR 'comprehensive executive function inventory':ti OR 'cefi adult self-report':ti OR 'conners adult adhd diagnostic interview for dsm-iv':ti OR 'caadid part 1':ti OR 'caadid part 2':ti OR 'caars–o:l':ti OR 'caars–s:l':ti OR 'caars-2 observer':ti OR 'conners adult adhd rating scales':ti OR 'caars-2 self-report':ti OR 'delis-kaplan executive function system':ti OR 'd-kefs':ti OR 'dot counting test':ti

OR 'grooved pegboard test kaufman test of educational achievement':ti OR 'ktea-3':ti OR 'nepsy-ii developmental neuropsychological battery':ti OR 'neuropsychological assessment battery':ti OR 'attention, language, memory, spatial,':ti) AND 'ex- ecutive functions modules':ti OR 'nih executive abilities–measures':ti) AND 'instruments for neurobehavioral evaluation':ti AND 're search':ti OR 'nih examiner':ti OR 'personality assessment inventory':ti OR 'promis sleep assessments pediatric parent proxy':ti OR 'repeatable battery for the assessment of neuropsychological status':ti OR 'rbans':ti OR 'rey-osterrieth complex':ti OR 'wechsler abbreviated scale of intelligence':ti OR 'wasi-2':ti OR 'wechsler adult intelligence scale':ti OR 'wais-4':ti OR 'wais-iv':ti OR 'wechsler individual achievement test':ti OR 'wiat-4':ti OR 'wechsler intelligence scale':ti OR OR 'wechsler memory scale':ti OR 'wms-4':ti OR 'wide range achievement test':ti OR 'wrat-5':ti OR 'adult adhd rating scale':ti OR 'adhd-rs':ti OR 'brown add scales':ti OR 'continuous performance tests':ti OR 'conners cpt':ti OR 'qb test':ti OR 'tova':ti OR 'wender utah adult adhd scale':ti OR 'diagnostic interview for adult adhd':ti
AND
"Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder")
OR
(#1 'attention deficit disorder with hyperactivity':ab,ti OR 'attention deficit hyperactivity disorder':ab,ti OR 'adhd':ab,ti OR 'attention deficit disorder':ab,ti 61194
#2 ((adult:ab,ti OR aged:ab,ti OR middle:ab,ti) AND aged:ab,ti OR young:ab,ti) AND adult:ab,ti OR adult:ab,ti OR adults:ab,ti 2231374
#3 ((('attention deficit and disruptive behavior disorders/diagnosis':ab,ti OR mass:ab,ti) AND screening:ab,ti OR questionnaires:ab,ti OR interviews:ab,ti) AND as:ab,ti AND topic:ab,ti OR psychometrics:ab,ti OR psychiatric:ab,ti) AND status:ab,ti AND rating:ab,ti AND scales:ab,ti OR 'diagnostic techniques and procedures':ab,ti OR 'referral and consultation':ab,ti OR questionnaire:ab,ti OR questionnaires:ab,ti OR screening:ab,ti OR screen:ab,ti OR scale:ab,ti OR instrument:ab,ti OR instruments:ab,ti OR interview:ab,ti OR interviews:ab,ti OR diagnosis:ab,ti OR diagnostic:ab,ti OR diagnosed:ab,ti OR measure:ab,ti OR test:ab,ti OR tests:ab,ti OR testing:ab,ti OR 'attention deficit disorder with hyperactivity/diagnostic imaging':ab,ti 11386521
#4 'sensitivity and specificity':ab,ti OR 'diagnostic errors':ab,ti OR sensitivity:ab,ti OR specificity:ab,ti OR (accura*:ab,ti AND (diagnos*:ab,ti OR classif*:ab,ti)) OR 'roc curve':ab,ti OR 'positive predictive value':ab,ti OR 'negative predictive value':ab,ti OR 'false positive':ab,ti OR 'false negative':ab,ti OR 'likelihood ratio':ab,ti 2280552
#5 #1 AND #2 AND #3 AND #4 814
#6 #5 AND [humans]/lim 787
#7 #6 AND ([article]/lim OR [article in press]/lim) 509)

APA PsycINFO

(((title: ("Adaptive Behavior Assessment System") OR title: ("ABAS-3") OR title: ("Advanced Clinical Solutions") OR title: ("Word Choice Test") OR title: ("Test of Premorbid Functioning") OR title: ("Social Cognition") OR title: ("Beck Anxiety Inventory") OR title: ("BAI") OR title: ("Beck Depression Inventory") OR title: ("BDI-2") OR title: ("Behavioral Assessment System for Children") OR title: ("Self-Report of Personality") OR title: ("BASC-3 SRP Adolescent") OR title: ("Behavioral Assessment System for Children") OR title: ("Parent Rating Scales") OR title: ("BASC-3 PRS Adolescent") OR title: ("BASC-3 SRP College") OR title: ("Teacher Rating

Scales") OR title: ("BASC-3 TRS Adolescent") OR title: ("Brown Executive Function/Attention Scales") OR title: ("Brown EF/A Self") OR title: ("California Verbal Learning Test") OR title: ("CVLT-3") OR title: ("Standard Form California Verbal" "CVLT-3 Brief") OR title: ("California Verbal Learning Test") OR title: ("CVLT-C") OR title: ("Childhood Autism Rating Scale") OR title: ("CARS-2") OR title: ("Childhood Autism Rating Scale") OR title: ("High-Functioning Version") OR title: ("CARS-2 HF") OR title: ("Clinical Evaluation of Language Fundamentals") OR title: ("CELF-5") OR title: ("Comprehensive Executive Function Inventory") OR title: ("CEFI Adult Observer") OR title: ("Comprehensive Executive Function Inventory") OR title: ("CEFI Adult Self-Report") OR title: ("Conners' Adult ADHD Diagnostic Interview for DSM-IV") OR title: ("CAADID Part 1") OR title: ("CAADID Part 2") OR title: ("CAARS–O:L") OR title: ("CAARS–S:L") OR title: ("CAARS-2 Observer") OR title: ("Conners' Adult ADHD Rating Scales") OR title: ("CAARS-2 Self-Report") OR title: ("Delis-Kaplan Executive Function System") OR title: ("D-KEFS") OR title: ("Dot Counting Test") OR title: ("Grooved Pegboard Test Kaufman Test of Educational Achievement") OR title: ("KTEA-3") OR title: ("NEPSY-II Developmental Neuropsychological Battery") OR title: ("Neuropsychological Assessment Battery") OR title: ("Attention, Language, Memory, Spatial, and Ex- ecutive Functions Modules") OR title: ("NIH Executive Abilities–Measures and Instruments for Neurobehavioral Evaluation and Re-search") OR title: ("NIH EXAMINER") OR title: ("Personality Assessment Inventory") OR title: ("PROMIS Sleep Assessments Pediatric Parent Proxy") OR title: ("Repeatable Battery for the Assessment of Neuropsychological Status") OR title: ("RBANS") OR title: ("Rey-Osterrieth Complex") OR title: ("Wechsler Abbreviated Scale of Intelligence") OR title: ("WASI-2") OR title: ("Wechsler Adult Intelligence Scale") OR title: ("WAIS-4") OR title: ("WAIS-IV") OR title: ("Wechsler Individual Achievement Test") OR title: ("WIAT-4") OR title: ("Wechsler Intelligence Scale ") OR title: ("Wechsler Memory Scale") OR title: ("WMS-4") OR title: ("Wide Range Achievement Test") OR title: ("WRAT-5") OR title: ("Adult ADHD Rating Scale") OR title: ("ADHD-RS") OR title: ("Brown ADD scales") OR title: ("Continuous Performance Tests") OR title: ("Conners CPT") OR title: ("QB Test") OR title: ("TOVA") OR title: ("Wender Utah Adult ADHD Scale") OR title: ("diagnostic interview for Adult ADHD")))
AND
((title: ("Attention Deficit Disorder with Hyperactivity") OR title: ("attention deficit hyperactivity disorder") OR title: ("ADHD") OR title: ("attention deficit disorder")) OR (abstract: ("Attention Deficit Disorder with Hyperactivity") OR abstract: ("attention deficit hyperactivity disorder") OR abstract: ("ADHD") OR abstract: ("attention deficit disorder"))))
OR
(((title: ("Attention Deficit Disorder with Hyperactivity") OR title: ("attention deficit hyperactivity disorder") OR title: ("ADHD") OR title: ("attention deficit disorder")) OR (abstract: ("Attention Deficit Disorder with Hyperactivity") OR abstract: ("attention deficit hyperactivity disorder") OR abstract: ("ADHD") OR abstract: ("attention deficit disorder"))) AND ((title: (Adult) OR title: (Aged) OR title: (Middle Aged) OR title: (Young Adult) OR title: (Adult) OR title: (Adults)) OR (abstract: (Adult) OR abstract: (Aged) OR abstract: (Middle Aged) OR abstract: (Young Adult) OR abstract: (Adult) OR abstract: (Adults))) AND ((title: ("Attention Deficit and Disruptive Behavior Disorders/diagnosis") OR title: (mass screening) OR title: (questionnaires) OR title: (Interviews as Topic) OR title: (Psychometrics) OR title: (Psychiatric Status Rating Scales) OR title: (diagnosis) OR title: ("Diagnostic Techniques and Procedures") OR title: ("Referral and Consultation") OR title: (questionnaire) OR title:

(questionnaires) OR title: (screening) OR title: (screen) OR title: (scale) OR title: (instrument) OR title: (instruments) OR title: (interview) OR title: (interviews) OR title: (diagnosis) OR title: (diagnostic) OR title: (diagnosed) OR title: (Measure) OR title: (test) OR title: (tests) OR title: (testing) OR title: ("Attention Deficit Disorder with Hyperactivity/diagnostic imaging")) OR (abstract: ("Attention Deficit and Disruptive Behavior Disorders/diagnosis") OR abstract: (mass screening) OR abstract: (questionnaires) OR abstract: (Interviews as Topic) OR abstract: (Psychometrics) OR abstract: (Psychiatric Status Rating Scales) OR abstract: (diagnosis) OR abstract: ("Diagnostic Techniques and Procedures") OR abstract: ("Referral and Consultation") OR abstract: (questionnaire) OR abstract: (questionnaires) OR abstract: (screening) OR abstract: (screen) OR abstract: (scale) OR abstract: (instrument) OR abstract: (instruments) OR abstract: (interview) OR abstract: (interviews) OR abstract: (diagnosis) OR abstract: (diagnostic) OR abstract: (diagnosed) OR abstract: (Measure) OR abstract: (test) OR abstract: (tests) OR abstract: (testing) OR abstract: ("Attention Deficit Disorder with Hyperactivity/diagnostic imaging"))) AND ((title: ("Sensitivity and Specificity") OR title: ("Diagnostic Errors") OR title: (sensitivity) OR title: (specificity) OR (title: (accura*) AND (title: (diagnos*) OR title: (classif*))) OR title: ("ROC curve") OR title: ("positive predictive value") OR title: ("negative predictive value") OR title: ("false positive") OR title: ("false negative") OR title: ("likelihood ratio")) OR (abstract: ("Sensitivity and Specificity") OR abstract: ("Diagnostic Errors") OR abstract: (sensitivity) OR abstract: (specificity) OR (abstract: (accura*) AND (abstract: (diagnos*) OR abstract: (classif*))) OR abstract: ("ROC curve") OR abstract: ("positive predictive value") OR abstract: ("negative predictive value") OR abstract: ("false positive") OR abstract: ("false negative") OR abstract: ("likelihood ratio"))) AND Population Group: Human AND Publication Type: Peer Reviewed Journal)

Cochrane Database of Systematic Reviews (CDSR)
(#1 ("Adaptive Behavior Assessment System" OR "ABAS-3" OR "Advanced Clinical Solutions" OR "Word Choice Test" OR "Test of Premorbid Functioning" OR "Social Cognition" OR "Beck Anxiety Inventory" OR "BAI" OR "Beck Depression Inventory" OR "BDI-2" OR "Behavioral Assessment System for Children" OR "Self-Report of Personality" OR "BASC-3 SRP Adolescent" OR "Behavioral Assessment System for Children" OR "Parent Rating Scales" OR "BASC-3 PRS Adolescent" OR "BASC-3 SRP College" OR "Teacher Rating Scales" OR "BASC-3 TRS Adolescent" OR "Brown Executive Function/Attention Scales" OR "Brown EF/A Self" OR "California Verbal Learning Test" OR "CVLT-3" OR "Standard Form California Verbal" "CVLT-3 Brief" OR "California Verbal Learning Test" OR "CVLT-C" OR "Childhood Autism Rating Scale" OR "CARS-2" OR "Childhood Autism Rating Scale" OR "High-Functioning Version" OR "CARS-2 HF" OR "Clinical Evaluation of Language Fundamentals" OR "CELF-5" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Observer" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Self-Report" OR "Conners' Adult ADHD Diagnostic Interview for DSM-IV" OR "CAADID Part 1" OR "CAADID Part 2" OR "CAARS–O:L" OR "CAARS–S:L" OR "CAARS-2 Observer" OR "Conners' Adult ADHD Rating Scales" OR "CAARS-2 Self-Report" OR "Delis-Kaplan Executive Function System" OR "D-KEFS" OR "Dot Counting Test" OR "Grooved Pegboard Test Kaufman Test of Educational Achievement" OR "KTEA-3" OR "NEPSY-II Developmental Neuropsychological Battery" OR "Neuropsychological Assessment Battery" OR "Attention, Language, Memory, Spatial, and Ex- ecutive Functions Modules" OR "NIH Executive Abilities–Measures and Instruments for Neurobehavioral Evaluation and Re-search" OR "NIH

EXAMINER" OR "Personality Assessment Inventory" OR "PROMIS Sleep Assessments Pediatric Parent Proxy" OR "Repeatable Battery for the Assessment of Neuropsychological Status" OR "RBANS" OR "Rey-Osterrieth Complex" OR "Wechsler Abbreviated Scale of Intelligence" OR "WASI-2" OR "Wechsler Adult Intelligence Scale" OR "WAIS-4" OR "WAIS-IV" OR "Wechsler Individual Achievement Test" OR "WIAT-4" OR "Wechsler Intelligence Scale" OR "Wechsler Memory Scale" OR "WMS-4" OR "Wide Range Achievement Test" OR "WRAT-5" OR "Adult ADHD Rating Scale" OR "ADHD-RS" OR "Brown ADD scales" OR "Continuous Performance Tests" OR "Conners CPT" OR "QB Test" OR "TOVA" OR "Wender Utah Adult ADHD Scale" OR "diagnostic interview for Adult ADHD"):ti,ab,kw (Word variations have been searched)
#2 ("Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder"):ti,ab,kw (Word variations have been searched)
#3 #1 AND #2)
OR
(#1
MeSH descriptor: [Attention Deficit Disorder with Hyperactivity] explode all trees
#2
("attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder"):ti,ab,kw (Word variations have been searched)
#3
#1 OR #2
#4
MeSH descriptor: [Adult] explode all trees
#5
MeSH descriptor: [Aged] in all MeSH products
#6
MeSH descriptor: [Middle Aged] explode all trees
#7
(Young Adult OR Adult OR Adults):ti,ab,kw
(Word variations have been searched)
#8
#4 OR #5 OR #6 OR #7
#9
MeSH descriptor: [Mass Screening] explode all trees
#10
MeSH descriptor:[Surveys and Questionnaires] explode all trees
#11
MeSH descriptor: [Interviews as Topic] explode all trees
#12
MeSH descriptor: [Psychometrics] explode all trees
#13
MeSH descriptor: [Psychiatric Status Rating Scales] explode all trees
#14
MeSH descriptor: [Diagnosis] this term only
#15

MeSH descriptor: [Diagnostic Techniques and Procedures] explode all trees
#16
MeSH descriptor: [Referral and Consultation] explode all trees
#17
("Attention Deficit and Disruptive Behavior Disorders" AND diagnosis):ti,ab,kw
(Word variations have been searched)
#18
("Attention Deficit and Disruptive Behavior Disorders" AND "diagnostic imaging"):ti,ab,kw
(Word variations have been searched)
#19
(questionnaire OR questionnaires OR screening OR screen OR scale OR instrument OR instruments OR interview OR interviews OR diagnosis OR diagnostic OR diagnosed OR Measure OR test OR tests OR testing):ti,ab,kw
#20
#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19
#21
("Sensitivity and Specificity" OR "Diagnostic Errors" OR sensitivity OR specificity OR (accura* AND (diagnos* OR classif*)) OR "ROC curve" OR "positive predictive value" OR "negative predictive value" OR "false positive" OR "false negative" OR "likelihood ratio"):ti,ab,kw
#22
#3 AND #8 AND #20 AND #21)

Campbell Collaboration

("Adaptive Behavior Assessment System" OR "ABAS-3" OR "Advanced Clinical Solutions" OR "Word Choice Test" OR "Test of Premorbid Functioning" OR "Social Cognition" OR "Beck Anxiety Inventory" OR "BAI" OR "Beck Depression Inventory" OR "BDI-2" OR "Behavioral Assessment System for Children" OR "Self-Report of Personality" OR "BASC-3 SRP Adolescent" OR "Behavioral Assessment System for Children" OR "Parent Rating Scales" OR "BASC-3 PRS Adolescent" OR "BASC-3 SRP College" OR "Teacher Rating Scales" OR "BASC-3 TRS Adolescent" OR "Brown Executive Function/Attention Scales" OR "Brown EF/A Self" OR "California Verbal Learning Test" OR "CVLT-3" OR "Standard Form California Verbal" "CVLT-3 Brief" OR "California Verbal Learning Test" OR "CVLT-C" OR "Childhood Autism Rating Scale" OR "CARS-2" OR "Childhood Autism Rating Scale" OR "High-Functioning Version" OR "CARS-2 HF" OR "Clinical Evaluation of Language Fundamentals" OR "CELF-5" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Observer" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Self-Report" OR "Conners' Adult ADHD Diagnostic Interview for DSM-IV" OR "CAADID Part 1" OR "CAADID Part 2" OR "CAARS–O:L" OR "CAARS–S:L" OR "CAARS-2 Observer" OR "Conners' Adult ADHD Rating Scales" OR "CAARS-2 Self-Report" OR "Delis-Kaplan Executive Function System" OR "D-KEFS" OR "Dot Counting Test" OR "Grooved Pegboard Test Kaufman Test of Educational Achievement" OR "KTEA-3" OR "NEPSY-II Developmental Neuropsychological Battery" OR "Neuropsychological Assessment Battery" OR "Attention, Language, Memory, Spatial, and Executive Functions Modules" OR "NIH Executive Abilities–Measures and Instruments for Neurobehavioral Evaluation and Re-search" OR "NIH EXAMINER" OR "Personality Assessment Inventory" OR "PROMIS Sleep Assessments Pediatric Parent Proxy" OR "Repeatable Battery for the Assessment of Neuropsychological

Status" OR "RBANS" OR "Rey-Osterrieth Complex" OR "Wechsler Abbreviated Scale of Intelligence" OR "WASI-2" OR "Wechsler Adult Intelligence Scale" OR "WAIS-4" OR "WAIS-IV" OR "Wechsler Individual Achievement Test" OR "WIAT-4" OR "Wechsler Intelligence Scale" OR "Wechsler Memory Scale" OR "WMS-4" OR "Wide Range Achievement Test" OR "WRAT-5" OR "Adult ADHD Rating Scale" OR "ADHD-RS" OR "Brown ADD scales" OR "Continuous Performance Tests" OR "Conners CPT" OR "QB Test" OR "TOVA" OR "Wender Utah Adult ADHD Scale" OR "diagnostic interview for Adult ADHD")
OR
("Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder")

PROSPERO (https://www.crd.york.ac.uk/prospero/)
(#1 ("Adaptive Behavior Assessment System" OR "ABAS-3" OR "Advanced Clinical Solutions" OR "Word Choice Test" OR "Test of Premorbid Functioning" OR "Social Cognition" OR "Beck Anxiety Inventory" OR "BAI" OR "Beck Depression Inventory" OR "BDI-2" OR "Behavioral Assessment System for Children" OR "Self-Report of Personality" OR "BASC-3 SRP Adolescent" OR "Behavioral Assessment System for Children" OR "Parent Rating Scales" OR "BASC-3 PRS Adolescent" OR "BASC-3 SRP College" OR "Teacher Rating Scales" OR "BASC-3 TRS Adolescent" OR "Brown Executive Function/Attention Scales" OR "Brown EF/A Self" OR "California Verbal Learning Test" OR "CVLT-3" OR "Standard Form California Verbal" "CVLT-3 Brief" OR "California Verbal Learning Test" OR "CVLT-C" OR "Childhood Autism Rating Scale" OR "CARS-2" OR "Childhood Autism Rating Scale" OR "High-Functioning Version" OR "CARS-2 HF" OR "Clinical Evaluation of Language Fundamentals" OR "CELF-5" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Observer"):TI
#2 ("Comprehensive Executive Function Inventory" OR "CEFI Adult Self-Report" OR "Conners Adult ADHD Diagnostic Interview for DSM-IV" OR "CAADID Part 1" OR "CAADID Part 2" OR "CAARS?OL" OR "CAARS?SL" OR "CAARS-2 Observer" OR "Conners Adult ADHD Rating Scales" OR "CAARS-2 Self-Report" OR "Delis-Kaplan Executive Function System" OR "D-KEFS" OR "Dot Counting Test" OR "Grooved Pegboard Test Kaufman Test of Educational Achievement" OR "KTEA-3" OR "NEPSY-II Developmental Neuropsychological Battery" OR "Neuropsychological Assessment Battery" OR "Attention, Language, Memory, Spatial, and Ex- ecutive Functions Modules" OR "NIH Executive Abilities?Measures and Instruments for Neurobehavioral Evaluation and Re-search"):TI
#3 ("NIH EXAMINER" OR "Personality Assessment Inventory" OR "PROMIS Sleep Assessments Pediatric Parent Proxy" OR "Repeatable Battery for the Assessment of Neuropsychological Status" OR "RBANS" OR "Rey-Osterrieth Complex" OR "Wechsler Abbreviated Scale of Intelligence" OR "WASI-2" OR "Wechsler Adult Intelligence Scale" OR "WAIS-4" OR "WAIS-IV" OR "Wechsler Individual Achievement Test" OR "WIAT-4" OR "Wechsler Intelligence Scale" OR "Wechsler Memory Scale" OR "WMS-4" OR "Wide Range Achievement Test" OR "WRAT-5" OR "Adult ADHD Rating Scale" OR "ADHD-RS" OR "Brown ADD scales" OR "Continuous Performance Tests" OR "Conners CPT" OR "QB Test" OR "TOVA" OR "Wender Utah Adult ADHD Scale" OR "diagnostic interview for Adult ADHD"):TI
#4 #3 OR #2 OR #1
#5 (MeSH DESCRIPTOR Attention Deficit Disorder with Hyperactivity EXPLODE ALL

TREES):TI
#6 MeSH DESCRIPTOR Attention Deficit Disorder with Hyperactivity EXPLODE ALL TREES
#7 ("Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder"):TI
#8 #7 OR #6
#9 #8 AND #4)
OR
#1 MeSH DESCRIPTOR Attention Deficit Disorder with Hyperactivity EXPLODE ALL TREES
#2 "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder"
#3 #2 OR #1
#4 MeSH DESCRIPTOR Aged, 80 and over EXPLODE ALL TREES
#5 MeSH DESCRIPTOR Adult EXPLODE ALL TREES
#6 MeSH DESCRIPTOR Middle Aged EXPLODE ALL TREES
#7 Young Adult OR Adult OR Adults
#8 #4 OR #5 OR #6 OR #7
#9 MeSH DESCRIPTOR Mass Screening EXPLODE ALL TREES
#10 "interviews as topics"
#11 psychometrics
#12 MeSH DESCRIPTOR Psychiatric Status Rating Scales EXPLODE ALL TREES
#13 MeSH DESCRIPTOR Diagnosis EXPLODE ALL TREES
#14 MeSH DESCRIPTOR diagnosis EXPLODE ALL TREES
#15 MeSH DESCRIPTOR diagnosis
#16 MeSH DESCRIPTOR Diagnostic Techniques and Procedures EXPLODE ALL TREES
#17 MeSH DESCRIPTOR Referral and Consultation EXPLODE ALL TREES
#18 attention deficit and disruptive behavior disorders
#19 "attention deficit and disruptive behavior disorders" AND diagnosis
#20 "attention deficit and disruptive behavior disorders" AND "diagnostic imaging"
#21 questionnaire OR questionnaires OR screening OR screen OR scale OR instrument OR instruments OR interview OR interviews OR diagnosis OR diagnostic OR diagnosed OR Measure OR test OR tests OR testing
#22 #9 OR #10 OR #11 OR #12 OR #15 OR #16 OR #17 OR #19 OR #20 OR #21
#23 "Sensitivity and Specificity" OR "Diagnostic Errors" OR sensitivity OR specificity OR (accura* AND (diagnos* OR classif*)) OR "ROC curve" OR "positive predictive value" OR "negative predictive value" OR "false positive" OR "false negative" OR "likelihood ratio"
#24 #3 AND #8 AND #22 AND #23

ECRI Guidelines Trust https://guidelines.ecri.org/
("'Adaptive Behavior Assessment System" OR "ABAS-3" OR "Advanced Clinical Solutions" OR "Word Choice Test" OR "Test of Premorbid Functioning" OR "Social Cognition" OR "Beck Anxiety Inventory" OR "BAI" OR "Beck Depression Inventory" OR "BDI-2" OR "Behavioral Assessment System for Children" OR "Self-Report of Personality" OR "BASC-3 SRP Adolescent" OR "Behavioral Assessment System for Children" OR "Parent Rating Scales" OR

"BASC-3 PRS Adolescent" OR "BASC-3 SRP College" OR "Teacher Rating Scales" OR "BASC-3 TRS Adolescent" OR "Brown Executive Function/Attention Scales" OR "Brown EF/A Self" OR "California Verbal Learning Test" OR "CVLT-3" OR "Standard Form California Verbal" "CVLT-3 Brief" OR "California Verbal Learning Test" OR "CVLT-C" OR "Childhood Autism Rating Scale" OR "CARS-2" OR "Childhood Autism Rating Scale" OR "High-Functioning Version" OR "CARS-2 HF" OR "Clinical Evaluation of Language Fundamentals" OR "CELF-5" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Observer" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Self-Report" OR "Conners' Adult ADHD Diagnostic Interview for DSM-IV" OR "CAADID Part 1" OR "CAADID Part 2" OR "CAARS–O:L" OR "CAARS–S:L" OR "CAARS-2 Observer" OR "Conners' Adult ADHD Rating Scales" OR "CAARS-2 Self-Report" OR "Delis-Kaplan Executive Function System" OR "D-KEFS" OR "Dot Counting Test" OR "Grooved Pegboard Test Kaufman Test of Educational Achievement" OR "KTEA-3" OR "NEPSY-II Developmental Neuropsychological Battery" OR "Neuropsychological Assessment Battery" OR "Attention, Language, Memory, Spatial, and Executive Functions Modules" OR "NIH Executive Abilities–Measures and Instruments for Neurobehavioral Evaluation and Re-search" OR "NIH EXAMINER" OR "Personality Assessment Inventory" OR "PROMIS Sleep Assessments Pediatric Parent Proxy" OR "Repeatable Battery for the Assessment of Neuropsychological Status" OR "RBANS" OR "Rey-Osterrieth Complex" OR "Wechsler Abbreviated Scale of Intelligence" OR "WASI-2" OR "Wechsler Adult Intelligence Scale" OR "WAIS-4" OR "WAIS-IV" OR "Wechsler Individual Achievement Test" OR "WIAT-4" OR "Wechsler Intelligence Scale" OR "Wechsler Memory Scale" OR "WMS-4" OR "Wide Range Achievement Test" OR "WRAT-5" OR "Adult ADHD Rating Scale" OR "ADHD-RS" OR "Brown ADD scales" OR "Continuous Performance Tests" OR "Conners CPT" OR "QB Test" OR "TOVA" OR "Wender Utah Adult ADHD Scale" OR "diagnostic interview for Adult ADHD'")
OR
("Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder"
FILTER: Patient Age
Adolescent (13 to 18 years), Adult (19 to 44 years), Middle Age(45 to 64 years), Aged(65 to 79 years), Aged (80 and over)

Guidelines International Network Library (G-I-N, https://guidelines.ebmportal.com/)
("Adaptive Behavior Assessment System" OR "ABAS-3" OR "Advanced Clinical Solutions" OR "Word Choice Test" OR "Test of Premorbid Functioning" OR "Social Cognition" OR "Beck Anxiety Inventory" OR "BAI" OR "Beck Depression Inventory" OR "BDI-2" OR "Behavioral Assessment System for Children" OR "Self-Report of Personality" OR "BASC-3 SRP Adolescent" OR "Behavioral Assessment System for Children" OR "Parent Rating Scales" OR "BASC-3 PRS Adolescent" OR "BASC-3 SRP College" OR "Teacher Rating Scales" OR "BASC-3 TRS Adolescent" OR "Brown Executive Function/Attention Scales" OR "Brown EF/A Self" OR "California Verbal Learning Test" OR "CVLT-3" OR "Standard Form California Verbal" "CVLT-3 Brief" OR "California Verbal Learning Test" OR "CVLT-C" OR "Childhood Autism Rating Scale" OR "CARS-2" OR "Childhood Autism Rating Scale" OR "High-Functioning Version" OR "CARS-2 HF" OR "Clinical Evaluation of Language Fundamentals" OR "CELF-5" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Observer" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Self-Report" OR

"Conners' Adult ADHD Diagnostic Interview for DSM-IV" OR "CAADID Part 1" OR "CAADID Part 2" OR "CAARS–O:L" OR "CAARS–S:L" OR "CAARS-2 Observer" OR "Conners' Adult ADHD Rating Scales" OR "CAARS-2 Self-Report" OR "Delis-Kaplan Executive Function System" OR "D-KEFS" OR "Dot Counting Test" OR "Grooved Pegboard Test Kaufman Test of Educational Achievement" OR "KTEA-3" OR "NEPSY-II Developmental Neuropsychological Battery" OR "Neuropsychological Assessment Battery" OR "Attention, Language, Memory, Spatial, and Ex- ecutive Functions Modules" OR "NIH Executive Abilities– Measures and Instruments for Neurobehavioral Evaluation and Re-search" OR "NIH EXAMINER" OR "Personality Assessment Inventory" OR "PROMIS Sleep Assessments Pediatric Parent Proxy" OR "Repeatable Battery for the Assessment of Neuropsychological Status" OR "RBANS" OR "Rey-Osterrieth Complex" OR "Wechsler Abbreviated Scale of Intelligence" OR "WASI-2" OR "Wechsler Adult Intelligence Scale" OR "WAIS-4" OR "WAIS-IV" OR "Wechsler Individual Achievement Test" OR "WIAT-4" OR "Wechsler Intelligence Scale" OR "Wechsler Memory Scale" OR "WMS-4" OR "Wide Range Achievement Test" OR "WRAT-5" OR "Adult ADHD Rating Scale" OR "ADHD-RS" OR "Brown ADD scales" OR "Continuous Performance Tests" OR "Conners CPT" OR "QB Test" OR "TOVA" OR "Wender Utah Adult ADHD Scale" OR "diagnostic interview for Adult ADHD")
OR
("Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder")

ClinicalKey

("Adaptive Behavior Assessment System" OR "ABAS-3" OR "Advanced Clinical Solutions" OR "Word Choice Test" OR "Test of Premorbid Functioning" OR "Social Cognition" OR "Beck Anxiety Inventory" OR "BAI" OR "Beck Depression Inventory" OR "BDI-2" OR "Behavioral Assessment System for Children" OR "Self-Report of Personality" OR "BASC-3 SRP Adolescent" OR "Behavioral Assessment System for Children" OR "Parent Rating Scales" OR "BASC-3 PRS Adolescent" OR "BASC-3 SRP College" OR "Teacher Rating Scales" OR "BASC-3 TRS Adolescent" OR "Brown Executive Function/Attention Scales" OR "Brown EF/A Self" OR "California Verbal Learning Test" OR "CVLT-3" OR "Standard Form California Verbal" "CVLT-3 Brief" OR "California Verbal Learning Test" OR "CVLT-C" OR "Childhood Autism Rating Scale" OR "CARS-2" OR "Childhood Autism Rating Scale" OR "High-Functioning Version" OR "CARS-2 HF" OR "Clinical Evaluation of Language Fundamentals" OR "CELF-5" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Observer" OR "Comprehensive Executive Function Inventory" OR "CEFI Adult Self-Report" OR "Conners' Adult ADHD Diagnostic Interview for DSM-IV" OR "CAADID Part 1" OR "CAADID Part 2" OR "CAARS–O:L" OR "CAARS–S:L" OR "CAARS-2 Observer" OR "Conners' Adult ADHD Rating Scales" OR "CAARS-2 Self-Report" OR "Delis-Kaplan Executive Function System" OR "D-KEFS" OR "Dot Counting Test" OR "Grooved Pegboard Test Kaufman Test of Educational Achievement" OR "KTEA-3" OR "NEPSY-II Developmental Neuropsychological Battery" OR "Neuropsychological Assessment Battery" OR "Attention, Language, Memory, Spatial, and Executive Functions Modules" OR "NIH Executive Abilities– Measures and Instruments for Neurobehavioral Evaluation and Re-search" OR "NIH EXAMINER" OR "Personality Assessment Inventory" OR "PROMIS Sleep Assessments Pediatric Parent Proxy" OR "Repeatable Battery for the Assessment of Neuropsychological Status" OR "RBANS" OR "Rey-Osterrieth Complex" OR "Wechsler Abbreviated Scale of

Intelligence" OR "WASI-2" OR "Wechsler Adult Intelligence Scale" OR "WAIS-4" OR "WAIS-IV" OR "Wechsler Individual Achievement Test" OR "WIAT-4" OR "Wechsler Intelligence Scale" OR "Wechsler Memory Scale" OR "WMS-4" OR "Wide Range Achievement Test" OR "WRAT-5" OR "Adult ADHD Rating Scale" OR "ADHD-RS" OR "Brown ADD scales" OR "Continuous Performance Tests" OR "Conners CPT" OR "QB Test" OR "TOVA" OR "Wender Utah Adult ADHD Scale" OR "diagnostic interview for Adult ADHD")
OR
("Attention Deficit Disorder with Hyperactivity" OR "attention deficit hyperactivity disorder" OR "ADHD" OR "attention deficit disorder"
FILTERS: Journal Articles, Guidelines)

Appendix B: List of Excluded Studies

Appendix C: Evidence Tables

Appendix Table A1: Evidence Table Combinations as Index Test

Study ID	Population	Combination Index Test	Results	Subgroup
Chen, 2021{#5100} N = 69 UK Specialty care	**Target:** ADHD patients in the period between 2014 and 2017 with demographics and a number of validated self-reported screening questionnaires and clinical interviews **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults in the NHS Trust who did not meet DSM-IV diagnostic criteria for ADHD and were assessed as part of standard mental health services **Female:** 34.8% **Age:** 33.01 (9.931) Min age: 18 Max age: 51 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** No COI	**Test description:** Combination of demographics, self-reported assessment, Conner's Adult ADHD Rating Scale (short version) with self and observer model, QbTest, and DIVA (Diagnostic Interview for AHDH in adults), 93 variables, decision tree analysis Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on structured clinical interviews using the Diagnostic Interview for ADHD in Adults (DIVA) and validated self-reported screening questionnaires collected from a National Health Service specialist mental health provider **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The highest achieved accuracy was 85.5%. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy 86 AUC 0.871 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Dvorsky, 2016{#10626} N = 86 US College	**Target:** Undergraduate students at a large public university who self-identified as having attention or concentration difficulties or a prior ADHD diagnosis, required consent for parental interviews, completed a comprehensive ADHD evaluation including structured diagnostic interviews, and met DSM-5 ADHD criteria based on both student and parent ratings **ADHD presentation:** inattentive : 55.9,combined : 44.1	**Test description:** Combination prediction model with BAARS parent and self rating of current and childhood ADHD diagnosis Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners' Adult ADHD	**Diagnostic accuracy summary:** Parent ratings of childhood inattention had the highest predictive validity (AUC 0.79), outperforming self-report (AUC 0.56). Self-reports had high sensitivity (89%) but low specificity (30%), leading to a high false-positive rate. The prediction model with both parent and student ratings of current symptoms and parent ratings of childhood symptoms accurately classified 88.9% of individuals who had a	**Subgroup analysis:** N/A

	Comorbidity: N/A **Other:** Undergraduate students at the same university who self-identified with attention or concentration difficulties but did not meet DSM-5 criteria for ADHD based on structured diagnostic interviews and parent ratings **Female:** 42.4% **Age:** 19.71 (2.72) Min age: 18 Max age: 27 **Age subgroup**: Adults **Ethnicity:** Other : ADHD: 0, non-ADHD: 3.7 % Hispanic or Latino : 8.5,Other : 11.1 % Black/African American : 6.8,Other : non-ADHD: 18.5 % White : 76.3,Other : non-ADHD: 55.6 % Multiracial : 8.5,Other : 11.1 Single center **Funding:** Unclear	Diagnostic Interview for DSM-IV, which included structured diagnostic interviews separately administered to students and their parents by trained graduate-level clinicians under supervision, requiring endorsement of at least five current symptoms in two or more settings and six childhood symptoms before high school **Diagnosed by:** Specialist (e.g., mental health) Psychologists **Timing:** Concurrent	diagnosis of ADHD and 63.3% of individuals who did not have a diagnosis. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Groom, 2016{#325} N = 57 UK College	**Target:** Adults who were clinically diagnosed with ADHD by a psychiatrist **ADHD presentation:** inattentive : 9.09,hyperactive : 3.03,combined : 75.76,N/A : 12.12 **Comorbidity:** N/A **Other:** Adults diagnosed with Asperger's syndrome as part of autism spectrum disorder by a psychiatrist **Female:** 39% **Age:** 31.64 (10.17) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** Integration of CAARS-E (Conners Adult ADHD Rating Scale - ADHD Index), the AQ10 (Autism Quotient - 10), and the QbTest (computerized Continuous Performance Test with motion tracking) to differentiate ADHD from Autism Spectrum Disorder Machine learning: No Validation dataset: Unclear **Reference standard:** Clinical diagnosis Participants diagnosed with ADHD by a psychiatrist establishing current and long-term diagnosis using DSM-5	**Diagnostic accuracy summary:** QbTotal yielded the highest AUC value 0.87 (classified as 'good'). ROCs indicate that at equivalent sensitivity of around 80%, QbTotal demonstrates superior specificity compared with CAARS-E in differentiating ADHD and autism spectrum disorder. CAARS-E AUC was .77 ('fair') in differentating ADHD and autism spectrum disorder. QbTest added to clinical ratings may improve the differentiation of ADHD and autism spectrum disorder in adults. Sensitivity 94% Specificity 84% PPV NPV LR+ LR- Accuracy 90 AUC	**Subgroup analysis:** N/A

		Diagnosed by: Specialist (e.g., mental health) Psychiatrist **Timing:** Prior diagnosis	**Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Kingston, 2013{#5349} N = 120 Canada Specialty care	**Target:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **ADHD presentation:** N/A **Comorbidity:** Other : Aggression dysregulation **Other:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **Female:** 0% **Age:** 32.6 (10.3) Min age: 18 Max age: 64 **Age subgroup**: Adults **Ethnicity:** Other : Aboriginal: 6.5% % Hispanic or Latino : 2.8 % Black/African American : 2.8 % White : 78.5 Single center **Funding:** Industry	**Test description:** Integration of ASRS-v1, CAARS-Self and CAARS-Observer, Brown ADD scale, and WURS in a discriminant function Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis ADHD diagnosis was determined based on DSM-IV-TR criteria following a comprehensive clinical interview and review of relevant available collateral information; interviews were conducted independently by two psychiatrists who were certified in forensic psychiatric practice; final group classification was based on consensus diagnoses and the inter-rater agreement was approximately 90% **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The integrated variables of multiple self reports and an observer report demonstrated particularly good classification accuracy, with high sensitivity (91%) and good specificity (82%). Sensitivity 91% Specificity 82% PPV NPV LR+ LR- Accuracy 86 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Nikolas, 2019{#10710} N = 246 US Specialty care	**Target:** Adults diagnosed with ADHD based on a comprehensive clinical interview and standardized psychiatric assessments, required to have symptom onset before age 16, met full DSM-5 diagnostic criteria, provided informant reports verifying symptoms, excluded if they had neurological conditions, learning disabilities, major psychiatric disorders other than depression/anxiety, or substance abuse **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with a diagnosed unipolar mood disorder (depression) and healthy controls without ADHD or mood disorders, recruited through advertisements, email listservs, and outreach to neuropsychological clinics, with controls matched approximately by age and sex to clinical groups **Female:** 60.6% **Age:** 24.8 (6.2) Min age: 18 Max age: 40 **Age subgroup**: Adults **Ethnicity:** % White : 83.7,Other : Control: 80, depressed: 86.5 Multicenter **Funding:** Industry	**Test description:** Combination of self/informant symptom ratings (BAARS-IV), family history, and reactiontime variability from TOVA (Test of Variables of Attention) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive clinical interview, standardized psychiatric assessment, and meeting full DSM-5 diagnostic criteria with verification of symptom onset before age 16 using self-report and informant ratings **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** While single test measures provided performed poorly in identifying ADHD participants, analyses revealed that a combined approach using self and informant symptom ratings, a positive family history of ADHD, and a reaction time variability measure correctly classified 87% of cases. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy 87.2 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Pettersson, 2018{#711} N = 108 Sweden Specialty care	**Target:** Adults referred for ADHD assessment, required availability of a collateral historian to provide information on childhood symptoms, excluded if treated with ADHD medications, had an IQ ≤ 70, or substance-related disorders **ADHD presentation:** inattentive : 21.7,hyperactive : 7.1,combined : 76.7 **Comorbidity:** N/A **Other:** Adults referred to the same specialty neuropsychological clinic for	**Test description:** Model with DIVA report, QbTest cardinal variable Acticity, QbTest cardinal variable Inattention, and CpT II Commission errors, combining neuropsychological tests, DIVA clinician report, and self report ASRS Screener Machine learning: No Validation dataset: No	**Diagnostic accuracy summary:** All instruments showed poor discriminative ability except for the DIVA, which showed a relatively good ability to discriminate between the groups (sensitivity 90.0; specificity 72.9). A logistic regression analysis model with the DIVA and measures of inattention, impulsivity, and activity from continuous performance tests (CPTs) showed a sensitivity of 90.0 and a specificity of 83.3. Sensitivity 90% Specificity 83%	**Subgroup analysis:** N/A

	assessment, did not meet the diagnostic criteria for ADHD, included individuals with other psychiatric conditions for comparison **Female:** 46.7% **Age:** 28.18 (9.09) Min age: 18 Max age: 55 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Reference standard:** Clinical diagnosis Clinical consensus decision by a multidisciplinary assessment team using clinical interviews, neuropsychological test results, self-report measures, collateral historian input, and DSM criteria **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	PPV NPV LR+ LR- Accuracy 87 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** **Admin time:** N/A	
Robeva, 2004{#10732} N = 12 US College	**Target:** Female college students with a current ADHD diagnosis, taking ADHD medication for at least three years, not on anxiety or depression medication, without significant health conditions affecting EEG recordings, diagnosed in childhood according to Utah standards **ADHD presentation:** combined : 100 **Comorbidity:** N/A **Other:** Female college students with no history of ADHD or disruptive behavioral disorders, never prescribed or taken stimulant medication, not on anxiety or depression medication, without significant medical conditions affecting EEG data collection, screened to confirm the absence of ADHD symptoms **Female:** 100% **Age:** 20.7 (1.5) Min age: 18 Max age: 22 **Age subgroup**: Young **Ethnicity:** N/A	**Test description:** Bayesian probability model integrated three diagnostic tools (WURS, ConsistencyIndex (EEG), Alpha Blockade Index (EEG) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a prior clinical diagnosis made during childhood following Utah criteria, confirmed through self-report screening using the Brown Attention-Deficit Disorder Scale and the ADHD Symptom Inventory, with additional verification that participants were currently prescribed and taking stimulant medication for ADHD management	**Diagnostic accuracy summary:** The procedure significantly improved the score separation between ADHD and non-ADHD groups. The final average probabilities for ADHD were 76% for the ADHD group and 8% for the control group. These probabilities correlated (r=.87) with the Brown ADD scale and (r=.84) with the ADHD-Symptom Inventory used for screening the participants. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy average probabilities for ADHD: 76%, controls: 8% (F= 12.005, p = 0.006) AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha	**Subgroup analysis:** N/A

	Single center **Funding:** Other	**Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Van Voorhees, 2011{#949} N = 349 US Specialty care	**Target:** Adults seeking evaluation for attention difficulties at an ADHD clinic diagnosed with DSM-IV **ADHD presentation:** inattentive : 8.9,combined : 33.1 **Comorbidity:** N/A **Other:** Adults seeking evaluation for attention difficulties at an ADHD clinic not diagnosed with ADHD **Female:** 38.5% **Age:** mean 32, median: 28 Min age: 18 Max age: 70 **Age subgroup**: Adults **Ethnicity:** Other : Race data were only available for 77.8% of the sample % Hispanic or Latino : 1.8 % Asian : 2.9 % White : 86.4 % Multiracial : 3.7 Single center **Funding:** Other	**Test description:** CAARS-LV combining self-report CAARS:S and observer-report CAARS-O Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on CAARS, CAADID, and structured clinical interview for DSM-IV (SCID by a doctoral-level clinician **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Later diagnosis	**Diagnostic accuracy summary:** Self- and observer-ratings on the CAARS provide clinically relevant data about attention problems in adults, but the instrument does not effectively distinguish between ADHD and other adult psychiatric disorders. Combining self- and observer-ratings decreased the scales' sensitivity. Sensitivity 43% Specificity 83% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** CAARS-S and CAARS-O (ratings from friends, parents, and spouses) Kappa ICC 0.11-0.37 **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Appendix Table A2: Evidence Table Self Report as Index Test

Study ID	Population	Self Report Index Test	Results	Subgroup
Aita, 2018{#16} N = 280 US Specialty care	**Target:** Individuals from one of two university-affiliated psychology training clinics, diagnosed with ADHD **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Mood/Anxiety Disorder group or Clinic No Diagnosis group: Individuals from one of two university-affiliated psychology training clinics, not diagnosed with ADHD Control group or ADHD Simulator group: Students were prospectively recruited from three southeastern universities **Female:** 45.1% Study 1 - ADHD group: 45.1%; ADHD Simulators group: 73.9%; Mood/Anxiety Disorder group: 65.0%; Clinic No Diagnosis group: 42.3%; Healthy Controls group: 69.2%; Study 2 - ADHD group: 43.8%; ADHD Simulators group: 73.9%; Mood/Anxiety Disorder group: 65.4%; Clinic No Diagnosis group: 37.8%; Healthy Controls group: 75.5% **Age:** 20.29 (1.87) Study 1 - ADHD group: 21.77 (3.99); ADHD Simulators group: 19.83 (1.54); Mood/Anxiety Disorder group: 22.71 (4.58); Clinic No Diagnosis group: 22.05 (5.07); Healthy Controls group: 19.18 (1.57) Study 2 - ADHD group: 22.33 (3.93); ADHD Simulators group: 19.83 (1.54); Mood/Anxiety Disorder group: 21.98 (4.26); Clinic No Diagnosis	**Test description:** PAI (Personality Assessment Inventory), a self-report personality measure comprised of 344 items on a 4-point scale with anchor points of false and very true; items are categorized into 4 scales that assess validity of responding, 11 clinical syndrome scales, 5 treatment scales, and 2 interpersonal scales Machine learning: No Validation dataset: Yes **Reference standard:** Clinical diagnosis All evaluations were conducted by doctoral graduate students in a clinical psychology program. Evaluations included a thorough clinical interview and all diagnoses were made under the supervision of a licensed psychologist. **Diagnosed by:** Researcher Doctoral graduate students in a clinical psychology program, under supervision of a licensed psychologist **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The new index's classification accuracy was superior to most existing PAI validity scales across groups. An item-level PAI algorithm had a sensitivity of 85% and specificity of 97% for identifying feigned ADHD. Sensitivity 92% Specificity % % PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	group: 22.80 (5.13); Healthy Controls group: 19.45 (1.35) Min age: 18 Max age: 25 **Age subgroup**: Adults **Ethnicity:** Other : Other Race: Study 1 - ADHD Simulators group: 4.3%; Clinic No Diagnosis group: 0.9%; Healthy Controls group: 3.8%; Study 2 - ADHD Simulators group: 4.3%; Clinic No Diagnosis group: 2.2%; Healthy Controls group: 3.8% Other : Study 1 - ADHD group: 5.8%; ADHD Simulators group: 7.2%; Mood/Anxiety Disorder group: 1.5%; Clinic No Diagnosis group: 1.8%; Healthy Controls group: 3.8%; Study 2 - ADHD group: 9.6%; ADHD Simulators group: 7.2%; Healthy Controls group: 5.7% Other : Study 1 - ADHD group: 10.1%; ADHD Simulators group: 10.1%; Mood/Anxiety Disorder group: 8.8%; Clinic No Diagnosis group: 10.8%; Healthy Controls group: 24.1%; Study 2 - ADHD group: 9.6%; ADHD Simulators group: 10.1%; Mood/Anxiety Disorder group: 5.8%; Clinic No Diagnosis group: 8.9%; Healthy Controls group: 9.4% Other : Study 1 - ADHD group: 1.4%; ADHD Simulators group: 5.8%; Mood/Anxiety Disorder group: 2.2%; Clinic No Diagnosis group: 1.8%; Healthy Controls group: 3.8%; Study 2 - ADHD group: 2.7%; ADHD Simulators: 5.8%; Mood/Anxiety Disorder group: 1.9%; Clinic No Diagnosis group: 4.4%; Healthy Controls group: 1.9% Other : Study 1 - ADHD group: 82.7%; ADHD Simulators group: 72.5%; Mood/Anxiety Disorder			

Study ID	Population	Self Report Index Test	Results	Subgroup
	group: 87.6%; Clinic No Diagnosis group: 84.7%; Healthy Controls group: 64.7%; Study 2 - ADHD group: 75.3%; ADHD Simulators: 72.5%; Mood/Anxiety Disorder group: 92.3%; Clinic No Diagnosis group: 84.4%; Healthy Controls: 79.2% Multicenter **Funding:** Other			
Bakare, 2020{#45} N = 69 UK Specialty care	**Target:** Participants aged between 30 and 63 were recruited from a series of patients referred to adult ADHD outpatient clinics, adults with moderate or severe learning disabilities, organic brain injury or poor command of English were excluded **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** There were 8 participants who were not diagnosed with ADHD and therefore were the healthy control group **Female:** 38.3% **Age:** 45 (6.95) Min age: 30 Max age: 63 **Age subgroup**: Middle age **Ethnicity:** N/A Multicenter **Funding:** No COI	**Test description:** WURS-brief (Wender Utah Rating Sclae); admistered together with the CAADID Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis ICD-10 diagnosis **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The WURS-brief had respectable sensitivity when compared with existing diagnostic tools Sensitivity 89% Specificity 11% % PPV 67 NPV 33 LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Agreement WURS-brief and DIVA rating Kappa 0.006 ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Bastiaens, 2017{#55} N = 140 US	**Target:** Adults with a diagnosis of ADHD and substance use diagnosis **ADHD presentation:** N/A **Comorbidity:** N/A	**Test description:** ASRS-5 (WHO Adult ADHD Self Report Scale Screener for DSM-5), dimensional scoring with 12 as threshold	**Diagnostic accuracy summary:** Both screeners performed equally with no significant difference between them, regardless of the scoring system used. The dimensional scoring method with a cutoff of 12/24 provided the best diagnostic	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
Specialty care	**Other:** Adults without a diagnosis of ADHD but a substance use diagnosis **Female:** 35.71% **Age:** 33.5 (8.3) Min age: 24 Max age: 44 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** No COI	Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a clinical psychiatric interview using DSM-5 criteria conducted by a child and adolescent psychiatrist **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	accuracy, achieving sensitivity and negative predictive value above 80%. Sensitivity 81% Specificity 71% % PPV 74 NPV 79 LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Chiasson, 2012{#141} N = 183 Canada Specialty care	**Target:** A retrospective study conducted on the profiles of all newly admitted SUD patients in a multidisciplinary rehabilitation center with the ASRS-v1.1 and were later assessed by a psychiatrist specialized in ADHD **ADHD presentation:** N/A **Comorbidity:** SUD : all with SUD **Other:** Family members of suspected ADHD patients were also interviewed to acquire collateral information on patient behavior patterns **Female:** % N/A **Age:** N/A Min age: Max age:	**Test description:** ASRS-v1.1 ADHD Self-Report Scale Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a psychiatric evaluation by a specialist using DSM-IV criteria, including collateral information from family members and consensus discussion with the clinical team. **Diagnosed by:** Specialist (e.g., mental health)	**Diagnostic accuracy summary:** The ASRS-v1.1 correctly identified all ADHD cases (100% sensitivity) but had a low specificity, leading to a high false positive rate, with only 26% of those screening positive diagnosed with ADHD by a psychiatrist. Sensitivity 100% Specificity % % PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:**	**Subgroup analysis:** Unclear

Study ID	Population	Self Report Index Test	Results	Subgroup
	Age subgroup: Age unclear **Ethnicity:** N/A Single center **Funding:** Other	**Timing:** Concurrent	Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Dakwar, 2012{#172} N = 102 US College	**Target:** Adults seeking outpatient treatment for cocaine dependence, recruited through advertisements, diagnosed with ADHD based on the Conners Adult ADHD Diagnostic Interview for DSM-IV (CAADID), aged 18 years or older, with no exclusion criteria based on comorbid psychiatric conditions or other substance use disorders **ADHD presentation:** inattentive : 2.9,hyperactive : 2,combined : 9.8 **Comorbidity:** SUD : The study population consisted entirely of adults seeking outpatient treatment for cocaine dependence **Other:** Adults seeking outpatient treatment for cocaine dependence, without a diagnosis of ADHD, recruited from the same specialty care setting, with no exclusion based on comorbid psychiatric conditions or other substance use disorders, serving as a comparison group to differentiate ADHD diagnosis **Female:** 17% **Age:** Min age: 18 Max age: 57 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 39 % Black/African American : 33.7	**Test description:** WURS or CAARS or ASRS-VI.I: WURS (Wender Utah Rating Scale), evaluates childhood ADHD symptoms based on Utah criteria, and the CAARS (Conners Adult ADHD Rating Scale), a DSM-IV-based tool assessing ADHD symptoms in adults; ASRS-V1.1 (Adult ADHD Self-Report Scale-Version 1.1), a brief 6-item tool developed by the WHO designed for quick ADHD screening in adults Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners Adult ADHD Diagnostic Interview for DSM-IV, a validated semistructured interview conducted by trained clinicians to assess symptoms, age of onset, pervasiveness, and impairment. **Diagnosed by:** Specialist (e.g., mental health)	**Diagnostic accuracy summary:** The CAARS emerged with the highest κ scores and positive predictive value, but the WURS outperformed the other instruments in regard to sensitivity (87.5%). The most sensitive conjunctions arose (96.0%) when all instruments were administered together, with a suggestive score on any single scale indicating the diagnosis, and when the WURS was administered alongside the CAARS. Sensitivity 96% WURS: 88; CAARS: 80; ASRS-V1.1: 61 Specificity 65% WURS: 75; CAARS: 91; ASRS-V1.1: 86% PPV 47.06 WURS: 52.5; CAARS: 74.07; ASRS-V1.1: 58.33 NPV 98.04 WURS: 95.08; CAARS: 93.06; ASRS-V1.1: 86.76 LR+ WURS: 3.55; CAARS: 8.46; ASRS-V1.1: 4.20 LR- 0.17 CAARS: 0.22; ASRS-V1.1: 0.46 Accuracy 68.99 CAARS: 74.32; ASRS-V1.1: 57.72 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	% White : 27.4 Single center **Funding:** Public funding	clinicians with either a PhD or MA in clinical psychology **Timing:** Concurrent	**Cost:** N/A **Admin time:** N/A	
De Quiros, 2001{#10619) N = 48 Multiple countries Specialty care	**Target:** Adults from a specialty clinic, met DSM-IV criteria for ADHD with at least 6 of 9 inattentive and/or hyperactive/impulsive symptoms, had retrospectively met full DSM-IV criteria for ADHD in childhood, had no other psychiatric disorder that could explain ADHD-like symptoms, had never been diagnosed with ADHD or received stimulant therapy, and were not taking psychoactive medications **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Age-matched adults without ADHD symptoms, recruited as controls, had no history of attention or behavior problems, were not taking psychoactive medications, and were evaluated in the same specialty care setting **Female:** 47.9% **Age:** 34 (11) Min age: 23 Max age: 45 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** APQ (Adult Problem Questionnaire), a 43-item self-rating scale that assesses distractibility, impulsivity, and behavioral control using a cutoff score of 2.5 on three key items; admistered together with the CHI (Conners Hyperactivity Index) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria by a behavioral neurologist through clinical interviews with the patient and, when feasible, their spouse or parents, with retrospective confirmation of childhood ADHD symptoms. **Diagnosed by:** Specialist (e.g., mental health) Behavioral neurologist **Timing:** Concurrent	**Diagnostic accuracy summary:** Discriminant analysis revealed the APQ correctly classified 83% of ADHD and 90% of controls correctly. Sensitivity 83% CHI: 81 Specificity 90% CHI: 90% PPV 91 CHI: 91 NPV 82 CHI: 80 LR+ 8.33 CHI: 8.13 LR- 0.19 CHI: 0.21 Accuracy 86 CHI: 85 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Dunlop, 2018{#212} N = 95 US Specialty care	**Target:** Adults diagnosed with Major Depressive Disorder based on DSM-IV criteria, scoring ≥15 on the Hamilton Depression Rating Scale, off psychiatric medications (except sedative/hypnotic) for one month prior, with ADHD diagnosis	**Test description:** ASRS-v1.1 (Adult ADHD Self-Report Scale v1.1) a self-administered questionnaire based on the 18 DSM-IV ADHD symptom criteria; Part A contains six questions used as the	**Diagnostic accuracy summary:** The ASRS-v1.1 demonstrated fair performance in identifying full-syndrome DSM-IV ADHD in adults with MDD, with sensitivity of 60%, specificity of 69%, PPV of 21.4%, NPV of 92.3%, and total classification accuracy of 67.5%. Sensitivity 60% (CI 14.7, 94.7) Specificity 68.6% (CI 50.7, 83.2) %	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) A high false-positive rate was observed in participants with Major Depressive Disorder, attributed to overlapping

Study ID	Population	Self Report Index Test	Results	Subgroup
	confirmed by structured interview and psychiatrist assessment **ADHD presentation:** inattentive : 40,combined : 60 **Comorbidity:** Depression : Major Depressive Disorder (MDD) as the primary diagnosis **Other:** The healthy control group consisted of adults without a DSM-IV mental illness diagnosis in the past year, no history of MDD or dysthymia, no psychotropic medication use, scoring ≤7 on the Hamilton Depression Rating Scale, recruited through the same specialty psychiatric care setting **Female:** 72.5% healthy control: 70.9 **Age:** 49.5 (8.1) Healthy control: 44.0 (11.5) Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** % Black/African American : 32.5 % White : 47.5 % Multiracial : 20 Single center **Funding:** Public funding	primary screening tool, with a threshold score of ≥4 indicating a positive ADHD screen; Part B contains 12 additional questions providing further insight into symptom severity but not used for diagnostic purposes Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the ADHD module of the Mini International Neuropsychiatric Interview (MINI), requiring symptoms to meet DSM-IV criteria, including onset before age 7 and functional impairment, confirmed through a structured clinician interview and assessment by a study psychiatrist **Diagnosed by:** Specialist (e.g., mental health) psychiatrists **Timing:** Concurrent	PPV 21.4 (CI 10.3, 39.4) NPV 92.3 (CI 80, 97.3) LR+ 1.91 LR- 0.58 Accuracy 67.5 (CI 50.9, 81.4) AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	symptoms such as inattention and anxiety, which are common in both conditions. Participants with a positive ADHD screen demonstrated greater functional impairments and higher levels of anxiety compared to those without ADHD, suggesting that functional impairments are more pronounced in those with comorbid conditions. The high false-positive rate was attributed to symptom overlap, particularly with inattention and anxiety in depressive disorders. MDD participants with ADHD symptomatology reported significantly higher levels of anxiety (Hamilton Anxiety Rating Scale) and rumination compared to those without ADHD, demonstrating the impact of comorbid conditions on symptom scores.
Dvorsky, 2016{#10626} N = 86 US College	**Target:** Undergraduate students at a large public university who self-identified as having attention or concentration difficulties or a prior ADHD diagnosis, required consent for parental interviews, completed a comprehensive ADHD evaluation including structured diagnostic interviews, and met DSM-5 ADHD	**Test description:** BAARS-IV (Barkley Adult ADHD Rating Scale-IV) for self-reported assessment of ADHD symptoms on a 4-point scale (0 = never or rarely to 3 = very often), cut off > 3 symptoms presence Machine learning: No Validation dataset: No	**Diagnostic accuracy summary:** Parent ratings of childhood inattention had the highest predictive validity (AUC 0.79), outperforming self-report (AUC 0.56). Self-reports had high sensitivity (89%) but low specificity (30%), leading to a high false-positive rate. The prediction model with both parent and student ratings of current symptoms and parent ratings of childhood symptoms accurately classified 88.9%	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	criteria based on both student and parent ratings **ADHD presentation:** inattentive : 55.9,combined : 44.1 **Comorbidity:** N/A **Other:** Undergraduate students at the same university who self-identified with attention or concentration difficulties but did not meet DSM-5 criteria for ADHD based on structured diagnostic interviews and parent ratings **Female:** 42.4% **Age:** 19.71 (2.72) Min age: 18 Max age: 27 **Age subgroup**: Adults **Ethnicity:** Other : ADHD: 0, non-ADHD: 3.7 % Hispanic or Latino : 8.5,Other : 11.1 % Black/African American : 6.8,Other : non-ADHD: 18.5 % White : 76.3,Other : non-ADHD: 55.6 % Multiracial : 8.5,Other : 11.1 Single center **Funding:** Unclear	**Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners' Adult ADHD Diagnostic Interview for DSM-IV, which included structured diagnostic interviews separately administered to students and their parents by trained graduate-level clinicians under supervision, requiring endorsement of at least five current symptoms in two or more settings and six childhood symptoms before high school **Diagnosed by:** Specialist (e.g., mental health) Psychologists **Timing:** Concurrent	of individuals who had a diagnosis of ADHD and 63.3% of individuals who did not have a diagnosis. Sensitivity 89% Specificity 30% % PPV 68 NPV 60 LR+ LR- Accuracy 61 AUC Total scores on student ratings of current symptoms = inattention: 0.56 (0.41, 0.71), hyperactivity: 0.51 (0.37, 0.64), impulsivity: 0.51 (0.37, 0.65) **Concordance:** N/A **Rater agreement:** BAARS-IV self report vs BAARS-IV parent ratings Kappa ICC 0.43 current hyperactivity: 0.31, current impulsivity: 0.32, retrospective children inattention: 0.42, retrospective childhood hyperactivity/impulsivity: 0.37 **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Erhardt, 1999{#5176} N = 51 US College	**Target:** Participants were individuals aged 18–60 years who self-reported ADHD symptoms based on DSM-IV criteria, were recruited through university clinics, and had no other primary psychiatric disorders. **ADHD presentation:** inattentive : 39,hyperactive : 17,combined : 44 **Comorbidity:** N/A	**Test description:** CAARS:S (Conners' Adult ADHD Rating Scale Self Report), a standardized questionnaire evaluating ADHD symptoms based on DSM-IV criteria, completed by participants to assess inattentive, hyperactive, and combined presentations with cutoff	**Diagnostic accuracy summary:** Sensitivity and specificity were high, with an overall diagnostic efficiency rate of 85%. Coefficient alphas ranged from .86 to .92. Median test-retest reliability for the four factors was .89. Sensitivity 82% Specificity 87.2% % PPV 86.49 NPV 82.93 LR+ 6.31 LR- 0.21	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) Sensitivity and specificity were consistent across ADHD presentationtypes (inattentive, hyperactive, and combined), with no significant differences observed between self-reported and clinically

Study ID	Population	Self Report Index Test	Results	Subgroup
	Other: Non-ADHD participants included neurotypical individuals without reported ADHD symptoms, matched for age and gender, and recruited from the same university setting for comparison. **Female:** 29% **Age:** 29.7 (7.8) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	scores applied for diagnostic evaluation Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through a structured clinical interview conducted by trained clinicians. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	Accuracy 84.62 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** 0.89 **Internal consistency:** Cronbach's alpha 0.86 - 0.92 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	diagnosed participants. Misdiagnosis rates were slightly higher for the inattentive subtype in self-reports compared to clinician diagnoses. Inter-rater reliability between self-report and clinical ratings was moderate, with higher concordance for combined presentation.
Faraone, 2010{#10633} N = 370 US Specialty care	**Target:** Adults recruited through psychiatric clinics and advertisements , met DSM-IV criteria for childhood-onset ADHD or had late-onset ADHD (met all criteria except age-at-onset), excluded if they had deafness, blindness, psychosis, inadequate English proficiency, or IQ <80 **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD adults recruited through advertisements, did not meet DSM-IV criteria for ADHD, included subthreshold ADHD participants and neurotypical controls, the setting was community-based rather than clinical **Female:** 13% **Age:** 34 (N/A) Min age: 18 Max age: 55 **Age subgroup**: Adults **Ethnicity:**	**Test description:** CBS (Current Behavior Scale), a 99-item questionnaire assessing ADHD-related behaviors with responses ranging from never to very often; Barkley's 9-item algorithm (derived from CBS and DSM-IV symptoms) is a self-report based on difficulties with attention, impulsivity, and organization Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria using the Structured Clinical Interview for DSM-IV Axis I Disorders and modules from the Schedule for Affective Disorders and Schizophrenia for School-	**Diagnostic accuracy summary:** Barkley's 9-item algorithm showed substantial diagnostic efficiency as a predictor of current DSM-IV diagnoses in adults. The best nine items and the best 18 items were not better than Barkley's 9-item algorithm. Sensitivity 92% calculated from AUC 0.8606) Specificity 99% calculated from AUC 0.8606)% PPV NPV LR+ LR- Accuracy AUC 0.8606 9 -item CBS algorithm: 0.8224, 18-item CBS algorithm: 0.8152 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A	**Subgroup analysis:** ADHD diagnosis (effect of different reference status),Age of diagnosis The study compared full ADHD, late-onset ADHD, subthreshold ADHD, and non-ADHD controls, finding that both full and late-onset ADHD groups exhibited higher psychiatric comorbidities, greater functional impairment, and more severe executive dysfunction tha The study stratified participants by full ADHD, late-onset ADHD, and subthreshold ADHD, finding that full ADHD and late-onset ADHD groups had similar levels of

Study ID	Population	Self Report Index Test	Results	Subgroup
	% White : 84 Multicenter **Funding:** Public funding	Age Children—Epidemiologic Version, administered by trained interviewers and reviewed by a diagnostic committee of board-certified child and adolescent psychiatrists or licensed psychologists (kappa 0.88 for ADHD) **Diagnosed by:** Specialist (e.g., mental health) Licensed psychologists **Timing:** Concurrent	**Cost:** N/A **Admin time:** N/A	psychiatric comorbidities, executive function deficits, and functional impairment, despite late
Gift, 2021{#299} N = 487 US Specialty care	**Target:** Adults who met DSM criteria for ADHD during adulthood with evidence of childhood ADHD symptoms, confirmed by intake assessments and clinical interviews, excluding individuals with incomplete information, contradictory diagnoses, or significant comorbidities such as personality disorders **ADHD presentation:** N/A **Comorbidity:** N/A : The study explicitly excluded participants with significant comorbidities, such as personality disorders or complex conditions. While participants with MDD or GAD were included in the comparison group, the ADHD target population did not systematical **Other:** Major depressive disorder or generalized anxiety disorder confirmed through DSM criteria and clinical interviews, as well as non-clinical controls without psychiatric diagnoses, including ADHD, recruited from the community, with balanced gender representation and age distribution **Female:** 33% male = 67%	**Test description:** WURS (Wender Utah Rating Scale) is a self-report questionnaire designed to assess DHD symptoms in adults by retrospectively evaluating childhood behaviors; it includes 25 core items (WURS-25) focusing on hyperactivity, inattention, and impulsivity, with responses rated on a 5-point Likert scale Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on structured clinical interviews conducted by mental health professionals using DSM criteria, evaluating current and childhood symptoms, functional impairments, and ruling out alternative diagnoses	**Diagnostic accuracy summary:** The WURS-25 produced good separation of ADHD subjects from normal controls with ROC (AUC = 0.974) and logistic regression (Sensitivity = 91%, Specificity = 92%). Conversely, the full WURS better separated ADHD subjects from psychiatric controls with both ROC (AUC = 0.995) and logistic regression (Sensitivity = 84%, Specificity = 94%). Use of the full WURS with its five factors proved more successful at distinguishing ADHD from MDD and GAD than did the WURS-25. I Sensitivity 62% For MDD/GAD vs ADHD (WURS-25): 62%; For Non-Clinical Control vs ADHD (WURS-25): 91% Specificity 86% For MDD/GAD vs ADHD (WURS-25): 86%; For Non-Clinical Control vs ADHD (WURS-25): 92%% PPV 73 For MDD/GAD vs ADHD (WURS-25): 73%; For Non-Clinical Control vs ADHD (WURS-25): 93% NPV 79 For MDD/GAD vs ADHD (WURS-25): 79%; For Non-Clinical Control vs ADHD (WURS-25): 90% LR+ 4.43 For the MDD/GAD vs ADHD (WURS-25): 4.43; For the Non-Clinical Control vs ADHD (WURS-25): 11.38 LR- 0.44 For the MDD/GAD vs ADHD (WURS-25): 0.44; For the Non-Clinical Control vs ADHD (WURS-25): 0.10	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Age: 31.2 (9.0) Min age: 20 Max age: 49 **Age subgroup**: Adults **Ethnicity:** N/A Unclear/Not reported **Funding:** No COI	**Diagnosed by:** Specialist (e.g., mental health) The study mentions that diagnoses were conducted by mental health professionals using structured clinical interviews and DSM criteria, indicating that specialists performed the diagnosis. **Timing:** Concurrent	Accuracy 77 For MDD/GAD vs ADHD (WURS-25): 77; For Non-Clinical Control vs ADHD (WURS-25): 91.5 AUC 0.838 For MDD/GAD vs ADHD (WURS-25): AUC = 0.838; For Non-Clinical Control vs ADHD (WURS-25): AUC 0.974 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Grogan, 2018{#323} N = 126 Ireland College	**Target:** Adults with clinical diagnosis of ADHD and ADHD with anxiety recruited from ADHD specialist clinic and support group websites **ADHD presentation:** inattentive : 42.3,combined : 30.8 **Comorbidity:** Anxiety **Other:** Adults with clinical diagnosis of anxiety alone recruited from specialist clinic and support group websites, control group included adults recruited from a university sample **Female:** 36% **Age:** 30.64 (7.56) Min age: 18 Max age: 44 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** CAARS (Conners Adult ADHD Rating Scale long version), a 66 item questionnaire, contains 8 subscales regarding inattentive, hyperactive and impulsive issues, and is completed through an online questionnaire form for ADHD evaluation with responses rated on a 4 point scale, cut-off criteria is T scores of 70STAI (State Trait Anxiety Inventory) is also an online questionnaire with 40-items and 2 subscales: state (current anxiety symptoms) and trait (general anxiety symptoms) Machine learning: No Validation dataset: No	**Diagnostic accuracy summary:** Most of CAARS subscales demonstrated low sensitivity and specificity in diagnosing and differentiating between ADHD and/or anxiety Sensitivity 53% For CAARS ADHD index Specificity 87.5% Total sample of CAARS ADHD index% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha 0.982 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) ADHD group and ADHD + anxiety group had statistically significant higher scores on CAARS subscale of inattention issues (p=0.001) compared to anxiety group alone however no significant differences amongst the ADHD groups seen

Study ID	Population	Self Report Index Test	Results	Subgroup
		Reference standard: Clinical diagnosis Participants diagnosed with ADHD by a multidisciplinary team which included a consultant psychiatrist and clinical psychologist **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Groom, 2016{#325} N = 57 UK College	**Target:** Adults who were clinically diagnosed with ADHD by a psychiatrist **ADHD presentation:** inattentive : 9.09,hyperactive : 3.03,combined : 75.76,N/A : 12.12 **Comorbidity:** N/A **Other:** Adults diagnosed with Asperger's syndrome as part of autism spectrum disorder by a psychiatrist **Female:** 39% **Age:** 31.64 (10.17) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** CAARS-E (Conners Adult ADHD Rating Scale-subscale E) Machine learning: No Validation dataset: Unclear **Reference standard:** Clinical diagnosis Participants diagnosed with ADHD by a psychiatrist establishing current and long-term diagnosis using DSM-5 **Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** QbTotal yielded the highest AUC value 0.87 (classified as 'good'). ROCs indicate that at equivalent sensitivity of around 80%, QbTotal demonstrates superior specificity compared with CAARS-E in differentiating ADHD and autism spectrum disorder. CAARS-E AUC was .77 ('fair') in differentating ADHD and autism spectrum disorder. QbTest added to clinical ratings may improve the differentiation of ADHD and autism spectrum disorder in adults. Sensitivity % Specificity % % PPV NPV LR+ LR- Accuracy AUC 0.77 fair **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
			Side effects: N/A **Cost:** N/A **Admin time:** N/A	
Harrison, 2019{#351} N = 201 Canada College	**Target:** University and college students who were diagnosed by clinical psychologists, passed symptom validity testing, and had CAARS scores below eight, they provided evidence to corroborate lifetime impairment, had selfreported deficits in keeping with observed and documented behavioral problems, and provided evidence from reliable collateral informants to confirm that their self-reported impairments were both present and severe **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Students with attention issues who did not meet ADHD criteria but passed the Word Memory Test, and symptom validity testing **Female:** 37.8% **Age:** 21.1 (4.6) Min age: 18 Max age: 22 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** CAARS:S ADHD Index (Conners' Adult ADHD Rating Scale Self Report), corresponds to DSM-IV symptoms Machine learning: No Validation dataset: Yes **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical psychologists **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The overall discriminant validity of the CAARS was 69%, and it had an unacceptably high false positive and false negative rate. At lower prevalence rates, a high score on the CAARS has only a 22% chance of accurately identifying individuals with ADHD. Sensitivity 14% Specificity 92% % PPV 47 NPV 68 LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Houston, 2011{#390} N = 343 US Primary Care	**Target:** Adults at least 18 years old, nonpsychotic, presenting to primary care **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults from primary care who did not meet ADHD diagnostic criteria **Female:** 60%	**Test description:** PDI-4 (Provisional Diagnostic Instrument-4), a 17-item screening tool designed to assess generalized anxiety disorder, major depressive episode, past/present mania, and adult ADHD; rating of symptom frequency, with a scoring system requiring at least	**Diagnostic accuracy summary:** A comparison of limited symptom-based versus full DSM-IV criteria-based diagnosis showed minimal differences in relative diagnostic accuracy. Sensitivities and specificitieswere 82% and 73% for ADHD. Sensitivity 82% (70, 90) GAD: 83 (63, 95), MDE: 80 (70, 88), Mania: 83 (63, 95) Specificity 73% (68, 79) GAD: 75 (70, 80), MDE: 80 (74, 84), Mania: 82 (77, 86)%	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Age: 50 (N/A) Min age: 18 Max age: **Age subgroup**: Age unclear **Ethnicity:** N/A Multicenter **Funding:** Industry	three of four symptom responses within a diagnostic category, and reported functional impairment for a provisional diagnosis Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the SCID (Structured Clinical Interview for DSM-IV) and the ACDS (Adult ADHD Clinician Diagnostic Scale) **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	PPV 42 (33, 51) GAD: 20 (13, 30), MDE: 58 (49, 67), Mania: 26 (17, 38) NPV 94 (91, 97) GAD: 98 (96, 100), MDE: 92 (87, 95), Mania: 98 (96, 100) LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Kessler, 2005{#442} N = 154 US Community	**Target:** Adults from the US National Comorbidity Survey Replication (NCS-R) were stratified into four groups based on self-reported childhood ADHD symptoms and persistence into adulthood, assessed using DSM-IV criteria, including those meeting full childhood ADHD criteria and reporting current symptoms. **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults from the general population who denied childhood ADHD symptoms or reported subthreshold symptoms without current persistence were drawn from a nationally representative community sample. **Female:** % N/A	**Test description:** The ASRS (Adult ADHD Self-Report Scale) includes 18 DSM-IV Criterion A symptom questions assessing inattentive and hyperactive-impulsive symptoms over the past six months using a 5-point Likert scale (Never, Rarely, Sometimes, Often, Very Often), with a 6-item short-form screener derived using logistic regression for optimal predictive accuracy. Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** The Adult ADHD Self-Report Scale (ASRS) Full 18-Item Scale distinguished ADHD from non-ADHD participants with 56.3% sensitivity, 98.3% specificity, 96.2% accuracy, and an AUC of 0.77. The ASRS 6-Item Screener outperformed the full version with 68.7% sensitivity, 99.5% specificity, 97.9% accuracy, and an AUC of 0.84. Sensitivity 56% Full 18- Item Scale: 56.3, 6-Item Screener: 68.7 Specificity 98.3% Full 18- Item Scale: 98.3, 6-Item Screener: 99.5% PPV NPV LR+ LR- Accuracy 96.2 Full 18- Item Scale: 96.2, 6-Item Screener: 97.9 AUC 0.77 Full 18- Item Scale: 0.77, 6-Item Screener: 0.84 **Concordance:**	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Age: N/A Min age: 18 Max age: 44 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** Public funding	Diagnosed with ADHD based on a semi-structured clinical interview using the ADHD Rating Scale and DSM-IV criteria **Diagnosed by:** Specialist (e.g., mental health) PhD clinical psychologists **Timing:** Concurrent	**Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** **Admin time:**	
Kessler, 2007{#443} N = 20011 US Primary Care	**Target:** Adults enrolled in a managed care plan in California and Georgia, excluding those receiving treatment for ADHD, screening for DSM-IV ADHD criteria using the ASRS Screener **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** The study included both ADHD and non-ADHD participants, using a structured ASRS Screener followed by a clinical interview for validation. **Female:** % N/A **Age:** N/A Min age: 18 Max age: N/A **Age subgroup**: Age unclear **Ethnicity:** N/A Multicenter **Funding:** Industry	**Test description:** ASRS (Adult ADHD Self-Report Scale Screener) is a six-question self-report screening tool evaluated both a dichotomous scoring approach (0–3 vs. 4–6) and a continuous scoring approach (0–24) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Adult ADHD Clinician Diagnostic Scale (ACDS v1.2). **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The diagnostic accuracy of the ASRS Screener in distinguishing adults with ADHD from those without ADHD in a managed care population, finding that the dichotomous scoring approach (cutoff 4–6) had a sensitivity of 39.1%, specificity of 88.3%, and AUC of 0.64, while the continuous scoring approach (cutoff 14+) improved sensitivity to 64.9%, specificity to 94.0%, and AUC to 0.79. Sensitivity % Dichotomous: 39.1, Continuous: 64.9 Specificity % Dichotomous: 88.3, Continuous: 94.0% PPV Dichotomous: 23.5, Continuous: 49.9 NPV Dichotomous: 94, Continuous: 96.7 LR+ LR- Accuracy Dichotomous: 84.1, Continuous: 91.5 AUC Dichotomous: 0.64, Continuous: 0.79 **Concordance:** The ASRS Screener was validated against clinical diagnoses made by mental health specialists using the Adult ADHD Clinician Diagnostic Scale (ACDS v1.2) **Rater agreement:** Agreement between self-reported ASRS scores and clinician-diagnosed ADHD using the ACDS v1.2 in a managed care population Kappa ICC	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
			Test-retest: Test-retest reliability was assessed using Pearson correlations at three time points: T1-T2 (0.63), T2-T3 (0.67), and T1-T3 (0.47), with retests conducted between 6 months to 1 year apart T1-T2: 0.63, T2-T3: 0.67, T1-T3: 0.47 The expected T1-T3 correlation (0.42) was close to the observed correlation (0.47) **Internal consistency:** Cronbach's alpha T1: 0.63, T2: 0.72, T3: 0.70 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** **Admin time:** Less than 2 minutes for completion and scoring .	
Kessler, 2010{#444} N = 345 US Community	**Target:** ADHD respondents meeting DSM-IV/ACDS criteria for ADHD **ADHD presentation:** inattentive : 60.8,inattentive_other : More predictive of adult persistence,hyperactive : 12.1,hyperactive_other : Lower persistence compared to inattentive,combined : 34.9,combined_other : Most common among those who had both subtypes in childhood **Comorbidity:** N/A **Other:** Adults without ADHD **Female:** % N/A **Age:** ADHD-Combined group: 34.34 (8.78), ADHD-Inattentive group 36.08 (11.60) Min age: 18 Max age: 44 **Age subgroup**: Adults **Ethnicity:** N/A	**Test description:** ASRS (AdultADHD Self-Report Scale), a structured questionnaire designed to assess DSM-IVADHD symptoms in adults, includes inattention and hyperactivity-impulsivity symptom items, with responses based on frequency ratings over the past six months Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria using ACDS **Diagnosed by:** Specialist (e.g., mental health) PhD-level clinical interviewers	**Diagnostic accuracy summary:** Almost half (45.7%) of individuals with childhood ADHD continued to meet full DSM-IV criteria for adult ADHD, with inattention persisting more strongly than hyperactivity-impulsivity. Executive functioning deficits were the most specific and consistent predictors of DSM-IV adult ADHD. Sensitivity 70% Specificity 93% % PPV NPV LR+ LR- Accuracy AUC 0.93 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A	**Subgroup analysis:** Age,Age of diagnosis,ADHD presentation,Setting Sensitivity and specificity of ADHD diagnoses were significantly higher in structured clinical settings compared to community-based settings, likely due to more thorough clinician-led evaluations and standardized diagnostic interviews. Functional impairme Executive functioning impairments were more predictive of ADHD persistence in older adults, while hyperactivity-impulsivity symptoms were more

Study ID	Population	Self Report Index Test	Results	Subgroup
	Multicenter **Funding:** Industry	trained by board-certified psychiatrists specializing in adult ADHD research **Timing:** Concurrent	**Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	prevalent in younger adults, suggesting age-related shifts in symptom expression and diagnostic criteria applica Sensitivity and specificity of ADHD diagnoses were higher in younger adults (18–30 years) compared to older adults (31–44 years), likely due to better recall of childhood symptoms and reduced cognitive decline in memory-based reporting. The inattention symptoms were more predictive of ADHD persistence into adulthood than hyperactivity-impulsivity symptoms, with 94.9% of persistent ADHD cases meeting inattention criteria, compared to only 34.6% meeting hyperactivity-impulsivity criteria.
Kingston, 2013{#5349} N = 120 Canada Specialty care	**Target:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **ADHD presentation:** N/A **Comorbidity:** Other : Aggression dysregulation	**Test description:** ASRS-v1.1 Part A, a scale of adult attention-deficit/hyperactivity disorder based on nosological criteria and pertain to frequency, rather than severity, of ADHD symptoms; Part A comprises 6 screening questions and is considered to be the most	**Diagnostic accuracy summary:** The integrated variables of multiple self reports and an observer report demonstrated particularly good classification accuracy, with high sensitivity (91%) and good specificity (82%). Sensitivity 76% (CI 63, 86) ASRS-v1.1 (Part B): .66 (.53–.78); Brown ADD Scale: .84 (.71–.92); CAARS-Self ADHD Index: .63 (.49–.75); WURS: .82 (.69–.91) Specificity 84% (CI 71, 92) ASRS-v1.1 (Part B): .93 (.82–.98); Brown ADD Scale: .73 (.59–.83); CAARS-Self ADHD Index: .91 (.79–.97); WURS: .69 (.54–.80)%	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Other: Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **Female:** 0% **Age:** 32.6 (10.3) Min age: 18 Max age: 64 **Age subgroup**: Adults **Ethnicity:** Other : Aboriginal: 6.5% % Hispanic or Latino : 2.8 % Black/African American : 2.8 % White : 78.5 Single center **Funding:** Industry	predictive of symptoms consistent with ADHD Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis ADHD diagnosis was determined based on DSM-IV-TR criteria following a comprehensive clinical interview and review of relevant available collateral information; interviews were conducted independently by two psychiatrists who were certified in forensic psychiatric practice; final group classification was based on consensus diagnoses and the inter-rater agreement was approximately 90% **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	PPV 83 (CI 70, 92) ASRS-v1.1 (Part B): .91 (.77–.97); Brown ADD Scale: .76 (.63–.85); CAARS-Self ADHD Index: .88 (.72–.95); WURS: .73 (.61–.83) NPV 77 (CI 64, 86) ASRS-v1.1 (Part B): .72 (.60–.82); Brown ADD Scale: .82 (.67–.91); CAARS-Self ADHD Index: .70 (.58–.80); WURS: .79 (.64–.89) LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** rater agreement between self-report measures (ASRS-v1.1, CAARS-Self, WURS, and Brown ADD Scale) and observer-rated measures (CAARS-Observer) Kappa ICC r 0.51 **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Kumar, 2011{#4185} N = 110 US Specialty care	**Target:** Adults recruited from psychiatric inpatient unit of a general hospital with a chart diagnosis of ADHD **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with different mental disorders recruited from psychiatric inpatient unit of a general hospital **Female:** 50% **Age:** 36.6 (11.1) Min age: 25 Max age: 49	**Test description:** CAARS-S:SV (Conners' Adult ADHD Rating Scales: Screening Version), 30-item self-report tool that screens for ADHD symptoms in adults, using a 4-point rating scale to assess the frequency of symptoms based on DSM-IV criteria, cut off point wasT score>70 Machine learning: No Validation dataset: N/A	**Diagnostic accuracy summary:** The CAARS-S–S: SV indicated adequate disrimination. The MINI ADHD module was most effective for identifying inpatients without ADHD. Sensitivity 83% (CI 36, 100) Specificity 69% (CI 59, 78) % PPV 14 (CI 5, 29) NPV 99 (CI 93, 100) LR+ LR- Accuracy 70 (CI 61, 78) AUC 0.75 (CI 0.6, 0.91) **Concordance:** N/A	**Subgroup analysis:** Age,Sex ADHD diagnosis based on CAARS-S or MINI were not correlated with age. ADHD diagnosis based on CAARS-S or MINI were not correlated with sex.

Study ID	Population	Self Report Index Test	Results	Subgroup
	Age subgroup: Adults **Ethnicity:** % Hispanic or Latino : 8 % Black/African American : 16 % White : 64 % Multiracial : 12,Other : other ethnic backgrounds Single center **Funding:** Unclear	**Reference standard:** Clinical diagnosis Chart diagnosis, diagnosed with ADHD by board certified psychiatrists after inpatient admission through DSM-IV-TR **Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Prior diagnosis	**Rater agreement:** Correlation self report CAARS-S:SV and MINI Kappa ICC r 0.58 **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Kwan, 2024{#490} N = 550 Canada College	**Target:** Community college or university students referred for assessments and diagnosed with ADHD through a comprehensive evaluation that included self-and-observer ratings, historical record reviews, and symptom validity tests **ADHD presentation:** inattentive : 66.7,hyperactive : 1,combined : 30.4 **Comorbidity:** N/A : currently reported academic difficulties **Other:** Students from the same educational settings referred for assessments but did not meet ADHD criteria, diagnosed based on a multi-method, multi-informant approach **Female:** 48% **Age:** mean 21.5 Min age: 17 Max age: 40 **Age subgroup**: Adults **Ethnicity:** Other : 3% Middle Eastern; 7% other % Black/African American : 6 % Asian : 12 % White : 72 Single center **Funding:** Other	**Test description:** CAARS-S:L (Conners' Adult ADHD Rating Scales–Self-Report: Long Version), a 66-itemquestionnaire to measure symptoms and behaviors associated with ADHD in adults. It uses a 4-point Likert scale (0 = not at all/never, 3 = very much/very frequently) and includes subscales such as Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability, and Self-Concept Problems. Additional scales align with DSM-IV criteria for Inattentive Symptoms, Hyperactive-Impulsive Symptoms, and Total ADHD Symptoms; tool is designed for screening but not for definitive diagnosis Machine learning: No Validation dataset: No	**Diagnostic accuracy summary:** Cutoffs of <54 (ADHD Symptoms Total subscale) and <63 (Inattentive Symptoms subscale) were also identified, both with a sensitivity of 0.95 or higher. The analysis found the ADHD Index to be a poor predictor of a negative ADHD diagnosis. Sensitivity 100% Specificity 10% % PPV 20 NPV 100 LR+ 1.11 LR- 0 Accuracy 27 AUC 0.767 (CI 0.721, 0.813) **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** ADHD presentation,Comorbidity (e.g. anxiety, depression) .The sensitivity for the inattentive subtype was 100% at a cutoff score of <54 on the Inattentive Symptoms subscale, with specificity at 25%. For the combined subtype, sensitivity and specificity data were not specifically stratified. Misdiagnosis was mor The comorbidities, particularly learning disabilities and anxiety, contributed to challenges in specificity, as these conditions often overlap with ADHD symptoms, increasing the likelihood of false positives. Sensitivity remained unaffected, maintaining high accuracy in identifying ADHD

Study ID	Population	Self Report Index Test	Results	Subgroup
		Reference standard: Clinical diagnosis Diagnosed with ADHD based on a multi-method, multi-informant assessment procedure including historical records, semi-structured clinical interviews using DSM criteria, self-and observer ratings of symptoms, and performance validity tests. **Diagnosed by:** Specialist (e.g., mental health) Clinical psychologists or supervised graduate students trained in ADHD assessment **Timing:** Concurrent		regardless of comorbidities. Functional impairment and standardized symptom scores were noted to be higher in participants with comorbid conditions
Lancaster, 2018{#495} N = 166 US College	**Target:** Adult clients who requested ADHD or learning disorder assessment at a university outpatient center, completed the Personality Assessment Inventory (PAI), and provided consent for their data to be used for research purposes; participants with valid PAI profiles and complete intelligence tests were included, with ADHD diagnoses made based on semistructured interviews and Conners' Adult ADHD Rating Scales **ADHD presentation:** inattentive : 42.27,hyperactive_other : 3.64,combined : 49.09 **Comorbidity:** N/A **Other:** Non-ADHD participants were adult clients seeking assessment at a university outpatient center who completed the Personality Assessment Inventory (PAI) and	**Test description:** PAI (Personality Assessment Inventory), includes 344 items rated on a four-point scale, measuring various psychological domains such as anxiety, depression, and impulsivity, with specific subscales examined for their association with ADHD symptoms Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on semistructured interviews, Conners' Adult ADHD Rating Scales (Self-Report and Observer Report when applicable),	**Diagnostic accuracy summary:** Adding the PAI scales to the criterion variables significantly improved the model's fit, with an overall classification accuracy of 75%. Sensitivity 55% Specificity 86% % PPV 65.22 NPV 88.79 LR+ 3.93 LR- 0.523 Accuracy 75 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) 20% of participants with ADHD had comorbid conditions, including anxiety (7%), depression (5%), and other disorders. Still, these comorbidities did not significantly affect the sensitivity (55%) or specificity (86%) of the PAI scales. Comorbidities such as anxiety and depression were associated with elevated scores on specific PAI subscales (ANX-C, DEP), which may overlap with ADHD symptoms and potentially contribute to

Study ID	Population	Self Report Index Test	Results	Subgroup
	intelligence testing, with no ADHD diagnosis determined based on semistructured interviews and Conners' Adult ADHD Rating Scales; they included individuals with various presenting concerns or no significant clinical conditions. **Female:** 61.45% **Age:** 24.39 (8.32) Min age: 18 Max age: 63 **Age subgroup**: Adults **Ethnicity:** % Black/African American : 11.45 % White : 83.13 Single center **Funding:** Unclear	and strict adherence to DSM criteria. **Diagnosed by:** Specialist (e.g., mental health) mental health clinicians **Timing:** Concurrent	**Cost:** N/A **Admin time:** N/A	misclassification. The findings highlight the importance of considering comorbid conditions during assessment to minimize misdiagnosis and improve diagnostic accuracy.
Lewandowski, 2008{#525} N = 534 US College	**Target:** College students who provided documentation to the university Office of Disability Services verifying a professional ADHD diagnosis, evidence of past and current impairment, patterns of symptoms across the lifespan, and substantial limitations in learning **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants were college students recruited from introductory psychology courses representing a neurotypical sample without reported ADHD diagnoses, spanning various academic years and demographic backgrounds **Female:** 39.47% **Age:** mean 19.2 Min age: 18 Max age: 49 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 4	**Test description:** ADHD Rating Scale with DSM-IV Checklist for ADHD (18 items reflecting ADHD symptoms, binary response option (rarely/never vs. often/always)) and additional items assessing academic and test-taking concerns Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on professional evaluation, evidence of past and current impairment, patterns of symptoms across the lifespan, and documentation of substantial limitation in learning submitted to the	**Diagnostic accuracy summary:** College students with ADHD reported significantly more ADHD symptoms and academic concerns than their peers without ADHD, but none of the 18 ADHD symptoms or six academic concerns were sensitive and specific to ADHD. Sensitivity (84%) and specificity (70%) of the self-report tool were calculated based on clinical diagnosis as the reference standard, indicating that self-reports were moderately effective at identifying individuals with ADHD but less accurate at ruling out false positives. Functional impairment (academic challenges) was also higher in the ADHD group, although similar complaints were noted among non-ADHD participants, indicating poor specificity for these academic concerns. Sensitivity 84% Specificity 70% % PPV 17.7 NPV 98.3 LR+ 2.8 LR- 0.23 Accuracy 71 AUC **Concordance:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	% Black/African American : 6.5 % Asian : 6 % White : 81 % Multiracial : 2.5 Single center **Funding:** Unclear	university Office of Disability Services **Diagnosed by:** Specialist (e.g., mental health) mental health clinician **Timing:** Concurrent	**Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** Approximately 15 minutes.	
Liu, 2023{#536} N = 955 Canada Specialty care	**Target:** Adults attending a tertiary mental health center who consented to participate in a retrospective study and completed the EarlyDetect questionnaire; ADHD diagnosis was confirmed by certified psychiatrists based on DSM-5 criteria **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults attending a tertiary mental health center for various mental health concerns, including major depressive disorder, generalized anxiety disorder, bipolar disorder, and alcohol use disorder, and were assessed by certified psychiatrists using DSM-5 criteria to confirm the absence of ADHD **Female:** 56.4% **Age:** 31.31 (10.66) Min age: 17 Max age: 76 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** EarlyDetect Questionnaire, comprehensive digital tool incorporating multiple clinical screening instruments, including ASRS-v1.1 Part A (Adult ADHD Self-Report Scale) to assess ADHD symptoms based on DSM criteria and the Sheehan Disability Scale to evaluate functional impairments; assesses mental health history, ADHD-related symptoms, and functional impairments, which were then used as features for machine learning-based ADHD screening Machine learning: Yes Validation dataset: Yes **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria through face-to-face assessment by certified psychiatrists blinded to the	**Diagnostic accuracy summary:** The ADHD classification model using composite scoring achieved a balanced accuracy of 0.788, a 2.1% increase over standalone ADHD screening. The classification model, including ADHD with comorbidity, was also successful (balanced accuracy = 0.712). Sensitivity 82% Specificity 75.3% % PPV 73.3 NPV 83.7 LR+ 3.33 LR- 0.24 Accuracy 78.8 AUC 0.86 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) The participants with ADHD and comorbid conditions (e.g., major depressive disorder, bipolar disorder, generalized anxiety disorder, alcohol use disorder) exhibited differences in classification accuracy. The model achieved a balanced accuracy of 0.712 in differentiating ADHD with comorbidities from ADHD-only cases, with lower sensitivity (61.3%) and higher specificity (81.0%) compared to ADHD-only classification. Functional impairment and specific symptom scores were more pronounced in ADHD participants with comorbidities, highlighting the impact of

Study ID	Population	Self Report Index Test	Results	Subgroup
		results of the EarlyDetect screening questionnaire. **Diagnosed by:** Specialist (e.g., mental health) Mental health clinician **Timing:** Concurrent		comorbid conditions on diagnostic outcomes and the importance of screening for these conditions to prevent misdiagnosis and optimize treatment.
Luty, 2009{#555} N = 107 UK Community	**Target:** Adults attending NHS community drug and alcohol services in South East England are able to provide informed consent and complete self-report questionnaires, excluding those unable to complete them due to illiteracy or acute agitation **ADHD presentation:** N/A **Comorbidity:** SUD : all teated for opiate dependence and alcohol use disorders **Other:** Adults attending NHS community drug and alcohol services in South East England without a confirmed diagnosis of ADHD, including individuals with substance use disorders or other mental health conditions, recruited from the same community care settings as the ADHD group **Female:** 37% **Age:** 37.8 (11.4) Min age: 18 Max age: 58 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** No COI	**Test description:** CAARS-S:L (Connors Adult ADHDRating Scale Self-report Long version), the WHO Adult ADHD Self-report Screener, and the Wender Utah Adult ADHD Scale, which assess ADHD symptoms in adults, with validation against DSM-IV diagnostic interviews with both the patient and a collateral informant Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through an interview with the patient and a collateral informant conducted by a trained psychiatrist. **Diagnosed by:** Specialist (e.g., mental health) Psychiatrist with qualifications such as Member of the Royal College of Psychiatrists **Timing:** Concurrent	**Diagnostic accuracy summary:** CAARS-S:L had the highest diagnostic accuracy with a cutoff of 91 of 198, yielding 97% sensitivity and 83% specificity. WHO-ASRS confirmed the optimal cutoff of 12 of 13, achieving 89% sensitivity and 83% specificity. WURS, though designed for childhood ADHD assessment, demonstrated 88% sensitivity and 70% specificity for diagnosing adult ADHD. Sensitivity 97% ASRS: 89; WURS: 88 Specificity 83% ASRS: 83; WURS: 70% PPV 78.49 ASRS: 76.99; WURS: 65.22 NPV 97.74 ASRS: 92.19; WURS: 90.12 LR+ 5.71 ASRS: 5.24; WURS: 2.93 LR- 0.036 ASRS: 0.13; WURS: 0.17 Accuracy 88.46 ASRS: 85.34; WURS: 77.02 AUC **Concordance:** N/A **Rater agreement:** Agreement between patient's self-report and the collateral informant's report Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Marchant, 2015{#571} N = 242	**Target:** Adults meeting DSM-IV or Utah Criteria for ADHD, with at least moderate impairment on the Clinical Global Impressions-Severity scale,	**Test description:** SR-WRAADDS (Self-Report Wender-Reimherr Adult ADHD Scale), a self-	**Diagnostic accuracy summary:** The Self-Report Wender-Reimherr Adult Attention Deficit Disorder Scale (SR-WRAADDS) distinguished adults with ADHD from normal controls with 97% sensitivity	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
US Setting varies	excluding those with major depressive disorder, panic disorder, bipolar disorder, schizophrenia, psychosis, or recent psychiatric hospitalization **ADHD presentation:** inattentive : 40 **Comorbidity:** N/A : Emotional Dysregulation Presentation, Inattentive Presentation **Other:** Couples recruited from community settings without a personal or family history of ADHD, recent Axis I disorders, or psychiatric hospitalization, while anxiety or depression trials involved participants with no ADHD diagnosis as confirmed through retrospective chart review and self-report scales; in addition, participants in anxiety and depression trials with some mental health concerns, even if they do not have a specific diagnosis of ADHD; the normative sample was selected with criteria to exclude ADHD, psychiatric disorders, or recent hospitalization **Female:** 29% **Age:** 33.7 (11.7) Min age: 18 Max age: 63 **Age subgroup**: Adults **Ethnicity:** % White : 87 Multicenter **Funding:** Industry	administered tool assessing 7 ADHD domains (attention difficulties, hyperactivity/restlessness, temper, affective lability, emotional over-reactivity, disorganization, impulsivity) based on the Utah Criteria for ADHD, cutoff score of ≥15 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical evaluation using the Wender-Reimherr Adult ADHD Scale and DSM-IV or Utah Criteria by trained mental health clinicians with moderate or greater impairment on the Clinical Global Impressions-Severity scale **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	and 89% specificity. When used to screen for ADHD in individuals with depression or anxiety, the SR-WRAADDS had 87% sensitivity and 49% specificity. The SR-WRAADDS successfully differentiated ADHD inattentive presentation from ADHD emotional dysregulation presentation with 72% agreement compared to the clinician-rated WRAADDS. Sensitivity 97% screening for ADHD depression or anxiety: 87 Specificity 89% screening for ADHD depression or anxiety: 49% PPV 91.8 NPV 95.9 LR+ LR- Accuracy 60% of anxiety/depression trial participants for whom evidence of ADHD was lacking had SR-WRAADDS scores above this threshold. AUC **Concordance:** N/A **Rater agreement:** Self reported SR-WRAADDS and investigator-rated WRAADDS Kappa ICC r 0.51 (p 0.001) **Test-retest:** N/A **Internal consistency:** Cronbach's alpha 0.78 split-half reliability r 0.92 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
McCann, 2004{#603} N = 82 US	**Target:** Adults presenting to a university-affiliated ADHD specialty clinic based on suspected ADHD, diagnosed with ADHD through a structured clinical interview incorporating DSM-IV criteria,	**Test description:** ARS, 25-item scale based on DSM-III-R criteria measuring inattention, hyperactivity, and impulsivity, scored on a 4-	**Diagnostic accuracy summary:** ADSA, ARS, and Symptom Inventory for ADHD were sensitive to the presence of ADHD in adults (correctly identifying 78-92% of patients with ADHD).	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression)

Study ID	Population	Self Report Index Test	Results	Subgroup
Specialty care	corroborating documents, and family interviews. **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults presenting to the same ADHD specialty clinic for evaluation but diagnosed with non-ADHD conditions through structured clinical interviews, including major depressive disorder, dysthymia, bipolar disorder, anxiety disorders, and other psychiatric conditions. **Female:** 44.7% **Age:** 37.5 (10.1) Min age: 18 Max age: 59 **Age subgroup**: Adults **Ethnicity:** % White : 96.3 Single center **Funding:** Unclear	point Likert scale (0 to 3), with a cutoff of 31; ADSA is a 54-item scale with multiple subscales, focusing on Attention-Focus/Concentration and Behavior-Disorganized Activity, using a cutoff corresponding to a T-score of 70 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a semistructured clinical interview incorporating DSM-IV criteria, corroborating documents such as school records and performance evaluations, and interviews with significant others. **Diagnosed by:** Specialist (e.g., mental health) Senior psychologist and board-certified psychiatrists **Timing:** Concurrent	A high proportion of individuals with non-ADHD diagnosis screened positive (incorrectly identifying between 36 – 67% of non-ADHD patients. Sensitivity 78% ARS: 60, ADSA: 81 Specificity 53.5% ARS: 58, ADSA: 46% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Mehringer, 2002{#611} N = 101 US Specialty care	**Target:** Adults aged approximately 18-50 years with a history of smoking or cocaine dependence, enriched for ADHD cases by recruiting from smoking and substance use populations, assessed for ADHD diagnosis based on DSM-IV criteria requiring evidence of symptoms in both childhood and adulthood, with no other psychiatric disorders explaining the symptomatology **ADHD presentation:** N/A	**Test description:** AHA (Assessment of Hyperactivity and Attention), an 18-item self-report pencil-and-paperquestionnaire based on DSM-IV criteria, assessing both childhood and adult ADHD symptoms, with a cutoff score of 4 for adult symptoms and 6 for childhood symptoms	**Diagnostic accuracy summary:** AHA results had sensitivity of 0.80, specificity of 0.60, PPV of 0.67, NPV of 0.75, kappa of 50 , AUC of 0.79 with odds ratio of 6.15. Sensitivity 80% Specificity 60% % PPV 67 NPV 75 LR+ LR- Accuracy AUC 0.79 0.70, 0.88 **Concordance:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Comorbidity: N/A **Other:** Adults without ADHD, including smokers and cocaine-dependent individuals recruited from specialized research settings, assessed for ADHD diagnosis but not meeting the DSM-IV criteria for childhood or adult ADHD. **Female:** 26% **Age:** 33.7 (9.7) Min age: 18 Max age: 50 **Age subgroup**: Adults **Ethnicity:** % White : 73 Single center **Funding:** Public funding	Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a semi-structured clinical interview similar to the Structured Clinical Interview for DSM-IV (SCID), requiring at least 6 inattentive or hyperactive/impulsive symptoms in both childhood and adulthood, with no other psychiatric disorders explaining the symptoms **Diagnosed by:** Specialist (e.g., mental health) psychologists & psychiatrists **Timing:** Concurrent	**Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Notzon, 2020{#667} N = 99 US Specialty care	**Target:** Adults aged 18 and older seeking treatment for cannabis use disorders, recruited from a specialty care setting **ADHD presentation:** N/A **Comorbidity:** SUD : cannabis use disorders **Other:** Non-ADHD participants were adults in the same specialty care setting seeking treatment for cannabis use disorders, serving as a comparison group to assess ADHD prevalence and the utility of screening tools **Female:** 26% **Age:** 35 (11) Min age: 18 Max age: 65 **Age subgroup**: Adults	**Test description:** WURS plus CAARS, WURS (Wender Utah Rating Scale), CAARS (Conners Adult ADHD Rating Scale), and ASRS (Adult ADHD Self-Report Scale) assess ADHD symptoms based on retrospective childhood behavior and current adult symptoms Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners Adult ADHD Diagnostic Interview for DSM-IV	**Diagnostic accuracy summary:** WURS and CAARS were combined to enhance diagnostic accuracy by pairing the high sensitivity of WURS with the high specificity of CAARS (sensitivity 0.71, specificity 0.95). WURS sensitivity 0.88, specificity 0.75; CAARS sensitivity 0.80, specificity 0.91; ASRS sensitivity 0.61, specificity 0.86. Sensitivity 71% WURS: 88; CAARS: 80; ASRS: 61 Specificity 95% WURS: 75; CAARS: 91; ASRS: 86% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:**	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Ethnicity: Other : 4 % Hispanic or Latino : 21 % Black/African American : 25 % Asian : 5 % White : 44 Single center **Funding:** Public funding	conducted by a mental health clinician **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Palmer, 2023{#685} N = 71 UK Community	**Target:** Autistic young adults recruited from a population-based cohort (SNAP) with parent-informed ADHD research diagnoses based on DSM criteria, including those with varying intellectual functioning **ADHD presentation:** inattentive : 30,hyperactive : 37.5,combined : 32.5 **Comorbidity:** Autism : Autism subgroup includes young autistic adults with varying intellectual functioning **Other:** Autistic young adults identified through the same cohort and screened as non-ADHD cases **Female:** 10.1% sample with self report: 11.3 **Age:** 23.1 (0.77) Min age: 21.33 Max age: 25.08 **Age subgroup**: Young **Ethnicity:** % White : 94.1 Single center **Funding:** Public funding	**Test description:** CAARS-S (Conners Adult ADHD Rating Scales Self Report) ADHD Index assessed ADHD symptoms with a cutoff of ≥56; adminstered together with the SDQ (Strengths and Difficulties Questionnaire), cutoff of ≥9 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the parent-informed Young Adult Psychiatric Assessment using DSM criteria and conducted by a trained researcher or clinician to ascertain symptom frequency, duration, intensity, and impairment. **Diagnosed by:** Researcher **Timing:** Concurrent	**Diagnostic accuracy summary:** Although the measures performed at or close to adequate levels (AUC was 0.66 to 0.79 for the parent report and 0.70 to 0.65 for the self-report), no single measure simultaneously met adequate thresholds for sensitivity and specificity in young adults with autism. Sensitivity 57% (CI 23, 81) SDQ>9: 28 Specificity 81% (CI 63, 92) SDQ>9: 100% PPV NPV LR+ LR- Accuracy AUC 0.70 CI (0.51, 0.90) SDQ 0.65 (CI 0.44-0.87) **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Reimherr, 2021{#757}	**Target:** Adults with a primary diagnosis of ADHD who met criteria	**Test description:** WURS (Wender Utah Rating	**Diagnostic accuracy summary:** WURS-25 (Total Score) distinguished adults with ADHD from	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
N = 485 US Specialty care	for both adult and childhood ADHD, assessed through intake questionnaires and interviews, with patients experiencing comorbidity or incomplete data excluded **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with primary diagnoses of major depressive disorder or generalized anxiety disorder, and a community control group consisting of neurotypical adults, all recruited from specialty care settings and reviewed through intake processes in clinical trials **Female:** 43% **Age:** 32.5 (8.7) Min age: 18 Max age: 59 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** No COI	Scale), including the WURS-25 and WURS-45 versions, is used to assess symptoms of ADHD and differentiate it from other conditions, with factor scores calculated from item averages and evaluated for diagnostic accuracy using ROC curves and logistic regression Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on intake questionnaires, interviews with a clinic psychiatrist, and review by several clinicians in accordance with diagnostic criteria **Diagnosed by:** Specialist (e.g., mental health) Clinic psychiatrist **Timing:** Concurrent	those with MDD/GAD with 62% sensitivity and 86% specificity. WURS-25 (Factor Scores) improved diagnostic accuracy in distinguishing ADHD from MDD/GAD with 74% sensitivity and 88% specificity. WURS-45 effectively differentiated ADHD from MDD/GAD with 80% sensitivity and 90% specificity, maintaining strong diagnostic separation with reduced redundancy. WURS-61 (Full Version) had the highest accuracy in distinguishing ADHD from MDD/GAD, with 84% sensitivity and 94% specificity. Sensitivity 74% WURS-25 total: 62; WURS-45: 80; WURS-61: 84 Specificity 88% WURS-25 total: 86; WURS-45: 90; WURS-61: 94% PPV 79 WURS-25 total: 73; WURS-45: 83; WURS-61: 88 NPV 85 WURS-25 total: 79; WURS-45: 88; WURS-61: 91 LR+ LR- Accuracy AUC 0.924 WURS-25 total: 0.838; WURS-45: 0.942 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Reyes, 2019{#762} N = 379	**Target:** Adults diagnosed with alcohol dependence recruited from inpatient and outpatient addiction	**Test description:** ASRS-v1.1 (Adult ADHD Self-Report Scale, Version 1.1),	**Diagnostic accuracy summary:** The positive predictive value (PPV) of the ASRS-v1.1 was 18.1% (95% CI = [12.4, 25.7]), and the negative	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
US Specialty care	treatment facilities, excluding those with psychotic disorders, unstable psychiatric or medical conditions, and those not meeting DSM-IV-TR criteria for alcohol dependence or presenting with contraindications to study medication **ADHD presentation:** N/A **Comorbidity:** SUD : all alcohol dependence **Other:** Adults with alcohol dependence who did not meet criteria for ADHD, recruited from addiction treatment settings, including both inpatient and outpatient care, with similar exclusion criteria to ensure comparability **Female:** 34.6% **Age:** 41.9 (11.7) Min age: 18 Max age: 80 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 2.4 % Black/African American : 1.6 % American Indian or Alaska Native : 0.3 % Asian : 0.5 % White : 95.2 Single center **Funding:** Public funding	a six-item self-administeredscreening tool with a cutoff score of ≥4 to identify symptoms consistent with ADHD diagnosis, focusing on symptoms experienced within the past six months Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Psychiatric Research Interview for Substance and Mental Disorders (PRISM), a semi-structured interview designed to differentiate primary psychiatric disorders from substance-induced effects using DSM criteria. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	predictive value (NPV) was 97.6% (95% CI = [94.9, 98.9]). The ASRS-v1.1 demonstrated a sensitivity of 79.3% (95% CI = [61.6, 90.2]) and a specificity of 70.3% (95% CI = [65.3, 74.8]). Sensitivity 79% 61.6, 90.2 Specificity 70.3% 65.3, 74.8 % PPV 18.1 12.4, 25.7 NPV 97.6 94.9, 98.9 LR+ LR- Accuracy AUC 0.75 0.67, 0.83 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Robeva, 2004{#10732} N = 12 US College	**Target:** Female college students with a current ADHD diagnosis, taking ADHD medication for at least three years, not on anxiety or depression medication, without significant health conditions affecting EEG recordings, diagnosed in childhood according to Utah standards	**Test description:** WURS (Wender Utah Rating Scale), a 61-item retrospective questionnaire witha cutoff score of 30 on the short form with higher cutoff values Machine learning: No Validation dataset: No	**Diagnostic accuracy summary:** The procedure significantly improved the score separation between ADHD and non-ADHD groups. The final average probabilities for ADHD were 76% for the ADHD group and 8% for the control group. These probabilities correlated (r=.87) with the Brown ADD scale and (r=.84) with the ADHD-Symptom Inventory used for screening the participants. Sensitivity %	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	ADHD presentation: combined : 100 **Comorbidity:** N/A **Other:** Female college students with no history of ADHD or disruptive behavioral disorders, never prescribed or taken stimulant medication, not on anxiety or depression medication, without significant medical conditions affecting EEG data collection, screened to confirm the absence of ADHD symptoms **Female:** 100% **Age:** 20.7 (1.5) Min age: 18 Max age: 22 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Other	**Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a prior clinical diagnosis made during childhood following Utah criteria, confirmed through self-report screening using the Brown Attention-Deficit Disorder Scale and the ADHD Symptom Inventory, with additional verification that participants were currently prescribed and taking stimulant medication for ADHD management **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	Specificity % % PPV NPV LR+ LR- Accuracy <0.5 probabilities with ADHD = 0.42, control = 0.54 (t = 4.1 p = 0.02) AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Roy-Byrne, 1997{#4485} N = 143 US Specialty care	**Target:** Adults presented for ADHD evaluation at a university-based specialty clinic, self-reported inattentiveness, disorganization, distractibility, or procrastination, hyperactivity-impulsivity complaints were variable, must have been able to pay a $385 fee and wait 1-2 months for an appointment **ADHD presentation:** N/A **Comorbidity:** Other : Major mood disorder **Other:** Adults seeking ADHD evaluation at the same specialty clinic who either did not meet ADHD criteria or had ambiguous ADHD features due to a lack of childhood history or confounding psychiatric/substance abuse comorbidity	**Test description:** WURS (Wender Utah Rating Scale) evaluating childhood symptoms; a cutoff score of 46 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria using a structured psychiatric interview **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** WURS distinguished ADHD from non-ADHD participants but had high false positives in psychiatric patients (40-60% of patients without ADHD also had high scores on the WURS). CPT showed group differences, but no diagnostic accuracy data were reported. WRAT-R identified lower reading scores in ADHD patients, suggesting learning disability associations but was not used for ADHD diagnosis. Sensitivity 72% Specificity % 61% specificity in clear non-ADHD sample and 39% in unclear sample% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) More ADHD patients had learning disabilities (37%) than Possible ADHD (15.7%) and Non-ADHD (13%). Lifetime major mood disorder was highly prevalent across all groups (>50%), but current substance use disorder was significantly more common in the Possible ADHD group (25%) compared to ADHD (8%) and Non-ADHD (6%) (p < .03).

Study ID	Population	Self Report Index Test	Results	Subgroup
	Female: 31.5% **Age:** 33.1 (9.7) Min age: 18 Max age: 64 **Age subgroup**: Adults **Ethnicity:** % White : 95 Single center **Funding:** Unclear		Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** Comprehensive exam that included self report: $385 in 1997 **Admin time:** N/A	
Singh, 2015{#857} N = 113 UK Specialty care	**Target:** Adults recruited from an inpatient psychiatric assessment facility, excluding those unable to give written informed consent due to acute mental illness, incapacity, or poor language skills, or those detained under the Mental Health Act **ADHD presentation:** inattentive : 14.2,hyperactive : 8.4,combined : 36.1 **Comorbidity:** N/A **Other:** Adults from an inpatient psychiatric assessment facility, primarily with diagnoses of depression, anxiety disorders, personality disorders, or other psychiatric conditions, excluding those with acute mental illness, incapacity, or poor language skills **Female:** 42% males: 69 **Age:** 34 (11.4) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** Other : 91% White British Single center **Funding:** No COI	**Test description:** IPDE-SQ, 59-item report screening questionnaire designed to assess personality traits across multiple domains (interpersonal relations, impulsivity, and mood regulation, 11-item subscale was derived from the IPDE-SQ to identify adult ADHD, cutoff score of 5 on the 11-item subscale Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the DSM-IV clinical interview conducted by a mental health clinician, assessing all 18 core criteria for ADHD subtypes and their associated impairment across multiple life domains	**Diagnostic accuracy summary:** An 11-item subscale from the IPDE-SQ shows potential as a screening instrument for ADHD in an adult psychiatric population. Sensitivity 84% Specificity 82% % PPV NPV LR+ LR- Accuracy AUC 0.873 0.805–0.942 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** ADHD presentation,Comorbidity (e.g. anxiety, depression),Any additional description/clarification of subgroup reported on this form : The study analyzed ADHD presentation (inattentive, hyperactive-impulsive, and combined) and comorbid conditions such as p Diagnostic accuracy did not significantly vary across ADHD presentations/subtypes (inattentive, hyperactive-impulsive, and combined). However, the study noted that combined type ADHD was the most frequently identified subtype, which could influence overal The study noted a high prevalence of comorbid

Study ID	Population	Self Report Index Test	Results	Subgroup
		Diagnosed by: Specialist (e.g., mental health) Psychiatrist **Timing:** Concurrent		conditions, including depression (51%), anxiety disorders (28%), and personality disorders (23%), among participants and the authors did highlight the potential for symptom overlap between ADHD and personality disorders, which could influence diagnostic accuracy and risk of misdiagnosis, but no quantitative subgroup analyses were provided.
Skirrow, 2013{#10752} N = 88 UK Specialty care	**Target:** Male adults meeting DSM-IV criteria for ADHD, recruited from a National Adult ADHD Clinic with no current or past Axis I or II comorbid psychiatric disorders (except for recurrent or current major depressive disorder), no history of substance abuse or frequent substance use, no neurological conditions, no IQ below 70, and no recent exposure to psychoactive medication (minimum washout period of 1 month for stimulants and 6 months for other psychoactive medication) **ADHD presentation:** inattentive : 19.5,combined : 80.5 **Comorbidity:** N/A **Other:** Male adults recruited from hospital volunteer databases, local community advertisements, and university settings, screened to ensure they did not meet ADHD criteria using the Barkley Adult ADHD Rating Scale, with no current or past psychiatric conditions, neurological conditions, substance	**Test description:** ALS-SF (Affective Lability Scale-Short Form) to measure emotional lability, which is often associated with ADHD Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a structured clinical interview using DSM-IV criteria conducted by a consultant adult psychiatrist specializing in ADHD, incorporating the CAADID (Conners Adult ADHD Diagnostic Interview for DSM-IV), confirming symptom onset and chronicity before age 7 and meeting criteria for at least six symptoms of hyperactivity-impulsivity	**Diagnostic accuracy summary:** For ALS-SF (AUC 91), a mean score of 1.86 corresponded to a sensitivity of 85 and a specificity of 81. Similar results were found for CNS-LS (AUC 93), with a mean score of 1.06, corresponding to a sensitivity of 88 and a specificity of 83. Sensitivity 85% CNS-LS 88 Specificity 85% CNS-LS 83% PPV NPV LR+ LR- Accuracy AUC 91 88-98 CNS-LS AUC 93, CI 85-97 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	abuse history, or frequent substance use **Female:** 0% **Age:** 28.5 (9.5) Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	and/or inattention in adulthood **Diagnosed by:** Specialist (e.g., mental health) Consultant adult psychiatrist specializing **Timing:** Concurrent	**Admin time:** N/A	
Solanto, 2004{#870} N = 93 US Specialty care	**Target:** Adults diagnosed with ADHD by clinical evaluation based on DSM-IV criteria, excluding individuals with neurological disorders, intellectual disabilities, or severe substance use disorders, and requiring stable medication use or no psychotropic medications during the assessment timeframe **ADHD presentation:** inattentive : 25.24,combined : 47.42 **Comorbidity:** N/A **Other:** Adults recruited from the same specialty care setting, with diagnoses of other psychiatric conditions such as anxiety, depression, or adjustment disorders, but who did not meet the diagnostic criteria for ADHD **Female:** % male: ADHD combined (24); ADHD Inattentive (18); Other Psychiatric (16) **Age:** ADHD combined 34.34 (8.78); ADHD Inattentive 36.08(11.60); Other psychiatric 44.39(10.35) Min age: 25 Max age: 60 **Age subgroup**: Adults **Ethnicity:** Other : ADHD combined (2.2); ADHD Inattentive (0); other psychiatric (0)	**Test description:** BADDS (Brown Attention-Deficit Disorder Scale), assesses executive and adaptive functioning across five clusters (Activation, Attention, Effort, Affect, and Memory), cutoff 50; adminstered together with CAARS (Conners Adult ADHD Rating Scale) , cutoff ≥65 for inattention, hyperactivity-impulsivity, and total ADHD scores Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through a comprehensive clinical interview conducted by experienced psychologists, supplemented with developmental history, school records, standardized test reports **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** For the Brown scales, sensitivity to a diagnosis of ADHD was 92% and specificity in identifying adults in the Other Psychiatric group was 33%, yielding an overall correct classification rate of 74%. For the CPT scores, sensitivity to a diagnosis of ADHD-Inattentive type was 47% and specificity was 86%, yielding an overall correct classification rate of 70%. The results indicate a need for closer examination of executive and adaptive functioning in adults with ADHD compared with those with internalizing disorders to identify features that could assist in differential diagnosis. Sensitivity 92% Specificity 33% % PPV 76 NPV 67 LR+ LR- Accuracy 74 AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha 0.93 Cluster-specific coefficient alphas ranged from 0.79 to 0.92 for the Brown ADD Scale **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A	**Subgroup analysis:** ADHD diagnosis (effect of different reference status),Age Sensitivity and specificity for self-report measures (e.g., Brown ADD Scale) were high when comparing ADHD-diagnosed participants to the general population but were less effective when distinguishing ADHD from other psychiatric conditions, with overlappin Age was inversely correlated with scores on the Brown ADD Scale for attention and effort, suggesting that older participants exhibited fewer ADHD-related symptoms, potentially reflecting developmental improvements in executive functioning.

Study ID	Population	Self Report Index Test	Results	Subgroup
	Other : ADHD combined (2.2); ADHD Inattentive (0); other psychiatric (4.3) Other : ADHD combined (11.1); ADHD Inattentive (3.8); other psychiatric (4.3) Other : ADHD combined (84.1); ADHD Inattentive (96.2); other psychiatric (91.3) Single center **Funding:** Unclear		**Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Ustun, 2017{#940} N = 637 US Setting varies	**Target:** Adults with ADHD (DSM-5); Sample 1: Household sample from the National Comorbidity Survey Replication, a national face-to-face survey; Sample 2: Managed care sample based on a telephone survey of subscribers to a large managed health care plan; Sample 3: Clinical sample included patients who were either obtaining a free evaluation through the Adult ADHD Program at NYU Langone based on mass media recruitment and referrals **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with no current ADHD symptoms; Sample 1: Household sample from the National Comorbidity Survey Replication (NCS-R), a national face-to-face survey; Sample 2: Managed care sample based on a telephone survey of subscribers to a large managed health care plan; Sample 3: Recruited from primary care waiting rooms near the NYU Langone campus **Female:** % N/A **Age:** ADHD group: 33.1 (11.4) years Min age: 18 Max age: 44	**Test description:** ASRS (Adult ADHD Self-Report Scale) screening scale developed to generate 1 fully structured question for each DSM-IV Criterion A1-A2 symptom of inattention and hyperactivity-impulsivity plus 11 non–DSM-IV symptoms of deficits in higher-level executive function believed to be relevant to adult ADHD similar to the Utah Criteria for adult ADHD, each question asked how often the symptom occurred over the past 6 months with responses of never, rarely, sometimes, often, and very often Machine learning: Yes Validation dataset: Yes **Reference standard:** Clinical diagnosis Clinical diagnoses of DSM-5 adult ADHD were made based on semistructured interviews using version 1.2 of the adult ADHD Clinical Diagnostic Scale.	**Diagnostic accuracy summary:** The new ADHD screening scale is short, easily scored, detects the vast majority of general population cases at a threshold that also has high specificity and PPV. Sensitivity 91% Pooled National Comorbidity Survey Replication and managed care development samples: 91.4%; NYU Langone validation sample: 91.9% Specificity 96% Pooled National Comorbidity Survey Replication and managed care development samples: 96.0%; NYU Langone validation sample: 74.0%% PPV 67.3 Pooled National Comorbidity Survey Replication and managed care development samples: 67.3; NYU Langone validation sample: 82.8% NPV LR+ LR- Accuracy AUC Pooled National Comorbidity Survey Replication and managed care development samples: 0.94; NYU Langone validation sample: 0.83 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Age subgroup: Adults **Ethnicity:** N/A Multicenter **Funding:** Other	**Diagnosed by:** Researcher The ACDS was administered in the NCS-R by 4 experi- enced PhD-level clinical interviewers who received 40 hours of training from 2 board certified psychiatrists specializing in adult ADHD research (L.A.A. and T.J.S.). Each interviewer had to complete 5 pr **Timing:** Prior diagnosis	**Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
van de Glind, 2013{#942} N = 1138 Multiple countries Setting varies	**Target:** Adults with ADHD recruited from an international ADHD in SUD prevalence study, seeking treatment for SUD during the study period, excluding those with inadequate language skills, severe physical or psychiatric issues, or who declined informed consent **ADHD presentation:** N/A **Comorbidity:** SUD : all treatment seeking **Other:** Adults without ADHD recruited from an international ADHD in SUD prevalence study, seeking treatment for SUD during the study period, excluding those with inadequate language skills, severe physical or psychiatric issues, or who declined informed consent **Female:** 26% **Age:** 35.7 (10.2) Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** Industry	**Test description:** ASRS (Adult ADHD Self-Report-Scale), a 6 item validated self-report screening tool designed for optimal alignment with clinical classifications, with scores ranging from 0 to 24 based on the sum of the first six item, cut off score 14 or more Machine learning: No Validation dataset: N/A **Reference standard:** Other Diagnosis of ADHD based on CAADID (Conners' ADHD Adult Diagnostic Interview for DSM-IV) **Diagnosed by:** Unclear/NR Principal investigator of study trained site co-ordinators and local addiction treatment professionals to use CAADID as reference standard in this study **Timing:** Later diagnosis	**Diagnostic accuracy summary:** Sensitivity was 84% and specificity 66%. Sensitivity 84% CI 76, 88)(Specificity 66% (CI 63, 69) % PPV 26 (CI 22, 30) NPV 97 (CI 96, 98) LR+ 2.44 LR- 0.19 Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** Agreement between the Adult ADHD Self-Report Scale (ASRS) at baseline (t1)and after 1–2 weeks (t2) in the same individuals 55% scored negative on the ASRS both at t1 and t2 and 29% scored positive at both time points; 8% scored positive at t1 but negative at t2 and 8% scored negative at t1 and positive at t2; findings indicate a stable result in 84% and a change of results in **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) Specificity was significantly higher in patients with alcohol use disorders at 0.76 compared to 0.56 in patients with drug use disorders, while sensitivity remained similar across both groups (0.80 and 0.85).

Study ID	Population	Self Report Index Test	Results	Subgroup
			Cost: N/A **Admin time:** N/A	
Van Voorhees, 2011{#949} N = 349 US Specialty care	**Target:** Adults seeking evaluation for attention difficulties at an ADHD clinic diagnosed with DSM-IV **ADHD presentation:** inattentive : 8.9,combined : 33.1 **Comorbidity:** N/A **Other:** Adults seeking evaluation for attention difficulties at an ADHD clinic not diagnosed with ADHD **Female:** 38.5% **Age:** mean 32, median: 28 Min age: 18 Max age: 70 **Age subgroup**: Adults **Ethnicity:** Other : Race data were only available for 77.8% of the sample % Hispanic or Latino : 1.8 % Asian : 2.9 % White : 86.4 % Multiracial : 3.7 Single center **Funding:** Other	**Test description:** CAARS:S (Conners' Adult ADHD Rating Scales, Self Rating, Long Version), 66-items rated on a 4-point scale (0 to 3) Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on CAARS, CAADID, and structured clinical interview for DSM-IV (SCID by a doctoral-level clinician **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Later diagnosis	**Diagnostic accuracy summary:** Self- and observer-ratings on the CAARS provide clinically relevant data about attention problems in adults, but the instrument does not effectively distinguish between ADHD and other adult psychiatric disorders. Combining self- and observer-ratings decreased the scales' sensitivity. Sensitivity 65% Specificity 61% % PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Self-reports (CAARS-S) and observer reports (CAARS-O including ratings from friends, parents, and spouses) Kappa ICC Ranged from r 0.24 ("distractible") through r 0.46 ("on the go/driven by a motor") **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Vizgaitis, 2023{#960} N = 122 US Setting varies	**Target:** Archival data from adults (>18years) seeking ADHD assessment, including those who consented to research participation, and completed a comprehensive ADHD assessment battery including the CAARS-S:L **ADHD presentation:** N/A	**Test description:** CAARS-S:L (Conners' Adult ADHD Rating Scale—Self-Report: Long Version), a 66-itemself-report tool rated on a 4-point scale, assessing ADHD symptoms across four primary and four composite subscales	**Diagnostic accuracy summary:** The CAARS-S:L may be useful for screening purposes in some cases but should not be the main method used for diagnostic purposes. Sensitivity 56% Specificity 62.1% % PPV 29.4 NPV 83.1 LR+	**Subgroup analysis:** Sex Stratified by gender, the outcome measures for the ADHD index were for males: sensitivity 53.3%, specificity 72.7%, PPV 40%, NPV 82.1%; and for females: sensitivity

Study ID	Population	Self Report Index Test	Results	Subgroup
	Comorbidity: N/A **Other:** Not ADHD, similar to ADHD group **Female:** 44.4% **Age:** 23 (7.81) Min age: 18 Max age: 67 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 3.3 % Black/African American : 1.6 % Asian : 11.5 % White : 75.4 % Multiracial : 8.2 Multicenter **Funding:** No COI	Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnoses made utilizing multiple self-report scales and interview assessments and confirmed by doctorate-level psychology trainees and a licensed psychologist **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Later diagnosis	LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	58.2%, specificity 52%, PPV 22.6%, NPV 83.9%.
Williamson, 2014{#4702} N = 76 US College	**Target:** Adults with a history of ADHD diagnosis confirmed by a mental health practitioner based on more than self-reported symptoms, to have received their diagnosis before age 18, and to have abstained from stimulant medication for 12 hours prior to the study **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical individuals without a history of diagnosed or suspected ADHD, learning disorders, neurological disorders, or psychological disorders, recruited from an introductory psychology participant pool or a university disability resource center, with a subset instructed to feign ADHD **Female:** 36.36% **Age:** 19.05 (1.29) Min age: 18 Max age: 23 **Age subgroup**: Young	**Test description:** WAIS-IV PSI (Wechsler Adult Intelligence Scale-IV Processing Speed) lower than 97, administered together with the Woodcock-Johnson III Test of Achievement, and the CTIP (Computerized Test of Information Processing) assessed cognitive abilities such as processing speed, reading fluency, and attention control under controlled conditions Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on evaluation by a mental health practitioner using clinical interviews, self-report symptom scales, and cognitive or	**Diagnostic accuracy summary:** Sensitivity of the WAIS-IV PSI was 65% for feigning ADHD, specificity for detecting ADHD decreased from 73% to 59% in a subgroup of participants with comorbidity. Performance validity tests such as the Test of Memory Malingering (TOMM), the Letter Memory Test (LMT), and the Nonverbal Medical Symptom Validity Test (NV-MSVT) were effective in differentiating both ADHD groups from normal participants feigning ADHD. Sensitivity 72% Specificity 72% in comorbid participants 59%% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
	Ethnicity: N/A Single center **Funding:** Unclear	neuropsychological testing, with diagnosis required to be established before age 18 **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Young, 2016{#1017} N = 392 UK Other	**Target:** All-male sample recruited from a UK prison through flyers and letters, interviewed through DIVA-2 (structured interview for ADHD) **ADHD presentation:** inattentive : 14.4,hyperactive : 13.1,combined : 18 **Comorbidity:** N/A **Other:** All-male sample without ADHD, recruited from a UK prison through flyers and letters **Female:** 0% **Age:** 30.3 (N/A) Min age: 28 Max age: 50 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** BAARS-IV (Barkley Adult ADHD Rating Scale) is a self-rating scale and assesses18 current and childhood ADHD symptoms, onset age, and impairment domains cutoff value > or = to 3 Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on Diagnostic Interview for ADHD in Adults (DIVA-2) interview by mental health professional **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The brief screening tool for ADHD demonstrated improved diagnostic accuracy among UK prison inmates, with a sensitivity of 0.82, specificity of 0.84, and overall accuracy of 0.84. This tool outperformed the original BAARS-IV scale, offering a more efficient and reliable method for identifying ADHD in correctional settings. Sensitivity 84% Based on BAARS-IV brief screening tool Specificity 82.2% Based on BAARS-IV brief screening tool% PPV NPV LR+ LR- Accuracy 83.6 Based on BAARS-IV brief screening tool AUC 0.89 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Self Report Index Test	Results	Subgroup
Young, 2023{#1016} N = 897 US Specialty care	**Target:** Adults seeking outpatient psychiatric care with suspected ADHD, aged 18-71, screened for ADHD using the ASSET-BS, Brown EF/A, and CAARS scales, inclusion required elevated T-scores across these measures and no validity flags on CAARS assessments **ADHD presentation:** combined : 74.26 **Comorbidity:** N/A **Other:** Adults seeking outpatient psychiatric care for a wide range of DSM-5 psychiatric conditions, including generalized anxiety disorder, major depressive disorder, social phobia, and bipolar disorders, assessed in a specialty care setting **Female:** % Study 1: 83.9; Study 2: 52; Study 3: 64 **Age:** Study 1: 32.02 (12.22); Study 2: 39.13 (12.54); Study 3: 30.26 (11.38) Min age: 18 Max age: 71 **Age subgroup**: Adults **Ethnicity:** Other : Study 3: 1.5 Other : study 2: 15 Other : Study 1: 5.8; Study 2: 9.5; Study 3: 3 Other : Study 2: 1.2 Other : Study 1: 4.5; Study 2: 3.9; Study 3: 3.7 Other info : Study 2: 0.3 Other : Study 1: 83.9; Study 2: 65.8; Study 3; 91 Other : Study 2: 2.9 Single center **Funding:** Unclear	**Test description:** ASSET-BS (ADHD Symptom and Side Effect Tracking - Baseline Scale), a 10-itemself-report screening tool designed to measure the impact of ADHD symptoms on daily functioning using a 6-point Likert scale Machine learning: No Validation dataset: Yes **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on elevated T-scores across the Brown Executive Function/Attention Scales, Conners' Adult ADHD Rating Scales, and clinician referral for ADHD evaluation, with no validity flags triggered on the CAARS assessments **Diagnosed by:** Specialist (e.g., mental health) Clinicians **Timing:** Concurrent	**Diagnostic accuracy summary:** The scale demonstrated effectiveness in screening for ADHD in a psychiatric outpatient population. Sensitivity 80% ASSET BS Factor at 4.04: 96.7 Specificity 80.2% ASSET BS Factor at 4.04: 65.9% PPV 57.14 ASSET BS Factor at 4.04: 47.54 NPV 92.6 ASSET BS Factor at 4.04: 98.39 LR+ LR- 0.13 - 0.49 Accuracy AUC 0.895 0.835 - 0.954 **Concordance:** N/A **Rater agreement:** self-reported ASSET-BS scores with CAARS Observer-Report scores Kappa ICC r 0.55 **Test-retest:** N/A **Internal consistency:** Cronbach's alpha 0.899 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** Sex Sensitivity and specificity of the ASSET-BS did not significantly vary between sexes.

Appendix Table A3: Evidence Table Peer Report as Index Test

Study ID	Population	Peer Report Index Test	Results	Subgroup
Dvorsky, 2016{#10626} N = 86 US College	**Target:** Undergraduate students at a large public university who self-identified as having attention or concentration difficulties or a prior ADHD diagnosis, required consent for parental interviews, completed a comprehensive ADHD evaluation including structured diagnostic interviews, and met DSM-5 ADHD criteria based on both student and parent ratings **ADHD presentation:** inattentive : 55.9,combined : 44.1 **Comorbidity:** N/A **Other:** Undergraduate students at the same university who self-identified with attention or concentration difficulties but did not meet DSM-5 criteria for ADHD based on structured diagnostic interviews and parent ratings **Female:** 42.4% **Age:** 19.71 (2.72) Min age: 18 Max age: 27 **Age subgroup**: Adults **Ethnicity:** Other : ADHD: 0, non-ADHD: 3.7 % Hispanic or Latino : 8.5,Other : 11.1 % Black/African American : 6.8,Other : non-ADHD: 18.5 % White : 76.3,Other : non-ADHD: 55.6 % Multiracial : 8.5,Other : 11.1 Single center **Funding:** Unclear	**Test description:** BAARS-IV (Barkley Adult ADHD Rating Scale-IV) used for parent ratings, 4-point scale (0 = never or rarely to 3 = very often), total sum score equal or larger than 25 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners' Adult ADHD Diagnostic Interview for DSM-IV, which included structured diagnostic interviews separately administered to students and their parents by trained graduate-level clinicians under supervision, requiring endorsement of at least five current symptoms in two or more settings and six childhood symptoms before high school **Diagnosed by:** Specialist (e.g., mental health) Psychologists **Timing:** Concurrent	**Diagnostic accuracy summary:** Parent ratings of childhood inattention had the highest predictive validity (AUC 0.79), outperforming self-report (AUC 0.56). Self-reports had high sensitivity (89%) but low specificity (30%), leading to a high false-positive rate. The prediction model with both parent and student ratings of current symptoms and parent ratings of childhood symptoms accurately classified 88.9% of individuals who had a diagnosis of ADHD and 63.3% of individuals who did not have a diagnosis. Sensitivity 60% Specificity 77% PPV 80 NPV 50 LR+ LR- Accuracy 63 AUC Total scores parent ratings of current symptoms = inattention: 0.68 (0.56, 0.81), hyperactivity: 0.50 (0.31, 0.61), impulsivity: 0.51 (0.37, 0.65); parent ratings of childhood symptoms = inattention: 0.78 (0.66, 0.89), hyperactivity/impulsivity: 0.54 (0.41, 0.67) **Concordance:** N/A **Rater agreement:** Parent ratings were compared against student self-reports for both current and childhood ADHD symptoms using Pearson correlations, intraclass correlations (ICCs), and mean differences. Kappa ICC Current inattention ICC 0.43, current hyperactivity ICC 0.31, current impulsivity ICC 0.32, retrospective children inattention ICC 0.42, retrospective childhood hyperactivity/impulsivity ICC 0.37 **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Unintended consequences:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Peer Report Index Test	Results	Subgroup
			Cost: N/A **Admin time:** N/A	
Palmer, 2023{#685} N = 71 UK Community	**Target:** Autistic young adults recruited from a population-based cohort (SNAP) with parent-informed ADHD research diagnoses based on DSM criteria, including those with varying intellectual functioning **ADHD presentation:** inattentive : 30,hyperactive : 37.5,combined : 32.5 **Comorbidity:** Autism : Autism subgroup includes young autistic adults with varying intellectual functioning **Other:** Autistic young adults identified through the same cohort and screened as non-ADHD cases **Female:** 10.1% sample with self report: 11.3 **Age:** 23.1 (0.77) Min age: 21.33 Max age: 25.08 **Age subgroup**: Young **Ethnicity:** % White : 94.1 Single center **Funding:** Public funding	**Test description:** CAARS-P (Conners Adult ADHD Rating Scales Peer Report) ADHD Index cutoff >56; administered together with ABC (Aberrant Behavior Checklist) Hyperactivity/Non-compliance subscale (a cutoff of ≥3) is a parent-reported tool designed to measure hyperactive and non-compliant behaviors in individuals with developmental disabilities Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the parent-informed Young Adult Psychiatric Assessment using DSM criteria and conducted by a trained researcher or clinician to ascertain symptom frequency, duration, intensity, and impairment. **Diagnosed by:** Researcher **Timing:** Concurrent	**Diagnostic accuracy summary:** Although the measures performed at or close to adequate levels (AUC was 0.66 to 0.79 for the parent report and 0.70 to 0.65 for the self-report), no single measure simultaneously met adequate thresholds for sensitivity and specificity in young adults with autism. Sensitivity 94% (CI 85, 100) ABC scale: 91% Specificity 57% (CI 34, 80) ABC scale: 42% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Unintended consequences:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Appendix Table A4: Evidence Table Neuropsychological Tests as Index Test

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
Adamou, 2022{#8} N = 69 UK Specialty care	**Target:** Adults over the age of 18 years with good comprehension of the English language, and IQ within normal range (>70), diagnosed with ADHD **ADHD presentation:** inattentive_other : ADHD group, N=45, M(SD) scores for DIVA - symptoms of attention-deficit: 8.6 (0.6),hyperactive_other : ADHD group, N=45, M(SD) scores for DIVA - symptoms of hyperactivity-impulsivity: 7.6 (1.8) **Comorbidity:** N/A **Other:** Adults over the age of 18 years with a good comprehension of the English language, and IQ within normal range (>70), not diagnosed with ADHD after full assessment in the study **Female:** 34.8% **Age:** 33 (9.9) Min age: 23 Max age: 42 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** QbTest, a continuous performance test measuring inattention, impulsivity, and hyperactivity combined with activity levels which are measured by an infrared motion tracking camera; consists of unconditional identical pair paradigm to avoid floor to ceiling effects; participants are asked to sit 1m from a monitor which the infrared motion tracking camera is attached to, and to hold a handheld responder; participants are instructed (by standardized instruction on the screen, and verbally) that there will be time for a 5-minute practice before they begin, and that accuracy and speed is the objective; consists of 600 stimuli presented on the monitor, each stimulus is present for 200ms, followed by an interval of 2000ms; stimulus consists of red or blue circles and squares; participants are instructed to only press the responder when the stimuli they see matches the previous stimuli in color or shape; attention is measured by number of correctly identified targets, reaction time, and variability of reaction time. Impulsivity is measured by incorrect responses, and hyperactivity is measured using the motion-tracking system using the infrared camera (captures movement by tracking a reflective headband); camera captures movement throughout the whole of the task at a frequency of 50 samples a second and with spatial resolution of 1/27 mm per infrared camera unit Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with DIVA interview by a doctor with expertise in ADHD and General Psychiatry	**Diagnostic accuracy summary:** The QbTest+ demonstrated 70% sensitivity and 43% specificity, failing to effectively differentiate between those diagnosed with ADHD and those without a diagnosis after full clinical assessment. Sensitivity 70% Specificity 43% PPV 60 NPV 54 LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 20 minutes on average.	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
		Diagnosed by: Specialist (e.g., mental health) **Timing:** N/A		
Biederman, 2017{#73} N = 60 US Specialty care	**Target:** Adults aged 18 to 55 years with a DSM-IV diagnosis of ADHD, onset of symptoms in childhood, persistence into adulthood, unmedicated for at least 1 week before the study, and no active symptoms of depression or anxiety **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Healthy adults aged 18 to 55 years without ADHD or other psychiatric disorders, recruited as controls to differentiate ADHD from neurotypical individuals in a specialty care setting **Female:** 23.33% **Age:** 30.06 (10.76) Min age: 18 Max age: 55 **Age subgroup**: Adults **Ethnicity:** % White : 82 Single center **Funding:** Industry	**Test description:** Go-NoGo task errors, participants were seated in a dimly lit room at a distance of 70 cm from a 17-inch CRT screen; Go stimuli were white English alphabetic letters appearing in equal proportions, the NoGo stimulus was a white x symbol, stimuli were presented on the center of a black background computer screen for 150 ms and were located between 2 vertical white lines, 10 trial practice block, analyzed reaction time, error rates (commission and misses) Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through clinical evaluation and ADHD module of the K-SAD-E conducted by clinicians with expertise in ADHD diagnosis and treatment **Diagnosed by:** Specialist (e.g., mental health) clinicians **Timing:** Concurrent	**Diagnostic accuracy summary:** EEG Brain Network Activation analysis demonstrated high diagnostic accuracy in distinguishing adults with ADHD from neurotypical controls, with an AUC of 0.92, sensitivity of 0.86, and specificity of 0.95 in the Go condition, and an AUC of 0.84, sensitivity of 0.76, and specificity of 0.91 in the NoGo condition. Neuropsychological tests alone showed no high discriminability for any of the indicators. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy AUC 0.67 hit RT: AUC 0.52; commission RT: AUC 0.43; percent misses: AUC 0.61; percent commission: AUC 0.64 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 12 minutes.	**Subgroup analysis:** N/A
Brunkhorst-Kanaan, 2020{#111}	**Target:** ADHD group composed of 94 (82.5%) patients who met the criteria for an ADHD diagnosis.	**Test description:** QbTest Machine learning: No Validation dataset: N/A	**Diagnostic accuracy summary:** The QbTest demonstrated limited clinical utility in differentiating adult ADHD from other psychiatric conditions, with hyperactivity	**Subgroup analysis:** Comorbidity (e.g.

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
N = 114 Germany Specialty care	**ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD-Group where adult ADHD was ruled out during the diagnostic process, consists of 20 patients (17.5%) **Female:** 42.6% **Age:** 34.7 (11.05) Min age: 23 Max age: 48 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Diagnostic Interview for Adult ADHD (DIVA 2.0), which assessed current and childhood ADHD symptoms, impairment in multiple domains of functioning, and additional childhood symptom information from the Wender Utah Rating Scale (WURS-K), with final diagnosis confirmed through clinical judgment **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	being the only parameter showing some discriminative ability (AUC = 0.65). Despite a sensitivity of 68% and specificity of 48%, its low accuracy suggests it is not a reliable standalone diagnostic tool in real-world outpatient settings. Sensitivity 68% Specificity 48% PPV NPV LR+ LR- Accuracy AUC 0.65 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 20 minutes	anxiety, depression),Setting The QbTest had poor sensitivity (68%) and specificity (48%) for ADHD diagnosis, with outcomes largely unaffected by participant comorbidities or clinical settings. The QbTest had poor sensitivity (68%) and specificity (48%) for ADHD diagnosis, with outcomes largely unaffected by participant comorbidities or clinical settings.
Cohen, 2007{#152} N = 58 US College	**Target:** Adults aged 19 to 25, recruited through psychology classes, the university disabilities office, and local medical offices; ADHD diagnosis confirmed via self-report, scores on the Conners' Adult ADHD Rating Scale exceeded 1.5 standard deviations above the mean on the DSM-IV Inattentive or Hyperactive-Impulsive Symptoms Scales, excluded if on psychoactive medication other than ADHD medication **ADHD presentation:** inattentive : 54,hyperactive : 4,combined : 43	**Test description:** Model combined Flicker task commission errors and C-CPT reaction time standard error; Flicker task measures change blindness and focused attention by requiring participants to detect changes between alternating images separated by a blank screen; metrics include the number of cycles needed to detect changes, variability, and accuracy; administered together with the C-CPT (Conners' Continuous Performance Test) assessing sustained attention, impulsivity, and response inhibition through a computerized go/no-go task in which participants respond to all letters except "X	**Diagnostic accuracy summary:** Integration showed a sensitivity of 75% and specificity of 80%. Flicker task did not demonstrate better discriminative utility than the C-CPT, although it supported the robust nature of change blindness. The CCPT exhibited only modest utility for discriminating performance in adults with and without ADHD, with weak sensitivity and moderate specificity. Sensitivity 75% Flicker Task: 57; C-CPT: 71	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Comorbidity: N/A **Other:** Neurotypical adults with no history of ADHD diagnosis, scores within one standard deviation on the Conners' Adult ADHD Rating Scale, recruited from the same university setting, excluded if on psychoactive medication **Female:** 68% **Age:** 20.46 (1.71) Min age: 19 Max age: 25 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Unclear	" to measure reaction time, omission and commission errors and variability in response Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on self-reported diagnosis, supported by scores exceeding 1.5 standard deviations above the mean on the DSM-IV Inattentive or Hyperactive-Impulsive Symptoms Scales using the Conners' Adult ADHD Rating Scale, confirmed through demographic screening and clinician evaluation. **Diagnosed by:** Specialist (e.g., mental health) clinician **Timing:** Concurrent	Specificity 80% Flicker Task: 87; C-CPT: 77 PPV 78 Flicker Task: 80; C-CPT: 74 NPV 77 Flicker Task: 68; C-CPT: 74 LR+ 3.75 Flicker Task: 4.38; C-CPT: 3.09 LR- 0.3125 Flicker Task: 0.49; C-CPT: 0.38 Accuracy 77.6 Flicker Task: 72; C-CPT: 74 AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 10-15 minutes, depending on participant performance	
Edebol, 2012{#220} Edebol, 2013{#221} N = 306 Sweden Specialty care	**Target:** Adults diagnosed with ADHD based on DSM-IV criteria, requiring symptoms to be present since childhood, assessed at neuropsychiatric clinics using clinical interviews, psychological testing, and QbTest-Plus, excluding those with clinically unstable psychiatric conditions **ADHD presentation:** inattentive : 3.8,combined : 88.7,N/A : 7.5 **Comorbidity:** N/A **Other:** Neurotypical adults without psychiatric diagnoses recruited from universities, workplaces, and music organizations, assessed using the QbTest-Plus in research or specialty care settings to serve as normative	**Test description:** QbTest-Plus plus motion tracking, objectively measures hyperactivity by tracking head movements during a 20-minute continuous performance test, records movement distance, frequency, and variability, providing a quantitative measure of motor activity that helps differentiate ADHD from non-ADHD participants Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through clinical interviews, psychological testing, self-report scales, and corroborative information from relatives conducted by mental health clinicians **Diagnosed by:** Specialist (e.g., mental health) mental health clinicians	**Diagnostic accuracy summary:** The QbTest-Plus, combining a Continuous Performance Test and Motion Tracking System, distinguished ADHD from non-ADHD normative participants with 87% sensitivity and 85% specificity. The Prediction of ADHD variable, developed from QbTest-Plus data, identified ADHD with 86% sensitivity and 83% specificity. Sensitivity 86% Specificity 83% PPV 57.32 NPV 95.43 LR+ 5.06 LR- 0.17 Accuracy AUC Concordance: N/A **Rater agreement:** Kappa ICC	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) Misdiagnosis was more common in individuals with overlapping symptoms of borderline personality disorder and bipolar disorder, reducing specificity to 36% in these subgroups. The QbTest-Plus was effective in differentiating

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	controls and differentiate from ADHD participants **Female:** 54.72% 2013: 54.55 **Age:** 35.89 (12.25) 2013: 33.35 (8.84) Min age: 18 Max age: 64 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** No COI	**Timing:** Concurrent	**Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 20 minutes	ADHD from normative participants, with functional impairment and standardized symptom scores aligning well with clinical diagnoses in specialty care settings.
Elbaum, 2020{#3913} N = 85 Israel College	**Target:** Undergraduate adult ADHD students with normal or corrected to normal vision without any learning disabilities and/or other neuropsychiatric issues **ADHD presentation:** hyperactive : 4.4,N/A : Assessed attention = 94.4% impulsivity= 4.82% **Comorbidity:** N/A **Other:** Healthy control undergraduate students with normal or corrected to normal vision without any potential ADHD indicators **Female:** 60.5% **Age:** 23.84 (2.28) Min age: 21 Max age: 26 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Industry	**Test description:** MOXO-dCPT (continuous performance test) uses varying visual and auditory distractors to simulate real-world challenges and assess performance, and it's integrated with the EyeLink 1000 eye tracker, monitoring eye movements with calibration performed for each participant before the task Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 by licensed clinician **Diagnosed by:** Other care provider (e.g., primary care physician) Licensed/trained clinician **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The findings indicate the utility of eye tracker-integrated CPTs and their enhanced diagnostic precision. Sensitivity 69% Specificity 69% PPV NPV LR+ LR- Accuracy AUC 0.78 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 18.7 minutes	**Subgroup analysis:** N/A
Emser, 2018{#234} N = 136 Germany	**Target:** Participants with ADHD were clinically referred, met DSM-IV criteria for ADHD (combined, inattentive, or hyperactive/impulsive subtype), had IQ ≥ 80, and were free	**Test description:** QbTest variables and motion tracker; Qb+ (Quantified BehaviorTest) is a continuous performance task (CPT) combined with motion tracking that evaluates sustained attention,	**Diagnostic accuracy summary:** The diagnostic accuracy using only objective data showed 79% accuracy. Predicting an ADHD diagnosis using both subjective and objective measures exceeded the accuracy	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
Specialty care	from other medical conditions causing inattention, hyperactivity, or impulsivity such as hyperthyroidism or brain disorders **ADHD presentation:** inattentive : 10.5,hyperactive : 2.6,combined : 81.6 **Comorbidity:** N/A **Other:** Non-ADHD participants were age- and gender-matched controls recruited from local universities and advertisements, had no established or suspected ADHD diagnosis or family history of ADHD, and were neurotypical without significant medical or psychiatric conditions **Female:** 34.2% children with ADHD: 30 **Age:** 35.1 (11.7) Min age: Max age: 63 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** No COI	impulsivity, and hyperactivity; the test captures metrics such as reaction time, omission/commission errors, and physical activity (distance traveled, area covered, micro-movements) using a motion tracking system; TAP (Test Battery of Attention) is a neuropsychological battery for assessing selective attention, divided attention, and sustained attention. Subtests include Go/NoGo Task assesses response inhibition and selective attention, Divided Attention Task evaluates the ability to process visual and auditory stimuli simultaneously, Sustained Attention Task measures attentiveness over a prolonged period Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical interviews conducted by experienced clinicians using DSM-IV criteria, including the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL) for children and the Wender Reimherr Interview (WRI) for adults **Diagnosed by:** Specialist (e.g., mental health) clinicians **Timing:** Concurrent	of objective measures for adults (89.5%) with the subjective variables proving to be the most relevant. Sensitivity 82% Specificity 76% PPV NPV LR+ LR- Accuracy 79 AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 20 minutes	
Galloway-Long, 2022{#283} N = 133 US College	**Target:** Adults met full DSM criteria for ADHD based on the Conners' Adult ADHD Diagnostic Interview, recruited from counties, required cross-situational severity and impairment based on standardized behavior rating scales, stimulant medication use was discontinued 24–48 hours prior, exclusions included sensorimotor disabilities, neurological disorders, autism, psychosis, non-stimulant ADHD	**Test description:** Go-No-Go task, percentage of failed inhibits on Go-No-Go Task; reaction time variability measures were used to assess inhibitory control and attention in ADHD; GNG task involved responding to frequent "go" stimuli while withholding responses to infrequent "no-go" stimuli and assessing response inhibition; reaction time variability, including standard deviation of reaction time ex-Gaussian parameters analyzed to determine	**Diagnostic accuracy summary:** Go-no-go percentage of failed inhibits successfully discriminated between adults with and without ADHD. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	medication use, and low estimated IQ **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults never diagnosed with or treated for ADHD, recruited from the same counties, reported fewer than two inattentive or hyperactive/impulsive symptoms and fewer than three total ADHD symptoms, exclusions matched those for the ADHD group including neurological disorders and low estimated IQ **Female:** % pre-school: 30.67, school-aged: 33.80, adult: 56.45 **Age:** 21.13 (1.80) Min age: Max age: 25 **Age subgroup**: Young **Ethnicity:** N/A Multicenter **Funding:** Public funding	cognitive processing differences between ADHD and non-ADHD participants Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Diagnostic Interview Schedule for Children version IV for children or the Conners' Adult ADHD Diagnostic Interview for adults, required cross-situational severity and impairment based on standardized behavior rating scales, including parent and teacher reports for children and self-report for adults, stimulant medication use was discontinued 24–48 hours prior, exclusions included neurological disorders, autism, psychosis, and low estimated IQ **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	AUC 0.73 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** Approximately 15 minutes.	
Groom, 2016{#325} N = 57 UK College	**Target:** Adults who were clinically diagnosed with ADHD by a psychiatrist **ADHD presentation:** inattentive : 9.09,hyperactive : 3.03,combined : 75.76,N/A : 12.12 **Comorbidity:** N/A **Other:** Adults diagnosed with Asperger's syndrome as part of autism spectrum disorder by a psychiatrist **Female:** 39% **Age:** 31.64 (10.17) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** N/A	**Test description:** QbTest is a computerized continuous performance test with infra-red motion tracking system, designed to assess attention, impulsivity, and activity levels; participants respond to stimuli on a screen while their movements are tracked, and scores are calculated based on attention accuracy, reaction time, and movement data, standardized against a normative sample Machine learning: No Validation dataset: Unclear **Reference standard:** Clinical diagnosis Participants diagnosed with ADHD by a psychiatrist establishing current and long-term diagnosis using DSM-5	**Diagnostic accuracy summary:** QbTotal yielded the highest AUC value 0.87 (classified as 'good'). ROCs indicate that at equivalent sensitivity of around 80%, QbTotal demonstrates superior specificity compared with CAARS-E in differentiating ADHD and autism spectrum disorder. CAARS-E AUC was .77 ('fair') in differentating ADHD and autism spectrum disorder. QbTest added to clinical ratings may improve the differentiation of ADHD and autism spectrum disorder in adults. Sensitivity 84% Specificity 80% PPV NPV LR+	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Single center **Funding:** Public funding	**Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Prior diagnosis	LR- Accuracy AUC 0.87 good Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** Approximately 20 minutes.	
Khan, 2022{#447} N = 317 US Specialty care	**Target:** Adults referred for outpatient neuropsychological evaluation for suspected or confirmed ADHD, reported English as their primary language, underwent a standardized diagnostic protocol including record review, clinical interview, and neuropsychological testing, and were evaluated for ADHD using DSM-5 criteria **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants included adults referred for neuropsychological evaluation who failed performance validity tests, with evaluations conducted in a specialty care setting focused on diagnostic clarification for conditions other than ADHD **Female:** 62.46% **Age:** 27.7 (6.67) Min age: 18 Max age: 60	**Test description:** SCWT WR raw (Stroop Color and Word Test word reading trial), SCWT assesses cognitive flexibility and processing speed through 3 trials: word reading, color naming, and color-word interference, cut of 75 or less Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria by a board-certified clinical neuropsychologist **Diagnosed by:** Specialist (e.g., mental health) Clinical neuropsychologist **Timing:** Concurrent	**Diagnostic accuracy summary:** The embedded validity indicators from the Stroop Color and Word Test were effective in determining validity status. Word Reading and Color Naming trials demonstrated acceptable classification accuracy (AUCs 0.750–0.794), with optimal cut scores of WR raw ≤75 (54% sensitivity, 89-90% specificity), WR T score ≤28 (54% sensitivity, 87-88% specificity), CN raw ≤57 (42% sensitivity, 90% specificity), and CN T score ≤30 (40% sensitivity, 90% specificity). Sensitivity 54% range for subscores 37 to 54% Specificity 89% range for subscores 82 to 90% PPV NPV LR+ LR- Accuracy AUC 0.775 range 0.75 to 0.79 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:**	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Age subgroup: Adults **Ethnicity:** Other : 5 % Black/African American : 24 % Asian : 10 % White : 46 Single center **Funding:** Unclear		**Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Kingston, 2013{#5349} N = 120 Canada Specialty care	**Target:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **ADHD presentation:** N/A **Comorbidity:** Other : Aggression dysregulation **Other:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **Female:** 0% **Age:** 32.6 (10.3) Min age: 18 Max age: 64 **Age subgroup**: Adults **Ethnicity:** Other : Aboriginal: 6.5% % Hispanic or Latino : 2.8 % Black/African American : 2.8 % White : 78.5 Single center **Funding:** Industry	**Test description:** IVA + Plus FSRCQ (Integrated Visual and Auditory Continuous Performance Test Full Scale Response Control Quotient), a computerized continuous performance test utilizing visual and auditory stimuli to assess response control; constant and sustained attention is required, as participants respond or inhibit their response to 500 counterbalanced trials; FSRCQ measures impulsivity and commission errors, normative quotient scores have a mean of 100 and a standard deviation of 15 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis ADHD diagnosis was determined based on DSM-IV-TR criteria following a comprehensive clinical interview and review of relevant available collateral information; interviews were conducted independently by two psychiatrists who were certified in forensic psychiatric practice; final group classification was based on consensus diagnoses and the inter-rater agreement was approximately 90% **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The integrated variables of multiple self reports and an observer report demonstrated particularly good classification accuracy, with high sensitivity (91%) and good specificity (82%). Sensitivity 30% (CI 17, 45) IVA + Plus (FSAQ): .39 (.29–.54) Specificity 74% (CI 58, 86) IVA + Plus (FSAQ): .69 (.53–.82) PPV 54 CI (33, 74) IVA + Plus (FSAQ): .57 (.38–.74) NPV 50 (CI 37, 63) IVA + Plus (FSAQ): .52 (.38–.65) LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
Kovner, 1998{#479} N = 29 US Specialty care	**Target:** Adults diagnosed with ADHD based on DSM-IV criteria, without known medical or neurological conditions that could account for ADHD symptoms, not on psychoactive medication, and evaluated independently by a psychiatrist, neurologist, and neuropsychologist using historical, questionnaire, and interview data **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with no ADHD diagnosis but with other non-psychotic psychiatric disorders (depression, generalized anxiety, narcissistic personality disorder) or pre-diagnosed learning disabilities, recruited from a specialty clinic setting and assessed similarly to the ADHD group using independent psychiatric, neurological, and neuropsychological evaluations **Female:** 26.32% **Age:** 33.1 (11.3) Min age: 18 Max age: 57 **Age subgroup**: Adults **Ethnicity:** % White : 100 Single center **Funding:** Unclear	**Test description:** Model with DGBR and HSST4MNR (Digits Backwards from Digit Span subtest of WAIS-R and mean reaction time from the 4h set of the Shifting Sets Test, based on Digit Span Subtest of the WAIS-R measured working memory and inhibitory control; Shifting Sets Test assessed cognitive flexibility, response inhibition, and reaction time; Continuous Performance Tests (Connors CPT and repeated stimuli CPT) evaluated sustained attention and impulse control; Recognition Memory Tests (Warrington Recognition Memory Test) gauged memory recall and incidental learning; Boston Naming Test evaluated language processing and naming ability; (WRAT-R (Wide Range Achievement Test-Revised) measured academic skills like reading, spelling, and arithmetic Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria following evaluation by a psychiatrist, neurologist, and neuropsychologist, incorporating clinical interviews, rating scales, historical records, and neuropsychological testing. **Diagnosed by:** Specialist (e.g., mental health) psychiatrist, neurologist, neuropsychologist **Timing:** Concurrent	**Diagnostic accuracy summary:** Three measures significantly ($p < 0.01$) distinguished the groups: Digits Backwards from the WAIS-R and two reaction time measures from a computerized task modeled after Luria's Competing Motor Programs. ROC curve analyses indicated that, in combination, these measures had greater than 90% accuracy for classifying ADHD and non-ADHD patients. Sensitivity % Specificity % PPV NPV LR+ LR- Accuracy AUC probability of classifying someone with ADHD and someone without ADHD was between 90.5 and 93.3% Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 2.5 hours	**Subgroup analysis:** N/A
Lev, 2022{#4209} N = 66 Israel College	**Target:** Adults with a previous ADHD diagnosis by a licensed clinician confirmed using the Structured Clinical Interview for DSM-5, undergraduate students with normal or corrected-to-normal vision, no significant	**Test description:** MOXO-dCPT with eye tracking measures included gaze duration at different areas of interest and task area of interest visit count Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** Integrating an eye tracker with CPTs is a feasible way of enhancing diagnostic precision and shows initial promise for clarifying the cognitive profile of ADHD patients. Sensitivity 76% Specificity 82%	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	neuropsychiatric comorbidities, and no learning disabilities **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical adults matched by age and education to the ADHD group, undergraduate students with normal or corrected-to-normal vision, no reported attentional impairments and a total score below 37 on the Adult ADHD Self-Report Scale (ASRS) **Female:** 67% **Age:** 23.3 (2.13) Min age: 20 Max age: 26 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Unclear	Diagnosed with ADHD by a licensed clinician using the Structured Clinical Interview for DSM-5 criteria and confirmation of prior diagnosis. **Diagnosed by:** Specialist (e.g., mental health) Clinician **Timing:** Concurrent	PPV NPV LR+ LR- Accuracy AUC 0,826 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 18.7 minutes	
Lovejoy, 1999{#544} N = 52 US Specialty care	**Target:** Adults diagnosed with ADHD based on DSM-IV criteria through clinical interview by a board-certified psychiatrist specializing in ADHD, currently taking stimulant medication (methylphenidate or dextroamphetamine) and reporting these medications as "very helpful" for addressing ADHD symptoms, with no comorbid psychiatric disorders, substance abuse history, or neurological disease, and with an estimated IQ of 85 or higher based on WAIS-R **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants were recruited from clinic waiting rooms in medical practice settings and through referrals, endorsed three or fewer ADHD symptoms on a DSM-	**Test description:** SNST (Stroop Neuropsychological Screening Test) evaluates cognitive inhibition and impulsivity; cutoff scores used were the 20th-21st percentile for ages 18-49 and 11th percentile for age 50 and above. Trail Making Test (Parts A and B) assesses attention, cognitive flexibility, and working memory; clinical cutoff was set at 1 standard deviation below the normative mean. California Verbal Learning Test (CVLT) measures verbal memory and organization; cutoff was set at 2 standard deviations below the normative mean for Short-Delay Free Recall. Controlled Oral Word Association Test (COWA) evaluates verbal fluency and executive functioning; clinical cutoff was below the 16th percentile. WAIS-R Freedom From Distractibility Factor (Digit Span and Arithmetic) assesses attention-concentration and working memory; cutoff	**Diagnostic accuracy summary:** Individual neuropsychological tests showed high positive predictive power (PPP) (83–100%), but negative predictive power (NPP) was lower. When considering the entire test battery, classification accuracy improved significantly Sensitivity % COWA: 58, CVLT: 38, SNST: 23, Trails A: 19, Trails B 96, WAIS:38 Specificity % COWA: 92, CVLT: 92, SNST: 1.0, Trails A: 1.0, Trails B 96, WAIS:1.0 PPV COWA: 88, CVLT: 83, SNST: 1.0, Trails A: 1.0, Trails B 86, WAIS:1.0 NPV COWA: 69, CVLT: 60, SNST: 57, Trails A: 55, Trails B 56, WAIS:62 LR+ LR- Accuracy AUC Concordance: N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	IV checklist, had no history of taking stimulant medications for attentional difficulties, and met the same criteria as ADHD participants for IQ (≥85), absence of psychiatric or neurological conditions, and absence of learning disabilities. **Female:** 50% **Age:** median 41 (N/A) Min age: 21 Max age: 55 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	was set at 1 standard deviation below the normative mean. Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a clinical interview by a board-certified psychiatrist specializing in ADHD using DSM-IV criteria and confirmation through participant self-report of sufficient symptoms for inattentive, hyperactive-impulsive, or combined subtype. **Diagnosed by:** Specialist (e.g., mental health) Psychiatrists **Timing:** Concurrent	**Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Mostert, 2015{#638} N = 265 Netherlands Specialty care	**Target:** Adults diagnosed with persistent ADHD present since childhood by a psychiatrist according to DSM-IV criteria, aged 18–60 years, with no psychosis, alcohol or substance addiction in the last six months, current major depression, IQ <70, neurological disorders, or sensorimotor disabilities, and no non-Caucasian ethnicity **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical adults recruited as healthy controls from the community and university settings, with no history of psychiatric or neurological disorders and no first-degree relatives with such disorders, matched for age, gender, and estimated IQ **Female:** 57.9% **Age:** mean 36 Min age: 18 Max age: 59 **Age subgroup**: Adults	**Test description:** Model with Digit span (forward), Flanker (total SD of RT), SAdots (SD series errors and response bias), Delay discounting (k100) and Time estimation (absolute median deviation from 1000ms) based on a battery of tasks assessing executive functioning (working memory, attention, inhibition), delay discounting, time estimation, and reaction time variability; tasks included measures like the WAIS-III Digit Span, Flanker Task, SART (Sustained Attention to Response Task), and a delay discounting task chosen to align with ADHD-related cognitive pathways; variability in errors, reaction times, and impulsivity parameters analyzed Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a structured Diagnostic Interview for ADHD in Adults (DIVA) and prior clinical diagnosis by a psychiatrist according to DSM-IV criteri	**Diagnostic accuracy summary:** A combined predictive model incorporating 6 neuropsychological measures achieved 82.1% specificity and 64.9% sensitivity in diagnosing ADHD. Sensitivity 65% Specificity 82% PPV NPV LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Ethnicity: N/A Single center **Funding:** Public funding	**Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Concurrent	**Admin time:** Prior to MRI: 1.5 hours; Post-MRI: 1 hour	
Nielsen, 2011{#659} N = 60 Denmark Specialty care	**Target:** Adults aged 18–43 referred to a regional outpatient psychiatric center for ADHD evaluation, diagnosed with ADHD based on DSM-IV-TR and ICD-10 criteria, with no prior ADHD medication use, including individuals with substance abuse, personality disorders, or other co-morbidities **ADHD presentation:** combined : 70 **Comorbidity:** N/A **Other:** Neurotypical adults aged 18–43 recruited from the urban community, matched by age and sex with the ADHD group, with no history of neuropsychiatric disorders or recent changes in daily habits **Female:** 53.3% **Age:** 28.3 (6.6) Min age: 18 Max age: 43 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** AQT (A Quick Test of Cognitive Speed) evaluates single- and dual-dimension naming speed (color, form, and color-form combination) and processing efficiency (overhead). Cutoffs of ≥60 seconds for dual-dimension naming and ≥6 seconds for processing efficiency were applied to differentiate adults with ADHD from neurotypical controls Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV-TR and ICD-10 criteria using a psychiatric interview and behavioral rating scales **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** When using fail criteria for dual-dimension naming (60s) and overhead (processing efficiency) (6s) together, the sensitivity was 93% and specificity 100%. Sensitivity 93% Specificity 100% PPV NPV LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Nikolas, 2019{#10710} N = 246 US Specialty care	**Target:** Adults diagnosed with ADHD based on a comprehensive clinical interview and standardized psychiatric assessments, required to have symptom onset before age 16, met full DSM-5 diagnostic criteria, provided informant reports verifying symptoms, excluded if they had neurological conditions, learning disabilities, major psychiatric disorders other than	**Test description:** TOVA ommission errors, cutoff <95 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive clinical interview, standardized psychiatric assessment, and meeting full DSM-5 diagnostic criteria with verification of symptom onset before age 16 using self-report and informant ratings	**Diagnostic accuracy summary:** While single test measures provided performed poorly in identifying ADHD participants, analyses revealed that a combined approach using self and informant symptom ratings, a positive family history of ADHD, and a reaction time variability measure correctly classified 87% of cases. Sensitivity 50% Specificity 85% PPV NPV	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	depression/anxiety, or substance abuse **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with a diagnosed unipolar mood disorder (depression) and healthy controls without ADHD or mood disorders, recruited through advertisements, email listservs, and outreach to neuropsychological clinics, with controls matched approximately by age and sex to clinical groups **Female:** 60.6% **Age:** 24.8 (6.2) Min age: 18 Max age: 40 **Age subgroup**: Adults **Ethnicity:** % White : 83.7,Other : Control: 80, depressed: 86.5 Multicenter **Funding:** Industry	**Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	LR+ LR- Accuracy AUC 0.66 Youden 0.35; AUC ranged from 0.43 for Numbers and letter efficiency to 0.55 for TOVA reaction time variability and TOVA omission errors. Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Pettersson, 2018{#711} N = 108 Sweden Specialty care	**Target:** Adults referred for ADHD assessment, required availability of a collateral historian to provide information on childhood symptoms, excluded if treated with ADHD medications, had an IQ ≤ 70, or substance-related disorders **ADHD presentation:** inattentive : 21.7,hyperactive : 7.1,combined : 76.7 **Comorbidity:** N/A **Other:** Adults referred to the same specialty neuropsychological clinic for assessment, did not meet the diagnostic criteria for ADHD, included individuals with other	**Test description:** Model with CPT II Commission errors, QbTest cardinal variable Inattention, and QbTest cardinal variable Activity Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Clinical consensus decision by a multidisciplinary assessment team using clinical interviews, neuropsychological test results, self-report measures, collateral historian input, and DSM criteria **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** All instruments showed poor discriminative ability except for the DIVA, which showed a relatively good ability to discriminate between the groups (sensitivity 90.0; specificity 72.9). A logistic regression analysis model with the DIVA and measures of inattention, impulsivity, and activity from continuous performance tests (CPTs) showed a sensitivity of 90.0 and a specificity of 83.3. Sensitivity 80% Specificity 67% PPV NPV LR+ LR- Accuracy	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	psychiatric conditions for comparison **Female:** 46.7% **Age:** 28.18 (9.09) Min age: 18 Max age: 55 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding		AUC 74.1 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 20 minutes	
Rogers, 2021{#4474} N = 147 US College	**Target:** Adults with a prior clinical diagnosis of ADHD, assessed using comprehensive psychological evaluations, with common comorbidities including major depressive disorder, learning disorders, and anxiety disorders **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Undergraduate students enrolled in psychology courses, with no history of ADHD or ADHD medication use, instructed to simulate ADHD symptoms for the purpose of the study **Female:** 54.8% **Age:** 25.59 (4.17) Min age: 18 Max age: 34 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 23.8 % Black/African American : 10.2 % Asian : 4.8 % White : 51 % Multiracial : 6.1 Single center	**Test description:** DS FE-90 (Digit Span) assessed with Matrix Reasoning WAIS-IV (Wechsler Adult Intelligence Scale–4h Edition), WAIS assesses cognitive functioning through subtests such as Digit Span, Matrix Reasoning, Visual Puzzles, and Coding, targeting attention, working memory, and processing speed Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on comprehensive psychological assessments conducted by clinicians, including clinical interviews and standardized testing to confirm the diagnosis. **Diagnosed by:** Specialist (e.g., mental health) Clinician mental specialist **Timing:** Concurrent	**Diagnostic accuracy summary:** Very large effect sizes (Cohen's ds from 1.66 to 1.90) differentiated between genuine and feigned ADHD. Two strategies (significantly below-chance performance and floor effect) showed strong promise if cross-validated for other feigning presentations. Sensitivity 73% Specificity 83% PPV NPV LR+ LR- Accuracy 81 AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Funding: Other		**Admin time:** N/A	
Schreiber, 1999{#818} N = 36 US Specialty care	**Target:** Adults diagnosed with ADHD by independent evaluations from a neuropsychologist and psychiatrist, fulfilling DSM-IV criteria, without comorbid psychiatric disorders or learning disabilities predominantly of the inattentive type, with assessments conducted before any medication trial. **ADHD presentation:** inattentive : 89,combined : 2 **Comorbidity:** N/A **Other:** The non-ADHD participants were neurotypical adults without a history of neurological, psychiatric, developmental, or learning disorders, matched to the ADHD group on gender, age, and education level, and recruited from university, community center, and clinical settings **Female:** 50% **Age:** 30.3 (10.4) Min age: 18 Max age: 52 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** BQSS (Boston Qualitative Scoring System) for the Rey-Osterrieth Complex Figure was used as the index test to assess executive functioning deficits in adults with ADHD. Performance was compared on configurable accuracy, planning, neatness, and perseveration measures Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on independent evaluations by a neuropsychologist and psychiatrist using DSM-IV criteria, supported by clinical interviews, Barkley's ADHD-IV Self-Report of Current Behavior, third-party confirmation of childhood symptoms, and neuropsychological assessments focusing on attention, executive functioning, and learning **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The BQSS may be a useful tool contributing to the neuropsychological evaluation of adults with ADHD, sensitivity was 75% and specificity 81%. Sensitivity 75% ROCF 36 point score: 68 Specificity 81% ROCF 36 point score: 71 PPV NPV LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** Inter-rater agreement between two trained research assistants who independently scored the ROCF using the Boston Qualitative Scoring System (BQSS). Kappa ICC The six summary scores had good to excellent reliability. **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Shepler, 2024{#4544} N = 140 US Specialty care	**Target:** Adults referred for evaluation at two private psychology practices, diagnosed with ADHD or ADHD comorbid with psychiatric disorders based on DSM-5 criteria, comprehensive clinical interviews, rating scales, and standardized test performance **ADHD presentation:** N/A	**Test description:** Model combining WMI, PSI and 5 CTMT trail scores; CTMT (Comprehensive Trail Making Test) to assess executive functioning, including tasks related to set-shifting, visual search, and attentional control, alongside the WMI (WAIS-IV Working Memory Index) and PSI (Processing Speed Index) to evaluate cognitive performance Machine learning: No	**Diagnostic accuracy summary:** Logistic regression analyses indicated that WMI and CTMT trail 5 scores were individually useful indicators in identifying the presence of ADHD. The best model showed a sensitivity of 67% and a specificity of 79%. Sensitivity 67% Specificity 79% PPV 75 NPV 72	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Comorbidity: N/A **Other:** Adults referred to the same private psychology practices, diagnosed with psychiatric disorders other than ADHD, including mood, anxiety, or learning disorders, using similar diagnostic criteria and assessments **Female:** 41.18% **Age:** 35.31 (14.18) Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** Other : 1.4 Albanian and Jamaican % Hispanic or Latino : 0.7 % Black/African American : 3.6 % Asian : 0.7 % White : 85 % Multiracial : 0.7 Multicenter **Funding:** Other	Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria, comprehensive clinical interviews, review of patient and informant rating scales, and performance on standardized neuropsychological tests. **Diagnosed by:** Specialist (e.g., mental health) Clinicians & clinical psychologists **Timing:** Concurrent	LR+ 3.19 LR- 0.42 Accuracy 73 AUC WMI: 0.708; PSI: 0.632; CTMT: 0.701 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Soederstroem, 2014{#869} N = 61 Sweden Specialty care	**Target:** Adults referred for neuropsychological assessment at a specialty clinic, aged 18 years or older, with suspected ADHD based on clinical evaluation and presenting with ADHD symptoms, excluding those on ADHD-specific medications during the assessment timeframe **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults referred to a specialty care neuropsychological clinic for psychiatric evaluation, including those with comorbid conditions such as mood disorders, anxiety disorders, and substance dependence but who did not meet the diagnostic criteria for ADHD **Female:** 56.1% 60%	**Test description:** Model with QbTest Plus QbInattention and QbActivity, assessed with QbTest Plus, a computerized neuropsychological test designed to assess ADHD symptoms Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical interviews, self-report scales, cognitive screening, and a general psychiatric assessment evaluating Axis I and II disorders **Diagnosed by:** Specialist (e.g., mental health) Mental health specialist **Timing:** Concurrent	**Diagnostic accuracy summary:** The discriminant validity of self-rating scales and the more objective measure of ADHD symptoms are poor and should be integrated generally with other sources of data. The combination function yielded an overall correct classification of 72.1% and the cross-validated classification showed the same result; the classification correctly identified 87.8% of the patients diagnosed with ADHD and 40.0% of the patients not diagnosed with ADHD. Sensitivity 88% Specificity 40% PPV NPV LR+ LR- Accuracy 72 AUC	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Age: 32.46 (8.99) Min age: 18 Max age: 54 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding		Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 20 minutes	
Solanto, 2004{#870} N = 93 US Specialty care	**Target:** Adults diagnosed with ADHD by clinical evaluation based on DSM-IV criteria, excluding individuals with neurological disorders, intellectual disabilities, or severe substance use disorders, and requiring stable medication use or no psychotropic medications during the assessment timeframe **ADHD presentation:** inattentive : 25.24,combined : 47.42 **Comorbidity:** N/A **Other:** Adults recruited from the same specialty care setting, with diagnoses of other psychiatric conditions such as anxiety, depression, or adjustment disorders, but who did not meet the diagnostic criteria for ADHD **Female:** % male: ADHD combined (24); ADHD Inattentive (18); Other Psychiatric (16) **Age:** ADHD combined 34.34 (8.78); ADHD Inattentive 36.08(11.60); Other psychiatric 44.39(10.35) Min age: 25 Max age: 60 **Age subgroup**: Adults	**Test description:** C-CPT (Conners Continuous Performance Test), a 14-minute computerized task where participants respond to non-target stimuli Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through a comprehensive clinical interview conducted by experienced psychologists, supplemented with developmental history, school records, standardized test reports **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** For the Brown scales, sensitivity to a diagnosis of ADHD was 92% and specificity in identifying adults in the Other Psychiatric group was 33%, yielding an overall correct classification rate of 74%. For the CPT scores, sensitivity to a diagnosis of ADHD-Inattentive type was 47% and specificity was 86%, yielding an overall correct classification rate of 70%. The results indicate a need for closer examination of executive and adaptive functioning in adults with ADHD compared with those with internalizing disorders to identify features that could assist in differential diagnosis. Sensitivity 47% Specificity 86% PPV NPV LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A	**Subgroup analysis:** ADHD diagnosis (effect of different reference status),Age Sensitivity and specificity for self-report measures (e.g., Brown ADD Scale) were high when comparing ADHD-diagnosed participants to the general population but were less effective when distinguishing ADHD from other psychiatric conditions, with overlappin Age was inversely correlated with scores on the Brown ADD Scale for attention and effort, suggesting that older participants

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Ethnicity: Other : ADHD combined (2.2); ADHD Inattentive (0); other psychiatric (0) Other : ADHD combined (2.2); ADHD Inattentive (0); other psychiatric (4.3) Other : ADHD combined (11.1); ADHD Inattentive (3.8); other psychiatric (4.3) Other : ADHD combined (84.1); ADHD Inattentive (96.2); other psychiatric (91.3) Single center **Funding:** Unclear		**Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 15 minutes	exhibited fewer ADHD-related symptoms, potentially reflecting developmental improvements in executive functioning.
Sollman, 2010{#872} N = 73 US College	**Target:** College students with a verifiable diagnosis of ADHD confirmed through neuropsychological or psychological evaluation, including corroborative interviews with parents or teachers, medication washout for 12 hours before testing, excluding those with comorbid learning disabilities, psychiatric or neurological conditions, or substance abuse **ADHD presentation:** inattentive : 20,hyperactive : 5,combined : 75 **Comorbidity:** N/A **Other:** College students recruited from the same university setting, divided into two groups: a normal honest-responding group with no history of ADHD or related disorders, and a feigning group instructed to simulate ADHD based on provided materials; participants were screened to exclude those with learning disabilities, psychiatric or neurological conditions, or substance abuse	**Test description:** C-CPT (Conner's Continuous Perofrmance Test-II) detectability index Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive clinical evaluation including neuropsychological testing, symptom self-report measures, corroborative interviews with parents or teachers, and confirmation of developmental origin of symptoms **Diagnosed by:** Specialist (e.g., mental health) Mental health clinicians **Timing:** Concurrent	**Diagnostic accuracy summary:** The detectability index in Connor's CPT-II had a sensitivity of 17% to detect ADHD and a specificity of 90% for feigning ADHD. Failing 1 or more, 2 or more, 3 or more, 4 or more cognitive feigning test indices lowered the sensitivity from 63 to 50, 47, and 35%, while the specificity increased from 82, to 93, 100%, and 100%. Indicates limited sensitivity in distinguishing ADHD from controls and susceptible to manipulation by feigning participants; results point to a need for a thorough evaluation of history, cognitive and emotional functioning, and the consideration of exaggerated symptomatology in the diagnosis of ADHD. Sensitivity 17% Specificity 90% PPV NPV LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** Kappa ICC	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Female: 44.8% **Age:** 19.40 (1.21) Min age: 18 Max age: 21 **Age subgroup**: Young **Ethnicity:** % Black/African American : 6.90 % Asian : 0 % White : 86.20 % Multiracial : 6.90 Single center **Funding:** Unclear		**Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Unal, 2019{#937} N = 44 Ireland N/A	**Target:** Aged between 18 and 65 years of age with minimum of 5 years of education and literate in English, diagnosed with ADHD **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Healthy volunteers aged between 18 and 65 years of age with minimum of 5 years of education and literate in English **Female:** 50% **Age:** ADHD group: 47.29 (9.03) years; Control group: 41.57 (11.42) years Min age: 18 Max age: 65 **Age subgroup**: Middle age **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** Stroop Test accuracy; assessed with Stroop Color-Word: Each run consisted of an initial screen where a fixation dot measuring 0.5 cm x 0.5 cm was shown at the centre of the screen for 745 milliseconds; second screen consisted of the stimulus: one colour word at a pixel resolution of 85 was shown at the centre of the screen; any of the four color words were shown randomly and the font colour of the words were either congruent to the word itself or incongruent; ratio of congruent to incongruent tasks was 1:3.; Stroop Plus Test: included an inhibitory stimulus in the form of an arrow pointing to a colored box, an arrow measuring 3 cm pointed randomly at any of the colored boxes during each run and acted as an inhibitory stimulus; participants were asked to respond to the font colour of the word and not the verbal content of the word as fast as they could by pressing the relevant key; the Stroop test was used to measure selective attention, the arrow was used to measure intentional motor inhibition, also known as executive inhibition; Perceptual Selectivity Test: The Stroop test was presented at the centre of the screen. Four shapes were shown 5 cm from the word stimulus and were positioned perpendicularly to it, 3 were similar and one	**Diagnostic accuracy summary:** Adults with ADHD have a longer response time and perform less accurately than controls. The data suggest that there is a use for objective visual attention tests in the diagnosis of adult ADHD. Sensitivity % Sensitivity is not available Specificity % Specificity is not available PPV NPV LR+ LR- Accuracy AUC 0.814 CI 0.679, 0.949) Stroop Test (response time): 0.810; Stroop Plus Test (accuracy): 0.723; Stroop Plus Test (response time): 0.724; Perceptual Selectivity Test (accuracy): 0.707; Perceptual Selectivity Test (response time): 0.783 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
		was different, the colour font of the shapes were either all blue or all yellow in 50% of runs, in the other 3 were yellow and 1 blue; in each of the shapes a line of 2.5 cm length was presented and was positioned at either -45 degrees , 0 degrees , 45 degrees or 90 degrees; participants were asked to respond to the orientation of the line in the odd shape by pressing the corresponding key and the hence colours were irrelevant to measure a subtype of selective attention known as perceptual selectivity Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD by using the Connor's Adult ADHD Diagnostic Interview for DSM-IV (CAADID) **Diagnosed by:** Unclear/NR **Timing:** Prior diagnosis	**Side effects:** N/A **Cost:** N/A **Admin time:** Total run time was 10 minutes.	
Wiig, 2012{#987} N = 134 Denmark Specialty care	**Target:** Adults referred to a regional outpatient psychiatric clinic for evaluation of possible ADHD, diagnosed with ADHD based on ICD-10 and DSM-IV criteria, including those with impaired academic achievement, difficulties with employment, and comorbidities such as substance abuse or mild personality disorders **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults referred to a regional outpatient psychiatric clinic, diagnosed with mild psychiatric disorders such as personality disorders, addiction, affective disorders, or obsessive-compulsive disorder based on ICD-10 criteria; this group excluded individuals with	**Test description:** AQT (A Quick Test of Cognitive Speed) was administered, involving tasks of single-dimension naming (color and form) and dual-dimension naming (color-form combinations) to measure processing speed and efficiency, with specific fail criteria applied for diagnosing ADHD Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on ICD-10 and DSM-IV criteria through structured psychiatric interviews conducted by a mental health clinician at a regional outpatient clinic. **Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Concurrent	**Diagnostic accuracy summary:** Results support AQT as a possible complement to psychiatric intake procedures to differentiate adults with ADHD from those with mild psychiatric disorders. Sensitivity 89% Specificity % PPV NPV LR+ LR- Accuracy AUC Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha The study mentions that AQT tests are highly reliable with test-retest	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	ADHD, autism spectrum disorder, organic brain disorders, or bipolar disorder **Female:** 43.8% **Age:** 31.14 (9.7) Min age: 17 Max age: 55 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear		reliability coefficients ranging from 0.91 to 0.95 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Woods, 2002{#999} N = 52 US Specialty care	**Target:** ADHD participants were recruited through their psychiatrists, met DSM-IV criteria, and were evaluated after a 12-hour medication break, and were excluded if they had other Axis I diagnoses, use of non-stimulant psychoactive medications, intellectual scores <85, substance abuse, neurological issues, learning disabilities, or ineffective response to stimulant medication **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Control group participants were included if they were matched to ADHD participants based on age and gender, with no more than three ADHD symptoms, no current or prior ADHD diagnosis, intellectual scores ≥85, and no history of substance abuse, no neurological disease, or head injury, or prior diagnosis or special education for a learning disability **Female:** 50% **Age:** 38.38 (9.27) Min age: 21 Max age: 55 **Age subgroup**: Adults	**Test description:** DII (Discrepancy Impairment Index) assesses cognitive impairment by comparing test performance to an individual's estimated IQ. It uses six scores: (1) COWA total, (2) CVLT Short-Delay Free Recall, (3) Color-Word from SNST, (4) TMT A time, (5) TMT B time, and (6) average age-adjusted WAIS-R FD index. DII scores are calculated by subtracting each test's T-score from the IQ estimate T-score, highlighting cognitive deficits beyond IQ, cutoff score >2 measures impaired. Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosis based on DSM-IV criteria by board certified psychiatrist **Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** These results support the consideration of discrepancies between intellectual ability and frontal/executive functioning in the assessment of adult ADHD. Sensitivity % 38 -100 Specificity % 23 - 100 PPV 57 - 100 NPV 44 - 100 LR+ LR- Accuracy AUC 0.8373 Concordance: N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuropsychological Tests as Index Test	Results	Subgroup
	Ethnicity: N/A Single center **Funding:** Unclear			

Appendix Table A5: Evidence Table Neuroimaging as Index Test

Study ID	Population	Neuroimaging as Index Test	Results	Subgroup
Amen, 2021{#28} N = 1135 US Specialty care	**Target:** Subjects drawn from the 8 Amen Clinics, Incorporated (a multidisciplinary group of psychiatric clinics that incorporate single-photon emission computed tomography (SPECT) neuroimaging into diagnostic assessment and treatment) branches, met the DSM-IV criteria for ADHD and no other diagnoses **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Did not meet criteria for any psychiatric condition and had not history of traumatic or toxic brain injury; recruited using local advertisements in newspapers and local colleges; met clinical criteria for a healthy brain subject based on authors' criteria that included the absence of current medical illnesses, brain trauma, family history of psychiatric illness, drug/alcohol abuse and no current or past evidence of behavioral or psychiatric issues as measured by a detailed clinical history, Minnesota Multiphase Personality Inventory and Structured Clinical Interview for Diagnosis for DSM-IV **Female:** 34% ADHD group: 34% female; control group: % female not reported **Age:** 37.7 (15.5) Min age: 22 Max age: 72 **Age subgroup**: Adults **Ethnicity:** Other : ADHD group: 33% non-caucasian; Control group: race information not reported Multicenter **Funding:** Unclear	**Test description:** Brain perfusion SPECT (single-photon emission computed tomography), photon emission was captured using a high-resolution Picker Prism 3000 triple-headed gamma camera with fan beam collimator with data collected in 128 × 128 matrices, yielding 120 images per scan separated by 3 degrees spanning 360 degrees; a low pass filter applied with a high cutoff and Chang attenuation correction; patients sat upright in a quiet, dimly lit room with open eyes, and the bolus was injected after 10 min, patients sat for an additional 10 minutes after Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Clinical diagnosis based on DSM-IV by specialists at the Amen Clinics, Incorporated branches **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** SPECT Functional Neuroimaging distinguishes adult ADHD patients without comorbidities from healthy controls with 100% sensitivity and specificity in post-hoc ROI analysis. Visual reads of SPECT scans showed 100% sensitivity and >97% specificity in distinguishing adult ADHD from healthy controls. Sensitivity 100% Specificity 97% PPV NPV LR+ LR- Accuracy AUC 97.6 **Concordance:** N/A **Rater agreement:** Rater agreement in visual interpretations of SPECT scans by multiple nuclear medicine physicians and radiologists by analyzing regional cerebral blood flow (rCBF) in 14 cortical and 7 subcortical regions Kappa 0.79 ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Neuroimaging as Index Test	Results	Subgroup
Chaim-Avancini, 2017{#134} N = 116 Brazil Specialty care	**Target:** Stimulant-naive men with ADHD **ADHD presentation:** inattentive : 53.7,combined : 46.2 **Comorbidity:** N/A **Other:** 66 healthy controls (44 men) **Female:** 22.39% non-ADHD: 33.33 **Age:** 27 (6.0) Min age: 18 Max age: 50 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** Structural MRI and diffusion tensor imaging, 1.5T Espree system Machine learning: Yes Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Structured Clinical Interview for DSM-IV and ADHD-related items from an adapted version of the Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS-E), requiring at least six inattention or hyperactivity/impulsivity symptoms persisting from childhood into adulthood with impairment in multiple domains. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The combination of T1-weighted MRI and DTI features achieved an AUC of 0.71 and diagnostic accuracy of 65.4% (P = 0.005) in a mixed-gender ADHD group. In a male-only ADHD subgroup, the combination of T1-weighted MRI and DTI features improved classification accuracy to 74% (P = 0.0001), with an AUC of 0.74. Sensitivity 65% 53, 78 Male only: 73 (62, 86) Specificity 86% Male only: 86 (77, 97) PPV 68.6 Male only: 79.1 NPV 63.1 Male only: 71.1 LR+ LR- Accuracy 65.4 Male only: 73.8 AUC 0.71 Male only: 0.74 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** Sex,Comorbidity (e.g. anxiety, depression) Classification performance was higher in the male-only subgroup compared to the mixed-gender subgroup, suggesting that male ADHD patients may have more significant neuroanatomical deviations from controls. However, the authors cautioned that they could not conduct female-only analyses due to the limited sample size, and the observed differences might reflect the use of a more homogeneous sample rather than actual sex-based differences in diagnostic performance. 12 ADHD participants and

Study ID	Population	Neuroimaging as Index Test	Results	Subgroup
				one healthy control (HC) had comorbid psychiatric diagnoses, but the impact of these comorbidities on diagnostic performance was not explicitly analyzed. However, the authors excluded participants with a history of major psychiatric disorders (except mild depression, anxiety disorders, and disruptive behavior disorders), as well as those with prior ADHD treatment, drug abuse, or general medical conditions.
Schneider, 2014{#812} N = 427 Canada Specialty care	**Target:** Participants with a clinical DSM-IV diagnosis of ADHD identified through community-based psychiatric clinics and offices, presenting with complex or refractory cases requiring further diagnostic clarification, ranging in age from teenage to geriatric **ADHD presentation:** N/A	**Test description:** 3D thresholded SPECT (Single Photon Emission Computed Tomography), which discards areas below 55% of maximum activity to evaluate regional cerebral blood flow; Tc99m radiotracers following baseline or concentration protocols, and	**Diagnostic accuracy summary:** 3D thresholded SPECT provides a stronger signal for ADHD detection in clinical settings and may outperform conventional SPECT in sensitivity while maintaining reasonable specificity. Sensitivity 54% (CI 46, 61) Conventional SPECT: 4% Specificity 76% (CI 71, 81) Conventional SPECT: 97%	**Subgroup analysis:** N/A

Study ID	Population	Neuroimaging as Index Test	Results	Subgroup
	Comorbidity: N/A **Other:** Individuals with various psychiatric and neuropsychiatric disorders other than ADHD, identified through the same community-based psychiatric clinics and offices, representing a mix of general and specialty care settings **Female:** 51% **Age:** 40.9 (15.7) Min age: 14 Max age: 82 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	a nuclear medicine physician interpreted them without formal blinding Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical DSM-IV criteria evaluated by community-based psychiatrists using clinical interviews and patient history **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	PPV NPV LR+ LR- Accuracy 67 computed AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** Approximately 15-20 minutes.	
Wang, 2013{#970} N = 46 China College	**Target:** Adults (>18 years) with combined lifetime ADHD **ADHD presentation:** inattentive : 73.9,hyperactive : 56.5 **Comorbidity:** N/A **Other:** Adults (>18years) without ADHD, gender and age matched to ADHD group **Female:** 21.7% **Age:** 31.54 (9.75) VC: 32.04 (9.23) Min age: 18 Max age: 35 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** No COI	**Test description:** Resting-state fMRI and structural MRI images were collected while subjects relaxed, with voxel-wise ReHO (regional homogeneity) used to extract signals as model inputs to differentiate patterns amongst ADHD group and control group Machine learning: Yes Validation dataset: Unclear **Reference standard:** Other Unclear how reference standard test was completed and unclear if conducted by appropriate clinician **Diagnosed by:** Unclear/NR **Timing:** N/A	**Diagnostic accuracy summary:** ADHD brain regions were more activated than normal controls during resting state. Linear support vector classifier can provide useful discriminative information of altered ReHo patterns for ADHD; and feature selection can improve the performances of classification. Sensitivity 87% Specificity 74% PPV NPV LR+ LR- Accuracy 80 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC Correlation coefficient map amongst the ADHD-RS scores and the most discriminative ReHo features, Inattentive scores showed positive correlation with ReHo, while	**Subgroup analysis:** N/A

Study ID	Population	Neuroimaging as Index Test	Results	Subgroup
			hyperactive/impulsive results demonstrated negative correlation with ReHo **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Yao, 2018{#1011} N = 251 China College	**Target:** Participants with ADHD were recruited from clinics at Peking University Sixth Hospital, quired to be drug-naïve for stimulants and psychotropic drugs, have an IQ score greater than 80, and be right-handed **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants were age-matched healthy controls recruited from local universities, with no history of psychiatric or neurological disorders, and were selected to match the ADHD participants in demographics and handedness **Female:** 25.3% **Age:** 25.93 (4.86) Min age: Max age: 34 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** Public funding	**Test description:** Resting state-fMRI (functional MRI) to analyze functional connectivity patterns across 246 brain regions, identifying potential biomarkers for ADHD diagnosis; a novel feature selection method, FS_RIEL, was applied to reduce dimensionality and improve classification accuracy Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical evaluation conducted by mental health clinicians at Peking University Sixth Hospital following established diagnostic criteria. **Diagnosed by:** Specialist (e.g., mental health) Mental health clinician **Timing:** Concurrent	**Diagnostic accuracy summary:** The method achieved 80% accuracy in distinguishing ADHD from healthy controls. Sensitivity 91% Specificity 65% PPV NPV LR+ LR- Accuracy 80 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Appendix Table A6: Evidence Table EEG as Index Test

Study ID	Population	EEG as Index Test	Results	Subgroup
Baghdassarian, 2018{#425} N = 108 Sweden Specialty care	**Target:** Adults, unmedicated for at least 24 hours prior to testing, meeting DSM-IV criteria through multidisciplinary assessment, with a history of ADHD symptoms from before age 7, excluding those with psychosis, schizophrenia, schizoaffective disorder, bipolar disorder, autism spectrum disorders, brain damage, epilepsy, ongoing substance misuse, or neurological disorders **ADHD presentation:** inattentive : 37.5,combined : 58.3 **Comorbidity:** N/A **Other:** Adults recruited from healthy controls (primarily students and hospital personnel, screened to rule out ADHD, psychosis, or prodromal syndromes) and patients with schizophrenia diagnosed using DSM-IV criteria; healthy controls were neurotypical with no significant mental or neurological conditions, while schizophrenia patients were stably medicated, had a duration of illness of at least one year, recruited from a psychosis hospital unit **Female:** 50% **Age:** 29.5 (8.1) Min age: 18 Max age: 50 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** Auditory Brainstem Response profiling tests using disease-specific traits derived from auditory waveform characteristics to differentiate ADHD from other conditions and healthy controls. The cutoff for a positive diagnosis was a disease index ≥50% Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria through multidisciplinary assessments conducted at a neuropsychiatric outpatient clinic using structured clinical interviews and consensus best estimate diagnosis methods **Diagnosed by:** Specialist (e.g., mental health) psychiatrists and neuropsychiatric experts **Timing:** Concurrent	**Diagnostic accuracy summary:** Profiling identified adult ADHD versus controls with a sensitivity of 87.5% and a specificity of 91.4%. 1/26 schizophrenia patients was a false positive for ADHD. Sensitivity 88% Specificity 91% PPV 80.8 NPV 94.6 LR+ 10.2 LR- 0.14 Accuracy 90.2 DOR 72.8 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** About 40 minutes.	**Subgroup analysis:** Sex The sensitivity for the ABR profiling test was lower in females (83.3%) compared to males (91.6%) for ADHD diagnosis.
Biederman, 2017{#73} N = 60 US Specialty care	**Target:** Adults aged 18 to 55 years with a DSM-IV diagnosis of ADHD, onset of symptoms in childhood, persistence into adulthood, unmedicated for at least 1 week before the study, and no active symptoms of depression or anxiety **ADHD presentation:** N/A **Comorbidity:** N/A	**Test description:** Event-related potential data to analyze brain activity patterns during Go/NoGo task, Go condition Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** EEG Brain Network Activation analysis demonstrated high diagnostic accuracy in distinguishing adults with ADHD from neurotypical controls, with an AUC of 0.92, sensitivity of 0.86, and specificity of 0.95 in the Go condition, and an AUC of 0.84, sensitivity of 0.76, and specificity of 0.91 in the NoGo condition.	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
	Other: Healthy adults aged 18 to 55 years without ADHD or other psychiatric disorders, recruited as controls to differentiate ADHD from neurotypical individuals in a specialty care setting **Female:** 23.33% **Age:** 30.06 (10.76) Min age: 18 Max age: 55 **Age subgroup**: Adults **Ethnicity:** % White : 82 Single center **Funding:** Industry	Diagnosed with ADHD based on DSM-IV criteria through clinical evaluation and ADHD module of the K-SAD-E conducted by clinicians with expertise in ADHD diagnosis and treatment **Diagnosed by:** Specialist (e.g., mental health) clinicians **Timing:** Concurrent	Neuropsychological tests alone showed no high discriminability for any of the indicators. Sensitivity 86% NoGo condition: 76 Specificity 95% NoGo condition: 91 PPV 0.93 NoGo condition: 0.90 NPV 0.85 NoGo condition: 0.80 LR+ LR- Accuracy AUC 0.92 NoGo condition: 0.84 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** 12 minutes.	
Hadas, 2021{#334} N = 108 Israel College	**Target:** Adults with confirmed diagnosis of ADHD without other co-morbidites, with no use of psychoactive medications 1 week prior to study **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Fit and healthy adults recruited from university advertisement boards or online newsletter **Female:** 33% **Age:** 26 (0.3) Min age: 17 Max age: 30 **Age subgroup**: Adults **Ethnicity:** N/A Single center	**Test description:** EEG (Electroencephalography) was recorded during transcranial magnetic stimulation (TMS) targeting the right prefrontal cortex and during the Stop Signal Task. Stop Signal Task assesses the response inhibition and cognitive control by responding to visual cues and pressing corresponding buttons, the test serves as a paradigm to elicit neural activity for EEG recordings. Machine learning: No Validation dataset: N/A	**Diagnostic accuracy summary:** Significant reductions in transcranial magnetic stimulation-evoked potentials (TEPs) and event-related potentials (ERPs) in individuals with ADHD compared to healthy controls, as well as significant correlations between ADHD severity and TEP. Sensitivity 88% Specificity 54% PPV NPV LR+ LR- Accuracy 72 AUC 0.73 **Concordance:** N/A	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
	Funding: Other	**Reference standard:** Clinical diagnosis Diagnosis of ADHD confirmed with psychiatrist through clinical interview **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** N/A	
Kaur, 2020{#436} N = 97 India College	**Target:** ADHD diagnosis is confirmed in 48 cases after clinical assessment, group must fulfill DSM-5 criteria of ADHD diagnosis **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Age matched adults not having ADHD or any other psychopathology **Female:** 20.3% **Age:** 20.3 (1.12) Min age: 19 Max age: 23 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** EEG (electroencephalography) to record brain activity from 19 scalp electrodes under 3 conditions: eyes-open, eyes-closed, and during the CPT (Continuous Performance Test); EEG signals were preprocessed to remove artifacts, and phase space reconstruction features were extracted to classify ADHD and control adults using machine learning classifiers Machine learning: Yes Validation dataset: Yes **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** EEG in the CPT condition provided the highest accuracy, achieving 100% sensitivity and specificity with the Neural Dynamic Classifier NDC; testing accuracy was 93.3% under the eyes-open, 90% under the eyes-closed, and 100% under the CPT condition. Sensitivity 100% Eyes open condition: 100, Eyes closed condition: 93.3, CPT: 100 Specificity 87% Eyes open condition: 86.7, Eyes closed condition: 86.7, CPT: 100 PPV NPV LR+ LR- Accuracy 93.3 Eyes open condition: 93.3, Eyes closed condition: 90, CPT: 100 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
			Labeling: N/A **Side effects** N/A **Cost:** N/A **Admin time:** 17 minutes	
Kiiski, 2020{#448} N = 134 Ireland Specialty care	**Target:** Adults with ADHD **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** 1st degree relatives of people with ADHD (18 siblings, 27 parents) and healthy controls recruited from the general population, support groups and a secondary mental health care service **Female:** 50% **Age:** 27.1 (10.4) Min age: 27 Max age: 38 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** Resting-state EEG functional connectivity to assess brain network differences, connectivity patterns analyzed across the delta, theta, alpha, beta, and gamma frequency bands using the weighted phase lag index and machine learning models Machine learning: Yes Validation dataset: Yes **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners' Adult ADHD Rating Scale and the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** While EEG connectivity could predict ADHD symptom severity, it did not reliably classify ADHD, 1st-degree relatives, and controls, with modest classification performance (AUC up to 0.669); EEG may serve as a neuromarker for ADHD symptoms, but its diagnostic utility remains limited due to variability in classification accuracy. Sensitivity 73% Eyes open (ADHD vs relatives 70.13, Control vs relatives 49), Eyes closed (ADHD vs relatives 69.12, Controls vs relatives 49) Specificity 37% Eyes open (ADHD vs relatives 46.41, Control vs relatives 63.1), Eyes closed (ADHD vs relatives 57.79, Controls vs relatives 63.1) PPV NPV LR+ LR- Accuracy AUC 0.575 Eyes open (ADHD vs relatives 0.578, Control vs relatives 0.548), Eyes closed (ADHD vs relatives 0.669, Controls vs relatives 0.617) **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
			Cost: N/A **Admin time:** 6 minutes	
Kim, 2021{#454} N = 79 Korea Specialty care	**Target:** Participants with ADHD (DSM-V) from the Department of Psychiatry in a hospital, all were drug-naïve **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Healthy controls with no history of disorders **Female:** 17.6% **Age:** 24.76 (7.02) Min age: 18 Max age: 45 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** Mismatch negativity sensor level plus source level, an event-related potential component representing pre-attentive auditory processing closely associated with cognitive status assessed via EEG (electroencephalography); source localization was performed using standardized low-resolution brain electromagnetic tomography to estimate cortical distributions of mismatch negativity activity in the frontal, temporal, and limbic lobes Machine learning: Yes Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on full DSM-V criteria evaluated by a board-certified psychiatrist specializing in adult ADHD **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The best classification performance showed an 81.01% accuracy, 82.35% sensitivity, and 80.00% specificity based on source activity features; results suggest that abnormal mismatch negativity reflects the adult ADHD patients' pathophysiological characteristics and might serve clinically as a neuromarker of adult ADHD. Sensitivity 82% Specificity 80% PPV NPV LR+ LR- Accuracy 81 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** About 10 minutes	**Subgroup analysis:** N/A
Mueller, 2011{#640} N = 167 Multiple countries Setting varies	**Target:** Adults aged 20-50 diagnosed with ADHD based on DSM-IV criteria, including combined, inattentive, or hyperactive-impulsive subtypes, unmedicated or off methylphenidate for 24 hours before testing, no history of neurological or systemic medical diseases, no psychosis symptoms, no significant head injuries	**Test description:** Event-related potentials recorded while participants performed a visual two-stimulus go/no-go task Machine learning: Yes Validation dataset: Yes **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** A classification accuracy of 91% using a 10-fold cross-validation approach to differentiate adult ADHD patients from controls based on independent ERP components. The predictive power of the SVM was validated with an independent ADHD	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
	ADHD presentation: inattentive : 56,hyperactive : 12,combined : 32 **Comorbidity:** N/A **Other:** Age- and sex-matched neurotypical adults from the community with no ADHD diagnosis, no significant head injuries, no neurological or systemic medical diseases, scoring below clinical significance on the Brief Symptom Inventory, and not receiving medication **Female:** 49.33% **Age:** 36.05 (8.42) Min age: 20 Max age: 50 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** Public funding	Diagnosed with ADHD based on DSM-IV criteria assessed through a structured clinical interview conducted by trained psychologists **Diagnosed by:** Specialist (e.g., mental health) Psychologists **Timing:** Concurrent	sample, achieving a classification accuracy of 94%. Sensitivity 91% Specificity 91% PPV NPV LR+ LR- Accuracy 91 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** N/A	
Mueller, 2020{#642} N = 328 Switzerland Setting varies	**Target:** Adults diagnosed with ADHD based on DSM-5 criteria, recruited via media advertisements, local psychiatrists, and ADHD associations, excluding those with IQ <80, neuropsychological performance quotient <75, history of brain injury requiring rehabilitation, epilepsy, primary mental disorders other than ADHD, or insufficient knowledge of German or French. **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical adults recruited through media, schools, companies, and associations, excluding those with psychiatric diagnoses or histories of psychotropic medication intake. **Female:** 49.72% **Age:** 34.54 (10.16) Min age: 18 Max age: 60	**Test description:** EEG/ERP measures to capture brain activity patterns analyzed using a machine-learning framework, incorporating spectral power, event-related potential amplitudes, and latencies Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria verified by a psychiatric specialist through clinical interviews, ADHD screening questionnaires, and structured diagnostic assessments.	**Diagnostic accuracy summary:** ADHD patients and healthy controls could be classified with a sensitivity of 75% to 83% and a specificity of 71% to 77%. In the analysis of the repeated measurements, sensitivity values of the selected logistic regression model remained high (72% and 76%), while specificity values slightly decreased over time (64% and 67%). Sensitivity 75% after 12 months: 72; after 24 months: 76 Specificity 77% after 12 months: 64; after 24 month: 67 PPV NPV LR+ LR- Accuracy	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
	Age subgroup: Adults **Ethnicity:** N/A Multicenter **Funding:** Other	**Diagnosed by:** Specialist (e.g., mental health) Psychiatric specialists **Timing:** Concurrent	AUC 0.84 after 12 months: 0.68; after 24 months: 0.72 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** Measurment 1 or 2 years later 0.623 CI (0.560, 0.683), good consistency of classification performance over time **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** 26 minutes	
Poil, 2014{#722} N = 49 Switzerland Specialty care	**Target:** Adults diagnosed with ADHD using clinical interviews and standardized diagnostic tools **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical individuals recruited from personal contacts and public science presentations, with no current or past neurological or psychiatric diagnoses, matched for demographic variables and IQ to the ADHD group **Female:** 55% 37.5 in larger group **Age:** 37.9 (11.3) Min age: Max age: 61 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Public funding	**Test description:** Resting-state EEG was recorded during a 2.5-minute eyes-closed session using 60 scalp electrode positions, analyzing spectral power and central frequency across delta, theta, alpha-1, alpha-2, beta, and gamma frequency bands to identify diagnostic biomarkers for ADHD. Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical interviews conducted by experienced psychiatrists for adults and the Kiddie-SADS-PL for children following standardized diagnostic criteria **Diagnosed by:** Specialist (e.g., mental health) psychiatrists **Timing:** Concurrent	**Diagnostic accuracy summary:** Support vector machine classification of ADHD adults versus controls yielded a notable cross validated sensitivity of 67% and specificity of 83% using power and central frequency from all frequency bands. Sensitivity 67% Specificity 83% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
			Side effects N/A **Cost:** N/A **Admin time:** 2.5 minutes	
Ponomarev, 2014{#729} N = 472 Multiple countries College	**Target:** Adults aged 20–50 with symptoms meeting modified DSM-IV criteria for ADHD (at least 4 inattention and/or hyperactivity/impulsivity symptoms in childhood and the past 6 months), no head injury or neurological/systemic medical diseases, mostly unmedicated, with 63 meeting full DSM-IV criteria and 33 classified as subclinical **ADHD presentation:** inattentive : 23.96,hyperactive : 7.29,combined : 68.75 **Comorbidity:** N/A **Other:** Neurotypical adults recruited from university students, research staff, and general community members, with no neurological or psychiatric conditions, average or better academic performance, no current medication or substance use, and normal mental and physical development **Female:** 48% **Age:** 36.4 (8.36) Min age: 20 Max age: 50 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** Other	**Test description:** EEG (electroencephalography) with group independent component analysis and current source density Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed based on DSM-IV criteria assessed through clinical interviews and ADHD questionnaires conducted by an independent psychiatrist, including retrospective recall of childhood symptoms and current symptomatology **Diagnosed by:** Specialist (e.g., mental health) psychiatrist **Timing:** Concurrent	**Diagnostic accuracy summary:** Spectral power of local EEG activity isolated by gICA or CSD in the fronto-central areas may be a suitable marker for discrimination of ADHD and healthy adults. Sensitivity 94% Specificity 90% PPV NPV LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Shahaf, 2012{#840} N = 26 Israel Specialty care	**Target:** Adults diagnosed with the combined subtype of ADHD based on DSM-IV criteria, aged-matched and gender-matched, right-handed, with normal hearing and vision or corrected-to-normal vision, screened to exclude co-morbid disorders such as depression, anxiety, substance abuse, or learning disabilities, with 24-hour medication washout for those receiving methylphenidate therapy	**Test description:** Brain network activation utilizes EEG-based neurophysiological markers to analyze event-related potentials (ERPs) to identify patterns of brain activity associated with ADHD Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** The ADHD group was more characterized by the process of exerting attention in the early monitoring stages of the No-go signal, while the controls were more characterized by the process of inhibiting the response to that signal. Sensitivity 84% Specificity 92% PPV 91.67 computed	**Subgroup analysis:** N/A

Study ID	Population	EEG as Index Test	Results	Subgroup
	ADHD presentation: N/A **Comorbidity:** N/A **Other:** Neurotypical adults without ADHD recruited as student volunteers from the Institute of Technology, who underwent comprehensive neurological and neuropsychological evaluation, matching the ADHD group in age and gender **Female:** % N/A **Age:** 29.2 (6.1) Min age: 18 Max age: 39 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	Diagnosed with ADHD based on DSM-IV criteria using clinical interviews, confirmed by fulfilling ADHD symptoms on the Conners Adult ADHD Rating Scales and excluding co-morbid disorders through comprehensive neurological and neuropsychological evaluation **Diagnosed by:** Specialist (e.g., mental health) Specialists in the Neuro-Cognitive Unit at Rambam Health Care Campus **Timing:** Concurrent	NPV 85.71 computed LR+ 11 computed LR- 0.17 computed Accuracy 88 computed AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects** N/A **Cost:** N/A **Admin time:** N/A	

Appendix Table A7: Evidence Table Biomarker as Index Test

Study ID	Population	Biomarker Index Test	Results	Subgroup
Andrikopoulos, 2024{#5003} N = 76 Greece Specialty care	**Target:** Adults diagnosed with ADHD based on DSM-5 criteria, aged 18 years or older, IQ above 70, proficient in Greek, willing and able to provide informed consent, without major psychiatric disorders, significant neurological conditions, or severe learning disabilities **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical adults without an ADHD diagnosis, recruited from the same specialty care setting, meeting the same inclusion criteria as the ADHD group except for the diagnosis, including being aged 18 years or older, IQ above 70, proficient in Greek, and without major psychiatric disorders, significant neurological conditions, or severe learning disabilities **Female:** 34.38% **Age:** 33.26 (12.18) Min age: 18 Max age: 59 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** Physiological data, including electrodermal activity, heart rate variability, and skin temperature, using a wrist-worn wearable device during neuropsychological evaluations; biomarkers were analyzed using machine learning algorithms Machine learning: Yes Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria through a semi-structured Diagnostic Interview for ADHD in Adults (DIVA) conducted by an experienced psychiatrist, complemented by a psychiatric examination and additional information from relatives. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** Results indicate that the SVM-based model yielded the optimal performance, achieving 81.6% accuracy, maintaining a balance between the experimental and control groups, with sensitivity and specificity of 81.4% and 81.9%, respectively Sensitivity 81% Specificity 82% PPV NPV LR+ LR- Accuracy 82 AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Grünblatt, 2012 {#327} N = 143 Germany College	**Target:** Adults diagnosed with ADHD recruited from outpatient clinics at the university's psychiatry department **ADHD presentation:** inattentive : 21.3,hyperactive : 5.5,combined : 70.4 **Comorbidity:** N/A **Other:** Control participants recruited from newspaper ad **Female:** 45.7% **Age:** 34.7 (1.61) non-ADHD 39.6 (9.49)	**Test description:** Gene expression levels of 4 ADHD-associated genes—SLC6A3, DRD5, TPH1, and SNAP25—in peripheral blood, cut-off point 0.69 Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** Combining the gene expression levels of SLC6A3, DRD5, TPH1, and SNAP25 as predictors in a regression model resulted in sensitivity and specificity of over 80% (ROC: max R^2 0.587, AUC 0.917, $p < 0.001$, 95% CI 0.900–0.985), distinguishing adult ADHD from healthy controls. Sensitivity 81% SLC6A3 70%, DRD5 75%, SNAP25 64, TPH1 78 Specificity 82% SLC6A3 65%, DRD5 63%, SNAP25 62%, TPH1 71%	**Subgroup analysis:** Age,Sex Odds ratio 0.95 (CI 0.92-0.99), p 0.018, AUC 0.646 in predicting ADHD. Odds ratio 0.63 (CI 0.29–1.35), p 0.229,

Study ID	Population	Biomarker Index Test	Results	Subgroup
	Min age: 24 Max age: 50 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	Diagnosed with ADHD by team of psychiatrists through retrospective assessment using DSM-IV during structured clinical interview **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	PPV NPV LR+ LR- Accuracy AUC 0.917 0.9-0.985 SLC6A3 0.694, DRD5 0.749, SNAP25 0.689, TPH1 0.812 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	AUC 0.558 in predicting ADHD.
Jimenez, 2021{#417} N = 144 Multiple countries Specialty care	**Target:** Adults aged 18-65 diagnosed with ADHD without mental retardation, fluent in Spanish or English, no history of head injury or neurological illness, assessed with DSM criteria and confirmed through psychiatric and psychological evaluations **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Age- and sex-matched clinical participants with conduct disorder diagnoses, recruited from specialty care clinics and hospitals in Spain and the UK **Female:** 37.5% **Age:** 29.4 (12.4) Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter	**Test description:** Eye tracker, measuring modulation in the angle of eye vergence during an attention task using the BGaze eye-tracking system; participants were to maintain fixation on a central point while responding to visual stimuli, and the vergence angle (convergence or divergence of the eyes) was calculated using gaze vector data; a random forest classifier analyzed signals based on differential patterns in their eye vergence responses Machine learning: Yes Validation dataset: Partially	**Diagnostic accuracy summary:** Eye Vergence Responses showed a diagnostic accuracy of 79%, with an AUC of 0.77, a false positive rate of 25%, and a false negative rate of 20.55%. Sensitivity 80% Specificity 83% PPV 4.8 NPV 56 LR+ 72.92 LR- 0.24 Accuracy 72.92 AUC 0.77 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A	**Subgroup analysis:** N/A

Study ID	Population	Biomarker Index Test	Results	Subgroup
	Funding: Public funding	**Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Diagnostic and Statistical Manual of Mental Disorders criteria through psychiatric and psychological evaluations, including a semistructured interview, assessment of symptom onset before 12 years of age, and persistence of dysfunction in at least two settings **Diagnosed by:** Specialist (e.g., mental health) Clinical psychiatrists and psychologists **Timing:** Concurrent	**Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** Approximately 15 minutes	
Selek, 2012{#832} N = 87 Turkey Specialty care	**Target:** Adults aged 18–45 years diagnosed with ADHD using Turgay's Turkish version of the DSM-IV Adult ADD/ADHD Diagnostic Screening and Rating Scale, free from stimulant or ADHD medications, without severe organic conditions, epilepsy, infectious diseases, excessive obesity, or use of antioxidant agents, and scoring below 2 on the Clinical Global Impression-Severity Scale. **ADHD presentation:** inattentive : 34,hyperactive : 16,combined : 38,combined_other : 12 **Comorbidity:** N/A **Other:** Non-ADHD participants were healthy adults from the same hospital, including doctors and staff, free of medication for at least six weeks, and without a history or family history of psychiatric disorders. **Female:** 30% **Age:** 24.7 (7.5)	**Test description:** Blood total oxidative status levels above 9.8575 mmol H_2O_2 Eqv./L Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on Turgay's Turkish version of the DSM-IV Adult ADD/ADHD Diagnostic Screening and Rating Scale conducted by two psychiatrists. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** ADHD can be predicted for TOS over 9.8575 mmol H2O2 Eqv./L level with 86% positive predictive value and 100% negative predictive value. Sensitivity % Specificity % PPV 86 NPV 100 LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A	**Subgroup analysis:** Age The study reported a positive correlation between age and oxidative biomarkers (TOS and OSI) in ADHD patients, suggesting that oxidative stress may increase with the duration of the disease, but this correlation was not observed in the control group.

Study ID	Population	Biomarker Index Test	Results	Subgroup
	Min age: 18 Max age: 45 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other		**Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Udal, 2024{#5685} N = 115 Norway Specialty care	**Target:** Adults referred to a psychiatric outpatient clinic for diagnostic assessment, excluding those with schizophrenia, psychotic disorders, ongoing drug abuse, rheumatic, orthopedic, or neurological disorders, or medications affecting motor function **ADHD presentation:** inattentive : 34.1,combined : 65.0 **Comorbidity:** N/A **Other:** Adults from a psychiatric outpatient clinic presenting with various psychiatric diagnoses or subthreshold ADHD symptoms but not meeting full diagnostic criteria for ADHD **Female:** 59.3% **Age:** 33.0 (9.9) Min age: 18 Max age: 66 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** MFNU (Motor Function Neurological Assessment), neuromuscular assessment, assessing neuroromuscular dysregulation, analyzing maximum summed problem score Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Mini International Neuropsychiatric Interview (MINI-plus) and/or the Diagnostic Interview for ADHD in Adults 2.0 (DIVA-2.0) following structured clinical interviews conducted by a physician or clinical psychologist. **Diagnosed by:** Specialist (e.g., mental health) Physician or clinical psychologist **Timing:** Concurrent	**Diagnostic accuracy summary:** A MFNU-TS cut-off score of 13.5 yielded a near 98% sensitivity for ADHD diagnosis, both when including and excluding those with subthreshold ADHD symptoms. Sensitivity 98% 98% when excluding subthreshold from control group Specificity 25% 77% when excluding subthreshold ADHD from control group PPV NPV LR+ LR- Accuracy Youden index 0.23 (0.74 when excluding subthreshold ADHD from control group) AUC 0.66 AUC 0.90 when excluding subthreshold ADHD from control group **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** ADHD diagnosis (effect of different reference status) Participants with subthreshold ADHD symptoms had MFNU scores similar to the ADHD group, suggesting possible diagnostic overlap and the need for further evaluation in these cases. No analysis was conducted to assess outcomes like functional impairment, long-term effects, or treatment acceptability based on clinical ADHD status.

Appendix Table A8: Evidence Table Clinician Tools as Index Test

Study ID	Population	Clinician Tools Index Test	Results	Subgroup
Kingston, 2013{#5349} N = 120 Canada Specialty care	**Target:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **ADHD presentation:** N/A **Comorbidity:** Other : Aggression dysregulation **Other:** Men who were assessed at an outpatient forensic psychiatric clinic; individuals are typically referred to this program when they are engaging in aggression or other difficulties associated with anger dysregulation (e.g., relationship breakdown) **Female:** 0% **Age:** 32.6 (10.3) Min age: 18 Max age: 64 **Age subgroup**: Adults **Ethnicity:** Other : Aboriginal: 6.5% % Hispanic or Latino : 2.8 % Black/African American : 2.8 % White : 78.5 Single center **Funding:** Industry	**Test description:** CAARS-O ADHD Index (Observer), a 66-item measure that contains 9 empirically-derived scales related to adult ADHD symptoms completed by psychiatrist Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis ADHD diagnosis was determined based on DSM-IV-TR criteria following a comprehensive clinical interview and review of relevant available collateral information; interviews were conducted independently by two psychiatrists who were certified in forensic psychiatric practice; final group classification was based on consensus diagnoses and the inter-rater agreement was approximately 90% **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The integrated variables of multiple self reports and an observer report demonstrated particularly good classification accuracy, with high sensitivity (91%) and good specificity (82%). Sensitivity 76% (CI 61, 86) Specificity 75% (CI 57, 87) PPV 80 (CI 66, 90) NPV 69 (CI 52, 82) LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Kumar, 2011{#4185} N = 110 US Specialty care	**Target:** Adults recruited from psychiatric inpatient unit of a general hospital with a chart diagnosis of ADHD **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults with different mental disorders recruited from psychiatric inpatient unit of a general hospital	**Test description:** MINI (International Neuropsychiatric Interview), a short, structured diagnostic interview designed to assess a range of different mental health disorders Machine learning: No Validation dataset: N/A	**Diagnostic accuracy summary:** The CAARS-S–S: SV indicated adequate disrimination. The MINI ADHD module was most effective for identifying inpatients without ADHD. Sensitivity 83% (CI 36, 100) Specificity 52% (CI 42, 62) PPV 9 (CI 3, 20) NPV 98 (CI 90, 100) LR+	**Subgroup analysis:** Age,Sex ADHD diagnosis based on CAARS-S or MINI were not correlated with age.

Study ID	Population	Clinician Tools Index Test	Results	Subgroup
	Female: 50% **Age:** 36.6 (11.1) Min age: 25 Max age: 49 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 8 % Black/African American : 16 % White : 64 % Multiracial : 12,Other : other ethnic backgrounds Single center **Funding:** Unclear	**Reference standard:** Clinical diagnosis Chart diagnosis, diagnosed with ADHD by board certified psychiatrists after inpatient admission through DSM-IV-TR **Diagnosed by:** Specialist (e.g., mental health) Psychiatrist **Timing:** Prior diagnosis	LR- Accuracy 54 (CI 44, 63) AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 10-25 minutes.	ADHD diagnosis based on CAARS-S or MINI were not correlated with sex.
Palma-Alvarez, 2023{#684} N = 1263 Multiple countries Setting varies	**Target:** Adults aged 18–65 years starting a new treatment episode in addiction treatment centers, screened for ADHD using the MINI-Plus ADHD module, with no severe cognitive impairment, substance intoxication, acute psychiatric crisis, or severe somatic problems, and who provided informed consent **ADHD presentation:** N/A **Comorbidity:** SUD : seeking treatment for SUD **Other:** Adults in addiction treatment centers who did not meet ADHD criteria based on the CAADID, with similar inclusion settings focusing on treatment for substance use disorders **Female:** 26.5% **Age:** mean 39.98 Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** Other : 9.7	**Test description:** MINI-Plus (Mini International Neuropsychiatric Interview), a structured diagnostic interview designed to assess psychiatric disorders, including ADHD, based on DSM-IV and ICD-10 criteria targeting core symptoms of inattention and hyperactivity-impulsivity, without differentiating ADHD subtypes Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on the Conners' Adult ADHD Diagnostic Interview for DSM-IV conducted by trained clinicians **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** Sensitivity of the MINI-Plus ADHD module was 74%, specificity was 91%. Sensitivity 75% (CI 68, 80) Specificity 91% (CI 90, 93) PPV 60 (CI 52, 65) NPV 95.6 (CI 95, 97) LR+ LR- Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Clinician Tools Index Test	Results	Subgroup
	% White : 90.3 Multicenter **Funding:** Public funding		**Admin time:** N/A	

Appendix Table A8: Evidence Table Feigning ADHD

Study ID	Population	Feigning ADHD	Results	Subgroup
Abramson, 2023{#5} N = 242 US College	**Target:** Adult patients referred for neuropsychological evaluation at an academic medical center from 2018 to 2021, with clinical diagnosis of ADHD based on the comprehensive protocol of the study **ADHD presentation:** inattentive : 45,combined : 55 **Comorbidity:** N/A **Other:** Adult patients referred for neuropsychological evaluation at an academic medical center from 2018 to 2021, constituting the invalid group (failed two or more criterion measures of the performance validity test) **Female:** 58% **Age:** 27.47 (6.89) Min age: 21 Max age: 35 **Age subgroup**: Adults **Ethnicity:** Other : 7% were other race/ethnicity % Hispanic or Latino : 22 % Black/African American : 12 % Asian : 9 % White : 49 Single center **Funding:** Other	**Test description:** DCT (Dot Counting Test), a freestanding performance validity test Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis (1) a full medical/psychiatric record review (including review of prior ADHD evaluations/diagnostic work ups, when available); (2) a semistructured clinical interview which systematically gathered all relevant background information (e.g., ADHD symptom onset/course and associated functional impairment; medical, psychiatric, substance use, developmental, academic, and psychosocial history) and thoroughly assessed formal DSM-5 ADHD diagnostic criteria as well as comorbid psychopathology; (3) administration of an ADHD symptom inventory (i.e., Clinical Assessment of Attention Deficit—Adult [CAT-A]), which contains embedded symptom validity scales to identify noncredible symptom reporting and provides objective, normative-based qualification of ADHD symptomatology in both childhood and adulthood; (4) administration of a standardized core neuropsychological test battery which comprehensively assessed examinees' cognition across all major cognitive domains; and (5) administration of a validity-controlled inventory of personality and psychopathology (i.e., Minnesota Multiphasic Personality Inventory-2-Restructured Form [MMPI-2-RF]) to	**Diagnostic accuracy summary:** Classification accuracy was excellent, with 54.3% sensitivity and 92% specificity at optimal cut-scores of ≥14 (rounded) and ≥13.38 (unrounded) Sensitivity 54% Specificity 92% PPV NPV +LR -LR Accuracy AUC 0.843 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** Brief administration time and scoring procedures.	**Subgroup analysis:** ADHD presentation,Comorbidity (e.g. anxiety, depression) A series of ANOVAs revealed nearly identical test performance between ADHD subtypes (i.e., predominately inattentive vs combined), suggesting that these clinical factors did not meaningfully affect DCT performance. A series of ANOVAs revealed nearly identical test performance between the presence/absence of comorbid psychopathology, suggesting that these clinical factors did not meaningfully affect DCT performance.

Study ID	Population	Feigning ADHD	Results	Subgroup
		objectively assess for active comorbid psychological symptoms **Diagnosed by:** Unclear/NR **Timing:** Prior diagnosis		
Aita, 2018{#16} N = 280 US Specialty care	**Target:** Individuals from one of two university-affiliated psychology training clinics, diagnosed with ADHD **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Mood/Anxiety Disorder group or Clinic No Diagnosis group: Individuals from one of two university-affiliated psychology training clinics, not diagnosed with ADHD Control group or ADHD Simulator group: Students were prospectively recruited from three southeastern universities **Female:** 45.1% Study 1 - ADHD group: 45.1%; ADHD Simulators group: 73.9%; Mood/Anxiety Disorder group: 65.0%; Clinic No Diagnosis group: 42.3%; Healthy Controls group: 69.2%; Study 2 - ADHD group: 43.8%; ADHD Simulators group: 73.9%; Mood/Anxiety Disorder group: 65.4%; Clinic No Diagnosis group: 37.8%; Healthy Controls group: 75.5% **Age:** 20.29 (1.87) Study 1 - ADHD group: 21.77 (3.99); ADHD Simulators group: 19.83 (1.54); Mood/Anxiety Disorder group: 22.71 (4.58); Clinic No Diagnosis group: 22.05 (5.07); Healthy Controls group: 19.18 (1.57) Study 2 - ADHD group: 22.33 (3.93); ADHD Simulators group: 19.83 (1.54); Mood/Anxiety Disorder group: 21.98 (4.26); Clinic No Diagnosis group: 22.80 (5.13); Healthy Controls group: 19.45 (1.35) Min age: 18 Max age: 25	**Test description:** PAI (Personality Assessment Inventory), a self-report personality measure comprised of 344 items on a 4-point scale with anchor points of false and very true; items are categorized into 4 scales that assess validity of responding, 11 clinical syndrome scales, 5 treatment scales, and 2 interpersonal scales Machine learning: No Validation dataset: Yes **Reference standard:** Clinical diagnosis All evaluations were conducted by doctoral graduate students in a clinical psychology program. Evaluations included a thorough clinical interview and all diagnoses were made under the supervision of a licensed psychologist. **Diagnosed by:** Researcher Doctoral graduate students in a clinical psychology program, under supervision of a licensed psychologist **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The new index's classification accuracy was superior to most existing PAI validity scales across groups. An item-level PAI algorithm had a sensitivity of 85% and specificity of 97% for identifying feigned ADHD. Sensitivity 46% Specificity % PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Age subgroup: Adults **Ethnicity:** Other : Other Race: Study 1 - ADHD Simulators group: 4.3%; Clinic No Diagnosis group: 0.9%; Healthy Controls group: 3.8%; Study 2 - ADHD Simulators group: 4.3%; Clinic No Diagnosis group: 2.2%; Healthy Controls group: 3.8% Other : Study 1 - ADHD group: 5.8%; ADHD Simulators group: 7.2%; Mood/Anxiety Disorder group: 1.5%; Clinic No Diagnosis group: 1.8%; Healthy Controls group: 3.8%; Study 2 - ADHD group: 9.6%; ADHD Simulators group: 7.2%; Healthy Controls group: 5.7% Other : Study 1 - ADHD group: 10.1%; ADHD Simulators group: 10.1%; Mood/Anxiety Disorder group: 8.8%; Clinic No Diagnosis group: 10.8%; Healthy Controls group: 24.1%; Study 2 - ADHD group: 9.6%; ADHD Simulators group: 10.1%; Mood/Anxiety Disorder group: 5.8%; Clinic No Diagnosis group: 8.9%; Healthy Controls group: 9.4% Other : Study 1 - ADHD group: 1.4%; ADHD Simulators group: 5.8%; Mood/Anxiety Disorder group: 2.2%; Clinic No Diagnosis group: 1.8%; Healthy Controls group: 3.8%; Study 2 - ADHD group: 2.7%; ADHD Simulators: 5.8%; Mood/Anxiety Disorder group: 1.9%; Clinic No Diagnosis group: 4.4%; Healthy Controls group: 1.9% Other : Study 1 - ADHD group: 82.7%; ADHD Simulators group: 72.5%; Mood/Anxiety Disorder group: 87.6%; Clinic No Diagnosis group: 84.7%; Healthy Controls group: 64.7%; Study 2 - ADHD group: 75.3%; ADHD Simulators: 72.5%; Mood/Anxiety			

Study ID	Population	Feigning ADHD	Results	Subgroup
	Disorder group: 92.3%; Clinic No Diagnosis group: 84.4%; Healthy Controls: 79.2% Multicenter **Funding:** Other			
Becke, 2023{#58} N = 117 Netherlands Setting varies	**Target:** Individuals with suspected ADHD referred to the Department of Psychiatry and Psychotherapy for clinical evaluation, clinical interviews, gathering corroborating evidence of ADHD-realted impairments via asking parents, partners and/or employees **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Students who reported low levels of ADHD symptoms currently and retrospectively, differing significantly from the ADHD group; in addition to a group randomly assigned to 3 simulation instructions (general instructions to feign ADHD with no additional information, symptom-coached simulators who were given the DSM diagnostic criteria of ADHD, and fully coached simulators who received information on both the neuropsychological assessment of ADHD and its diagnostic criteria **Female:** 33% **Age:** 32 (12) Min age: 20 Max age: 44 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** WAFS CE (Perceptual and Attention Functions selective attention assessed with VTS (Vienna Test System), a computerized neuropsychological test battery for assessing cognitive functions in adult ADHD evaluates cognitive domains such as attention and executive functions, aiding in diagnosing ADHD and treatment planning Machine learning: Yes Validation dataset: Yes **Reference standard:** Other Diagnosed with ADHD based on clinical interviews conducted by two experienced professionals using the Diagnostic and Statistical Manual of Mental Disorders criteria, with corroborating evidence gathered from parents, partners, and employers when available **Diagnosed by:** Researcher **Timing:** Later diagnosis	**Diagnostic accuracy summary:** Although all ensured at least 90% specificity in the ADHD Group, sensitivity differed significantly between tests, ranging from 0% to 64.9%. Tests of selective attention, vigilance, and inhibition were most useful in detecting the instructed simulation of adult ADHD, whereas figural fluency and task switching lacked sensitivity. Sensitivity 65% Specificity 91% PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Berger, 2021{#68} N = 189	**Target:** Undergraduate students with a diagnosis of ADHD confirmed using the Structured Clinical Interview for DSM-5 (SCID-5) excluded if they had neurological or psychiatric disorders	**Test description:** MOXO-d-CPT Machine learning: No Validation dataset: Yes	**Diagnostic accuracy summary:** Simulators performed significantly worse on all MOXO-d-CPT indices than healthy controls and ADHD patients. Three MOXO-d-CPT indices (attention, hyperactivity, impulsivity) and a	**Subgroup analysis:** ADHD diagnosis (effect of different reference status) No significant differences were found between

Study ID	Population	Feigning ADHD	Results	Subgroup
Israel Setting varies	**ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Healthy controls who feigned ADHD and histoic healthy controls **Female:** 63.83% **Age:** 23.79 (2.17) Min age: 18 Max age: 65 **Age subgroup**: Adults **Ethnicity:** N/A Multicenter **Funding:** No COI	**Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a previous diagnosis by a licensed clinician (psychiatrist and/or clinical psychologist) following DSM-5 criteria, confirmed using the Structured Clinical Interview for DSM-5 (SCID-5-RV) upon study entry **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	scale combining these indices showed adequate discriminative capacity. Sensitivity 62% Specificity 91% PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** 18.2 minutes	archival and prospective data of ADHD patients.
Cook, 2016{#157} N = 86 US College	**Target:** Adults aged 18 and older referred to a university psychology training clinic for neuropsychological evaluation for concerns about ADHD and/or a learning disability, with exclusion criteria including any self-reported history of neurological illness or injury **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Adults without ADHD, referred to a university psychology training clinic for evaluation of mood disorders, anxiety disorders, adjustment disorders, substance abuse disorders, learning disabilities, or schizotypal personality disorders, with the setting	**Test description:** CII (Conner's Adult Attention Deficit/Hyperactivity Rating Scale Infrequency Index) to identify non-credible symptom reporting Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive neuropsychological evaluation conducted by advanced graduate students under the supervision of a licensed psychologist in a university psychology training clinic **Diagnosed by:** Specialist (e.g., mental health) licensed psychologist **Timing:** Concurrent	**Diagnostic accuracy summary:** The CII was 52% sensitive to extreme scores on CAARS DSM symptom subscales (with 97% specificity) and 20%-36% sensitive to invalid responding on MMPI-2-RF validity scales (with near 90% specificity), providing further evidence for the interpretation of the CII as an indicator of non-credible ADHD symptom report. However, the CII detected only 18% of individuals who failed a standalone performance validity test (WordMemoryTest), with 87.8% specificity, and was not accurate in detecting non-credible performance using embedded digit span cutoffs. Sensitivity % Specificity % PPV NPV +LR	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	being a specialty care clinic focused on neuropsychological assessment Participants with non-credible performance were identified with a Word Memory Test **Female:** 53% **Age:** 22 (5) Min age: 18 Max age: 42 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 2.3 % Black/African American : 3.5 % Asian : 3.5 % White : 79.1 Single center **Funding:** No COI		-LR Accuracy AUC 0.87 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Courrege, 2019{#164} N = 402 US N/A	**Target:** Participants with ADHD were included based on self-reported history of ADHD diagnosis, diagnosis details such as the type of professional providing the assessment, methods used in the assessment (interviews, symptom questionnaires, cognitive testing), and whether medication was prescribed **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants included control individuals without ADHD, individuals who suspected they had ADHD but were not diagnosed, and simulators instructed to feign ADHD symptoms; participants were recruited online through Amazon Mechanical Turk and completed assessments in a general, non-clinical setting **Female:** 60.37% **Age:** 36 (13.0) Min age: 19 Max age: 75	**Test description:** ASIS INF (Infrequency Scale), tool to assess self-reported ADHD symptoms and identify exaggeration or feigned responses, scales map onto DSM-5 criteria and items designed to detect symptom infrequency Machine learning: No Validation dataset: Partially **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on self-reported history, details of prior assessments including the type of professional conducting the evaluation, methods such as interviews, symptom questionnaires, cognitive testing, and whether medication was prescribed. **Diagnosed by:** Specialist (e.g., mental health) psychiatrists, physicians, and psychologists **Timing:** Concurrent	**Diagnostic accuracy summary:** Demonstrated strong sensitivity (.79–.86) and excellent specificity (.89) in detecting feigned ADHD symptoms compared to a sample of individuals self-reporting a history of ADHD diagnosis. Using a malingering base rate of 29%, the ASIS INF scale achieved a positive predictive value of .71–.79 and a negative predictive value of .92–.93, indicating strong diagnostic accuracy in differentiating simulated from genuine ADHD. Sensitivity 79% Specificity 89% PPV 71 NPV 92 +LR 7.18 -LR 0.24 Accuracy 79 AUC 0.92 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha 0.96	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Age subgroup: Adults **Ethnicity:** % Hispanic or Latino : 4 % Black/African American : 5.3 % American Indian or Alaska Native : 1.3 % Asian : 5.3 % White : 81.5 % Multiracial : 2.6 Single center **Funding:** Unclear		**Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Finley, 2023{#263} Finley, 2024{#264} N = 599 US Specialty care	**Target:** Adults referred to a Midwestern academic medical center for neuropsychological evaluation diagnosed with ADHD based on clinical interview, self-reported symptom questionnaires, and neurocognitive testing according to DSM criteria **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Individuals with psychiatric disorders such as depression, anxiety, PTSD, or no mental health diagnosis, evaluated in a specialty care setting; exclusions: intellectual disabilities, major neurocognitive disorders, severe mental illnesses, or invalid inconsistency scores to ensure accurate comparisons; participants were categorized into invalid and valid performance groups determined by scores from empirical perfoormance validity indicators **Female:** 62% **Age:** 28.12 (6.85) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 22	**Test description:** Integrating 6 indicators - Combination of WAIS-IV Symbol Search age-corrected scaled score (equal to smaller than 6), WAIS-IV Coding age-corrected scaled score (equal or smaler than6), WAIS-IV Letter-Number Sequencing age-corrected scaled score (equal or smaller than 7), SCWT Word Reading T-score (equal or smaller than 25), TMT-B T-score (equal or smaller than 34), and Lexical Fluency FAS T-score (equl or smaller than 34); cut-off failing 2; adminstered together with the NI (Negative Impression), IF (Infrequency), and PI (Positive Impression) scales Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical interview, review of medical and academic records, symptom questionnaires, and neurocognitive testing following DSM criteria. Comprehensive neuropsychological evaluation and symptom questionnaires were key components of the diagnostic process. Both studies	**Diagnostic accuracy summary:** AUC was 0.86 for the integrated neuropsychological test indicators. Self report results varied (Negative Impression scale ≤51; 30% sensitivity / 90% specificity; Infrequency scale ≥4; 18% sensitivity / 90% specificity; Positive Impression scale ≥27; 36% sensitivity / 90% specificity).{#264} Sensitivity 60% NI scale: 30%, IF scale: 18% Specificity 91% NI scale: 90%, IF scale: 90% PPV NPV +LR -LR Accuracy AUC 0.86 IF cale: 0.64 and 0.58 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	% Black/African American : 15 % Asian : 10 % White : 47 % Multiracial : 6 Single center **Funding:** Unclear	used a clinical diagnosis from mental health clinicians as the gold standard. **Diagnosed by:** Specialist (e.g., mental health) psychiatrists, neuropsychologists, and psychologists **Timing:** Concurrent	**Admin time:** N/A	
Fuermaier, 2016{#5208} N = 329 Germany Setting varies	**Target:** Adults diagnosed with ADHD referred by psychiatrists or neurologists meeting DSM-IV criteria confirmed by psychiatric interviews, scoring above cutoff on standardized self-report scales, and demonstrating objective impairments and multiple informant support for diagnosis **ADHD presentation:** inattentive : 43.21,hyperactive : 2,combined : 54.9 **Comorbidity:** N/A **Other:** Non-ADHD participants recruited through public announcements and word-of-mouth, selected to match the ADHD participants in age, gender, and intellectual functioning; in addition, undergraduate students were randomly assigned to a control group, a naive simulation group, a symptom-coached simulation group, or a test-coached simulation group **Female:** 41% **Age:** 34.0 (11.3) Min age: 18 Max age: 56 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Other	**Test description:** Embedded Figures Test developed for the detection of feigned ADHD in adulthood Machine learning: No Validation dataset: Yes **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical interviews according to DSM-IV criteria, including retrospective assessment of childhood symptoms and evidence from multiple informants such as employer and partner reports **Diagnosed by:** Specialist (e.g., mental health) Mental health specialists **Timing:** Concurrent	**Diagnostic accuracy summary:** The EFT (Embedded Figures Test) developed in the study demonstrated strong performance in distinguishing between individuals with genuine ADHD and those feigning ADHD. The test showed high sensitivity (88%) and specificity (90%). The EFT demonstrated excellent discriminatory power with an AUC of 0.948 Sensitivity 88% Specificity 90% PPV 89.8 NPV 88.2 +LR 8.8 -LR 0.13 Accuracy 89 AUC 0.948 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Harrison, 2007{#349} N = 142	**Target:** College or university students diagnosed using DSM-IV ADHD criteria with evidence of childhood and current impairments corroborated by collateral informants showing	**Test description:** Integrated CAARS, Reading Fluency subtest, and the 2 Processing Speed subtests from the Woodcock Johnson Psychoeducational Battery-III	**Diagnostic accuracy summary:** There was 75% correct classification across all groups. ADHD symptoms can be easily fabricated, with individuals feigning ADHD scoring higher on self-report measures (CAARS) and	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
Canada College	substantial academic or life impairments **ADHD presentation:** inattentive : 47.2,hyperactive : 52.8 **Comorbidity:** N/A **Other:** University undergraduates without ADHD, including a group instructed to simulate ADHD symptoms (Faking group) and a control group instructed to perform tasks honestly (Honest Normals) **Female:** 54.17% **Age:** 22.90 (7.01) Min age: 17 Max age: 22 **Age subgroup**: Young **Ethnicity:** Other : Student population self-identify as visible minorities Single center **Funding:** Public funding	Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-IV criteria, including objective evidence of childhood impairment, self-reported symptoms consistent with observed and documented behavioral problems, and confirmation from reliable collateral informants **Diagnosed by:** Specialist (e.g., mental health) Clinical psychologists **Timing:** Concurrent	performing worse on cognitive tests (WJPB-III) than genuine ADHD participants. Sensitivity % Specificity % PPV NPV +LR -LR Accuracy 75 normal group 80%, ADHD group 78%, faking group 66% correct classifications AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Harrison, 2016{#347} N = 608 Canada College	**Target:** Students (>17 years) diagnosed with ADHD by a clinical psychologist using DSM-IV, participants recruited from community colleges or universities **ADHD presentation:** inattentive : 24.6,hyperactive : 5.2,combined : 20.7,combined_other : exaggerating, feigning responses **Comorbidity:** N/A **Other:** Students seeking treatment diagnosed with other non-ADHD mental disorders eg anxiety, depression **Female:** 44.4% **Age:** 21.4 (4.5)	**Test description:** E-CAARS (Experimental Conners' Adult ADHD Rating Scale) added 18 items on atypical dissociative symptoms to capture exaggerated symptom responses, items were adapted to a 4-point CAARS format, cutoff >3 Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Students diagnosed with ADHD by a clinical psychologist following DSM-IV criteria **Diagnosed by:** Specialist (e.g., mental health) Psychologist	**Diagnostic accuracy summary:** While the tool demonstrated high specificity (97%), reducing false positives, its low sensitivity (24%) resulted in a significant number of false negatives, limiting its effectiveness in accurately identifying all true ADHD cases. Sensitivity 24% Specificity 97% PPV 0.58 NPV 0.88 +LR -LR Accuracy 86 correct classification rate AUC **Concordance:** N/A **Rater agreement:** Kappa ICC	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Min age: 17 Max age: 38 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	**Timing:** Prior diagnosis	**Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Harrison, 2019{#350} N = 331 Canada College	**Target:** Post-secondary students (age >17) from an archival database who sought assessment for ADHD as a possible cause of their reported difficulties, completed the PAI, were administered a PVT, and consented to their data being used for research **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Post-secondary students with no formal diagnosis served as no diagnosis group, students with primary mental health Non-ADHD condition served as clinical controls, students with definite malingering condition were in the malingering group **Female:** 36.9% **Age:** 21.9 (5.3) Min age: 17 Max age: 57 **Age subgroup**: Adults **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** PAI (Personality Assessment Inventory), a 344-item self-report inventory rated on a four-point Likert scale, providing clinicians with data on four validity scales, 11 clinical scales, five treatment scales, and two interpersonal scales to detect feigned ADHD Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on all five criteria listed in DSM-IV or DSM-5 depending on the date assessed, with assessments conducted by a licensed clinical psychologist or a graduate student trainee under supervision. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** The two proposed PAI algorithms were found to have poor positive predictive value (.19 and .17). Self-report validity measures from the Connors' Adult Attention Rating Scale, and the Negative Impression Management scale on the PAI returned more positive results. Sensitivity % Specificity % PPV 19 NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Harrison, 2020{#4042}	**Target:** Emerging adults referred to a university-based ADHD screening clinic, no prior ADHD diagnosis,	**Test description:** TOVA ACS score, Test of Variables of Attention as an embedded performance validity	**Diagnostic accuracy summary:** Of all TOVA and CAARS measures, the Attention Comparison Score had excellent	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
N = 245 Canada College	seeking evaluation due to self-reported difficulties, completed measures including TOVA and CAARS, provided evidence of substantial impairment in multiple major life activities prior to age 12 and currently **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Emerging adults referred to a university-based ADHD screening clinic, presenting with self-reported ADHD-like symptoms but not meeting diagnostic criteria for ADHD, included those with good effort scores on validity testing and no substantial impairments documented in major life activities **Female:** 49.4% **Age:** 20.4 (1.8) Min age: 17 Max age: 24 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Unclear	measure administered with CAARS measures Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive clinical assessment including a semi-structured interview, retrospective symptom ratings from parents/caregivers, review of childhood report cards, documentation of substantial impairment in major life activities prior to age 12 and currently, and evaluation of DSM criteria. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	discrimination; AUC was 0.797 (sensitivity 90%, specificity 47%). Commission Errors in the first half of the TOVA also showed good AUC and specificity but not sensitivity (AUC 0.818, sensitivity 12%, specificity 92%). Results support the use of the TOVA as an embedded performance validity measure in assessing late adolescents/emerging adults and support previous findings that symptom reports alone cannot distinguish credible from noncredible ADHD presentation. Sensitivity 47% Specificity 90% PPV NPV +LR -LR Accuracy AUC 0.797 **Concordance:** N/A **Rater agreement:** Rater agreement between self-reported ADHD symptoms (CAARS) and objective performance validity measures (TOVA) Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Khan, 2022{#447} N = 317 US	**Target:** Adults referred for outpatient neuropsychological evaluation for suspected or confirmed ADHD, reported English as their primary language, underwent a standardized diagnostic protocol including record review, clinical interview, and	**Test description:** SCWT (Stroop Color and Word Test) assesses cognitive flexibility and processing speed through 3 trials: word reading, color naming, and color-word interference; the word reading and color naming trials were used as	**Diagnostic accuracy summary:** The embedded validity indicators from the Stroop Color and Word Test were effective in determining validity status. Word Reading and Color Naming trials demonstrated acceptable classification accuracy (AUCs 0.750–0.794), with optimal cut scores of WR	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
Specialty care	neuropsychological testing, and were evaluated for ADHD using DSM-5 criteria **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants included adults referred for neuropsychological evaluation who failed performance validity tests, with evaluations conducted in a specialty care setting focused on diagnostic clarification for conditions other than ADHD **Female:** 62.46% **Age:** 27.7 (6.67) Min age: 18 Max age: 60 **Age subgroup**: Adults **Ethnicity:** Other : 5 % Black/African American : 24 % Asian : 10 % White : 46 Single center **Funding:** Unclear	embedded performance validity tests to evaluate the validity of neuropsychological test performance Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria by a board-certified clinical neuropsychologist **Diagnosed by:** Specialist (e.g., mental health) Clinical neuropsychologist **Timing:** Concurrent	raw ≤75 (54% sensitivity, 89-90% specificity), WR T score ≤28 (54% sensitivity, 87-88% specificity), CN raw ≤57 (42% sensitivity, 90% specificity), and CN T score ≤30 (40% sensitivity, 90% specificity). Sensitivity % Specificity % PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Lee Booksh, 2010{#4198} N = 166 US College	**Target:** Patients diagnosed with ADHD with clinical data obtained from previous archival records **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Undergraduate students (>18years) enrolled in psychology courses without ADHD or learning disabilities, screened for the absence of ADHD symptoms and neurological issues, were assigned to either control or simulated ADHD group **Female:** 70% **Age:** 21.11 (3.1) ADHD group age data collected retrospectively	**Test description:** WMT (Word Memory Test) evaluates verbal memory and effort by presenting participants with a list of 20 word pairs on a computer to learn Machine learning: No Validation dataset: N/A **Reference standard:** Clinical diagnosis Diagnosed with ADHD by a psycho-educational team in a university psychological clinic **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** Simulators successfully feigned ADHD symptoms on a retrospective self-report measure. Knowledge of ADHD was unrelated to objective attentional measure performance. Participants who simulated ADHD on some objective measures (i.e., specific Wechsler Adult Intelligence Scale–III [WAISIII] subtests) showed similar performance to the clinical ADHD comparison sample. Sensitivity % There are no established cut scores or validity measures specific to ADHD assessment that provide guidance on specificity or sensitivity values Specificity % There are no established cut scores or validity measures specific to ADHD	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Min age: 18 Max age: 29 **Age subgroup**: Adults **Ethnicity:** Other : This is the ethnicity data for ADHD group % Black/African American : 5.4 % Asian : 1.8 % White : 93 Single center **Funding:** Unclear		assessment that provide guidance on specificity or sensitivity values PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Marshall, 2010{#580} N = 268 US Specialty care	**Target:** Patients referred for ADHD assessment in a neuropsychological clinic, without neurological conditions, head injuries, learning disabilities, psychiatric disorders other than depression or anxiety, substance abuse dependence, or physical illnesses causing cognitive deficits, minimum estimated IQ of 70. **ADHD presentation:** inattentive : 45,combined : 36 **Comorbidity:** N/A **Other:** Patients referred for ADHD assessment in a neuropsychological clinic, many had other mental health conditions such as depression or anxiety, and it included patients suspected of exaggerating sypmptoms **Female:** 39% **Age:** 27.8 (9.1)	**Test description:** Word Memory test immediate recall Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on cognitive testing, behavior rating scales, and clinical interviews. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The Word Memory test immediate recall and consistency score (both 64%), TOVA omission errors (63%) and reaction time variability (54%), CAT-A infrequency scale (58%), and b Test (47%) had good sensitivity as well as at least 90% specificity. Sensitivity % Specificity % PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Min age: 17 Max age: 55 **Age subgroup**: Adults **Ethnicity:** Other : 10 % Hispanic or Latino : 10 % Black/African American : 9 % American Indian or Alaska Native : 10 % Asian : 10 % White : 81 Single center **Funding:** Unclear		**Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Morey, 2019{#4321} N = 368 US College	**Target:** Participants self-identified with ADHD, with 32 indicating a current diagnosis, 29 having received medication for ADHD, and 21 holding a current prescription **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants included neurotypical college students without a history of ADHD, divided into standard instruction and feigning instruction groups, recruited from an introductory psychology pool in a university setting **Female:** 56.3% **Age:** 18.9 (1.38) Min age: 18 Max age: 21 **Age subgroup**: Young **Ethnicity:** % Hispanic or Latino : 22.3 % Black/African American : 6.3 % Asian : 7.6 % White : 55.2 Single center **Funding:** Unclear	**Test description:** Integrating PAI (Personality Assessment Inventory) and TOAD (Tests of Attentional Distraction) NIM (Negative Impression) T-score ≥ 64 or TOAD total error rate ≥ 3.67% Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Self-reported ADHD diagnosis confirmed by a history of medication use, with clinically significant attention problems inferred from a Conners Adult Attention Rating Scale-Self-Report Short Version raw score of 21 **Diagnosed by:** Unclear/NR **Timing:** Concurrent	**Diagnostic accuracy summary:** Moderate to large effects differentiating the feigning group from control participants, both ADHD and non-ADHD, were observed for both the TOAD and PAI indicators. The disjunction rule enhances sensitivity beyond that of the invididual procesures at the expense of a decrease in specificity. Sensitivity 59% Specificity 82% PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
Musso, 2016{#646} N = 779 US College	**Target:** Individuals aged 17-29, diagnosed with ADHD or ADHD with comorbid mood or anxiety disorders, who completed psychoeducational evaluations and were assessed using the PAI validity indices **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Undergraduate college students aged 18-25 with no reported history of psychopathology, learning disorders, or ADHD, recruited from a large university in the Southeastern United States for extra credit, and instructed to simulate ADHD symptoms or respond honestly **Female:** 33.6% **Age:** 21.16 (2.91) Student volunteers: 19.72 (1.42) Min age: 17 Max age: 29 **Age subgroup**: Adults **Ethnicity:** Other : Clinical:1.7; student volunteers: 2.9 % Hispanic or Latino : Clinical: 3.2; student volunteers: 5.8 % Black/African American : Clinical: 9.1; student volunteers: 7.9 % Asian : Clinical: 1.5; student volunteers: 4.2 % White : Clinical: 84.4; student volunteers:79.2 Multicenter **Funding:** Other	**Test description:** Effort tests and symptoms validity tests are designed to assess the credibility of self-reported symptoms and detect intentional exaggeration or malingering. These tests include validity indices like the Negative Impression Management (NIM), Malingering Index (MAL), and Rogers Discriminant Function (RDF) in the Personality Assessment Inventory (PAI), which evaluate patterns of responses to identify response distortion. By analyzing these patterns, the tests differentiate genuine cases of ADHD from individuals feigning symptoms, providing insights into the reliability of the reported symptoms. Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive psychoeducational evaluation conducted by clinical graduate students under the supervision of a licensed psychologist using diagnostic criteria and standardized assessments **Diagnosed by:** Specialist (e.g., mental health) clinical graduate students under the supervision of a licensed psychologist **Timing:** Concurrent	**Diagnostic accuracy summary:** The alternative cutoff scores of ≥77 on the Negative Impression Management (NIM) scale, ≥ three on the Malingering Index (MAL), and ≥ one on the Rogers Discriminant Function (RDF) yielded excellent specificity in all groups and sensitivities of 33, 30, and 20%, respectively. Sensitivity 33% MAL cutoff >3: 22; RDF cutoff>1: 30 Specificity 98% MAL cutoff >3: 98; RDF cutoff>1: 96 PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Comparing self-reported ADHD symptoms (from ADHD simulators) to clinical data from individuals diagnosed with ADHD Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** Comorbidity (e.g. anxiety, depression) Misdiagnosis was more likely in individuals presenting with comorbid psychiatric conditions, such as anxiety or depression, as these conditions impacted validity index scores, potentially complicating differentiation between genuine and feigned symptoms.
Phillips, 2023{#5528} N = 317 US	**Target:** Adults referred for outpatient neuropsychological evaluation of known or suspected ADHD, the majority were actively enrolled college students, diagnoses based on DSM-5 criteria using a standardized multimodal diagnostic assessment protocol including history, symptom	**Test description:** RAVLT (Rey Auditory Verbal Learning Test), which assesses verbal/auditory learning and memory, and the BVMT-R (Brief Visuospatial Memory Test-Revised), which evaluates visuospatial learning and memory; incorporates embedded PVTs (performance validity tests) to	**Diagnostic accuracy summary:** These memory-based RAVLT and BVMT-R PVTs were able to accurately identify invalid neuropsychological test performance among adults undergoing evaluation for ADHD, regardless of whether diagnostic criteria for ADHD were met.	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
Specialty care	questionnaires, clinical interviews, and neuropsychological testing **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants included adults referred for outpatient neuropsychological evaluation who had attention complaints but did not meet DSM-5 criteria for ADHD, evaluated in a specialty care setting using standardized diagnostic protocols; participants were classified as having valid or invalid test performance based on performance validity tests **Female:** 61% **Age:** 27.7 (6.67) Min age: 18 Max age: 45 **Age subgroup**: Adults **Ethnicity:** Other : 5 % Hispanic or Latino : 24 % Black/African American : 15 % American Indian or Alaska Native : 46 % Asian : 10 Single center **Funding:** Unclear	differentiate valid from invalid cognitive performance among individuals referred for ADHD evaluation Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on DSM-5 criteria using a standardized multimodal diagnostic assessment protocol that included a detailed clinical interview, review of medical and psychiatric history, symptom questionnaires, and a comprehensive neuropsychological test battery **Diagnosed by:** Specialist (e.g., mental health) Primary care physicians or psychiatrists **Timing:** Concurrent	Sensitivity 43% FC<14: 49; RD<5: 35 Specificity 90% FC<14: 92; RD<5: 90 PPV NPV +LR -LR Accuracy AUC 0.74 FC<14: 0.70; RD<5: 0.63 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Potts, 2022{#733} N = 68 US College	**Target:** Participants with ADHD were young adults recruited from undergraduate psychology classes, diagnosed by a qualified professional with symptom onset before age 12, elevated scores on CAT-A symptom indexes, and substantial impairment in academic, occupational, or social areas **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants included young adults from the same	**Test description:** MARS Symptom Validity Index 4 (SV-index 4) and CAT-A Infrequency Scale evaluate whether individuals over-endorse unlikely ADHD symptoms. Effort tests like the Word Memory Test (WMT) assess cognitive performance consistency to detect poor effort. Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis	**Diagnostic accuracy summary:** The MARS SV index-4 demonstrated higher sensitivity rates for simulated malingering (61.8%) at close to optimal specificity (88.2%) compared to two published tests (which had sensitivity <42% at specificity >90%). Sensitivity 61.8% Specificity 88.2% PPV 84 NPV 69.8 +LR 5.2 -LR 0.4 Accuracy 75	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	undergraduate psychology classes instructed to feign ADHD symptoms for the purpose of the study **Female:** 52.9% **Age:** 18.82 (1.51) Min age: 18 Max age: 20 **Age subgroup**: Young **Ethnicity:** % Black/African American : 5.9 % Asian : 5.9 % White : 88.2 % Multiracial : 5.9 Single center **Funding:** Other	Diagnosed with ADHD based on a semi-structured clinical interview conducted by a licensed psychologist or advanced doctoral student, confirming diagnosis by a qualified professional with symptom onset before age 12, substantial functional impairment, and elevated T-scores on relevant symptom scales **Diagnosed by:** Specialist (e.g., mental health) Psychologist **Timing:** Concurrent	AUC 0.79 0.68, 0.90 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Quinn, 2003{#745} N = 60 US College	**Target:** Undergraduate students previously diagnosed with ADHD by a trained psychiatrist using DSM criteria, aged 17-29 years, recruited through a university disability office, with 50% currently prescribed stimulant medication but refraining for at least 12 hours prior to testing **ADHD presentation:** inattentive : 25,combined : 75 **Comorbidity:** N/A **Other:** Neurotypical undergraduate psychology students aged 17-29 years, randomly assigned to control or simulated malingerer conditions, with the control group instructed to perform accurately and malingerers instructed to convincingly simulate ADHD symptoms **Female:** 50% **Age:** 19.8 (N/A) Min age: 17 Max age: 29 **Age subgroup**: Adults **Ethnicity:** % Black/African American : 25	**Test description:** Self-report (ADHD Behavior Checklist) and neuropsychological testing (IVA CPT) are both used to assess feigning ADHD. Malingerers successfully faked symptoms on self-reports but overcompensated on the IVA CPT. The IVA CPT served as a symptom validity test by detecting inconsistencies in response patterns and reaction times. Effort testing was implicit in the CPT, as malingerers exhibited unnatural response patterns. Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD by a trained psychiatrist using DSM criteria, based on clinical interview and self-report questionnaire, with some participants previously assessed using a Continuous Performance Test (CPT) **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** IVA CPT could not be faked on 81% of its scales. The CPT's impairment index results revealed: sensitivity 94%, specificity 91%, PPP 88%, NPP 95%. Sensitivity 94% Specificity 91% PPV 88 NPV 95 +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	% Asian : 12.5 % White : 62.5 Single center **Funding:** Unclear			
Ramachandran, 2019{#750} N = 637 US College	**Target:** Participants self-reported a previous formal diagnosis of ADHD, provided details about the approximate time since diagnosis, the time spent during their first diagnostic appointment, and the specialty of the diagnosing practitioner; individuals with ADHD were asked to respond honestly to the survey **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD participants were either instructed to respond honestly (control group) or to feign ADHD symptoms (malingering group); the sample consisted of college students recruited from a university setting, primarily through online recruitment and incentives such as class credit **Female:** 49.5% **Age:** 20.9 (2.12) Min age: 18 Max age: 25 **Age subgroup**: Young **Ethnicity:** Other : 1.9 % Hispanic or Latino : 1.9 % Black/African American : 5.8 % Asian : 1.9 % White : 85.4 % Multiracial : 2.9 Single center **Funding:** Other	**Test description:** SAMS (Subtle ADHD Malingering Screener), a symptom validity test designed to detect individuals feigning ADHD symptoms through a self-report measure, evaluates effort and response patterns by distinguishing genuine ADHD from exaggerated or fabricated symptom presentations; administered together with the PAI (Personality Assessment Inventory cut offNIM score 92 or above or a Malingering Index score greater than 3) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on self-reported previous formal diagnosis, time since diagnosis, duration of the diagnostic appointment, and the specialty of the diagnosing practitioner **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** SAMS showed a sensitivity of 90% and specificity of 80%. The PAI was found to have a sensitivity of 51% and a specificity of 89%; the PAI's rate of false positives (10.8%) was somewhat lower than the SAMS, but the rate of false negatives (49.0%) was much higher. Sensitivity 90.3% Specificity 80.1% PPV NPV +LR -LR Accuracy AUC 0.901 **Concordance:** N/A **Rater agreement:** N/A Kappa ICC **Test-retest:** N/A **Internal consistency:** Cronbach's alpha N/A **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A
Robinson, 2023{#777}	**Target:** Adults referred to a university-affiliated clinic for psychoeducational evaluations due to concerns related to ADHD and/or specific learning disorder, assessed using the Conners	**Test description:** CPT-3 (Conners Continuous Performance Test-3) as an embedded validity indicator (EVI) to detect non-credible responders, which includes individuals feigning ADHD.	**Diagnostic accuracy summary:** Receiver operating characteristic curves (ROC) revealed that 5/9 individual indicators and 2/4 composite indicators met a minimally acceptable classification accuracy of ≥ 0.70	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
N = 201 US College	Continuous Performance Test-3 and multiple performance validity tests, credible participants failed 0 performance validity tests **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Non-ADHD included individuals referred to the same specialty care clinic for psychoeducational evaluations who did not meet the ADHD diagnostic criteria, including those assessed for specific learning disorders or presenting with unrelated concerns, non-credible participants failed 2 or more paerformance validity tests **Female:** 72.4% **Age:** 23.04 (6.80) Min age: 18 Max age: 50 **Age subgroup**: Adults **Ethnicity:** Other : 83.3 % Hispanic or Latino : 23 % Black/African American : 66.7 % Asian : 66.7 % White : 81.4 Single center **Funding:** Other	PVTs used as the reference standard, with non-credible responders classified based on failure of ≥2 PVTs. The CPT-3 indicators, including omissions, commissions, detectability, variability, and reaction time measures, are tested for their ability to distinguish between credible and feigned ADHD presentations. CPT-3 indicators classification accuracy threshold (AUC ≥ 0.70) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a combination of clinical interview, psychoeducational evaluation, and performance validity tests using established cutoffs to classify credible and non-credible responders **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	(AUC = 0.43–0.78). Individual (0.16–0.45) and composite indicators (0.23–0.35) demonstrated low sensitivity when using cutoffs that maintained specificity ≥90%. Sensitivity 45% Omission>59: 38; Commission>68: 36; VAR>59: 38; HRTSD>65: 29 Specificity 90% Omission>59: 91; Commission>68: 90; VAR>59: 90; HRTSD>65: 91 PPV NPV +LR -LR Accuracy AUC 0.76 Omission>59: 0.72; Commission>68: 0.70; VAR>59: 0.78; HRTSD>65: 0.73 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Rogers, 2021{#4474} N = 147 US College	**Target:** Adults with a prior clinical diagnosis of ADHD, assessed using comprehensive psychological evaluations, with common comorbidities including major depressive disorder, learning disorders, and anxiety disorders **ADHD presentation:** N/A **Comorbidity:** N/A	**Test description:** WAIS-IV (Wechsler Adult Intelligence Scale–4h Edition) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on comprehensive psychological assessments conducted by clinicians, including clinical interviews and	**Diagnostic accuracy summary:** Very large effect sizes (Cohen's ds from 1.66 to 1.90) differentiated between genuine and feigned ADHD. Two strategies (significantly below-chance performance and floor effect) showed strong promise if cross-validated for other feigning presentations. Sensitivity % Specificity % PPV NPV	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Other: Undergraduate students enrolled in psychology courses, with no history of ADHD or ADHD medication use, instructed to simulate ADHD symptoms for the purpose of the study **Female:** 54.8% **Age:** 25.59 (4.17) Min age: 18 Max age: 34 **Age subgroup**: Adults **Ethnicity:** % Hispanic or Latino : 23.8 % Black/African American : 10.2 % Asian : 4.8 % White : 51 % Multiracial : 6.1 Single center **Funding:** Other	standardized testing to confirm the diagnosis. **Diagnosed by:** Specialist (e.g., mental health) Clinician mental specialist **Timing:** Concurrent	+LR -LR Accuracy 64 AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Smith, 2017{#863} N = 129 US College	**Target:** Participants were diagnosed with ADHD through a semi-structured clinical interview assessing DSM criteria, including at least six symptoms of inattention or hyperactivity/impulsivity in at least two domains, evidence of impairment before age 7, and verification through self-reported ADHD diagnosis or treatment history **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** The non-ADHD participants included individuals recruited from a university psychology research pool, with one group instructed to simulate ADHD symptoms and another consisting of archival clinical records of individuals evaluated for ADHD at a university psychology training clinic **Female:** 45.5% male: 54.5 **Age:** 18.77 (1.19)	**Test description:** PAI (Personality Assessment Inventory) to evaluate symptom validity and distinguish between genuine and simulated ADHD presentations Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a semi-structured clinical interview assessing DSM criteria, including six or more symptoms of inattention or hyperactivity/impulsivity in at least two domains, evidence of impairment before age 7, and verification through self-reported diagnosis or treatment history. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The PAI may be informative as an indicator of potentially exaggerated or malingered symptom presentation, but alternative cut scores for symptom validity indicators may be necessary to maximize its utility in these particular types of psychological evaluations. Sensitivity 68% NIM(27.3); INF(65.2); MAL(76.2); PIM(56.1) Specificity 83% NIM(95.2); INF(74.6); MAL(60.9); PIM(68.3) PPV NPV +LR -LR Accuracy AUC 0.75 (CI 0.67, 0.83) NIM(0.75); INF(0.75); MAL(0.64); PIM(0.67) **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:**	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Min age: 18 Max age: 23 **Age subgroup**: Young **Ethnicity:** Other : Simulation sample: 7.60; clinical comparison procedure: 4.5; archival comparison sample: 4.9 Other : Simulation sample: 21.20; clinical comparison procedure: 4.5; archival comparison sample:26.8 Other : Simulation sample: 1.5; archival conparison sample: 22 Other : Simulation sample: 69.70; clinical comparison: 90.0; archival comparisoon: 46.3 Single center **Funding:** Unclear		Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Sollman, 2010{#872} N = 73 US College	**Target:** College students with a verifiable diagnosis of ADHD confirmed through neuropsychological or psychological evaluation, including corroborative interviews with parents or teachers, medication washout for 12 hours before testing, excluding those with comorbid learning disabilities, psychiatric or neurological conditions, or substance abuse **ADHD presentation:** inattentive : 20,hyperactive : 5,combined : 75 **Comorbidity:** N/A **Other:** College students recruited from the same university setting, divided into two groups: a normal honest-responding group with no history of ADHD or related disorders, and a feigning group instructed to simulate ADHD based on provided materials; participants were screened to exclude those with learning disabilities, psychiatric or neurological conditions, or substance abuse **Female:** 44.8% **Age:** 19.40 (1.21)	**Test description:** CAARS and Conner's Continuous Performance Test-II Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on a comprehensive clinical evaluation including neuropsychological testing, symptom self-report measures, corroborative interviews with parents or teachers, and confirmation of developmental origin of symptoms **Diagnosed by:** Specialist (e.g., mental health) Mental health clinicians **Timing:** Concurrent	**Diagnostic accuracy summary:** The detectability index in Connor's CPT-II had a sensitivity of 17% to detect ADHD and a specificity of 90% for feigning ADHD. Failing 1 or more, 2 or more, 3 or more, 4 or more cognitive feigning test indices lowered the sensitivity from 63 to 50, 47, and 35%, while the specificity increased from 82, to 93, 100%, and 100%. Indicates limited sensitivity in distinguishing ADHD from controls and susceptible to manipulation by feigning participants; results point to a need for a thorough evaluation of history, cognitive and emotional functioning, and the consideration of exaggerated symptomatology in the diagnosis of ADHD. Sensitivity 63% Specificity 83% PPV 78.6 NPV 69.3 +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	Min age: 18 Max age: 21 **Age subgroup**: Young **Ethnicity:** % Black/African American : 6.90 % Asian : 0 % White : 86.20 % Multiracial : 6.90 Single center **Funding:** Unclear		**Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Spencely, 2022{#883} N = 150 US College	**Target:** College students with a prior professional diagnosis of ADHD confirmed through self-report at multiple time points and reporting significant current symptoms of inattention or hyperactivity/impulsivity (at least four symptoms meeting the 95th percentile per a symptom checklist) **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** College students recruited from a midsized public university and a midsized private college, including a control group instructed to respond honestly and a simulated group instructed to feign ADHD symptoms for the purpose of the study **Female:** 50% **Age:** 19.93 (2.32) Min age: 18 Max age: 23 **Age subgroup**: Young **Ethnicity:** % Hispanic or Latino : 3.3,Other : Simulated ADHD group: 8.3, control group: 3.3 % Black/African American : 3.3,Other : Simulated ADHD group: 13.3, control group: 21.7	**Test description:** Medical Symptom Validity Test admistered together with the Woodcock-Johnson IV Tests of Cognitive Abilities (assess neuropsychological functioning, specifically focusing on processing speed, working memory, and cognitive efficiency as potential markers for feigned ADHD)and the BAARS (Barkley Adult ADHD Rating Scale IV) Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD by a physician, psychologist, or other health professional, confirmed through participant self-report at multiple time points and screened with an ADHD symptom checklist to ensure significant current symptoms at or above the 95th percentile **Diagnosed by:** Specialist (e.g., mental health) The ADHD diagnosis was made by a physician, psychologist, or other mental health professional **Timing:** Prior diagnosis	**Diagnostic accuracy summary:** Medical Symptom Validity Test showed the best performance (AUC 0.89, sensitivity 0.78, specificity 97). Several processing speed and working memory scores from the WJ-IV effectively identified students feigning ADHD, detecting at least 50% of those students at score cutoffs that also maintained specificity of 90% or more, close to the efficiency of the standalone PVT. The study found that individuals simulating ADHD showed significantly lower scores on these measures compared to those with genuine ADHD, suggesting that working memory and processing speed deficits may help detect feigned ADHD. Sensitivity 78% Specificity 97% PPV NPV +LR -LR Accuracy AUC 0.89 0.58 - 0.89 **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
	% Asian : 0,Other : Simulated ADHD group: 5, control group: 5 % Native Hawaiian or Pacific Islander : 0,Other info : Simulated ADHD group: 1.7, control group: 0 % White : 76.7,Other : Simulated ADHD group: 63.3, control group: 58.3 % Multiracial : 16.7,Other : Simulated ADHD group: 8.3, control group: 11.7 Multicenter **Funding:** Other		**Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	
Suhr, 2011{#10849} N = 1297 US College	**Target:** Participants self-reported a prior ADHD diagnosis were either university students participating in research or individuals seeking psychological evaluation at a university clinic; inclusion required documented evidence of childhood impairment, clinically significant current ADHD symptoms from multiple sources, and passing a cognitive validity test **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** University students without an ADHD diagnosis and individuals seeking psychological evaluation at a university clinic; categorized into a psychological control group (diagnosed with/treatment for a non-ADHD psychological condition or meeting criteria for a psychological disorder) and a normal control group (no history of ADHD or psychological disorders) **Female:** 47% **Age:** mean 19 Min age: 18 Max age: 59 **Age subgroup**: Adults **Ethnicity:** N/A Single center	**Test description:** CII (CAARS Infrequency Index) for detecting feigned ADHD (not for diagnosing ADHD). A cutoff score of ≥21 was used to indicate noncredible symptom reporting Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical evaluation, including evidence of childhood impairment, clinically significant current ADHD symptoms from multiple sources (self-report, behavioral observation, collateral report), and passing a cognitive validity test. **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** The CAARS Infrequency Index (CII) demonstrated moderate sensitivity (30-80%) but high specificity (>99%) in identifying feigned ADHD based on extreme scores on CAARS DSM-IV subscales, with an AUC of 0.92 for the Hyperactive/Impulsive Subscale (F). The Word Memory Test (WMT) showed low sensitivity (24%) but high specificity (95%) for detecting noncredible cognitive performance, distinguishing individuals feigning ADHD from those with genuine ADHD. Sensitivity 24% CAARS Inattentive Subscale E: 30, Hyperactive/Impulsive Subscale F: 80 Specificity 95% CAARS Inattentive Subscale E: >99, CAARS Hyperactive/Impulsive Subscale F: 93 PPV NPV +LR -LR Accuracy 67 Distinguishing feigned ADHD vs. genuine ADHD AUC CAARS E scores: 0.78, CAARS F scores: 0.92 **Concordance:** **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha	**Subgroup analysis:** Sex Males scored higher than females on the CAARS Infrequency Index (CII), suggesting possible gender differences in noncredible symptom reporting and a higher cutoff (≥22) may be needed for males to maintain specificity.

Study ID	Population	Feigning ADHD	Results	Subgroup
	Funding: Other		0.86 **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** **Admin time:**	
Williamson, 2014{#4702} N = 76 US College	**Target:** Adults with a history of ADHD diagnosis confirmed by a mental health practitioner based on more than self-reported symptoms, to have received their diagnosis before age 18, and to have abstained from stimulant medication for 12 hours prior to the study **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Neurotypical individuals without a history of diagnosed or suspected ADHD, learning disorders, neurological disorders, or psychological disorders, recruited from an introductory psychology participant pool or a university disability resource center, with a subset instructed to feign ADHD **Female:** 36.36% **Age:** 19.05 (1.29) Min age: 18 Max age: 23 **Age subgroup**: Young **Ethnicity:** N/A Single center **Funding:** Unclear	**Test description:** WAIS-IV PSI (Wechsler Adult Intelligence Scale-IV Processing Speed Index) lower than 97, administered together with the Woodcock-Johnson III Test of Achievement, and the CTIP (Computerized Test of Information Processing) assessed cognitive abilities such as processing speed, reading fluency, and attention control under controlled conditions Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on evaluation by a mental health practitioner using clinical interviews, self-report symptom scales, and cognitive or neuropsychological testing, with diagnosis required to be established before age 18 **Diagnosed by:** Specialist (e.g., mental health) **Timing:** Concurrent	**Diagnostic accuracy summary:** Sensitivity of the WAIS-IV PSI was 65% for feigning ADHD, specificity for detecting ADHD decreased from 73% to 59% in a subgroup of participants with comorbidity. Performance validity tests such as the Test of Memory Malingering (TOMM), the Letter Memory Test (LMT), and the Nonverbal Medical Symptom Validity Test (NV-MSVT) were effective in differentiating both ADHD groups from normal participants feigning ADHD. Sensitivity 65% Specificity % PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A

Study ID	Population	Feigning ADHD	Results	Subgroup
Young, 2011{#1015} N = 69 US College	**Target:** Adults aged 18-25 diagnosed with ADHD through clinical interviews, third-person symptom reports, intelligence and achievement measures, personality questionnaires, behavior checklists, and team-based faculty-supervised assessments at a campus psychological assessment center **ADHD presentation:** N/A **Comorbidity:** N/A **Other:** Clinical sample and neurotypical adults recruited from a university setting, including a control group with no history of ADHD or psychological disorders and a malingering group instructed to feign ADHD symptoms **Female:** % MMPI-2: 21 female (30.4%); WAIS-III 9 female (26.5%) **Age:** MMPI-2: 18.97 (1.29); WAIS-III: 20.29 (1.87) Min age: 18 Max age: 25 **Age subgroup**: Young **Ethnicity:** Other : MMPI-2: 1% Other : MMPI-2: 20%; WAIS-III: 3% Other info : MMPI-2: 6% Other : MMPI-2: 72%; WAIS-III: 97% Single center **Funding:** Unclear	**Test description:** The MMPI-2 (Minnesota Multiphasic Personality Inventory-2), focusing on validity scales such as the Infrequency-Psychopathology (Fp), Fake Bad Scale (FBS), Response Bias Scale (RBS), and Henry–Heilbronner Index (HHI) to detect response bias and differentiate between genuine ADHD and malingering. Machine learning: No Validation dataset: No **Reference standard:** Clinical diagnosis Diagnosed with ADHD based on clinical interviews, third-person symptom reports, intelligence and achievement measures, personality questionnaires, and behavior checklists conducted by a faculty-supervised assessment team at a campus psychological assessment center **Diagnosed by:** Specialist (e.g., mental health) Faculty-supervised assessment team at a campus psychological assessment center **Timing:** Concurrent	**Diagnostic accuracy summary:** The MMPI-2 offers a number of validity indices that may assist in detecting individuals attempting to feign ADHD. Sensitivity 59% Sensitivity was calculated for the Infrequency-Psychopathology (Fp) scale at a cutoff of ≥5, which showed the highest balance of sensitivity and specificity among the MMPI-2 validity scales. Specificity 94% Specificity was calculated for the Fp scale at a cutoff of ≥5 and was the highest among scales evaluated in this study. PPV NPV +LR -LR Accuracy AUC **Concordance:** N/A **Rater agreement:** Compares self-reported ADHD symptoms with MMPI-2 validity scale results Kappa ICC **Test-retest:** **Internal consistency:** Cronbach's alpha **Misdiagnosis impact:** N/A **Diagnosis impact:** N/A **Labeling:** N/A **Side effects:** N/A **Cost:** N/A **Admin time:** N/A	**Subgroup analysis:** N/A