European Psychiatry 44 (2017) 104–124

Contents lists available at ScienceDirect

European Psychiatry

journal homepage: http://www.europsy-journal.com

Review

Measurement properties of screening and diagnostic tools for autism spectrum adults of mean normal intelligence: A systematic review

A. Baghdadli [a,*], F. Russet [a], L. Mottron [b]

[a] Autism Resources Centre of Languedoc-Roussillon, University Hospital of Montpellier, CHRU de Montpellier–Child and adolescent University Psychiatry Department, hôpital de la Colombière, 291, avenue du Doyen-Giraud, 34295 Montpellier cedex 5, France
[b] Autism Center of Excellence, hôpital Rivière des prairies & CIUSSS du Nord, université de Montréal, 7070, boulevard Perras, H1E1A4 Montréal, QC, Canada

ARTICLE INFO

Article history:
Received 15 February 2017
Received in revised form 27 April 2017
Accepted 30 April 2017
Available online 17 May 2017

Keywords:
Autism
Screening
Diagnosis
Systematic review

ABSTRACT

Background: The autism spectrum (AS) is a multifaceted neurodevelopmental variant associated with lifelong challenges. Despite the relevant importance of identifying AS in adults for epidemiological, public health, and quality of life issues, the measurement properties of the tools currently used to screen and diagnose adults without intellectual disabilities (ID) have not been assessed.
Objectives: This systematic review addresses the accuracy, reliability, and validity of the reported AS screening and diagnostic tools used in adults without ID.
Methods: Electronic databases and bibliographies were searched, and identified papers evaluated against inclusion criteria. The PRISMA statement was used for reporting the review. We evaluated the quality of the papers using the COSMIN Checklist for psychometric data, and QUADAS-2 for diagnostic data. For the COSMIN assessment, evidence was considered to be strong when several methodologically good articles, or one excellent article, reported consistent evidence for or against a measurement property. For the QUADAS ratings, evidence was considered to be "satisfactory" if at least one study was rated with a low risk of bias and low concern about applicability.
Results: We included 38 articles comprising 32 studies, five reviews, and one book chapter and assessed nine tools (three diagnostic and six screening, including eight of their short versions). Among screening tools, only AQ-50, AQ-S, and RAADS-R and RAADS-14 were found to provide satisfactory or intermediate values for their psychometric properties, supported by strong or moderate evidence. Nevertheless, risks of bias and concerns on the applicability of these tools limit the evidence on their diagnostic properties. We found that none of the gold standard diagnostic tools used for children had satisfactory measurement properties.
Conclusion: There is limited evidence for the measurement properties of the screening and diagnostic tools used for AS adults with a mean normal range of measured intelligence. This may lessen the validity of conclusions and public health decisions on an important fraction of the adult autistic population. This not only justifies further validation studies of screening and diagnostic tools for autistic adults, but also supports the parallel use of self-reported information and clinical expertise with these instruments during the diagnostic process.

1. Introduction

Autism Spectrum (AS) is a multifaceted neurodevelopmental condition for which the diagnosis is stable throughout its development [1]. Estimates of the prevalence of AS are regularly revised, following the recurrent updating of diagnostic guidelines. The prevalence of the autistic spectrum is estimated to be approximately 1%, with higher rates in men (1.8%) than in women (0.2%) [2,3]. AS is characterized in children, as well as in adults, by the coexistence of atypical communication and social interaction, with restricted and repetitive activities or behaviors (DSM-5) [4]. Current diagnosis criteria cover various symptom severity, language/speech, and intellectual levels. The reported proportion of autistic adults without intellectual disability (ID) is 50% [5], but this plausibly represents a conservative estimate, due to measurement issues [6], speech delay, or conversely, the absence of adaptive issues, limiting case ascertainment [7]. These individuals are frequently identified as "High Functioning", a potentially

* Corresponding author.
E-mail address: a-baghdadli@chu-montpellier.fr (A. Baghdadli).

http://dx.doi.org/10.1016/j.eurpsy.2017.04.009
0924-9338/

misleading label, confusing the level of adaptation with measured intelligence, and missing autistic people of normal intelligence with limited use of speech. Autistic adults of normal intelligence in these populations often manifest strong adaptive deficits that, in combination with poor societal adaptation, severely limit their socio-professional status [8,9]. They are often identified late, resulting in a 'lost generation' of adult autistic people [10]. Some may never even be diagnosed [2,11–13], with uncertain consequences for their well-being. Diagnostic challenges in adult autistics arise from a decreased magnitude of symptoms and atypicality with age and the presence of comorbid psychiatric conditions [14–17], and overlap between the signs of AS and those of other psychiatric or neurodevelopmental conditions [18]. One major obstacle to better identifying AS among adults is the lack of robust screening and diagnostic tools, as emphasized by most international guidelines [19–22].

In the absence of a biological gold standard method for the diagnosis of AS, its identification remains clinical, and requires multidisciplinary assessments from multiple sources [23,24]. The current state-of-the-art of adult AS diagnosis relies on self-report or informant questionnaires, observation guides, and clinical interviews. *Screening tools* are typically used to determine whether an individual is at risk for having AS and/or to justify a more formal assessment [25]. However, they are also used during the diagnostic process by primary care professionals or researchers with limited clinical expertise in the general or at-risk populations [20,26]. Screening tools designed for adults focus on the core symptoms of AS in different contexts, particularly among people referred for a medical diagnosis in clinical settings, notably in psychiatry units [20,27]. *Diagnostic tools* are used after a positive screening test to determine the presence or absence of AS when an individual displays signs of this condition. Trained professionals usually administer them during a multidisciplinary assessment [19,20,28]. They are more comprehensive, but also more time-consuming, and their use requires greater clinical expertise.

The choice of a diagnostic or screening tool depends, among several factors, on its measurement properties [29]: reliability, accuracy (or validity), sensitivity, specificity, and generalizability to the population for which they are intended to be used [20,30,31]. Systematic reviews of research evaluating tools devoted to the diagnosis of AS may guide clinicians and researchers in the selection of the best tools. The previous reviews of AS screening or diagnostic tools cover the entire range of age and/or IQ's [32–35]. Two reviews focus specifically on tools to assess AS adults without ID, but one is not systematic [36], and the other does not fully explore their measurement properties [37]. The aim of this study was to evaluate the measurement properties of the tools used for the screening and diagnosis of AS in adults without ID, focusing on their psychometric measurements and diagnostic accuracy.

2. Materials and methods

2.1. Search strategy

We performed a literature search for articles published in English or French in PsycInfo-Esbco (Psycinfo, Eric, PsycARTICLES, Psychology and Behavioral Sciences Collection), PubMed, Web of science (Web of Science™ Core Collection, KCI-Korean Journal Database, MEDLINE®, SciELO Citation Index), Cochrane Library, Science Direct, and Springer Linkin. The search was conducted in May 2016 and updated in September 2016, without limitation on the publication year. We followed the PRISMA standards, a 27-item checklist, and a four-phase process including identification, screening, eligibility, and inclusion of studies [38]. The keywords used were: "adult*", Diagnos*/Screen*, "Tool"/ "scale"/ "questionnaire", AS/autism*/asperger*. The algorithm used in each database was (diagnos* OR screen* OR assess*) AND (autism* OR AS OR Asperger*) AND adult* AND (tool* OR scale* OR questionnaire). We applied this algorithm to Abstracts for PsycInfo-Esbco; to all fields for PubMed, to Title, Abstract and Keywords for the Cochrane Library and Science Direct; to Topic for the Web of science; and to all the words for SpringerLink. We performed a complementary search using the reference lists of the studies selected for the review. Additionally, we searched the "grey literature" via Internet (Google and Google Scholar), according to the same keywords used in the database search. When a paper was not available, the authors were contacted via ResearchGate. The screening and selection processes are detailed in Fig. 1.

2.2. Inclusion criteria

Inclusion criteria for the selected papers were:

- documentation of AS screening and diagnostic tools focused primarily on AS core signs;
- reporting at least reliability, validity, or diagnostic accuracy of AS screening and diagnostic tools;
- a mean age of over 18 years at study entry, and a mean IQ over 70 for at least half of the participant sample. Papers in which the chronological age and intellectual level of their participants were not reported were excluded;
- having their participants defined through a "best estimate" diagnosis of autism, atypical autism, Asperger Syndrome, or PDD-NOS (Pervasive Developmental Disorder-Not Otherwise Specified), according to ICD-10 or DSM-IV criteria or to ASD DSM-5 criteria and a multidisciplinary assessment. The use of specific diagnostic tools, such as the Autism Diagnostic Observation Schedule-Generic (ADOS-G), or the Autism Diagnostic Interview-Revised (ADI-R), was not required. Autistic individuals with another physical or mental health condition were included;
- minimum sample size of 10 per group [20].

2.3. Data extraction

Two authors (AB &FR) read abstracts, and selected them if they were broadly consistent with the inclusion criteria. If consensus was not reached, the abstracts were set aside for further evaluation. Then, AB reviewed full-text articles of the selected abstracts against the inclusion criteria. Data were extracted from full-text articles (FR) and reviewed (AB), with regular verification and discussion to ensure consistency. Data extraction from full texts was organized into the following sections:

- tools (authors; type (e.g. questionnaire); targeted population; short description);
- information about each article (author(s); year; sample characteristics/sample size, age, cognitive level, gender, control groups);
- psychometric properties, including reliability (internal consistency, test-retest reliability, inter-rater reliability), construct validity (content validity, internal validity, criterion validity), and diagnostic validity (sensitivity, specificity, and accuracy measured by the Area Under the ROC Curve [AUC].

2.4. Data analysis

Measurement properties were independently assessed according to thresholds reported in the literature, such as satisfactory, intermediate, unsatisfactory, or no information available [25,27,29,39–49]. Sensitivity was considered to be satisfactory if the value

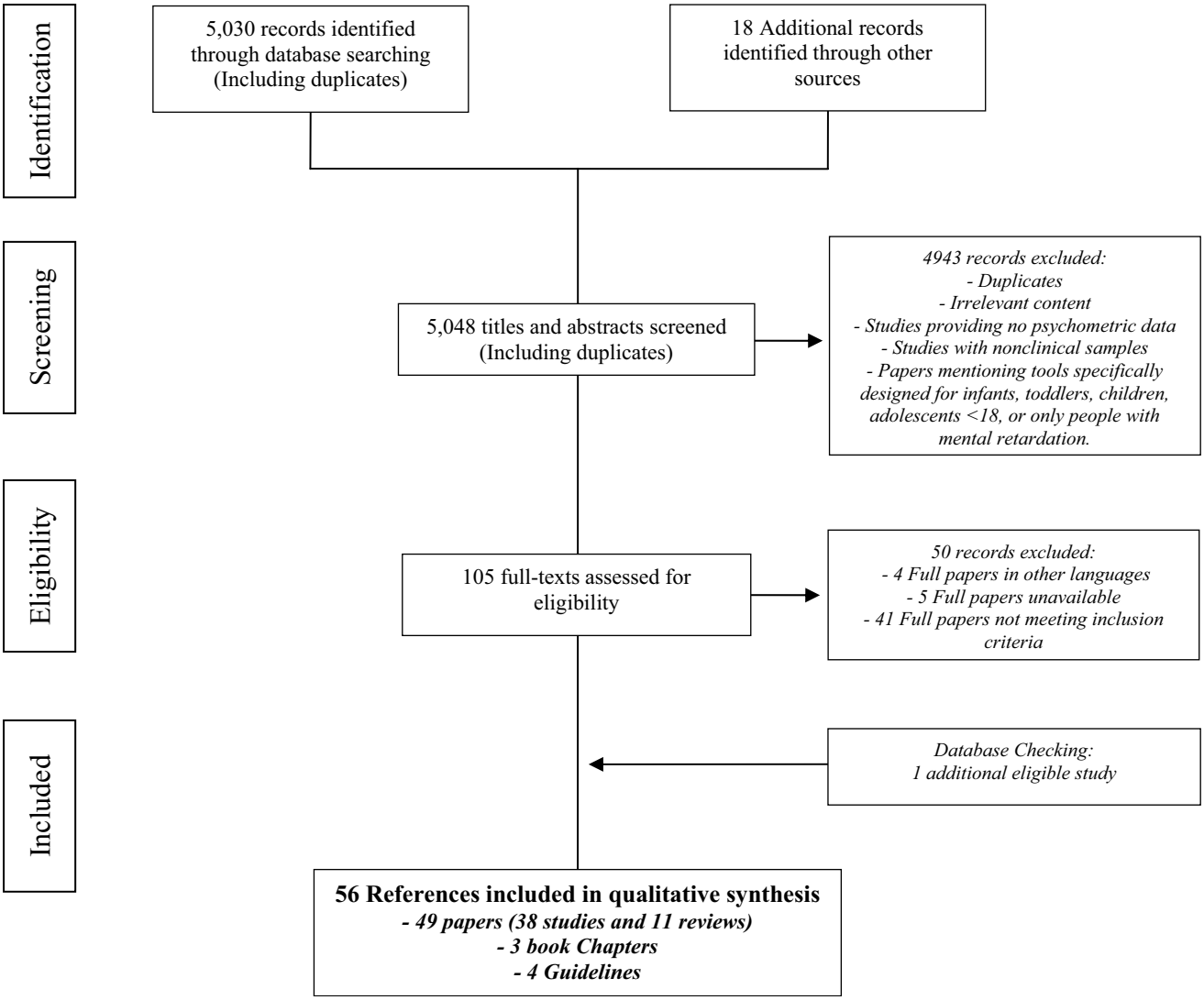

Fig. 1. Screening and selection process according to PRISMA flowchart for systematic review.

was > 70% and internal consistency if Cronbach's alpha was equal to or greater than 0.70 [27,46]. The methodological quality of studies exploring sensitivity, specificity, and AUC was assessed using Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [50]. For each study, four items related to risks of bias and three related to concerns about applicability were rated as high risk/concern, low risk/concern, or unclear. An overall judgment was then provided for each study as follows:

- if a study was judged to be "low" in all domains related to bias or applicability, the overall judgment was "low risk of bias" or "low concern regarding applicability" for this study;
- if a study was judged to be "high" or "unclear" in one or more domains, it was judged to be "at risk of bias" or to have "concerns regarding applicability" [50].

The methodological quality of studies exploring other psychometric properties was assessed using COSMIN [51]. For each property, the methodological quality was rated as excellent, good, fair, or poor, based on the lowest item awarded. Finally, the overall methodological quality was synthesized for each measurement property and for each tool. For the synthesis of the COSMIN assessment, evidence was considered to be "strong" when several methodologically good articles, or one excellent article, found consistent evidence for or against a measurement property; "moderate" for several methodologically fair studies, or one good study; "limited" for one study of fair quality; and otherwise "conflicting evidence" or "unknown" [52]. For the synthesis of QUADAS ratings, evidence was considered to be "satisfactory" if at least one study was rated with low risk of bias and low concern about applicability, or evidence that was considered to be "unsatisfactory" in the other cases.

3. Results

3.1. Reference selection - systematic literature search

Of the 5030 abstracts identified, we rejected 4942 publications at abstract review, and assessed 88 available full texts. Of these, 50 did not meet the inclusion criteria, mainly because of age and IQ issues. We finally selected a total of 38 papers, comprising 32 studies, five reviews, and one book chapter. We identified 17 additional references through other resources, including 11 full papers (five studies and six reviews), two book chapters, and four guidelines. Database screening added one additional study. Finally, 56 references were included in our review: 49 articles (including 38 studies), three book chapters, and four guidelines.

3.2. Identification of studies examining measurement properties

We identified three diagnostic tools and six screening tools (plus eight of their short versions) used for adults without ID. The selected tools (* in the reference section) were three observation-rating scales, one clinician questionnaire, three self-report questionnaires, and two multiple-source questionnaires (with a self-report version and an informant version). Among these nine tools, three were designed for diagnosis (Autism Diagnostic Observation Schedule–Generic or ADOS-G, Autism Mental State Exam or AMSE, and Adult Asperger Assessment or AAA). Their scoring and interpretation can only be performed by clinicians with diagnostic expertise. Six of the tools were specifically designed for AS adults without apparent ID. Two had a section for adults without ID (ADOS-Module 4 and Social Responsiveness Scale 2nd ed.–Adult form or SRS-A), and one was designed for individuals, irrespective of their age or IQ (AMSE). An abbreviated description of these tools is shown in Table 1.

A total of 38 studies specifically explored the measurement properties of the nine tools (including six assessing multiple tools):

- two related to the AAA;
- four to ADOS-G/ADOS-2 Module 4;
- one to the AMSE;
- 23 to the Autism Spectrum Quotient, or AQ, and its short-versions;
- one to the Adult Social Behavior Questionnaire or ASBQ;
- two to the Autism Spectrum Disorder in Adults Screening questionnaire or ASDASQ;
- three to the Ritvo Asperger and Autism Diagnostic Scale-Revised or RAADS-R, and one to a short-version;
- one to the Sensory Reactivity in Autism Spectrum or SR-AS;
- three to the SRS-A/SRS2-A, and short-versions (see Table 2 for sample characteristics of studies).

Measurement properties were tabulated for each study. Among studies including a control group (92%), approximately 69% included individuals with other psychiatric conditions, whereas approximately 31% included only typical individuals. The sample sizes of the AS groups in the 38 studies were always limited (71% of the studies recruited less than 100 individuals).

3.3. Quality assessment of the studies

The quality rating according to the COSMIN Checklist, was mostly "fair" for 25 studies, mostly "poor" for five studies, and heterogeneous for the rest. We assessed only internal consistency, reliability, content validity, structural validity, and hypothesis testing, because there were no available data in the selected articles concerning measurement error, responsiveness to change, or cross-cultural validity (Table 3). We explored the quality of the 36 studies that assessed diagnostic accuracy, according to QUADAS-2 criteria. All studies had a risk of bias, mostly related to patient selection (77%), and Index-test (69%). Overall, 61% of studies were judged to have concerns about applicability, and 95%, a high concern related to the patient selection item (Table 4).

3.4. Measurement properties

Table 5 gives a breakdown of the quality ratings for each of the following tools, based on COSMIN and QUADAS-2 criteria.

3.4.1. Diagnostic tools

3.4.1.1. ADOS-Module 4. The Autism Diagnostic Observation Schedule–Generic (ADOS-G), and its revised version, the Autism Diagnostic Observation Schedule–Second Edition (ADOS-2) are clinical rating scales with different modules according to age and language level. Module 4, a 1–2 hour interview, was specifically designed for verbally fluent adolescents and adults. Four articles about the original and revised algorithms of the Module 4 of Autism Diagnostic Observation Schedule–Generic (ADOS-G) were identified. In the articles about the original algorithm of ADOS-G [62,73], internal consistency ($\alpha = 0.87$ total score on whole sample) [73], and discriminant validity (except for Schizophrenia) [73] were satisfactory, but evidence was limited according to the COSMIN criteria. There was also limited evidence for convergent validity (the correlations with the PANNS were −0.59 for socialization, and 0.12 for communication) [73]. Moreover, satisfactory reliability was only reported for inter-rater reliability, with inconsistent evidence. Content and structural validity were not explored. Finally, evidence was satisfactory for a good AUC, but unsatisfactory for the heterogeneous sensitivity (0.68 to 1) and specificity (0.73 to 0.86). In the articles about the revised algorithm [74,75], there were no data about reliability and content validity, but evidence was strong for satisfactory structural validity and mixed internal consistency (Social Affect $\alpha = 0.84$, RRBs $\alpha = 0.61$) [74], and limited for discriminant validity, except for Schizophrenia [75]. Sensitivity (0.61 to 0.905) and specificity (0.50 to 0.822) were variable, with limited evidence.

3.4.1.2. Adult Asperger Assessment (AAA). This tool was designed to assist in the diagnosis of AS in adults without ID. It includes two self-report questionnaires and a clinical interview with the proband and/or an informant. Two articles examined its measurement properties. Diagnostic validity appeared to be satisfactory (sensitivity = 0.92 and specificity =1.00), but evidence was unsatisfactory. Psychometric properties were examined only through discriminant validity [53], and its structural validity was not confirmed [72].

3.4.1.3. Autism Mental State Exam (AMSE). This short clinical observation scale was designed to be used by clinicians with expertise in AS for diagnosis in various clinical settings. Its properties were examined by one article. Diagnostic accuracy was satisfactory (AUC = 0.97, sensitivity = 91%, specificity = 93%) [56], but with unsatisfactory evidence. Psychometric properties were not studied, except reliability, for which satisfactory results were found, but with conflicting evidence.

3.4.2. Screening tools

3.4.2.1. Autism Spectrum Quotient-50 (AQ-50) and short-versions. - This short self-report questionnaire was specifically designed for adults without ID, and to assess autistic traits in the general population. It is also used for screening. Different shorter versions have also been developed over time. Properties of the *AQ-50* were examined in 18 articles. None of the articles provided full data about content validity. Evidence was strong for intermediate to satisfactory internal consistency [79,88], and satisfactory test-retest reliability [79,89]. In addition, discriminant validity was satisfactory and explored in clinical and community control groups [79,87]. The values for convergent validity were heterogeneous, for which the evidence was moderate: high correlation ($r = 0.80$) with the SRS2-AS [89], very low ($r = 0.18$) with the ADI-R [85], or absent with the EQ [81]. The structural validity of a five factor-structure for the AQ50 was not confirmed, and this finding was supported by moderate evidence only [79]. Values related to diagnostic data were variable, with significant risks of bias and/or concerns about applicability, and small sample sizes: accuracy ranged from 0.72 to 0.90, AUC from 0.647 to 0.99, sensitivity from 0.75 to 0.95, and specificity 0.52 to 0.97.

There are shorter versions of the AQ. Four papers explored the properties of the *AQ-10*. Moderate evidence was found for

Table 1
Description of the 9 identified instruments (and 8 short versions) for screening/diagnosis of AS in adults without ID.

Instruments	Authors and year of publication	Type of instrument and purpose: Screening *(S)* Help to diagnosis *(HD)*	Target population	Short description
Adult Asperger Assessment (AAA)	Baron Cohen, Wheelwright, Robinson & Woodbury-Smith (2005) [53]	Self-report questionnaires + Guide to clinical interview (with patient and/or informant) HD	"Adults" with average IQ minimum	Two step PC-based instrument AAA template: four sections + a final section with prerequisites. Section A = "Qualitative impairment in social interaction", Section B = "Restricted, repetitive, and stereotyped patterns of behavior, interests, and activities". Section C = "Qualitative impairments in verbal or non-verbal communication", and Section D = "Impairments in imagination"; 1st step: AQ + EQ administrated to patient; 2nd step: clinical interview. To validate the information provided by AQ and EQ by collecting examples from the patient and an informant. To check the other symptoms and prerequisites *To meet criteria for a DSM-IV diagnosis of AS, patients must have two or more symptoms from Section A and one or more symptoms from other Sections. They also need to meet the pre-requisites in Sections E*
Autism Diagnostic Observation Schedule Generic (ADOS-G)* *Or* Second Edition (ADOS-2)** *MODULE 4* *Original and Revised algorithms*	Lord, Rutter, DiLavore, & Risi (2000)[a] [54] Lord, Rutter, et al. (2012)[b] [55]	Semi-structured, standardized, observational assessment tool HD	Verbally fluent adolescents and adults	The ADOS-2 assesses communication, social interaction, play, and restricted and repetitive behavior Module 4 contains the socio-emotional questions of the ADOS, along with interview items about daily living, additional tasks, and other items (imagination, sensory features, RRB. . .). 3-point scale from 0 (*no evidence of abnormality related to autism*) to 2 (*definite evidence*) *Criteria for original algorithm: for classification of AD or ASD, an individual must meet thresholds for the communication domain, social interaction domain, and the summation of these two domains (but not for the restricted and repetitive behaviors domain). Important: ADOS-G classification and the overall diagnosis may not be congruent and must then be distinguished *Criteria for revised algorithm: classification is based on the SARRB domain, which combines social, communication, and restricted behavior items. One single cut-off score, for ASD
Autism Mental Status Exam (AMSE)	Grodberg et al. (2014) [56]	Observation clinician rating scale HD	All ages High Functioning	Brief diagnostic tool covering social, communicative, and behavioral functioning in people with ASD Eight items: eye contact, interest in others, pointing, language, pragmatics, stereotype, preoccupations, and unusual sensitivities

Table 1 (*Continued*)

Instruments	Authors and year of publication	Type of instrument and purpose: Screening *(S)* Help to diagnosis *(HD)*	Target population	Short description
Adult Social Behavior Questionnaire (ASBQ)	Horwitz (2016) [57]	Questionnaires S or HD	Adults (Validated on people without ID)	Questionnaire as a self-report version (ASBQ-SR) and a version for spouses, parents, or other informants (ASBQ-OR) 44 items in each questionnaire, covering six domains: reduced contact, reduced empathy, reduced interpersonal insight and TOM, violation of social conventions, insistence on sameness, sensory stimulation & motor stereotypes Answer options: described behavior 'clearly applies to you/the subject' (score 2), "infrequently describes you/the subject" (score 1) or "does not describe you/the subject" (score 0)
Autism Spectrum Disorder in Adults Screening questionnaire (ASDASQ)	Nylander and Gillberg (2001) [58]	Clinician rating scale S	Adults Psychiatric out patients	Questionnaire completed by a clinician according to observations of patients' behavior Nine symptom/impairment-orientated questions concerning diagnostic issues + One question relating to previous contact with child and adolescent psychiatric services
Autism Spectrum Quotient (AQ-50) Original version	Baron-Cohen, Wheelwright, Skinner, Martin & Clubley (2001) [59]	Self-Questionnaires S	Adults IQ in normal or high range	50 questions covering five domains: social skill, attention switching, attention to detail, communication, imagination Answer options: definitely agree, slightly agree, slightly disagree, definitely disagree. Half the items worded to produce a "disagree" response and half an "agree" response in an ASD patient. Score tending to 50 indicates high autistic traits
The Short Autism Spectrum Quotient (AQ-10)	Allison, Auyeung, Baron-Cohen (2012) [60]			10 questions derived from AQ-50–same answer options
Autism Quotient-Short form (AQ-S or AQ28)	Hoesktra et al. (2011) [61]			28 questions derived from AQ-50–same answer options
Autism Quotient–20 (AQ-20)	Brugha et al. (2012) [62]			20 questions derived from AQ-50–same answer options
Autism Quotient Japanese version–21 (AQ-J-21)	Kurita, Koyama & Osada (2005) [63]			21 questions derived from AQ-50–same answer options
Autism Quotient–39 (AQ-39)	Lau, Kelly & Peterson (2013) [64]			39 questions derived from AQ-50–same answer options
Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R)	Ritvo et al. (2011) [65]	Self-questionnaire S or HD	Adults With average or above average intelligence	80 items covering four domains: social relatedness; circumscribed interests; language; sensorimotor and stereotypes. Answer options: present now and when young (before 16 years), only now, only when young (before 16 years), never 64 symptom-based questions and 16 non-symptom-based responses, scored in reverse order The higher the score, the higher the risk of ASD

Table 1 (*Continued*)

Instruments	Authors and year of publication	Type of instrument and purpose: Screening *(S)* Help to diagnosis *(HD)*	Target population	Short description
RAADS-14 screen	Erikkson, Andersen & Bejerot (2013) [66]	Self-questionnaire S		14 items selected based on the RAADS-R. Same answer options and scoring rules. Only one item with reversed score
Sensory Reactivity in Autism Spectrum (SR-AS)	Elwin, Schroder, Ek & Kjellin (2015) [67]	Self-questionnaire S	Adults With no intellectual disabilities	32 Items in four subscales: High Awareness/Hyper-reactivity (14 items); Low Awareness/Hypo-reactivity (10 items); Sensory interest four items); motor (four items) Response format: a four-point Likert-type scale ranging from 0 (totally disagree) to 3 (totally agree), with all questions worded positively. A high total score indicates a high frequency of sensory reactivity
Social Responsiveness Scale 2nd edition-Adult Form (SRS-Adult or SRS-A)	Constantino & Gruber (2012) [68]	Questionnaire (self or other report administration) S or HD	Adult With average or above average intelligence	65 items in five domains: social awareness, social cognition, social communication, social motivation, restricted interests and repetitive behavior, and overall total score Two possible administrations: self-report or parent-, other relative-, and spouse-report Rating: on a four-point Likert scale (0–3) that ranges from "not true" to "almost always true" Total score between 0 and 175, 175 indicating high degrees of social impairment
SRS2-AS30	Duku et al. (2013) [69]	Questionnaires (self or other report administration) S	Adult (> 16) With average or above average intelligence	30 items based on the SRS-2
SRS2-AS11	Kanne et al. (2009) [70] Reiersen et al. (2008) [71]			11 items based on the SRS-2

[a] ADOS-G.
[b] ADOS-2.

Table 2
Sample characteristics of identified studies.

Instruments	Study references	Population and subgroups	Age	Cognitive level	% Males or sex ratio
AAA	Baron-Cohen et al. (2005) [53]	*N* = 42 patients referred to Asperger Diagnostic Clinic 37 AS (DSM-IV) without ID 5 Non-AS	M = 34.1 yrs (10.6)	No other information	86% in total (92% for AS without ID)
	Kuenssberg & McKenzie (2011) [72]	*N* = 153 AS (DSM-IV-TR) without ID *No control group*	17–75 years, M = 33 (11)	No other information	72%
ADOS-G/ADOS-2 *Module 4 original algorithm*	Bastiaansen et al. (2011) [73]	*N* = 38 AS (DSM-IV-TR) *N* = 18 SCH *N* = 16 Psychopathy *N* = 21 Controls	18–66 yrs, M = 31.82 yrs (11.24) 19–61 yrs, M = 37 yrs (10.73) 23–60 yrs, M = 39 yrs (10.67) 21–53 yrs, M = 34.24 yrs (9.14)	QI, M = 101.14 (14.67) QI, M = 89.17 (13.89) QI, M = 92.73 (16.10) QI, M = 97.19 (16.37)	100%
	Brugha et al. (2012) [62]	Community sample *N* = 618 assessed by ADOS-G	Age = 16–75 yrs+ *N* = 110 aged 16–24 yrs *N* = 518 aged > 24 yrs	General population (2% mental retardation expected)[a]	M/F = 1:1
ADOS-G/ADOS-2 *Module 4 revised algorithm*	Hus & Lord (2014) [74]	*N* = 177 Autistics/*N* = 170 Other-DSM IV-TR AS	*M* = 20.12 yrs (6.30)/21.14 yrs (7.79)	IQ ≥ 70 (*n* = 303) VIQ ≥ 85 (n = 259)	80%
		90 Comparison group: Non-AS clinical referrals (mood and/or anxiety disorders, ADHD, ODD. . .) + Healthy Controls	13.33–62.25 yrs, *M* = 25.17 yrs (12.35)	VIQ, *M* = 91.94 (27.37) *(WASI, Differential Ability Scales, Ravens' Progressive Matrices)*	
	De Bildt (2015) [75]	*Idem Bastiaansen et al. (2011)*			
AMSE	Grodberg et al. (2014) [56]	*N* = 55 self-referred patients *N* = 25 (46%) with clinical diagnosis of ASD (DSM-5) *N* = 30 non-ASD (mood/anxiety disorder, BP, ADHD, SCH)	18–45 yrs, *M* = 28.90 yrs (8.29)	"Verbally fluent" (no IQ assessment)	Not reported
ASBQ	Horwitz (2016) [57]	Total sample: patients attending outpatient mental health center *N* = 249 AS (DSM-IV) 3 Non-AS-Clinical groups: ADHD (*N* = 34)/Mood Disorder (*N* = 59)/SCH (*N* = 21) *N* = 30 Healthy controls	17–87 yrs *M* = 32 yrs (12) *M* = 26 yrs (9)/*M* = 42 (12)/ *M* = 41 (8) *M* = 19 yrs (0.4)	Total sample: "No patients with ID" (with no other information)	81% 84%/27%/100% 60%
ASDASQ	Nylander & Gillberg (2001) [58]	*N* = 1323 patients of outpatient clinic, including *N* = 19 with diagnosis of AS (DSM-IV, ICD-10, Gillberg criteria)	22–60 yrs	1 case of mental retardation in 19 screened patients with AS	41%
	Chang et al. (2003) [76]	*N* = 660 patients of outpatient clinic, including 4 with diagnosis of AS (DSM-IV)	15–93 yrs, *M* = 39 yrs (15.2)	No case of mental retardation on 4 screened patients with AS–4 cases in 22 patients with high scores on ASDASQ	56.70%
Autism Spectrum Quotient (AQ-50) *Original version*	Baron-Cohen et al. (2001) [59]	*N* = 58 AS (DSM-IV) without ID *N* = 174 general population (randomly selected)/840 students/16 Mathematicians	*M* = 31.6 (11.8) 16.5–58.3 yrs *M* = 37 (7.7)/*M* = 37 (2.9)/ *M* = 17.4 (1)	15 AS Individuals randomly selected for checking IQ with WAIS-R (all > 85)	77.60% 44%/54%/94%
	Baron-Cohen et al. (2005) [53]	*N* = 42 patients referred to Asperger Diagnostic Clinic 37 AS (DSM-IV) without ID 5 Non Autism Spectrum	*M* = 34.1 yrs (10.6)	No other information	86% in total (92% for AS without ID)
	Kurita, Koyama & Osada (2005) [63]	*N* = 25 AS (DSM-IV) *N* = 215 Controls from general population	*M* = 24.2 yrs (5.5) *M* = 30.4 yrs (5.8)	19 AS full-scale IQ *M* = 101.6 (15.4) *(WAIS-R)* + 6 other AS "judged to have normal intelligence"	96% 40%

Table 2 (*Continued*)

Instruments	Study references	Population and subgroups	Age	Cognitive level	% Males or sex ratio
	Woodbury-Smith et al. (2005) [77]	*N* = 100 consecutive referrals to Asperger Diagnostic Clinic Including in the end: *N* = 73 AS (DSM-IV)/27 Non-AS	18–69 yrs (median = 32) *M* = 35.6 yrs (6.63)/*M* = 26.2 yrs (9.39)	"People with a history of mental retardation specifically excluded from assessment"	4 M:1F
Autism Spectrum Quotient (AQ-50) *Original version*	Wakabayashi et al. (2006) [78]	*N* = 57 AS (DSM-IV) without ID *N* = 194 Controls from general population/ *N* = 1050 Students	18–57 yrs, *M* = 26.9 (7.88) *M* = 33.6 (6.2)/*M* = 20.3 (1.9)	AS = "assumed to have an IQ in the normal range because all of them had completed senior-high school"	77.20% 53%/53%
	Hoekstra et al. (2008) [79]	Controls: *N* = 961 students + *N* = 302 General population 3 DSM-IV Clinical groups: *N* = 12 AS without ID; *N* = 12 OCD; *N* = 12 generalized anxiety disorder	*M* = 21.19 (3.69)/*M* = 35.68 (6.33) 19–57 years	Only patients who had successfully completed an educational degree	Not reported 83% in all groups
	Ketelaars et al. (2008) [80]	*N* = 36 referrals to Autism Centre: 15 AS (DSM-IV)/21 Non AS *N* = 369 patients from General Outpatient Clinic	18–24 yrs, *M* = 22 (5)/18–55 yrs, *M* = 27 (9) 14–72 yrs, *M* = 35 (11)	IQ > 80 *(no tool specified)*	80%/86% 86%
	Lepage et al. (2009) [81]	*N* = 23 AS (DSM-IV) *N* = 50 science students + *N* = 50 humanities students	18–55 yrs, *M* = 27.91 (10.70) *M* = 22.86 (4.42)/*M* = 21.96 (2.65)	"All AS reported an IQ in the normal range"	65% 68%/70%
	Sizoo et al. (2009) [82]	*N* = 76 AS (DSM-IV) *N* = 53 ADHD	*M* = 34.1 yrs (11.9) *M* = 32.1 yrs (11.4)	Exclusion criteria: IQ < 80 *(WAIS versions and Groninger Intelligence Test)*	82% 66%
	Naito et al. (2010) [83]	*N* = 51 AS (DSM-IV-TR) *N* = 46 SCH	17–44 yrs, *M* = 28.8 (9.4) 17–57 yrs, *M* = 34.1 (9.6)	FSIQ range 70–122 (WAIS-R or WAIS-III)	78% 50%
	Wouters & Spek (2011) [84]	*N* = 21 Autism (DSM-IV-TR) *N* = 21 SCH *N* = 21 Non clinical	*M* = 44 yrs (11.7) *M* = 40.9 yrs (8.1) *M* = 40.8 yrs (9.9)	Verbal comprehension Index > 80 *M* = 110.5 (13.6) (WAIS III)	100%
	Bishop & Seltzer (2012) [85]	*N* = 65 Community sample of patients with AS (DSM-IV)	18–52.6 yrs, *M* = 24.97 (8.22)	22% with FSIQ < 70 (*M* = 89.61; range: 35–142) *(Wide Range Intelligence Test)*	75%
	Booth et al. (2013) [86]	*N* = 149 AS (DSM-IV-TR) without ID *N* = 134 Controls from general population	For *N* = 140 age range: 17–75 yrs 17–65 yrs (*M* = 29.6)	No other information University Community	Around 2.5 M:1F 25%
	Broadbent, Galic & Stokes (2013) [87]	*N* = 104 ASD (DSM-5) *N* = 129 Typicals	*M* = 33.12 yrs (14.04) *M* = 27.28 yrs (8.06)	All IQ > 70 *(no tool specified)*	56% 36%
	Pisula et al. (2013) [88]	*N* = 60 AS (ICD-10) *N* = 60 controls/*N* = 2819 students	17–44 yrs, *M* = 2.10 (5.87) 18–30 yrs, *M* = 22.57 (6.27)	Exclusion if IQ < 90 *(WAIS-R, polish version)*	65% 65%/47%
	Nishiyama et al. (2014) [89]	*N* = 64 AS (DSM-IV) *N* = 3147 non clinical (workers + students)	16–65 yrs *M* = 25.5 (11.1) 17–72 yrs, *M* = 30.4 (11.8)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70% 39.10%
	Lugnegård, Hallerbäck & Gillberg (2015) [90]	*N* = 51 Asperger Syndrome (DSM-IV-TR)	*M* = 27.1 yrs (4.1)	All judged to be in the normal range of IQ	66.60%
		N = 36 SCH/; *N* = 49 non clinical	*M* = 29.1 yrs (4.2)/*M* = 28.6 yrs (9.2)		63.8%/39%
	Zangh et al. (2016) [91]	*N* = 32 AS (DSM-IV-TR) *N* = 37 SCH/; *N* = 38 OCD	All > 18 yrs, *M* = 19.41 (3.88) *M* = 20.95 yrs (3.67)/*M* = 21.29 (3.15)	IQ, *M* = 102.3 (14.4) IQ, *M* = 106.6 (16.8)/*M* = 103.2 (11.1)	81% males 81% males/81% males
		N = 38 Healthy	*M* = 21.32 yrs (3.32)	IQ, *M* = 108.9 (12.9) (*Standardized Raven Test*)	79% males
AQ-39	Lau, Kelly & Peterson (2013) [64]	*N* = 141 AS (DSM-IV) *N* = 314 Non AS (possibly no disorder + other disorders)	*M* = 40.56 yrs (7.80) *M* = 40.74 yrs (8.84)	High school graduate or university degree	30.5% Males 22.6% Males

Table 2 (*Continued*)

Instruments	Study references	Population and subgroups	Age	Cognitive level	% Males or sex ratio
AQ-10	Allison, Auyeung & Baron-Cohen (2012) [60]	Validation on: $N = 419$ Non-Clinical subjects	$M = 32.93$ yrs (12.20)		Not reported
		$N = 225$ AS (DSM-IV) without ID	$M = 35.62$ yrs (13.04)	No other information	
	Booth et al. (2013) [86]	$N = 149$ AS (DSM-IV-TR) without ID	For $N = 140$ age range: 17–75 yrs	No other information	Around 2.5 M:1F
		$N = 134$ Controls from general population	17–65 yrs ($M = 29.6$)	University Community	25%
	Nishiyama et al. (2014) [89]	$N = 64$ AS (DSM-IV)	16–65 yrs, $M = 25.5$ (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		$N = 3147$ non clinical (workers + students)	17–72 yrs, $M = 30.4$ (11.8)		39.10%
	Sizoo et al. (2015) [92]	$N = 210$ referrals for AS assessment, including $N = 139$ diagnosed with AS (no diagnostic classification specified)	$M = 39.4$ yrs (12.50)	No intellectual impairment (exclusion criteria)	75.70%
		$N = 63$ non clinical from general population	$M = 39.3$ yrs (13.80)		58.70%
AQ-S or AQ28	Hoesktra et al. (2011) [61]	$N = 274$ with a formal DSM-IV diagnosis of Asperger	$M = 35.37$ yrs (13.05)	No other information	57%
		$N = 1263$ (302 Dutch General population/961 students)	$M = 21.19$ yrs (3.69)/$M = 35.68$ (6.33)		40%
		$N = 1121$ Dutch general population/; $N = 1838$ English students	$M = 45.63$ (14.74)/$M = 20.90$ (2.47)		43%/40%
	Booth et al. (2013) [86]	$N = 149$ DSM-IV diagnosis of Autism or Asperger syndrome	For $N = 140$ age range: 17–75 yrs	No other information	Around 2.5 M:1F
		$N = 134$ Controls from general population	17–65 yrs ($M = 29.6$)	University Community	25%
	Kuennsberg et al. (2014) [93]	$N = 148$ AS (DSM-IV-TR) without ID	17–62 yrs, $M = 33.3$ (10.7)	No other information	72% Males
	Nishiyama et al. (2014) [89]	$N = 64$ AS (DSM-IV)	16–65 yrs, $M = 25.5$ (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		$N = 3147$ non-clinical (workers + students)	17–72 yrs, $M = 30.4$ (11.8)		39.10%
	Sizoo et al. (2015) [92]	$N = 210$ referrals for AS assessment, including $N = 139$ diagnosed with AS (no diagnostic classification specified)	$M = 39.4$ yrs (12.50)	No intellectual impairment (exclusion criteria)	75.70%
		$N = 63$ non clinical from general population	$M = 39.3$ yrs (13.80)		58.70%
AQ-20	Brugha et al. (2012) [62]	Community sample $N = 618$ assessed by ADOS-G	Age = 16–75 yrs+ $N = 110$ aged 16–24 yrs $N = 518$ aged > 24 yrs	General population (2% mental retardation expected)[a]	M/F = 1:1
	Nishiyama et al. (2014) [89]	$N = 64$ AS (DSM-IV)	16–65 yrs, $M = 25.5$ (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		$N = 3147$ non clinical (workers + students)	17–72 yrs, $M = 30.4$ (11.8)		39.10%
AQ-J-21	Kurita, Koyama & Osada (2005) [63]	$N = 25$ AS (DSM-IV)	$M = 24.2$ yrs (5.5)	19 AS full-scale IQ $M = 101.6$ (15.4) *(WAIS-R)* + 6 other AS "judged to have normal intelligence"	96%
		$N = 215$ Controls from general population	$M = 30.4$ yrs (5.8)		40%
	Nishiyama et al. (2014) [89]	$N = 64$ AS (DSM-IV)	16–65 yrs, $M = 25.5$ (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		$N = 3147$ non clinical (workers + students)	17–72 yrs, $M = 30.4$ (11.8)		39.10%
Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R)	Ritvo et al. (2011) [65]	$N = 201$ AS (DSM-IV-TR)	Age > 18 yrs, $M = 31.45$	IQ, $M = 119$ (IQ > 80) *(WAIS or WASI)*	72.10%
		$N = 302$ other disorders (SP, SCH, BP, Depression, OCD, GAD, PTSD...)	$M = 42.04$	IQ, $M = 112$	44.40%
		$N = 276$ No disorder	$M = 41.51$	IQ, $M = 116$	41.30%

Table 2 (*Continued*)

Instruments	Study references	Population and subgroups	Age	Cognitive level	% Males or sex ratio
	Andersen et al. (2011) [94]	*N* = 71 AS (DSM-IV-TR)	26–62 yrs, *M* = 31 (9)	IQ > 70 for all included subjects *(no tool specified)*	51%
		N = 197 Control group: 174 Typicals + 13 other disorders *(Sch Pers Dis, ADHD, SAD, depression…)*	19–75 yrs, *M* = 34 (13)		80:116
	Sizoo et al. (2015) [92]	*N* = 210 referrals for AS assessment, including *N* = 139 diagnosed with AS (no diagnostic classification specified)	*M* = 39.4 yrs (12.50)	No intellectual impairment (exclusion criteria)	75.70%
		N = 63 non clinical from general population	*M* = 39.3 yrs (13.80)		58.70%
RAADS-14 Screen	Erikkson, Andersen & Bejerot (2013) [66]	Phase II: 18 items pilot version		Exclusion of mental retardation	
		N = 58 AS (DSM-IV)	*M* = 33.7 yrs (11.5)		58%
		N = 139 other disorders/*N* = 590 No disorder/ *N* = 43 ADHD	*M* = 35.6 yrs (12.4)/45.0 (10.9)/ 37.0 (12.2)		45%/21%/42%
		Phase III: discriminatory properties of Raads-14 *N* = 77 AS (DSM-IV)	*M* = 35.2 yrs (10.9)/		43%
		N = 301 ADHD/*N* = 370 other disorders *(Mood Dis, Anxiety Dis, Psychotic Dis…)*	*M* = 32.6 yrs (12.0)/32.8 (9.8)		40%/37%
SR-AS	Elwin et al. (2015) [67]	*N* = 71 AS (ICD-10)	18–44 yrs = 81.7%	Exclusion criteria = intellectual disability	36.60%
		N = 162 Controls from general population	18–44 yrs = 76.5%	42.60%	
SRS-Adult *Informant Administration*	Bolte (2012) [95]	*N* = 20 AS (ICD-10)	18–36 yrs, *M* = 27.5 (6.5)	Full IQ *M* = 112.8 (16.4) *(Wechsler Intelligence Scales for Adults)*	75%
		N = 62 Other disorders/; = 163 No disorder	28–56 yrs, *M* = 40.0 (6.6)/19–79 yrs		55%/45%
	Takei et al. (2014) [96]	*N* = 65 AS (DSM-IV-TR)	19–51 yrs, *M* = 27.3 (7.7)	Individuals judged to have normal intellectual functioning	67.70%
		N = 78 Other disorders (depressive disorder; SCH, BP, anxiety)	20–59 yrs, *M* = 34.8 (10.6)		50%
		N = 592 students	19–59 yrs		41.60%
SRS-Adult *Self-report Administration*	Nishiyama et al. (2014) [89]	*N* = 64 AS (DSM-IV)	16–65 yrs, *M* = 25.5 (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		N = 3147 non-clinical (workers + students)	17–72 yrs, *M* = 30.4 (11.8)		39.10%
SRS2-AS30 *(short form)*	Nishiyama et al. (2014) [89]	*N* = 64 AS (DSM-IV)	16–65 yrs, *M* = 25.5 (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		N = 3147 non-clinical (workers + students)	17–72 yrs, *M* = 30.4 (11.8)		39.10%
SRS2-AS11 *(short form)*	Nishiyama et al. (2014) [89]	*N* = 64 AS (DSM-IV)	16–65 yrs, *M* = 25.5 (11.1)	IQ > 80 *(WISC-III or WAIS-R, Japanese versions; Tanaka–Binet Intelligence Scale)*	66.70%
		N = 3147 non-clinical (workers + students)	17–72 yrs, *M* = 30.4 (11.8)		39.10%

ADHD: Attention Deficit Hyperactivity Disorder; AS: Autism Spectrum; ASD: Autism Spectrum Disorder; AU: Childhood Autism; BP: Bipolar Disorder; FSIQ: Full Scale Intellectual Quotient; GAD: Generalized Anxiety Disorder; HFA: High Functioning Autism; ID: intellectual disability; LD: learning Disability; OCD: obsessive-compulsive disorder; PDDNOS: Pervasive Development Disorder Not Otherwise Specified; SCH: Schizophrenia; Schy Pers Dis: Schyzotypal personality disorder; SAD: Social Anxiety Disorder; SP: Social Phobia; VIQ: Verbal Intellectual Quotient.

[a] Comment from authors of the review.

Table 3
Methodological quality of studies according to COSMIN checklist.

Instrument	Study references	Internal Consistency	Reliability (Test-Retest & Interrater)	Content Validity	Structural Validity	Hypotheses testing
AAA	Baron-Cohen et al. (2005) [53]	–[a]	–	–	–	Poor *statistics*[b]
	Kuenssberg & McKenzie (2011) [72]	–	–	–	Poor *sample*	–
ADOS-G/ADOS-2 *Module 4 original algorithm*	Bastiaansen et al. (2011) [73]	Fair *missing items*	Poor *only one measurement*	–	–	Fair *missing items/hypothesis/ comparator*
	Brugha et al. (2012) [62]	–	–	–	–	Poor *comparator*
ADOS-G/ADOS-2 *Module 4 revised algorithm*	Hus & Lord (2014) [74]	Excellent	–	–	Excellent	
	De Bildt (2015) [75]	–	–	–	–	Fair *missing items*
AMSE	Grodberg et al. (2014) [56]	–	Poor *only one measurement*	Unknown *(Pilot study?)*	–	–
ASBQ	Horwitz (2016) [57]	Fair *missing items*	–	–	Fair *missing items*	Fair *missing items*
ASDASQ	Nylander & Gillberg (2001) [58]	Poor *unidimensionality*	Fair *missing items/sample*	–	–	–
	Chang et al. (2003) [76]	–	Fair *missing items/statistics*	–	–	–
Autism Spectrum Quotient (AQ-50) *Original version*	Baron-Cohen et al. (2001) [59]	Poor *unidimensionality*	Poor *sample*	Unknown		Fair *missing items/Hypothesis*
	Baron-Cohen et al. (2005) [53]	–	–	–	–	Fair *missing items/sample*
	Kurita, Koyama & Osada (2005) [63]	Poor *domain/sample*	Poor *sample*	–	–	Fair *missing items/hypothesis*
	Woodbury-Smith et al. (2005) [77]	–	–	–	–	Fair *missing items/hypothesis*
	Wakabayashi et al. (2006) [78]	Poor *unidimensionality*	Fair *missing items/statistics*	–	–	Fair *missing items/hypothesis*
	Hoekstra et al. (2008) [79]	Good	Good[c]	–	Good	Good
	Ketelaars et al. (2008) [80]	–	–	–	–	Fair *missing items*
	Lepage et al. (2009) [81]	Poor *unidimensionality*	–	–	Poor no factor analysis	Fair *missing items*
	Sizoo et al. (2009) [82]	–	–	–	–	Fair *missing items/hypothesis*
	Naito et al. (2010) [83]	Poor *unidimensionality*	–	–	–	Fair *missing items*
	Wouters & Spek (2011) [84]	–	–	–	–	Fair *missing items and comparator*
	Bishop & Seltzer (2012) [85]	–	–	–	–	Fair *comparator*
	Booth et al. (2013) [86]	–	–	–	–	Fair *missing items*
	Broadbent, Galic & Stokes (2013) [87]	Good	Poor *sample*	–	–	Good
Autism Spectrum Quotient (AQ-50) *Original version*	Pisula et al. (2013) [88]	Good	Fair *statistics*	–	–	Fair *hypothesis*
	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	Poor *statistics*	Fair *comparator/hypothesis*
	Lugnegård, Hallerbäck & Gillberg (2015) [90]	–	–	–	–	Fair *missing items/hypothesis*
	Zhang et al. (2016) [91]	–	–	–	–	Fair *missing items*
AQ-39	Lau, Kelly & Peterson (2013) [64]	Fair *missing items*	Fair *missing items*	–	Fair *missing items*	Fair *missing items/hypothesis*

Table 3 (*Continued*)

Instrument	Study references	Internal Consistency	Reliability (Test-Retest & Interrater)	Content Validity	Structural Validity	Hypotheses testing
AQ-10	Allison, Auyeung & Baron-Cohen (2012) [60]	Poor *unidimensionality*	–	–	–	Fair *missing items/hypothesis*
	Booth et al. (2013) [86]	–	–	–	–	Fair *missing items*
	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	P *statistics*	Fair *hypothesis*
	Sizoo et al. (2015) [92]	Poor *unidimensionality*	–	–	–	*Fair missing items/hypothesis*
AQ-S ou AQ28	Hoesktra et al. (2011) [61]	Good	–	–	Good	Fair *hypothesis*
	Booth et al. (2013) [86]	–	–	–	–	Fair *missing items*
	Kuennsberg et al. (2014) [93]	Good	–	–	Good	–
	Nishiyama et al. (2014) [89]	Poor *no domain*	Good	–	Poor *statistics*	Fair *hypothesis*
	Sizoo et al. (2015) [92]	Poor *no domain*	–	–	–	*Fair missing items/hypothesis*
AQ-20	Brugha et al. (2012) [62]	–	–	–	–	*Fair missing items/hypothesis*
	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	Poor *statistics*	Fair *hypothesis*
AQ-J-21	Kurita, Koyama & Osada (2005) [63]	Poor *unidimensionality*	Poor *sample*	–	–	Fair *missing items/hypothesis*
	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	Poor *statistics*	Fair *hypothesis*
Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R)	Ritvo et al. (2011) [65]	F *missing items*	Fair *sample/statistics*	Good	Fair *missing items*	Fair *missing items/hypothesis*
	Andersen et al. (2011) [94]	F *missing items*	Poor *sample*	–	–	Fair *hypothesis/sample*
	Sizoo et al. (2015) [92]	Poor *no domain*	–	–	–	Fair *missing items/hypothesis*
RAADS-14 Screen	Erikkson, Andersen & Bejerot (2013) [66]	Good	–	Good	Good	Fair *hypothesis*
SR-AS	Elwin et al. (2015) [67]	Fair *missing items*	–	Good	Fair *missing items*	–
SRS-Adult *Informant Administration*	Bolte (2012) [95]	Poor *unidimensionality/sample*	–	–	–	Poor *comparators*
	Takei et al. (2014) [96]	Good	–	–	Good	Poor *comparators*
SRS-Adult *Self-Report Administration*	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	Poor *statistics*	Fair *hypothesis*
SRS2-AS30 *(short form)*	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	Poor *statistics*	Fair *hypothesis*
SRS2-AS11 *(short form)*	Nishiyama et al. (2014) [89]	Poor *unidimensionality*	Good	–	Poor *statistics*	Fair *hypothesis*

[a] Property was not explored in the study.
[b] The reason(s) leading to a "Poor" or "Fair" rating is/are mentioned in italics.
[c] For self-questionnaires, interrater measurement was considered to be non-applicable.

Table 4
QUADAS-2 assessment of diagnostic accuracy studies.

Instrument	Study references	Risk of bias				Overall judgement[b]	Applicability concerns			Overall judgement[b]
		Patient Selection	Index test	Reference standard	Flow and Timing		Patient Selection	Index test	Reference standard	
AAA	Baron-Cohen et al. (2005) [53]	☺[a]	☹	☹	☺	*At risk*	☺	☺	☺	*Concerns*
	Kuenssberg & McKenzie (2011) [72]	N/A[c]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ADOS-G/ADOS-2 *Module 4 original algorithm*	Bastiaansen et al. (2011) [73]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
	Brugha et al. (2012) [62]	☺	☹	☹	☹	*At risk*	☹	☺	☹	*Concerns*
ADOS-G/ADOS-2 *Module 4 revised algorithm*	Hus & Lord (2014) [74]	☹	☺	☺	☺	*At risk*	☺	☺	☺	*Low concern*
	De Bildt (2015) [75]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
AMSE	Grodberg et al. (2014) [56]	😐	☺	☺	☺	*At risk*	☺	☺	☺	*Low concern*
ASBQ	Horwitz (2016) [57]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ASDASQ	Nylander & Gillberg (2001) [58]	☺	☹	☹	☺	*At risk*	☺	☺	☹	*Concerns*
	Chang et al. (2003) [76]	☺	☺	☹	☹	*At risk*	☺	☺	☺	*Low concern*
Autism Spectrum Quotient (AQ-50) *Original version*	Baron-Cohen et al. (2001) [59]	☹	☹	☺	☺	*At risk*	☹	☹	☺	*Concerns*
	Baron-Cohen et al. (2005) [53]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Kurita, Koyama & Osada (2005) [63]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Woodbury-Smith et al. (2005) [77]	☺	☹	😐	☺	*At risk*	☺	☺	☺	*Low concern*
	Wakabayashi et al. (2006) [78]	☹	☹	☺	☺	*At risk*	☹	☹	☺	*Concerns*
	Hoekstra et al. (2008) [79]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Ketelaars et al. (2008) [80]	😐	☺	☺	☺	*At risk*	☺	☺	☺	*Low concern*
	Lepage et al. (2009) [81]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Sizoo et al. (2009) [82]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Naito et al. (2010) [83]	☹	☹	☺	☺	*At risk*	☹	☹	☺	*Concerns*
	Wouters & Spek (2011) [84]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Bishop & Seltzer (2012) [85]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Booth et al. (2013) [86]	☹	☺	☺	☹	*At risk*	☹	☺	☺	*Concerns*
	Broadbent, Galic & Stokes (2013) [87]	☹	☹	☺	☺	*At risk*	☹	☹	☺	*Concerns*
	Pisula et al. (2013) [88]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
Autism Spectrum Quotient (AQ-50)	Lugnegård, Hallerbäck & Gillberg (2015) [90]	☹	☺	☹	☺	*At risk*	☹	☺	☹	*Concerns*
	Zhang et al. (2016) [91]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
AQ-39	Lau, Kelly & Peterson (2013) [64]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 4 (*Continued*)

Instrument	Study references	Risk of bias				Overall judgement[b]	Applicability concerns			Overall judgement[b]
		Patient Selection	Index test	Reference standard	Flow and Timing		Patient Selection	Index test	Reference standard	
AQ-10	Allison, Auyeung & Baron-Cohen (2012) [60]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Booth et al. (2013) [86]	☹	☺	☺	☹	*At risk*	☹	☺	☺	*Concerns*
	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Sizoo et al. (2015) [92]	☹	☹	☹	☹	*At risk*	☺	☺	☺	*Low concern*
AQ-S ou AQ28	Hoesktra et al. (2011) [61]	☹	☹	☺	☹	*At risk*	☹	☺	☺	*Concerns*
	Booth et al. (2013) [86]	☹	☹	☺	☹	*At risk*	☹	☺	☺	*Concerns*
	Kuennsberg et al. (2014) [93]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Sizoo et al. (2015) [92]	☺	☺	☺	☺	*At risk*	☺	☺	☺	*Low concern*
AQ-20	Brugha et al. (2012) [62]	☺	☹	☹	☺	*At risk*	☹	☺	☹	*Concerns*
	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
AQ-J-21	Kurita, Koyama & Osada (2005) [63]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R)	Ritvo et al. (2011) [65]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
	Andersen et al. (2011) [94]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
	Sizoo et al. (2015) [92]	☺	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
RAADS-14 Screen	Erikkson, Andersen & Bejerot (2013) [66]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
SR-AS	Elwin et al. (2015) [67]	☹	☹	☺	☺	*At risk*	☹	☺	☺	*Concerns*
SRS-Adult *Informant Administration*	Bolte (2012) [95]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
	Takei et al. (2014) [96]	☹	☹	☺	☺	*At risk*	☺	☺	☺	*Low concern*
SRS-Adult *Self-report Administration*	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
SRS2-AS30 *(short form)*	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*
SRS2-AS11 *(short form)*	Nishiyama et al. (2014) [89]	☹	☺	☺	☺	*At risk*	☹	☺	☺	*Concerns*

[a] ☺: low risk of bias/low concern about applicability; 😐: unclear risk of bias/unclear concern about applicability; ☹: high risk of bias/high concern about applicability.

[b] Quadas criteria for overall judgement of risks of bias and concerns about applicability [51]: "If a study is judged as "low" on all domains relating to bias or applicability then it is appropriate to have an overall judgment of "low risk of bias" or "low concern regarding applicability" for that study. If a study is judged "high" or "unclear" on one or more domains then it may be judged "at risk of bias" or as having "concerns regarding applicability".

[c] N/A: Not applicable (no diagnostic accuracy study).

Table 5
Synthesis for each instrument: qualitative assessment of extracted data weighted by methodological quality assessment of studies (COSMIN and QUADAS-2).

Instrument	Psychometric properties and overall COSMIN rating					Diagnostic Validity (Sensibility (Sens)/Specificity (Spec)/Area Under the Curve (AUC)) and overall QUADAS-2 judgement
	Internal Consistency	Reliability	Content Validity	Structural Validity	Hypotheses testing	
AAA	*N/A*[a]	*N/A*	*N/A*	No confirmation of suggested structure, *with conflicting evidence*[b]	Satisfactory Sens/Spec, *with unsatisfactory evidence*[c]	Satisfactory Sens/Spec, *with unsatisfactory evidence*[c]
ADOS-G/ADOS-2 *Module 4 original algorithm*	Satisfactory, *with limited evidence*	Satisfactory, *with conflicting evidence*	*N/A*	N/A	Mixed convergent validity (depending on domains), with limited evidence Satisfactory discriminant validity (except for SCH), *with limited evidence*	Satisfactory AUC, mixed results for Sens (0.68 to 1) and Spec (0.73 to 0.86), *with unsatisfactory evidence*
ADOS-G/ADOS-2 *Module 4 revised algorithm*	Satisfactory for social domain - Unsatisfactory for RRBs, *with strong evidence*	*N/A*	*N/A*	Satisfactory, with *strong evidence*	Satisfactory discriminant validity (except for SCH), *with limited evidence*	Mixed results for Sens (0.61 to 0.905) and Spec (0.50 to 0.822), *with unsatisfactory evidence*
AMSE	*N/A*	Satisfactory, *with conflicting evidence*	*Unknown (Pilot study?)*	*N/A*	*N/A*	Satisfactory AUC, Sens, and Spec, *with unsatisfactory evidence*
ASBQ	Self report and other version: Satisfactory (Total scores) and Intermediate (domain score), *with limited evidence*	*N/A*	*N/A*	Satisfactory, with *limited evidence*	Discriminant validity: Satisfactory (variable size effect depending on disorders), *with limited evidence*	N/A
ASDASQ	Satisfactory, with *conflicting evidence*	Mostly satisfactory, *with moderate evidence*	*N/A*	*N/A*	*N/A*	Satisfactory Sens and Spec, *with unsatisfactory evidence*
AQ-50 *Original version*	Intermediate to Satisfactory (Total score), Unsatisfactory to Satisfactory (Domains), *with strong evidence*	Satisfactory test-Retest, *with strong evidence*	*Unknown*	Mixed, *with moderate evidence*	Discriminant validity: Satisfactory, *with strong evidence* Convergence: mixed result, *with moderate evidence*	Mixed Accuracy (0.72 to 0.90) Mixed AUC (0.647 to 0.99) Mixed Sens (0 to 0.992) Mixed Spec (0.52 to 0.97) *with unsatisfactory evidence*
AQ-39	Domains: Intermediate to Satisfactory, *with limited evidence*	Satisfactory test-Retest, *with limited evidence*	*N/A*	Satisfactory, *with limited evidence*	Discriminant validity: Satisfactory, *with limited evidence*	N/A
AQ-10	Mixed results, *with conflicting evidence*	Satisfactory test-Retest, *with moderate evidence*	*N/A*	Satisfactory, *with conflicting evidence*	Discriminant validity: Satisfactory, *with moderate evidence* Convergence: Satisfactory, *with limited evidence*	Mixed results for AUC (0.650 to 0.951), Sens (0.62 to 0.88) and Spec (0.66 to 0.91), *with unsatisfactory evidence*
AQ-S or AQ28	Satisfactory for total score - Unsatisfactory to Satisfactory for domains *with strong evidence*	Satisfactory test-Retest, with moderate evidence	*N/A*	Satisfactory, *with strong evidence*	Discriminant validity: Satisfactory, *with moderate evidence*	Mixed results for AUC (0.653 to 0.97), Sens (0.57 to 0.97) and Spec (0.70 to 0.82), *with unsatisfactory evidence*

Table 5 (*Continued*)

Instrument	Psychometric properties and overall COSMIN rating					Diagnostic Validity (Sensibility (Sens)/Specificity (Spec)/Area Under the Curve (AUC)) and overall QUADAS-2 judgement
	Internal Consistency	Reliability	Content Validity	Structural Validity	Hypotheses testing	
AQ-20	Unsatisfactory, *with conflicting evidence*	Satisfactory test-Retest, *with moderate evidence*	*N/A*	Intermediate, *with conflicting evidence*	Discriminant validity: Satisfactory, *with limited evidence* Convergence: Unsatisfactory, *with conflicting evidence*	Satisfactory AUC and Sens, low Spec, *with unsatisfactory evidence*
AQ-J-21	Intermediate, *with conflicting evidence*	Satisfactory test-Retest, *with moderate evidence*	*N/A*	Intermediate, *with conflicting evidence*	Discriminant validity: Satisfactory, *with moderate evidence*	Satisfactory AUC, Accuracy, Sens and Spec, *with unsatisfactory evidence*
RAADS-R	Satisfactory (except for Language domain), *with moderate evidence*	Satisfactory test-Retest, *with limited evidence*	Satisfactory, *with moderate evidence*	Mixed result, *with limited evidence*	Discriminant validity: Satisfactory, *with moderate evidence* Convergence: Satisfactory, *with moderate evidence*	Satisfactory accuracy Mixed AUC (0.674 to 0.96) Satisfactory Sens Mixed Spec (0.77 to 1.00) *with unsatisfactory evidence*
RAADS-14 Screen	Satisfactory for total score - Intermediate for domains, *with moderate evidence*	*N/A*	Satisfactory, *with strong evidence*	Mixed result according to groups, *with moderate evidence*	Discriminant validity: Satisfactory, *with limited evidence*	Satisfactory AUC (with ADHD, other than AHDH or no disorder) Satisfactory Sens Mixed Spec (0.46 to 0.64 with disorders; 0.95 with no disorder), *with unsatisfactory evidence*
SR-AS	Satisfactory, *with limited evidence*	*N/A*	Satisfactory, *with strong evidence*	Satisfactory, *with Limited evidence*	*N/A*	Satisfactory AUC, *with unsatisfactory evidence*
SRS-Adult *Informant Administration*	Satisfactory, *with moderate evidence*	*N/A*	*N/A*	Satisfactory, *with moderate evidence*	Discriminant validity: Satisfactory, *with conflicting evidence* Convergence: mixed results according to tools, *with conflicting evidence*	Satisfactory AUC and Sens Satisfactory global Spec, but mixed if according to gender, *with unsatisfactory evidence*
SRS-Adult *Self-report Administration*	Satisfactory, *with conflicting evidence*	Satisfactory, *with moderate evidence*	*N/A*	*N/A*	Discriminant validity: Satisfactory, *with limited evidence* Convergence: Satisfactory, *with limited evidence*	Satisfactory AUC, *with unsatisfactory evidence*
SRS2-AS30 *(short form)*	Satisfactory, *with conflicting evidence*	Satisfactory, *with moderate evidence*	*N/A*	Satisfactory, *with conflicting evidence*	Discriminant validity: Satisfactory, *with limited evidence*	Satisfactory AUC, *with unsatisfactory evidence*
SRS2-AS11 *(short form)*	Satisfactory, *with conflicting evidence*	Satisfactory, *with moderate evidence*	*N/A*	Satisfactory, *with conflicting evidence*	Discriminant validity: Satisfactory, *with limited evidence*	Satisfactory AUC, *with unsatisfactory evidence*

AUC: area under curve; RRBs: repetitive and restricted behaviours; SCH: schizophrenia; Sens: sensibility; Spec: specificity.

[a] *N/A*: No Data Available.

[b] Synthesis of the COSMIN assessment [52]: evidence is considered to be "strong" when several methodologically good articles, or one excellent article, found consistent evidence for or against a measurement property "moderate" for several methodologically fair studies, or one good study; or "limited" for one study of fair quality; other cases are rated as "with conflicting evidence".

[c] Synthesis of the QUADAS assessment: the authors of the review considered that evidence for diagnostic data was "satisfactory" if one or more studies were rated with Low Risk of bias and Low concern about applicability; evidence was considered to be "unsatisfactory" in the other cases.

satisfactory test-retest reliability (ICC = 0.75) [90] and discriminant validity [60,89,92]. However, there were no data about validity content, and evidence about internal consistency and structural validity was contradictory. Data about diagnostic validity were inconsistent, with unsatisfactory evidence: AUC (0.650 to 0.951), sensitivity (0.62 to 0.88), and specificity (0.66 to 0.91). There were five studies on the properties of the *AQ-S*. Two showed satisfactory internal consistency ($\alpha = 0.84$ and $\alpha = 0.86$, total score) and satisfactory structural validity, with strong evidence [61,93]. Moderate evidence was also found for satisfactory test-retest reliability (ICC $r = 0.84$) [89] and discriminant validity [61,92]. Diagnostic accuracy could not be confirmed, due to the heterogeneous results and unsatisfactory evidence: AUC (0.653 to 0.97), sensitivity (0.57 to 0.97), and specificity (0.70 to 0.82). Articles about the properties of *AQ-20* [62,89] and *AQ-J-21* [63,89] showed conflicting evidence about internal reliability and internal structure. Satisfactory values for the AQ20 and AQJ21 (ICC of 0.70 and 0.80, respectively) were reported for test-retest with moderate evidence [89]. Limited evidence was found for the AQ20 concerning discriminant validity, and there were low correlations with the ADOS-Module 4 ($r = 0.24$) [62]. For the AQ-J-21, moderate evidence was found for discriminant validity. Satisfactory values were reported for diagnostic validity for AQ-20 and AQ-J-21, but with high risks of bias and concerns about applicability. Finally, satisfactory internal consistency, test-retest reliability, structural validity, and discriminant validity were found for the *AQ-39*, but the evidence was limited, and there were no data related to diagnostic properties [64].

3.4.2.2. Ritvo Asperger and Autism Diagnostic Scale-Revised (RAADS-R) and short-version. The RAADS-R is an 80 item self-report questionnaire for autistic adults, with a mean normal range of measured intelligence, used for screening or to assist in diagnosis. The properties of the *RAADS-R* were examined in three articles [65,92,94]. Mixed results were found, with moderate evidence for content validity and internal consistency (respectively, three of four domains with $\alpha > 0.80$, and two of four domains with $\alpha > 0.80$) [65,94]. Satisfactory test-retest reliability ($r = 0.987$) [65] and structural validity [65] were reported with limited evidence. Moderate evidence was reported for both discriminant and convergent validity. Thus, scores were significantly higher in the AS group than in clinical control groups [65,92,94], and the correlations were 0.96 with the SRS-Adult [65] and 0.84 with the AQ-50 [94]. Accuracy and sensitivity were satisfactory, but AUC (0.674 to 0.96) and specificity (0.77 to 1.00) were heterogeneous. All psychometric properties (except reliability) of the *RAADS-14* were explored by one article [66], in which strong evidence was found for satisfactory content validity, and moderate evidence for satisfactory internal consistency. Various data were found for structural validity. Sensitivity and AUC were satisfactory, but the evidence was limited by risks of bias.

3.4.2.3. Social Responsiveness Scale, 2nd edition–Adult form (SRS-A or SRS2A) and short-versions. The SRS-A or SRS2-A is the adult section of the SRS-2nd edition (19 to 89 years), assessing aspects of social behaviors related to AS. It can contribute to the diagnosis, irrespective of the cognitive level. The adult form includes a questionnaire usable as a self-report or with an informant. Two articles examined the measurement properties of the SRS-Adult for the informant-version [95,96], and one for the self-report version [89]. Evidence for the informant-version was moderate for satisfactory internal consistency ($\alpha = 0.96$) [96] and structural validity [96]. By contrast, evidence for discriminant validity [95] and convergent validity [95,96] was heterogeneous. Reliability and content validity were not explored. Satisfactory diagnostic data (AUC, sensitivity and specificity over 0.80) were found, but the evidence was unsatisfactory [95]. Evidence for the self-report version was moderate for satisfactory test-retest reliability (ICC = 0.91), and limited for discriminant validity and convergent validity ($r = 0.72$ with AQ-50) [89]. Satisfactory internal consistency and data on structural validity were also found, but the evidence was conflicting. Diagnostic validity was only explored through an AUC, which seemed to be satisfactory, but the evidence was unsatisfactory. The properties of the *SRS2-AS30* and *SRS2-AS11* were each examined by one article [89], in which results about internal consistency, internal structure and discriminant validity were satisfactory, but with conflicting evidence. Moderate evidence was found for satisfactory test-retest reliability (the ICC for the SRS2-AS30 and SRS2-AS11 was 0.87 and 0.70, respectively). In addition, diagnostic validity was examined only through AUC values that were satisfactory (0.88 for the two versions), but results were limited by concerns about applicability and risks of bias.

3.4.3. Other tools

The *ASBQ* (Adult Social Behavior Questionnaire) designed to assist in diagnosis was adapted from the Children Social Behavior Questionnaire. This brief, easy to administer, tool includes a self-report questionnaire and a questionnaire for an informant. The *SR-AS* (Sensory Reactivity in Autism Spectrum) is a 32-item self-report questionnaire in line with DSM-5, designed to assess the degree of sensory reactivity in autistic adults. The properties of the ASBQ and the SR-AS, were each examined by one article. There were no data about the diagnostic accuracy for the *ASBQ* [57], and evidence was limited for intermediate internal consistency, good structural validity, and discriminant validity. For the SR-AS [67], there was strong evidence for satisfactory content validity, but limited evidence for satisfactory internal consistency and structural validity. Reliability and discriminant validity of the SR-AS were not explored. The AUC was found to be satisfactory, but with risks of bias and concern about applicability.

The *ASDASQ* (Autism Spectrum Disorder in Adults Screening Questionnaire) is a nine-item clinician questionnaire, useful for screening. Two articles examined the measurement properties of the ASDASQ. Diagnostic validity appeared to be satisfactory (sensitivity = 0.895 to 1.00 and specificity = 0.962 to 0.97), but evidence was unsatisfactory. Satisfactory internal consistency was found, but with conflicting evidence [58], and satisfactory reliability with moderate evidence [58,76], inter-rater correlation (e.g. Spearman's $r = 0.82$ ($P < 0.001$), and test-retest reliability (e.g. Spearman's $r = 0.83$, [$P < 0.001$]). There were no data concerning other measurement properties.

4. Discussion

This systematic review appraised the measurement properties of the tools currently used for the screening and diagnosis of AS in adults without ID. We identified nine tools, plus eight of their short-versions, that have been studied for their psychometric properties and diagnostic accuracy in samples of autistic adults with a mean normal range of measured intelligence.

4.1. Diagnostic tools for adult autistics without intellectual disability: measurement issues and recommendations for further studies

None of the tools designed for the detection or confirmation of an AS diagnosis among adults had satisfactory measurement properties. The overall level of evidence of the included studies was unsatisfactory due to their poor methodological quality and small sample sizes. Content and structural validity were rarely fully examined. Measurement error, responsiveness to change, and

cross-cultural validity, were not studied, although they are important aspects for judging the robustness of these tools. The high number of missing items, and the lack of randomization in patient selection were an important source of bias and a strong limitation of the robustness of the tools. We also often found high risks of bias in association with patient selection: participants of the studies were not representative of those referred to clinical settings. No screening or diagnosis tool reviewed in this paper had satisfactory measurement properties corroborated by strong evidence. Of the nine selected tools (plus their eight short-versions), only the AQ-50, AQ-S, RAADS-R, and RAADS-14, which are all screening tools, had mostly satisfactory or intermediate values (more than three) for their psychometric properties. Evidence for the diagnostic properties was limited by high risks of bias and/or concerns about their applicability. Finally, the validation data collected for Module 4 of the ADOS did not establish that its psychometric qualities were sufficiently robust for use in adults. The validation studies of the ADI-R and DISCO identified by our systematic review were not performed on adult samples, although they can theoretically be used for anyone, regardless of their age or cognitive level. When they were, the proportion of adults without disabilities in their population sample was under 50%. Validation studies should follow the use of screening tools, and include the combined use of screening and diagnostic tools and clinical assessment.

Future studies should aim to improve the study design, analysis, and reporting of measurement properties in AS among adults without ID, with particular attention to patient selection [92] and comparison groups, such as including control groups with psychiatric or neurodevelopmental disorders. Practitioners and researchers should use tools validated in samples representative of individuals in the referred clinical settings [20] to confirm their conclusions on their specificity toward the most frequent differential diagnoses of AS. Diagnostic tools should differentiate AS from common psychiatric diagnoses, as atypical autistic social behaviors may be confused with those evident in axis 2 personality disorders, other mental health or neurodevelopmental conditions, or communication disorders [57,73,75,84,90]. Validation studies, mostly performed with males, probably fail to detect subtler female cases. Future studies should include women with AS to identify possible specific thresholds or items [97–100]. The impact of ageing should also be considered, despite the well-established notion that AS may result in diverse outcomes throughout life, including a less obvious clinical picture [101–105]. Validation studies should also include participants with a wide range of IQs, even beyond ID. For example, differential diagnoses of gifted individuals may represent a specific challenge. Furthermore, studies should be conducted in various socio-cultural contexts [78,106].

The clinical implications of these results are important. None of the gold standard tools, such as ADI and ADOS, should be used alone during the diagnostic process [107]. Prevalence conclusions based on these instruments may be misleading, resulting in over-estimation [108], underestimation [109], or misinterpretation of the sex ratio [110]. The use of standardized tools has an obvious benefit in terms of making research populations uniform. However, used in isolation for diagnostic purposes to advance current science, they may not be superior to expert multi-disciplinary team personnel. Non-standardized historical, and proxy-report information may also be fruitfully added to these instruments to determine a diagnosis [20,29,111]. A critical issue is that diagnostic tools constructed based on categorical assumptions, applied to prototypical autistic phenotypes, are at risk of setting an excessively high diagnostic threshold. This may be detrimental for adults, especially when they are intelligent, women, and elderly individuals, and may miss clinical features that the individuals themselves are able to detect. In contrast, autistic adults with typical speech levels, without ID, can accurately and reliably describe their own difficulties in self-report questionnaires [59,77,112,113], even when low insight limits their answers [57,80,94,114]. RAADS, a self-report questionnaire and one of the screening tools with the best psychometric properties, is particularly useful among adults without ID. Its use should therefore not be limited to situations when proxies are not available. Conversely, the poor specificity of diagnostic and screening instruments toward non-autistic conditions, which also modify socially oriented behaviors and the variety of interests, risks over-inclusivity. The clinical judgment of experts exposed to large autistic populations may still be better than a standardized tool in autism, even for preschool and school-age, prototypical, Kanner-type, "frank autism" [115]. This is even more true for adults of a mean normal range of measured intelligence.

4.2. Limitation

This systematic review had methodological limitations, notably the inclusion of articles published only in English, or French. The exclusion of tools for which studies did not provide detailed descriptions of sample characteristics can also be considered to be a limitation. The use of the very strict quality rating, combining COSMIN and QUADAS-2 criteria, may also have led to under-rating the quality of the studies. In addition, we did not include several tools used in autistic adults without ID (e.g. Gilliam Asperger Disorder Scale/GADS, and Gilliam Autism Rating Scale/GARS) in our review, because their properties were explored in samples in which the proportion of adults was too low according to our inclusion criteria [36,47]. Furthermore, other tools usable for the screening and diagnosis of autism, regardless of age or cognitive level (e.g. ADI-R), have not been included in our review, because their psychometric properties have never been examined in samples of adults without ID, or because chronological age in the sample was not specified (e.g. Child Autism Rating Scale-2nd edition/CAR2, Asperger Syndrome Disorder Interview/ASDI).

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgements

We especially thank Colette Boy for her help in the literature research, Cécile Michelon for her careful reading of the paper, and Michelle Dawson for her comments on a previous version of this paper.

References

[1] Woolfenden S, Sarkozy V, Ridley G, Williams K. A systematic review of the diagnostic stability of autism spectrum disorder. Res Autism Spectr Disord 2012;6(1):345–54.
[2] Brugha TS, McManus S, Bankart J, Scott F, Purdon S, Smith J, et al. Epidemiology of autism spectrum disorders in adults in the community in England. Arch Gen Psychiatry 2011;68:459–65.
[3] Volkmar FR, Reichow B, McPartland JC. Adolescents and adults with autism spectrum disorders. New York, NY, US: Springer Science + Business Media; 2014.
[4] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM5), 5th ed. Washington, DC: American Psychiatric Press; 2013.
[5] National Audit Office. Supporting people with autism through adulthood. Great Britain: Stationery Office; 2009.
[6] Dawson M, Soulières I, Gernsbacher MA, Mottron L. The level and nature of autistic intelligence. Psychol Sci 2007;18:657–62.
[7] Kim YS, Leventhal BL, Koh YJ, Fombonne E, Laska E, Lim EC, et al. Prevalence of autism spectrum disorders in a total population sample. Am J Psychiatry 2011;168(9):904–12.

[8] Aussilloux C, Baghdadli A. Évolution du concept et actualité clinique du syndrome d'Asperger. Rev Neurol (Paris) 2008;164(5):406–13.
[9] Lehnhardt F-G, Gawronski A, Pfeiffer K, Kockler H, Schilbach L, Vogeley K. The investigation and differential diagnosis of Asperger syndrome in adults. Dtsch Arztebl Int 2013;110(45):755–62.
[10] Lai M-C, Baron-Cohen S. Identifying the lost generation of adults with autism spectrum conditions. Lancet Psychiatry 2015;2:1013–27.
[11] Happé F, Charlton RA. Aging in autism spectrum disorders: a mini-review. Gerontology 2012;58(1):70–8.
[12] Stoddart K, Burke L, King R. Asperger syndrome in adulthood: a comprehensive guide for clinicians. WW Norton & Company; 2012.
[13] Luciano CC, Keller R, Politi P, Aguglia E, Magnano F, Burti L, et al. Misdiagnosis of high function autism spectrum disorders in adults: an Italian case series. Autism Open Access 2014. http://dx.doi.org/10.4172/2165-7890.1000131.
[14] Tantam D. Psychological disorder in adolescents and adults with Asperger syndrome. Autism 2000;4(1):47–62.
[15] Ghaziuddin M. Mental health aspects of autism and Asperger syndrome. London, England: Jessica Kingsley Publishers; 2005.
[16] Trammell B, Wilczynski S, Dale B, McIntosh D. Assessment and differential diagnosis of comorbid condition in adolescents and adults with autism spectrum disorders. Psychol Sch 2013;50(9):936–46.
[17] Takara K, Kondo T. Autism spectrum disorder among first-visit depressed adult patients: diagnostic clues from backgrounds and past history. Gen Hosp Psychiatry 2014;36(6):737–42.
[18] Cath DC, Ran N, Smit JH, van Balkom AJ, Comijs HC. Symptom overlap between autism spectrum disorder, generalized social anxiety disorder and obsessive- compulsive disorder in adults: a preliminary case-controlled study. Psychopathology 2008;41(2):101–10.
[19] Haute Autorité de Santé (HAS). Autisme et autres troubles envahissants du développement: diagnostic et évaluation chez l'adulte. - Recommandations. HAS; 2011. p. 23.
[20] National Collaborating Centre for Mental Health (NICE). The NICE guideline on recognition, referral, diagnosis and management of adults on the autism spectrum. The British Psychological Society and the Royal College of Psychiatrists; 2012.
[21] Interagency Autism Coordinating Committee. 2010 Strategic Plan for Autism Spectrum Disorder Research. NIH Publication; 2010.
[22] Interagency Autism Coordinating Committee. 2011 Strategic Plan for Autism Spectrum Disorder Research. NIH Publication; 2011.
[23] Howlin P. Assessment instruments for Asperger syndrome. Child Psychol Psychiatry Rev 2000;5(3):120–9.
[24] Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EHJ, Dawson G, et al. Practice parameter: screening and diagnosis of autism: report of the quality standards subcommittee of the American Academy of Neurology and the Child Neurology Society (English). Neurology 2000;55(4):468–79.
[25] Baer L, Blais MA. Handbook of clinical rating scales and assessment in psychiatry and mental health. Springer; 2010.
[26] Nachshen J, Garcin N, Moxness K, Tremblay Y, Hutchison P, Lachance A, et al. Screening assessment and diagnosis of autism spectrum disorders in young children: Canadian best practice guidelines. Montreal, Quebec: Miriam Foundation; 2008.
[27] Volkmar F, Siegel M, Woodbury-Smith M, King B, McCracken J, State M. Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry 2014;53:237–57.
[28] Ozonoff S, Goodlin-Jones BL, Solomon M. Evidence-based assessment of autism spectrum disorders in children and adolescents. J Clin Child Adolesc Psychol 2005;34(3):523–40.
[29] Falkmer T, Anderson K, Falkmer M, Horlin C. Diagnostic procedures in autism spectrum disorders: a systematic literature review. Eur Child Adolesc Psychiatry 2013;22(6):329–40.
[30] American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 1999.
[31] Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284–90.
[32] Sappok T, Heinrich M, Underwood L. Screening tools for autism spectrum disorders. Adv Autism 2015;1:12–29.
[33] Charman T, Gotham K. Measurement issues: screening and diagnostic instruments for autism spectrum disorders–lessons from research and practise. Child Adolesc Ment Health 2013;18(1):52–63.
[34] Campbell JM. Diagnostic assessment of Asperger's disorder: a review of five third- party rating scales. J Autism Dev Disord 2005;35(1):25–35.
[35] Charak DA, Stella JL. Screening and diagnostic instruments for identification of autism spectrum disorders in children, adolescents, and young adults: a selective review. Assess Eff Interv 2001;27(1–2):5–17.
[36] Stoesz BM, Montgomery JM, Smart SL, Hellsten L-AM. Review of five instruments for the assessment of Asperger's disorder in adults. Clin Neuropsychol 2011;25(3):376–401.
[37] Loureiro D, Pio-Abreu J, Machado A, Gonçalves R, Cerejeira J. Instruments for the assessment of autism spectrum disorders in adults without intellectual disabilities: a systematic review. Eur Psychiatry 2015;30(Suppl. 1):1857.
[38] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Plos Med 2009;6(7):e1000100. http://dx.doi.org/10.1371/journal.pmed.1000100.
[39] Falissard B. Mesurer la subjectivité en santé. Perspectives méthodologiques et statistiques. Paris: Masson; 2001.
[40] Laveault D, Grégoire J. Introduction aux théories des tests : en psychologie et en sciences de l'éducation. De Boeck Supérieur; 2002.
[41] Goodwin LD, Leech NL. The meaning of validity in the new standards for educational and psychological testing: implications for measurement courses. Meas Eval Couns Dev 2003;36:181–92.
[42] Falissard B. Comprendre et utiliser les statistiques dans les sciences du vivant. Masson eds; 1996.
[43] Hu L, Bentler P. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. Psychol Methods 1998;3:424–53.
[44] Chan EKH. Standards and guidelines for validation practices: development and evaluation of measurement instruments. In: Zumbo BD, Chan EKH, editors. Validity and Validation in Social, Behavioral, and Health Sciences. Springer International Publishing Switzerland; 2014. p. 9–24.
[45] Nunnally J, Bernstein I. Psychometric Theory. San Francisco: McGraw-Hill; 1994.
[46] Glascoe FP. Screening for developmental and behavioral problems. Ment Retard Dev Disabil Res Rev 2005;11(3):173–9.
[47] Volkmar FR, Paul R, Rogers SJ, Pelphrey KA. Handbook of autism and pervasive developmental disorders diagnosis, development, and brain mechanisms. John Wiley & Sons; 2014.
[48] Maruish M. The use of psychological testing for treatment planning and outcome assessment, 2. London: Lawrence Erlbaum Associates; 2014.
[49] Fletcher RH, Fletcher SW, Fletcher GS. Clinical epidemiology: the essentials. Lippincott Williams & Wilkins; 2012.
[50] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529–36.
[51] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19(4):539–49.
[52] De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge University Press; 2011.
[53] Baron-Cohen S, Wheelwright S, Robinson J, Woodbury-Smith M. The Adult Asperger Assessment (AAA): a diagnostic method. J Autism Dev Disord 2005;35(6):807–19.
[54] Lord C, Risi S, Lambrecht L, Cook Jr E, Leventhal B, DiLavore P, et al. The Autism Diagnostic Observation Schedule—Generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 2000;30(3):205–23.
[55] Lord C, Rutter M, DiLavore P, Risi S, Gotham K, Bishop S. Autism Diagnostic Observation Schedule, 2nd edition ADOS-2, Los Angeles, CA: Western Psychological Corporation; 2012.
[56] Grodberg D, Weinger PM, Halpern D, Parides M, Kolevzon A, Buxbaum JD. The autism mental status exam: sensitivity and specificity using DSM-5 criteria for autism spectrum disorder in verbally fluent adults. J Autism Dev Disord 2014;44(3):609–14.
[57] Horwitz EH, Schoevers RA, Ketelaars CEJ, Kan CC, van Lammeren AMDN, Meesters Y, et al. Clinical assessment of AS in adults using self- and other-report: psychometric properties and validity of the Adult Social Behavior Questionnaire (ASBQ). Res Autism Spectr Disord 2016;24:17–28.
[58] Nylander L, Gillberg C. Screening for autism spectrum disorders in adult psychiatric out-patients: a preliminary report. Acta Psychiatr Scand 2001;103(6):428–34.
[59] Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. J Autism Dev Disord 2001;31(1):5–17.
[60] Allison C, Auyeung B, Baron-Cohen S. Toward brief 'red flags' for autism screening: the Short Autism Spectrum Quotient and the Short Quantitative Checklist for autism in toddlers in 1,000 cases and 3,000 controls. J Am Acad Child Adolesc Psychiatry 2012;51(2). 203–212.e7.
[61] Hoekstra RA, Vinkhuyzen AA, Wheelwright S, Bartels M, Boomsma DI, Baron-Cohen S, et al. The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). J Autism Dev Disord 2011;41(5):589–96.
[62] Brugha TS, McManus S, Smith J, Scott FJ, Meltzer H, Purdon S, et al. Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. Psychol Med 2012;42(3):647–56.
[63] Kurita H, Koyama T, Osada H. Autism-Spectrum Quotient-Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. Psychiatry Clin Neurosci 2005;59(4):490–6.
[64] Lau WYP, Kelly AB, Peterson CC. Further evidence on the factorial structure of the Autism Spectrum Quotient (AQ) for adults with and without a clinical diagnosis of autism. J Autism Dev Disord 2013;43(12):2807–15.
[65] Ritvo RA, Ritvo ER, Guthrie D, Ritvo MJ, Hufnagel DH, McMahon W, et al. The Ritvo Autism Asperger Diagnostic Scale-Revised (RAADS-R): a scale to assist the diagnosis of autism spectrum disorder in adults-an international validation study. J Autism Dev Disord 2011;41(8):1076–89.

[66] Eriksson JM, Andersen LM, Bejerot S. RAADS-14 Screen: validity of a screening tool for autism spectrum disorder in an adult psychiatric population. Mol Autism 2013;4(1):49.
[67] Elwin M, Schröder A, Ek L, Kjellin L. Development and pilot validation of a sensory reactivity scale for adults with high functioning autism spectrum conditions: sensory reactivity in Autism Spectrum (SR-AS). Nord J Psychiatry 2016;70(2):103–10.
[68] Constantino JN, Gruber CP. Social responsiveness scale, second edition (SRS-2), Torrance (États-Unis): Western psychological services; 2012.
[69] Duku E, Vaillancourt T, Szatmari P, Georgiades S, Zwaigenbaum L, Smith I, et al. Investigating the measurement properties of the social responsiveness scale in preschool children with autism spectrum disorders. J Autism Dev Disord 2013;43(4):860–8.
[70] Kanne SM, Christ SE, Reiersen AM. Psychiatric symptoms and psychosocial difficulties in young adults with autistic traits. J Autism Dev Disord 2009;39(6):827–33.
[71] Reiersen AM, Constantino JN, Grimmer M, Martin NG, Todd RD. Evidence for shared genetic influences on self-reported ADHD and autistic symptoms in young adult australian twins. Twin Res Hum Genet 2008;11(6):579–85.
[72] Kuenssberg R, McKenzie K. Confirmatory factor analysis of the Adult Asperger Assessment: the association of symptom domains within a clinical population. Res Dev Disabil 2011;32(6):2321–9.
[73] Bastiaansen JA, Meffert H, Hein S, Huizinga P, Ketelaars C, Pijnenborg M, et al. Diagnosing autism spectrum disorders in adults: the use of Autism Diagnostic Observation Schedule (ADOS) module 4. J Autism Dev Disord 2011;41(9):1256–66.
[74] Hus V, Lord C. The Autism Diagnostic Observation Schedule, module 4: revised algorithm and standardized severity scores. J Autism Dev Disord 2014;44(8):1996–2012.
[75] De Bildt A, Sytema S, Meffert H, Bastiaansen JACJ. The Autism Diagnostic observation schedule, module 4: application of the revised algorithms in an independent, well-defined, Dutch sample ($n = 93$). J Autism Dev Disord 2015;46:21–30.
[76] Chang HL, Juang YY, Wang WT, Huang CI, Chen CY, Hwang YS. Screening for autism spectrum disorder in adult psychiatric outpatients in a clinic in Taiwan. Gen Hosp Psychiatry 2003;25(4):284–8.
[77] Woodbury-Smith MR, Robinson J, Wheelwright S, Baron-Cohen S. Screening adults for Asperger syndrome using the AQ: a preliminary study of its diagnostic validity in clinical practice. J Autism Dev Disord 2005;35(3):331–5.
[78] Wakabayashi A, Baron-Cohen S, Wheelwright S, Tojo Y. The Autism-Spectrum Quotient (AQ) in Japan: a cross-cultural comparison. J Autism Dev Disord 2006;36(2):263–70.
[79] Hoekstra RA, Bartels M, Cath DC, Boomsma DI. Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. J Autism Dev Disord 2008;38(8):1555–66.
[80] Ketelaars C, Horwitz E, Sytema S, Bos J, Wiersma D, Minderaa R, et al. Brief report: adults with Mild Autism Spectrum Disorders (ASD): scores on the Autism Spectrum Quotient (AQ) and comorbid psychopathology. J Autism Dev Disord 2008;38(1):176–80.
[81] Lepage J-F, Lortie M, Taschereau-Dumouchel V, Theoret H. Validation of French- Canadian Versions of the empathy quotient and autism spectrum quotient. Can J Behav Sci 2009;41(4):272–6.
[82] Sizoo BB, van den Brink W, Gorissen-van Eenige M, Koeter MW, van Wijngaarden-Cremers PJ, van der Gaag RJ. Using the autism-spectrum quotient to discriminate Autism Spectrum Disorder from ADHD in adult patients with and without comorbid substance use disorder. J Autism Dev Disord 2009;39(9):1291–7.
[83] Naito K, Matsui Y, Maeda K, Tanaka K. Evaluation of the validity of the Autism Spectrum Quotient (AQ) in differentiating high-functioning autistic spectrum disorder from schizophrenia. Kobe J Med Sci 2010;56(3):E116–24.
[84] Wouters SGM, Spek AA. The use of the Autism-Spectrum Quotient in differentiating high-functioning adults with autism, adults with schizophrenia and a neurotypical adult control group. Res Autism Spectr Disord 2011;5(3):1169–75.
[85] Bishop SL, Seltzer MM. Self-reported autism symptoms in adults with autism spectrum disorders. J Autism Dev Disord 2012;42(11):2354–63.
[86] Booth T, Murray AL, McKenzie K, Kuenssberg R, O'Donnell M, Burnett H. Brief report: an evaluation of the AQ-10 as a brief screening instrument for ASD in adults. J Autism Dev Disord 2013;43(12):2997–3000.
[87] Broadbent J, Galic I, Stokes MA. Validation of autism spectrum quotient adult version in an Australian sample. Autism Res Treat 2013;2013:1–7.
[88] Pisula E, Kawa R, Szostakiewicz L, Lucka I, Kawa M, Rynkiewicz A. Autistic traits in male and female students and individuals with high functioning autism spectrum disorders measured by the Polish version of the Autism-Spectrum Quotient. Plos One 2013;8(9):e75236.
[89] Nishiyama T, Suzuki M, Adachi K, Sumi S, Okada K, Kishino H, et al. Comprehensive comparison of self-administered questionnaires for measuring quantitative autistic traits in adults. J Autism Dev Disord 2014;44(5):993–1007.
[90] Lugnegard T, Hallerback MU, Gillberg C. Asperger syndrome and schizophrenia: overlap of self-reported autistic traits using the Autism-spectrum Quotient (AQ). Nord J Psychiatry 2015;69(4):268–74.
[91] Zhang L, Sun Y, Chen F, Wu D, Tang J, Han X, et al. Psychometric properties of the Autism-Spectrum Quotient in both clinical and non-clinical samples: Chinese version for mainland China. BMC Psychiatry 2016;16:213.
[92] Sizoo BB, Horwitz EH, Teunisse JP, Kan CC, Vissers C, Forceville E, et al. Predictive validity of self-report questionnaires in the assessment of autism spectrum disorders in adults. Autism 2015;19(7):842–9.
[93] Kuenssberg R, Murray AL, Booth T, McKenzie K. Structural validation of the abridged Autism Spectrum Quotient-Short Form in a clinical sample of people with autism spectrum disorders. Autism 2014;18(2):69–75.
[94] Andersen LMJ, Näswall K, Manouilenko I, Nylander L, Edgar J, Ritvo RA, et al. The Swedish version of the Ritvo autism and Asperger diagnostic scale: Revised (RAADS-R). A validation study of a rating scale for adults. J Autism Dev Disord 2011;41(12):1635–45.
[95] Bölte S. Brief Report: the Social Responsiveness Scale for Adults (SRS-A): initial results in a German cohort. J Autism Dev Disord 2012;42(9):1998–9.
[96] Takei R, Matsuo J, Takahashi H, Uchiyama T, Kunugi H, Kamio Y. Verification of the utility of the social responsiveness scale for adults in non-clinical and clinical adult populations in Japan. BMC Psychiatry 2014;14:1.
[97] Attwood T. The complete guide to Asperger's syndrome. London, England: Jessica Kingsley Publishers; 2007.
[98] Lai M-C, Lombardo MV, Pasco G, Ruigrok ANV, Wheelwright SJ, Sadek SA, et al. A behavioral comparison of male and female adults with high functioning autism spectrum conditions. Plos One 2011;6:e20835.
[99] Gould J, Ashton-Smith J. Missed diagnosis or misdiagnosis? Girls and women on the autism spectrum. Good Autism Practice (GAP) 2011;12:34–41.
[100] Andersson GW, Gillberg C, Miniscalco C. Pre-school children with suspected autism spectrum disorders: do girls and boys have the same profiles? Res Dev Disabil 2013;34(1):413–22.
[101] Amaral D, Geschwind D, Dawson G. Autism spectrum disorders. Oxford University Press; 2011.
[102] Van Niekerk MEH, Groen W, Vissers CTWM, van Driel-de Jong D, Kan CC, Voshaar RCO. Diagnosing autism spectrum disorders in elderly people. Int Psychogeriatr 2011;23(5):700–10.
[103] Mukaetova-Ladinska E, Perry E, Baron M, Povey C. Ageing in people with autistic spectrum disorder. Int J Geriatr Psychiatry 2012;27(2):109–18.
[104] Shea V, Mesibov GB. Age-related issues in the assessment of autism spectrum disorders. In: Goldstein S, Naglieri JA, Ozonoff S, editors. Assessment of autism spectrum disorders. New York, NY, US: Guilford Press; 2009. p. 117–37.
[105] Howlin P, Savage S, Moss P, Tempier A, Rutter M. Cognitive and language skills in adults with autism: a 40-year follow-up. J Child Psychol Psychiatry 2014;55(1):49–58.
[106] Freeth M, Sheppard E, Ramachandran R, Milne E. A cross-cultural comparison of autistic traits in the UK, India and Malaysia. J Autism Dev Disord 2013;43(11):2569–83.
[107] Mazefsky CA, McPartland JC, Gastgeb HZ, Minshew NJ. Brief report: comparability of DSM-IV and DSM-5 ASD research samples. J Autism Dev Disord 2013;43(5):1236–42.
[108] Pantelis PC, Kennedy DP. Estimation of the prevalence of autism spectrum disorder in South Korea, revisited. Autism 2016;20(5):517–27.
[109] Frazier TW, Youngstrom EA, Speer L, Embacher R, Law P, Constantino J, et al. Validation of proposed DSM-5 criteria for autism spectrum disorder. J Am Acad Child Psy 2012;51(1):28–40.
[110] Constantino JN, Zhang Y, Frazier T, Abbacchi AM, Law P. Sibling recurrence and the genetic epidemiology of autism. Am J Psychiatry 2010;167(11):1349–56.
[111] Möricke E, Buitelaar JK, Rommelse NNJ. Do we need multiple informants when assessing autistic traits? The degree of report bias on offspring, self, and spouse ratings. J Autism Dev Disord 2016;46:164–75.
[112] Cederlund M, Hagberg B, Billstedt E, Gillberg IC, Gillberg C. Asperger syndrome and autism: a comparative longitudinal follow-up study more than 5 years after original diagnosis. J Autism Dev Disord 2008;38(1):72–85.
[113] Ritvo RA, Ritvo ER, Guthrie D, Yuwiler A, Ritvo MJ, Weisbender L. A scale to assist the diagnosis of autism and Asperger's disorder in adults (RAADS): a pilot study. J Autism Dev Disord 2008;38(2):213–23.
[114] McPartland JC, Klin A, Volkmar FR. Asperger syndrome: assessing and treating high-functioning autism spectrum disorders:. Guilford Publications; 2014.
[115] De Marchena A, Miller J. "Frank" presentations as a novel research construct and element of diagnostic decision-making in autism spectrum disorder. Autism Res 2017;10:653–62. http://dx.doi.org/10.1002/aur.1706.

Amaria Baghdadli is a professor of child and adolescent psychiatry. She is the head of the Child Psychiatry University Department and the Autism Resources Center of the CHU of Montpellier (France). Her research focuses on outcome, diagnosis, and intervention in AS using a life-span approach.

Frederick Russet is psychologist in the Child and Adolescent Psychiatry Department of the CHU of Montpellier. He is conducting a PhD in association with the Autism Resources Center of Montpellier.

Laurent Mottron is a researcher and clinician in autism, and a professor of Psychiatry at the Université de Montréal (Montréal, Québec, Canada). He holds the Marcel and Rolande Gosselin research chair on cognitive neuroscience in autism. His research is focused on autistic perception, intelligence, and strength-informed intervention.