ORIGINAL PAPER

The Autism Mental Status Exam: Sensitivity and Specificity Using DSM-5 Criteria for Autism Spectrum Disorder in Verbally Fluent Adults

David Grodberg · Paige M. Weinger · Danielle Halpern · Michael Parides · Alexander Kolevzon · Joseph D. Buxbaum

Abstract The phenotypic heterogeneity of adults suspected of autism spectrum disorder (ASD) requires a standardized diagnostic approach that is feasible in all clinical settings. The autism mental status exam (AMSE) is an eight-item observational assessment that structures the observation and documentation of social, communicative and behavioral signs and symptoms of ASD. Previous findings indicate high classification accuracy when compared to the autism diagnostic observation schedule in a non-stratified population of high-risk patients suspected of having ASD. This protocol investigates the sensitivity and specificity of AMSE scores using DSM-5 criteria for ASD in a sample of high-risk verbally fluent adults. Findings indicate an optimized sensitivity of 0.91 and a specificity of 0.93 for this group. Because of its high clinical utility, the AMSE holds promise as a diagnostic assessment tool that can support one's clinical diagnosis of ASD in high-risk adults.

Introduction

Diagnostic assessment for autism spectrum disorder (ASD) has evolved dramatically over the last two decades, particularly due to the creation of standardized diagnostic instruments such as the autism diagnostic observation schedule (ADOS, Lord et al. 1999) and the autism diagnostic interview-revised (ADI-R, Rutter et al. 2003). However, clinicians continue to face diagnostic challenges, especially during the course of evaluating adults for ASD. Adults may present with less obvious impairments and ASD signs and symptoms may become eclipsed by other comorbid conditions such as social anxiety disorder,

D. Grodberg (✉) · D. Halpern · A. Kolevzon
Department of Psychiatry, Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, One Gustave Levy Place, Box 1230, New York, NY, USA
e-mail: david.grodberg@mssm.edu

D. Halpern
e-mail: Danielle.halpern@mssm.edu

A. Kolevzon
e-mail: Alexander.kolevzon@mssm.edu

P. M. Weinger
Pediatrics Department, Mailman Center for Child Development, University of Miami Miller School of Medicine, Miami, FL, USA
e-mail: Paige.weinger@gmail.com

M. Parides
Department of Health Evidence and Policy, Mount Sinai Center for Biostatistics, Icahn School of Medicine at Mount Sinai, New York, NY, USA
e-mail: Michael.parides@mssm.edu

J. D. Buxbaum
Departments of Psychiatry, Genetics and Genomic Sciences, and Neuroscience, Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY, USA
e-mail: Joseph.buxbaum@mssm.edu

obsessive compulsive disorder, and schizoaffective disorder (Ritvo et al. 2011). In addition, lower functioning adults' intellectual impairments may also obscure the clinical picture of ASD.

Epidemiologic studies indicate the prevalence of ASD in adult populations is approximately one percent, and closely parallels that found in children (Brugha et al. 2011). The wider recognition of ASD in the adult population reflects clinicians' increased knowledge as well as a growing availability of evidence-based treatments and research protocols involving clinical trials, genetics and neuroimaging (Bastiaansen et al. 2011). The steadily increasing numbers of adult patients presenting for ASD evaluations range from those who received a reliable diagnosis during childhood to those who more recently came to suspect ASD. Further challenges emerge when elderly parents or caregivers are not available to furnish a developmental history. Given the heterogeneity of ASD, particularly in adult populations, there is a critical need to develop a standardized and streamlined protocol that can be used to properly identify, phenotype and diagnose this underserved and understudied population.

The current gold standard ASD diagnostic assessment model for adult patients includes a clinical exam guided by the *diagnostic and statistical manual of mental disorders* (5th ed.; *DSM-5*; American Psychiatric Association 2013), supported by two standardized assessments: ADOS and ADI-R. When scores on both the ADOS and ADI-R meet classification cutoffs, the accuracy of a clinical diagnosis of ASD is significantly increased (Risi et al. 2006). This three-tiered approach, which is generally used in research settings, standardizes ASD case ascertainment. However, the significant time and resources required to administer such comprehensive assessments hampers their feasibility and utility in under-resourced clinical settings. As a result, clinicians often rely on their own examination without the support of standardized measures. In addition, many clinicians use self- or caregiver-report rating scales in the absence of a standardized and validated observational assessment. While rating scales for higher functioning adults are widely available (e.g., Social Responsiveness Scale (Constantino and Gruber 2005), social-communication questionnaire (Rutter et al. 2003), they may be susceptible to patients' over- or under-reporting of their symptoms (Johnson et al. 2009).

The autism mental status exam (AMSE) was developed to address the lack of standardized observational assessment for ASD in underserved and under-resourced clinical populations and provides promising implications for the diagnosis of ASD in adults. The AMSE is a brief diagnostic observational tool that structures the way clinicians observe and document social, communicative and behavioral functioning in people with ASD. The AMSE is comprised of eight operationalized items, which include: eye contact, interest in others, pointing, language, pragmatics, stereotypy, preoccupations, and unusual sensitivities. To minimize clinical burden, the AMSE capitalizes on the wealth of information provided by direct observation of patients during a physician's routine clinical exam. It does not add extra work, but rather structures the way data is observed and documented. While the AMSE is predominantly an observational assessment, it also provides opportunity to record clinical information that is reported by the patient or caregiver, regardless of whether the behavior is observed during the exam. This integration of reported data into an observational assessment is necessary, as has been demonstrated by studies examining the validity of the ADOS and ADI-R, which indicated that the combination of observed signs and reported symptoms provides the best diagnostic accuracy (Risi et al. 2006).

While an expert's subjective clinical impressions remain a cornerstone of any diagnostic process and indeed improve the stability of ASD diagnoses, it remains difficult to quantify certainty of an ASD diagnosis, especially in complex adult cases. In such cases, there is a clear need for validated observational assessment, such as the ADOS. However, in community based and public health settings, where a gold standard evaluation is not feasible, the lack of a brief, developmentally appropriate, validated observational assessment may contribute to delays and uncertainty in diagnostic assessments. The AMSE is intended to fill this practice gap and its preliminary psychometric properties hold promise to guide clinical judgment in the context of practice based diagnostic decision making.

Importantly, the AMSE's intended use is for clinicians with competence in ASD diagnosis who do not have the capacity to administer standardized observational assessments. For example, physicians such as psychiatrists and neurologists with expertise in ASD diagnosis may use the AMSE to support their clinical judgment. Although less experienced clinicians may have the capacity to quickly learn how to administer the AMSE, scoring and interpretation can only by performed by clinicians who have expertise in making an ASD diagnosis in the first place. Training on the AMSE is provided through an online training curriculum with a digital manual and practice assessments using animated video simulations of clinical exams. The training protocol is currently under investigation, however, preliminary results indicate that inter-rater reliability can be achieved quickly.

Preliminary validation of the AMSE in a large, high-risk, sample that ranged in age from toddlers to adults indicated excellent inter-rater reliability and strong classification accuracy when compared to the ADOS (Grodberg et al. 2012). Specifically, a score of ≥ 5 resulted in a specificity of 94 % and a sensitivity of 81 %. While these

results were promising, there remains a need to determine optimal cutoffs based on criteria such as age, intellectual ability, and level of language. The current study examines the validity of the AMSE in a population of verbally fluent adults. While future studies will replicate this protocol in subgroups of children and adolescents, this study was driven by the critical need for standardized diagnostic assessment in adult populations. The use of the AMSE is particularly promising given that all items on the AMSE are included in the DSM-5 criteria for ASD. To date, there are no known brief standardized observational assessment tools that have been validated to the DSM-5 criteria for ASD. Therefore, this study examines the clinical utility of the AMSE in the diagnostic assessment of verbally fluent adults who are suspected of ASD in order to determine the optimal AMSE cutoff score in predicting independent clinical diagnosis of ASD based on DSM-5 criteria.

Methods

Participants

Fifty consecutive patients between the ages of 18 and 45 ($M = 28.90$, $SD = 8.29$) received comprehensive autism-focused diagnostic assessments as part of the IRB approved assessment protocol at the Seaver Autism Center for Research and Treatment at the Icahn School of Medicine at Mount Sinai. Informed consent was obtained from all participants. Participants were at higher-risk for ASD than the general population as they were self-referred adults who either carried a previous diagnosis of ASD or suspected a diagnosis of ASD. Participants were all verbally fluent.

Materials

Autism Mental Status Exam

The *AMSE* (Grodberg et al. 2012) is an 8-item observational assessment that prompts the examiner to observe and document patients' social, communicative and behavioral functioning in the context of a routine clinical examination. The AMSE was developed by psychiatrists with autism expertise and is intended to guide clinical judgment in the context of diagnostic decision making. Each item is scored on a 0–2 scale with possible total scores ranging from 0 to 16; higher scores reflect greater severity. Social items must be observed during the clinical exam but communication and behavioral items can be reported or observed. Three items—pragmatics of language, encompassing preoccupations, and unusual sensitivities—prompt the examiner to specify whether the item is reported or observed. In these three items, the score is weighted if the item is observed. An online training curriculum offers a digital manual and video simulation of clinical examinations based on individuals of varying ages and levels of functioning. A summary of AMSE items and scoring guidelines can be found in Table 1.

Autism Diagnostic Observation Schedule

Autism diagnostic observation schedule (Lord et al. 1999) is a semi-structured observational assessment that is used to assess the presence of autism symptomatology within two domains: social-communication and repetitive, restricted behaviors. The ADOS module 4, which is intended for verbally fluent adults, was administered to all participants in this sample by trained clinicians.

Table 1 Summary of AMSE items and scoring guidelines

Item\score	0	1	2
Eye contact[a]	>3 s	Fleeting	None
Interactions[a]	Spontaneous	Passive	None
Pointing[a]	Can point/gesture	Only follows point	None
Language[b]	Complex sentences	Undeveloped sentences, phrases, single words	None
Pragmatics[b,c]	No impairment	Reported impairment	Observed impairment
Repetitive behaviors[b]	None	Compulsive-like behaviors	Stereotypy
Preoccupations[b,c]	None	Reported preoccupations	Observed preoccupations
Sensitivities[b,c]	None	Reported symptoms	Observed signs

Complete scoring instructions are available online at: www.autismmentalstatusexam.com

[a] Social items must be observed by the examiner

[b] Communicative and behavioral items may be reported or observed

[c] Pragmatics, preoccupations, and sensitivities are weighted if observed

Autism Diagnostic Interview-Revised

The *ADI-R* (Rutter et al. 2003) is a structured caregiver interview that is used to diagnose ASD. The ADI-R probes for current behaviors and a developmental history consistent with autism symptomatology based on questions in the following domains: early development, communication, reciprocal social interaction, and repetitive, restricted patterns of behavior. When feasible, the ADI-R was administered and scored by a research reliable clinician.
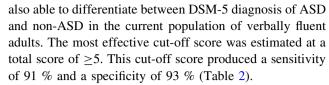
Procedure

Each participant first received a clinical evaluation by a psychiatrist with extensive experience in the diagnosis of ASD. The clinical evaluation included administration of the AMSE. Participants were then administered an ADOS-G, module 4, by an independent psychologist at the center who was blind to the AMSE score and blind to the psychiatrists' diagnostic impressions. An ADI-R was administered, when feasible, in 36 % of the sample. All ADOS and ADI-R administrations were performed and scored by reliable raters. In order to ascertain clinical diagnosis in a way that is sufficiently independent from the AMSE score, a best estimate clinical diagnosis (BECD) protocol was implemented in which the supervising psychologist at the Center reviewed the full ADOS protocol, the ADI-R protocol, and the psychiatric history, which included the chief complaint, history of present illness, and past psychiatric, medical, and developmental histories. In some cases, the BECD clinician communicated directly with the actual ADOS or ADI-R examiner to gather additional information. The BECD clinician was unable to gather additional information from the psychiatrist administering the AMSE, and did not have access to the AMSE score. DSM-5 criteria were then used to guide the BECD clinician's diagnostic formulation of ASD versus non-ASD.

Results

Forty-six percent of participants met BECD for ASD using DSM-5 criteria. A caregiver provided history in 38 % of cases and 10 % of participants carried a previous diagnosis of ASD. Diagnostic accuracy of the AMSE was assessed by the nonparametric measure of area under a receiver-operating characteristic (ROC) curve. The ROC curve analysis was used to determine the optimal cut-off for the AMSE compared to the BECD based on DSM-5. Area under the ROC curve (AUC) was 0.97 [95 % confidence interval (CI) 0.93–1.00].

This indicates that the same cut-off score previously determined to be optimal in a larger non-stratified sample is also able to differentiate between DSM-5 diagnosis of ASD and non-ASD in the current population of verbally fluent adults. The most effective cut-off score was estimated at a total score of ≥5. This cut-off score produced a sensitivity of 91 % and a specificity of 93 % (Table 2).

The relationship between AMSE total scores and ADOS classification was also examined. Fifty percent of participants met criteria for ASD on the ADOS. Area under the ROC curve (AUC) was 0.89 (95 % CI 0.79–0.99) and a score of ≥5 produced an optimal cut-off with a sensitivity of 84 % and a specificity of 92 %.

The ability of the AMSE to correctly classify patients is excellent as summarized by the AUC (0.97). The small sample size precluded a strategy of creating a training set to derive a classification rule, and a validation set to evaluate it. However, we addressed validity through a resampling (bootstrap) methodology based on a leave-one-out cross-validation approach. Similar to the initial determination of the AUC, validation based on ROC curves generated from logistic regression derived estimates of diagnosis, with utility and validity of the model summarized as the AUC. The "validation AUC" of 0.94 ± 0.3 denotes excellent validity.

Autism mental status exam total scores for participants meeting BECD DSM-5 ASD criteria ranged from 4 to 9, while total scores for participants who did not meet BECD DSM-5 ASD criteria ranged from 0 to 5. On the ADOS, AMSE total scores for participants meeting ADOS classification for ASD ranged from 1 to 9, while AMSE total scores for participants who did receive an ADOS classification of ASD ranged from 0 to 5. Of those participants who did not meet DSM-5 ASD criteria, all but one received other Axis I diagnoses. Table 3 displays AMSE scores and diagnoses for all participants.

An item-level analysis indicated that certain items held greater value than others in determining AMSE scores. Specifically, all adults in the sample received scores of 0 (no impairment) on the language and pointing items. In

Table 2 Sensitivity and specificity of AMSE cut-off scores for a best estimate clinical diagnosis of ASD based on DSM-5 criteria

AMSE cut-off	Sensitivity	Specificity
1	1.00	0.04
2	1.00	0.15
3	1.00	0.44
4	1.00	0.74
5	0.91	0.93
6	0.44	1.00
7	0.26	1.00
8	0.13	1.00
9	0.04	1.00

Table 3 Axis I diagnoses and AMSE scores assigned to non-ASD patients

Patient	Diagnosis	AMSE score
1	Generalized anxiety disorder	0
2	Schizophrenia	4
3	Major depressive disorder	3
4	No diagnosis	2
5	Major depressive disorder; generalized anxiety disorder	4
6	ADHD; major depressive disorder; generalized anxiety disorder	1
7	Generalized anxiety disorder	5
8	Major depressive disorder; substance abuse	1
9	Social anxiety disorder	2
10	Generalized anxiety disorder	4
11	Generalized anxiety disorder	3
12	Obsessive compulsive disorder	2
13	ADHD	3
14	ADHD; generalized anxiety disorder	3
15	Mood disorder, NOS	2
16	Generalized anxiety disorder	2
17	Social phobia	2
18	Post traumatic stress disorder; ADHD	3
19	Depressive disorder NOS; ADHD	4
20	Panic disorder; eating disorder, NOS	2
21	Bipolar II disorder	4
22	Generalized anxiety disorder	2
23	Major depressive disorder	3
24	Generalized anxiety disorder, PTSD	5
25	Major depressive disorder; eating disorder, NOS	3
26	Anxiety disorder, NOS	5
27	General anxiety disorder; depressive disorder, NOS	3

addition, all participants displayed, at minimum, fleeting eye contact and passive responding. Of items that were scored based on reported or observed symptoms, pragmatics was most often observed, while preoccupations were most often reported. Furthermore, unusual sensitivities were reported by 60 % of the sample, but were never observed during AMSE administration.

Discussion

The AMSE is the first mental status exam structured to examine signs and symptoms of ASD. It provides experienced physicians with an innovative approach that does not interfere with or add burden to the clinical exam. This study represents the first examination of the AMSE's clinical utility in a high-risk sample of verbally fluent adults. The present study used DSM-5 criteria to examine the effectiveness of the AMSE's ability to differentiate between patients who meet criteria for ASD and those who do not. Results from this exploratory study determined that a total score of ≥ 5 is the optimal AMSE cutoff for verbally

fluent adults suspected of ASD. This cutoff score was also determined to be the optimal cutoff when comparing AMSE total scores to ADOS classifications. Findings provided further support for our previous work, which indicated that an AMSE total score of ≥ 5 best predicted ADOS classification in a high-risk sample that was not stratified by age (Grodberg et al. 2012). The excellent sensitivity and specificity found in this study supports both the clinical utility of the AMSE as well as its ability to support ASD diagnoses in the context of adult research protocols. Furthermore, this study establishes the foundation for more extensive validation in this under-resourced and under-studied population.

While this study demonstrates the AMSE's high classification accuracy, several limitations must be addressed. First, the results may be inflated by the cohort's relatively high level of functioning. As expected, participant symptomatology was mild to moderate and AMSE total scores did not exceed a score of 9, out of a possible 16, for any participant. In addition, certain items were not relevant to this population (e.g., language, pointing). We expect that a cohort including non-verbal individuals with ASD, with or

without intellectual impairment, would have higher AMSE scores, and the shape of the ROC curve would reflect a more gradual change in sensitivity and specificity as the cutoff approaches a score of 5. Importantly, while the high functioning nature of the cohort affects the shape of the ROC curve, it does not affect the ultimate clinical utility of the cutoff of ≥5. Another potential limitation is that the psychiatrist's clinical notes (e.g., chief complaint, history of present illness, psychiatric, medical and developmental history) were reviewed by the BECD clinician. It could be argued that these data may be vulnerable to bias as a subject's lower AMSE score may influence the psychiatrist to under-report ASD symptoms in the clinical notes. To address this potential bias, we used the independent ADOS and ADI-R scores to support the final DSM-5 diagnosis of ASD. In addition, the BECD clinician did not have access to AMSE scores and was unable to consult with the psychiatrist who administered the AMSE when determining diagnoses. Furthermore, participants were referred to the Center because they were suspected of having ASD and are thus at higher-risk than the general population. Given the relatively small sample size, the generalizability of results remains limited and future studies must examine the use of the AMSE in larger samples that are further stratified by level of cognitive and adaptive functioning. In addition, the two psychiatrists who administered the AMSE had expertise in the clinical evaluation of ASD and also achieved 100 % inter-rater accuracy using 4 test cases. Considering our false positive rate of 7 %, it remains critical that in cases where a clinician does not have expertise in ASD diagnosis or in cases that have a complicated clinical presentation, the AMSE must not substitute for more comprehensive diagnostic assessment. In such cases referral for further work up must be made. Importantly, the level of sensitivity and specificity found in this study is dependent on reliable use of the AMSE. Further research is needed to determine the AMSE's psychometric properties amongst clinicians with varying levels of experience and expertise. We strongly recommend that the AMSE not be used by clinicians without specific diagnostic expertise in ASD. A current protocol will examine the effectiveness of a training curriculum across different mental health specialties and levels of autism clinical experience. Future studies must also examine the AMSE's clinical utility in general adult psychiatric clinics to further assess discriminatory characteristics of the AMSE.

Results from this study demonstrate that the AMSE offers a standardized diagnostic observational assessment that can be used to quickly and accurately support ASD diagnoses in verbally fluent adults suspected of ASD. While the AMSE cannot be used independently, given the lack of validated ASD assessment measures for adults, and the complexity of assessing ASD in these individuals, the AMSE holds promise as a streamlined brief observational assessment that can guide clinical judgment when considering a DSM-5 diagnosis of ASD. More accurate diagnosis in adult populations is necessary to promote appropriate referrals to evidence-based treatments as well as psychiatric, psychological, vocational and educational services.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., et al. (2011). Diagnosing autism spectrum disorders in adults: the use of autism diagnostic observation schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders, 41*(9), 1256–1266.

Brugha, T. S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., et al. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Archives of General Psychiatry, 68*(5), 459–465.

Constantino, J. N., & Gruber, C. P. (2005). *Social responsiveness scale (SRS)*. Los Angeles: Western Psychological Services.

Grodberg, D., Weinger, P. M., Kolevzon, A., Soorya, L., & Buxbaum, J. D. (2012). Brief report: the autism mental status examination: development of a brief autism-focused exam. *Journal of Autism and Developmental Disorders, 42*(3), 455–459.

Johnson, S. A., Filliter, J. H., & Murphy, R. R. (2009). Discrepancies between self- and parent-perceptions of autistic traits and empathy in high functioning children and adolescents on the autism spectrum. *Journal of Autism and Developmental Disorders, 39*(12), 1706–1714.

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule manual*. Los Angeles: Western Psychological Services.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., et al. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(9), 1094–1103.

Ritvo, R. A., Ritvo, E. R., Guthrie, D., Ritvo, M. J., Hufnagel, D. H., McMahon, W., et al. (2011). 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. *Journal of Autism and Developmental Disorders, 41*(8), 1076–1089.

Rutter, M., Bailey, A., & Lord, C. (2003a). *Social communication questionnaire*. Los Angeles: Western Psychological Services.

Rutter, M., LeCouteur, A., & Lord, C. (2003b). *Autism diagnostic interview-revised*. Western Psychological Services: Los Angeles.