JAMA Psychiatry | Original Investigation

The World Health Organization Adult Attention-Deficit/Hyperactivity Disorder Self-Report Screening Scale for *DSM-5*

Berk Ustun, MS; Lenard A. Adler, MD; Cynthia Rudin, PhD; Stephen V. Faraone, PhD; Thomas J. Spencer, MD; Patricia Berglund, MBA; Michael J. Gruber, MS; Ronald C. Kessler, PhD

Invited Commentary page 527

IMPORTANCE Recognition that adult attention-deficit/hyperactivity disorder (ADHD) is common, seriously impairing, and usually undiagnosed has led to the development of adult ADHD screening scales for use in community, workplace, and primary care settings. However, these scales are all calibrated to *DSM-IV* criteria, which are narrower than the recently developed *DSM-5* criteria.

OBJECTIVES To update for *DSM-5* criteria and improve the operating characteristics of the widely used World Health Organization Adult ADHD Self-Report Scale (ASRS) for screening.

DESIGN, SETTING, AND PARTICIPANTS Probability subsamples of participants in 2 general population surveys (2001-2003 household survey [n = 119] and 2004-2005 managed care subscriber survey [n = 218]) who completed the full 29-question self-report ASRS, with both subsamples over-sampling ASRS-screened positives, were blindly administered a semistructured research diagnostic interview for *DSM-5* adult ADHD. In 2016, the Risk-Calibrated Supersparse Linear Integer Model, a novel machine-learning algorithm designed to create screening scales with optimal integer weights and limited numbers of screening questions, was applied to the pooled data to create a *DSM-5* version of the ASRS screening scale. The accuracy of the new scale was then confirmed in an independent 2011-2012 clinical sample of patients seeking evaluation at the New York University Langone Medical Center Adult ADHD Program (NYU Langone) and 2015-2016 primary care controls (n = 300). Data analysis was conducted from April 4, 2016, to September 22, 2016.

MAIN OUTCOMES AND MEASURES The sensitivity, specificity, area under the curve (AUC), and positive predictive value (PPV) of the revised ASRS.

RESULTS Of the total 637 participants, 44 (37.0%) household survey respondents, 51 (23.4%) managed care respondents, and 173 (57.7%) NYU Langone respondents met *DSM-5* criteria for adult ADHD in the semistructured diagnostic interview. Of the respondents who met *DSM-5* criteria for adult ADHD, 123 were male (45.9%); mean (SD) age was 33.1 (11.4) years. A 6-question screening scale was found to be optimal in distinguishing cases from noncases in the first 2 samples. Operating characteristics were excellent at the diagnostic threshold in the weighted (to the 8.2% *DSM-5*/Adult ADHD Clinical Diagnostic Scale population prevalence) data (sensitivity, 91.4%; specificity, 96.0%; AUC, 0.94; PPV, 67.3%). Operating characteristics were similar despite a much higher prevalence (57.7%) when the scale was applied to the NYU Langone clinical sample (sensitivity, 91.9%; specificity, 74.0%; AUC, 0.83; PPV, 82.8%).

CONCLUSIONS AND RELEVANCE The new ADHD screening scale is short, easily scored, detects the vast majority of general population cases at a threshold that also has high specificity and PPV, and could be used as a screening tool in specialty treatment settings.

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Ronald C. Kessler, PhD, Department of Health Care Policy, Harvard Medical School, 180A Longwood Ave, Boston, MA 02115 (kessler@hcp.med.harvard.edu).

JAMA Psychiatry. 2017;74(5):520-526. doi:10.1001/jamapsychiatry.2017.0298
Published online April 5, 2017. Last corrected on October 2, 2019.

Attention-deficit/hyperactivity disorder (ADHD) is a commonly occurring childhood-onset disorder often persisting into adulthood.[1] Adult ADHD is associated with a high prevalence of secondary comorbid mental disorders,[2] work impairment,[3] accidents with injuries,[4] and early mortality,[5] but is typically undetected and untreated[6] despite evidence of treatment effectiveness.[7] Accordingly, there is interest in screening for adult ADHD in primary care[8] and workplace settings.[9] However, published versions of existing screening scales are still calibrated to *DSM-IV* criteria.[10] Prevalence should increase using *DSM-5* criteria given the reduced number of required symptoms and later required age of onset. We present a new version of the widely used Adult ADHD Self-Report Screening Scale (ASRS)[11,12] updated for *DSM-5* criteria based on analyses in a national household survey, a survey of managed care subscribers, and a specialty clinical sample.

Key Points

Question Can a brief screening scale based on patient responses to structured questions detect *DSM-5* adult attention-deficit/hyperactivity disorder in the general population?

Findings A new machine-learning algorithm was used to build a screening scale from responses to 6 questions in the World Health Organization Adult Attention-Deficit/Hyperactivity Disorder Self-Report Scale using optimal integer scoring rules. The scale had excellent cross-validated concordance with blinded clinical diagnoses of *DSM-5* adult attention-deficit/hyperactivity disorder.

Meaning The new scale is short, easily scored, and can detect the vast majority of adult attention-deficit/hyperactivity disorder cases in the general population with high sensitivity and specificity, discriminating well among patients presenting for evaluation and specialty treatment.

Methods

Samples

National Comorbidity Survey Replication

The household sample came from the National Comorbidity Survey Replication (NCS-R), a national face-to-face survey performed from February 5, 2001, to April 7, 2003.[13] Childhood ADHD was assessed retrospectively among respondents aged 18 to 44 years. Respondents were divided into 4 sampling strata for an ASRS clinical reappraisal follow-up telephone survey, those denying any childhood symptoms, those with childhood symptoms not meeting full criteria, those with childhood ADHD denying current symptoms, and those with childhood ADHD and current symptoms. The 29-question ASRS and a blinded, semistructured diagnostic interview (both described below) were administered to quota samples in each stratum in a follow-up telephone survey. Interviews were tape-recorded for quality control. Each respondent received $25 for participation. Verbal informed consent was obtained before interviews. These recruitment and consent procedures were approved by the human subjects committees of the University of Michigan and Harvard Medical School. The 119 completed clinical reappraisal interviews were weighted to adjust for differential sampling across strata. More details about study design are reported elsewhere.[11]

Managed Care Sample

The managed care sample was based on a 2004-2005 telephone survey of subscribers to a large managed health care plan[14] that included the *DSM-IV* ASRS screening scale.[12] A subsample of respondents, with oversampling of screened positives, was reinterviewed 6 months later. The full ASRS, followed by the same semistructured diagnostic interview as in the NCS-R, was then administered to a subsample of 218 second-survey respondents, with oversampling of stably screened positives. Recruitment and informed consent procedures were the same as in the NCS-R and were approved with a Health Insurance Portability and Accountability Act waiver by the institutional review board of New York University (NYU) Langone School of Medicine. The completed interviews were weighted to adjust for oversampling. More details about study design are reported elsewhere.[12]

NYU Langone Sample

The 2014-2015 clinical sample included 300 patients who were either obtaining a free evaluation through the Adult ADHD Program at NYU Langone based on mass media recruitment and referrals (n = 193), assessed between January 26, 2011, and September 7, 2012, or were controls from primary care waiting rooms near the NYU Langone campus (n = 107), assessed between September 16, 2015, and February 26, 2016. The full ASRS tool and same blinded, semistructured diagnostic interview as in the NCS-R and managed care samples were administered either face-to-face (patients seen at the ADHD program) or by telephone (patients recruited in primary care waiting rooms). Incentives for participation were the free evaluation to those recruited for evaluation and a $25 incentive to those recruited in primary care waiting rooms. Written informed consent was obtained before the interviews were administered. Recruitment and consent procedures were approved and a Health Insurance Portability and Accountability Act waiver was granted by the institutional review board of NYU Langone School of Medicine.

The ASRS Screening Scale Item Pool

The ASRS screening scale was developed by 2 board-certified psychiatrists (L.A.A. and T.J.S.) working with a World Health Organization (WHO) advisory group to generate 1 fully structured question for each *DSM-IV* Criterion A1-A2 symptom of inattention and hyperactivity-impulsivity plus 11 non-*DSM-IV* symptoms of deficits in higher-level executive function believed to be relevant to adult ADHD.[15] Many of the latter items were similar to the Utah Criteria for adult ADHD.[16] Each question asked how often the symptom occurred over the past 6 months with responses of never, rarely, sometimes, often, and very often.

The Clinical Interview

Clinical diagnoses of *DSM-5* adult ADHD were made based on semistructured interviews using version 1.2 of the Adult ADHD

Clinical Diagnostic Scale (ACDS),[17,18] a widely used research diagnostic interview.[10] The interview began with a retrospective assessment of childhood ADHD and then assessed recent (past 6 months) symptoms. Both childhood- and adult-specific prompts were used to ensure adequate exploration of symptom severity and breadth. The *DSM-5* and ACDS diagnoses required 6 to 9 childhood and 5 to 9 current adult *DSM-5* A1 or A2 symptoms (*DSM-5* Criterion A), at least 1 symptom prior to age 12 (Criterion B), some ADHD-related impairment in 2 or more domains of functioning in the past 6 months (Criterion C), and clinically significant ADHD-related impairment over the same time period (Criterion D). Criterion E (symptoms do not occur exclusively during the course of a pervasive developmental disorder or psychotic disorder and are not better accounted for by another mental disorder) was evaluated indirectly by asking clinical interviewers to probe to confirm that symptoms were not due to anxiety, mood, or substance disorders, but comorbid disorders were not explicitly assessed. Attention-deficit/hyperactivity disorder not otherwise specified was not indicated as a diagnosis. The *DSM-5* requirement of impairment before age 12 years was not operationalized.

The ACDS was administered in the NCS-R by 4 experienced PhD-level clinical interviewers who received 40 hours of training from 2 board certified psychiatrists specializing in adult ADHD research (L.A.A. and T.J.S.). Each interviewer had to complete 5 practice interviews with symptom ratings matching those of the trainers before beginning the interviews. All production interviews were tape-recorded, reviewed by a clinical supervisor, and referred to the trainers (L.A.A. and T.J.S.) for final consensus diagnoses when supervisor-interviewer discrepancies occurred. The ACDS interviews in the managed care sample were administered by 6 PhD-level clinical psychologists or MA-level social workers experienced in administering the ACDS in clinical studies and trained by 1 of the investigators (T.J.S.). The ACDS interviews in the NYU Langone sample were administered by 2 clinical psychology trainees (a PhD candidate with an MA, and an MA candidate with a BA), both trained by one of the investigators (L.A.A.). Ongoing group calibration meetings and reviews of a sample of tapes were used in all studies to ensure that symptom ratings on the ACDS conformed to best practices. In the NCS-R, where the supervisor was not an expert in adult ADHD, we took the additional steps of having 1 such expert (L.A.A.) attend all group calibration meetings and of having the supervisor review 100% of the tapes and refer all discrepancies with interviewer ratings to the psychiatrists (L.A.A. and T.J.S.) for discussion and resolution.

Developing the New Screening Scale

Analysis of deidentified data was approved by the institutional review board of NYU Langone School of Medicine and was carried out in 2016. We aimed to create a *DSM-5* ASRS screening scale with the same form as its *DSM-IV* counterpart: that is, a simple additive scale with a limited number of variables and integer scoring (ie, each response option a whole number in the range of 0-5) for quick calculation. Because there is no standard methodology to create such scales, existing tools are built manually by combining restricted and/or stepwise logistic regression with ad hoc rounding.[11] However, a new machine-learning algorithm, RiskSLIM (Risk-Calibrated Supersparse Linear Integer Model),[19] a risk-calibrated version of the SLIM[20] algorithm, now makes it possible to develop short integer-scored screening scales more rigorously by using modern optimization techniques that find a best-fitting logistic regression model with a fixed number of screening questions and optimal integer scoring of each response option to predict clinical outcomes.

RiskSLIM models were estimated by creating 4 nested dichotomous predictors for each ASRS question (eg, 1 for a response of never, 1 for responses of rarely or more, 1 for responses of sometimes or more) and estimating a separate model constrained to use no more than 8 questions (and as many nested dichotomies within questions as useful) constraining signs of coefficients for the never dichotomies to be negative and for the remaining responses to be positive. The specific questions were selected to minimize logistic loss + C_0 × (number of nonzero coefficients), where C_0 is a regularization parameter. The number of questions was chosen to optimize 10-fold cross-validated (10-CV) area under the receiver operating characteristic curve (area under the curve [AUC]; the proportion of times a randomly selected case will have a higher screening scale score than a randomly selected noncase)[21] and calibration accuracy (CAL) (square root of weighted mean squared error between predicted and observed ADHD prevalence across all scores on the screening scale).[22]

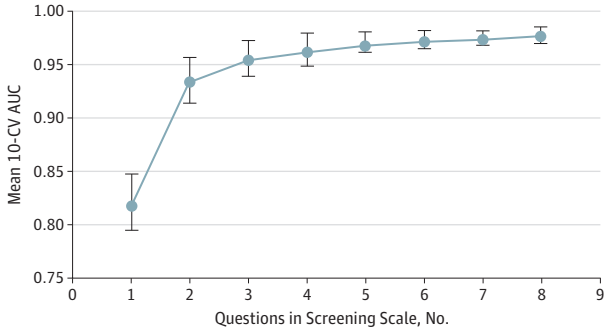
We report 10-CV estimates of operating characteristics to evaluate likely out-of-sample performance. Because SEs are downwardly biased when model selection is based on regularization, we plotted minimum and maximums of 10-CV AUCs to evaluate precision. To ensure that the final screening scale was appropriate for use in general population samples and was based on enough cases with clinical diagnoses to have an acceptable event-per-variable ratio for up to 8 screening questions,[23] we pooled NCS-R and managed care samples for estimation. Since both samples included extreme weights that may influence variation in 10-CV estimates, we transformed the data before analysis by dividing the original weights in each sample, which summed to the observed sample size, by the smallest weight, rounding each transformed weight to the nearest whole number and duplicating records for each respondent the number of times equal to the respondent's rounded weight. Once the best model was selected, coefficients were applied to both the general population and clinical samples and individual-level predicted probabilities were compared with clinical diagnoses to calculate sensitivity, specificity, positive predictive value (PPV), and AUC.

Results

Prevalence of *DSM-5*/ACDS Adult ADHD

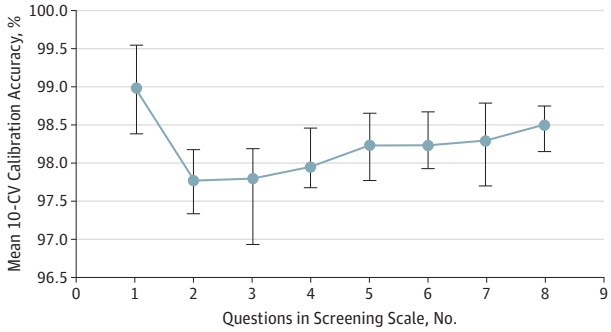
The weighted (to adjust for oversampling) prevalence of *DSM-5*/ACDS adult ADHD was 6.5% in the NCS-R sample, 9.2% in the managed care sample, and 8.2% in the pooled combination of the 2 samples. The unweighted prevalence was 37.0% in the NCS-R sample (n = 44 of *DSM-5*/ACDS cases), 23.4% in

Figure 1. Ten-fold Cross-Validated (10-CV) Area Under the Curve (AUC) vs the Number of Questions in the Screening Scale



Pooled National Comorbidity Survey Replication and managed care samples (n = 337). The reported range represents the highest and lowest values of mean AUC across the 10 separate folds for 8 questions.

Figure 2. Ten-fold Cross-Validated (10-CV) Mean Calibration Accuracy vs the Number of Questions in the Screening Scale



Pooled National Comorbidity Survey Replication and managed care samples (n = 337). The reported range represents the highest and lowest values of mean calibration accuracy across the 10 separate folds for 8 questions.

the managed care sample (n = 51), and 57.7% in the NYU Langone sample (n = 173). Of the respondents who met *DSM-5* criteria for adult ADHD, 123 were male (45.9%); mean (SD) age was 33.1 (11.4) years.

Choosing the Optimal Number of Screening Questions

Inspection of 10-CV AUC showed a monotonic increase as the number of questions increased, from approximately 0.82 for a model with 1 question to more than 0.95 for models with 3 to 8 questions and more modest increases after 3 questions (**Figure 1**). The likelihood of good out-of-sample performance is indicated by a narrow AUC range across 10-CV replicates. In contrast, inspection of 10-CV CAL showed a nonmonotonic pattern, with the highest CAL in the degenerate case of the model with only 1 screening question (degenerate in that the vast majority of the respondents did not have the disorder and model fit was driven by classifying the 91.8% of true noncases as such), decreasing CAL in models with 2 to 4 screening questions and increasing CAL in models with 5 to 6 screening questions (**Figure 2**). Based on these results, we focused on the 6-question model because it achieved best-in-class AUC and CAL. A 3-question model was a close competitor but lacked the granularity near the clinical threshold that further analysis (reported below) found for the 6-question model.

The questions in the final scale include (1) *DSM-5* symptom of inattention (Criterion A1c: does not listen when spoken to directly), (2) non–*DSM-5* symptoms of executive dysfunction (puts things off to last minute, depends on others to keep life in order), and (3) *DSM-5* symptoms of hyperactivity and impulsivity (Criterion A2b, leaves seat inappropriately; Criterion A2d, difficulty playing quietly/leisure time; Criterion A2g, blurts out answers) (**Table 1**). A 0 score is assigned to all never responses, but scores for higher responses vary, with the possible range of 0 to 25. As noted above, scores are weakly monotonic within questions by construction but vary within and across questions in other ways, reflecting differing associations with clinical diagnoses. Total screening scores summed across all responses have a 0 to 25 range.

Table 1. Questions in the Optimal RiskSLIM *DSM-5* ASRS Screening Scale[a]

1. How often do you have difficulty concentrating on what people say to you, even when they are speaking to you directly? (*DSM-5* A1c)
2. How often do you leave your seat in meetings or other situations in which you are expected to remain seated? (*DSM-5* A2b)
3. How often do you have difficulty unwinding and relaxing when you have time to yourself? (*DSM-5* A2d)
4. When you're in a conversation, how often do you find yourself finishing the sentences of the people you are talking to before they can finish them themselves? (*DSM-5* A2g)
5. How often do you put things off until the last minute? (Non-*DSM*)
6. How often do you depend on others to keep your life in order and attend to details? (Non-*DSM*)

Abbreviations: ADHD, attention-deficit/hyperactivity; ASRS, Adult ADHD Clinical Diagnostic Scale; RiskSLIM, Risk-Calibrated Supersparse Linear Integer Model.

[a] Response categories are never, rarely, sometimes, often, and very often. The never response option is scored 0 for all questions; the highest scores are 6 for question 3, 5 for questions 1 and 2, 4 for question 5, 3 for question 6, and 2 for question 4, resulting in a scale with scores in the range of 0 of 25.

Operating Characteristics of the New Screening Scale

The optimal threshold for maximizing the AUC in the pooled NCS-R and managed care samples was 14 or higher (AUC, 0.94), which screened 11.2% of respondents and captured 91.4% of *DSM-5*/ACDS cases (sensitivity), with false-positive rates (1 − specificity) of 4.0% and 67.3%, respectively, of screened positives being *DSM-5*/ACDS cases (PPV) of ADHD (**Table 2**). Performance was substantially worse when the operating threshold was lowered to 13 or higher, which resulted in a small increase in sensitivity (92.6%) at the expense of a major overestimation of prevalence (21.8%), yielding substantially worse PPV (34.9%). These results suggest that the 14 or higher operating threshold would most likely be preferred for screening purposes.

Screening scale performance was investigated separately in the NYU Langone sample. The optimal threshold of 14 or higher in the general population sample captured 91.9% of NYU Langone *DSM-5*/ACDS cases. The false-positive rate at that threshold (26.0%) was considerably higher than that in the general population sample, the AUC (0.829) was lower, and the PPV was high (82.8%). As in the general population sample,

Table 2. Operating Characteristics of the *DSM-5* ASRS Screening Scale

Score Threshold	Predicted Prevalence, %	Sensitivity, %[a]	Specificity, %[b]	AUC[c]	PPV, %[d]
Pooled NCS-R and managed care development samples[e]					
≥13 vs 0-12	21.8	92.6	84.5	0.89	34.9
≥14 vs 0-13	11.2	91.4	96.0	0.94	67.3
≥15 vs 0-14	9.3	87.0	97.6	0.92	76.5
≥16 vs 0-15	7.8	84.1	99.0	0.92	88.2
≥17 vs 0-16	7.0	79.2	99.5	0.89	93.2
NYU Langone validation sample[f]					
≥13 vs 0-12	69.3	93.1	63.0	0.78	77.4
≥14 vs 0-13	64.0	91.9	74.0	0.83	82.8
≥15 vs 0-14	59.3	89.0	81.1	0.85	86.5
≥16 vs 0-15	53.3	82.7	86.6	0.85	89.4
≥17 vs 0-16	47.0	76.3	92.9	0.85	93.6

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; ASRS, Adult ADHD Self-Report Scale; AUC, area under the curve; NCS-R, National Comorbidity Survey Replication; NYU Langone, New York University Langone Medical Center Adult ADHD Program; PPV, positive predictive value.

[a] The proportion of *DSM-5* and Adult Clinical ADHD Diagnostic Scale (ACDS) cases classified as cases by the screener at the threshold.

[b] The proportion of *DSM-5*/ACDS noncases classified as noncases by the screener at the threshold.

[c] The probability that a randomly selected *DSM-5*/ACDS case would have a higher screening scale score than a randomly selected noncase at the threshold, with ties on the screening scale assigned a predicted probability of 0.500.

[d] The proportion of respondents with screening scale scores above the threshold that meet *DSM-5*/ACDS criteria for the disorder.

[e] Sample of 337 participants; weighted prevalence, 8.2%.

[f] Sample of 300; participants; prevalence, 57.7%.

performance in the NYU Langone participants decreased markedly at the 13 or higher operating threshold owing to substantially increased false-positives (37.0%) and modestly increased sensitivity (93.1%). Thus, the 14 or higher threshold represents the most liberal threshold to consider for screening in a clinical sample.

Sensitivity at each threshold yielding a prevalence estimate close to the observed diagnostic prevalence was very similar in the general population and clinical samples, despite significantly lower *DSM-5*/ACDS prevalence in the former than latter samples (8.2% vs 57.7%). In contrast, specificity was considerably higher in the general population than clinical sample due to the much higher proportion of patients in the clinical sample with a meaningful number of symptoms not meeting full diagnostic criteria. The PPV was higher in the clinical than general population sample due to higher *DSM-5*/ACDS prevalence, resulting in most screened positives being true positives in the clinical sample despite the comparatively low specificity.

Comparing the *DSM-IV* and *DSM-5* ASRS Screening Scales

Given the strong correlations between responses to ASRS questions, it is possible that responses to the 6 questions in the original *DSM-IV*/ACDS screening scale[11,12] could discriminate between *DSM-5*/ACDS cases and noncases with good accuracy. If so, this would be useful for researchers who have already collected data using the *DSM-IV* ASRS screening scale. To investigate this possibility, we applied RiskSLIM to build a screening scale using only the responses to the 6 questions in the original *DSM-IV*/ACDS screening scale and assessed the operating characteristics of this updated *DSM-IV*/ACDS scale for predicting *DSM-5* adult ADHD by optimizing response option scores (**Table 3**). The results showed that this updated *DSM-IV* scale underperforms relative to the *DSM-5* version but still detects most *DSM-5* cases (sensitivity, 84.2% in the general population samples and 79.8% in the NYU Langone sample) at a threshold with a relatively low false-positive rate (1 – specificity, 10.5% in the general population sample and 10.2% in the NYU Langone sample), but at the expense of a highly upwardly biased prevalence estimate.

Discussion

The new *DSM-5* ASRS screening scale showed excellent operating characteristics, given that it correctly classified nearly all people who met diagnostic criteria for ADHD in clinical interviews both in the general population (where prevalence was low and cases were often mild) and in a specialty treatment setting (where prevalence was high and cases were often severe). In particular, 67.3% of screened positives in the general population and more than 80% in the specialty treatment sample were confirmed as having the disorder at a threshold that had sensitivity above 90% and yielded few false-positive results. These good operating characteristics mean that the ASRS scale could be used as a practical screen for *DSM-5* adult ADHD despite a number of the screening scale questions not being *DSM-5* symptoms, since the scale captures nearly all *DSM-5* cases above the screening threshold with high enough PPV that a full clinical assessment of *DSM-5* criteria could be carried out in screened positives to distinguish individuals with ADHD from those without ADHD requiring the evaluation of many of those without the disorder.

Limitations

The analysis was limited by the general population training samples being relatively small and restricted either to people

Table 3. Operating Characteristics of the *DSM-IV* ASRS Screening Scale Optimally Scored With RiskSLIM to Predict *DSM-5*/ACDS ADHD

Score Threshold	Predicted Prevalence, %	Sensitivity, %[a]	Specificity, %[b]	AUC[c]	PPV, %[d]
Pooled NCS-R and managed care development samples[e]					
≥9 vs 0-8	32.9	94.1	72.6	.83	23.5
≥10 vs 0-9	24.9	89.7	80.9	.85	29.5
≥11 vs 0-10	16.6	84.2	89.5	.87	41.7
≥12 vs 0-11	6.0	57.9	98.7	.78	79.6
≥13 vs 0-12	4.3	44.4	99.3	.72	84.7
NYU Langone validation sample[f]					
≥9 vs 0-8	63.3	90.2	72.2	.81	82.1
≥10 vs 0-9	56.3	85.0	82.7	.84	87.0
≥11 vs 0-10	50.3	79.8	89.8	.85	91.4
≥12 vs 0-11	45.3	74.6	94.5	.85	94.8
≥13 vs 0-12	34.0	58.4	99.2	.79	99.0

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; ASRS, Adult ADHD Self-Report Scale; AUC, area under the curve; NCS-R, National Comorbidity Survey Replication; NYU Langone, New York University Langone Medical Center Adult ADHD Program; PPV, positive predictive value.

[a] The proportion of *DSM-5* and Adult Clinical ADHD Diagnostic Scale (ACDS) cases classified as cases by the screener at the threshold.

[b] The proportion of *DSM-5*/ACDS noncases classified as noncases by the screener at the threshold.

[c] The probability that a randomly selected *DSM-5*/ACDS case would have a higher screening scale score than a randomly selected noncase at the threshold, with ties on the screening scale assigned a predicted probability of 0.500.

[d] The proportion of respondents with screening scale scores above the threshold that meet *DSM-5*/ACDS criteria for the disorder.

[e] Sample of 337 participants; weighted prevalence, 8.2%.

[f] Sample of 300; participants; prevalence, 57.7%.

aged 18 to 44 years (NCS-R) or to members of a managed care health plan. Even though the strong cross-validated results and good performance in an independent clinical sample reduce concerns about these limitations, it would be useful to validate the screening scale in other samples, such as primary care and workplace case-finding samples and subsamples of patients with presenting symptoms consistent with adult ADHD. Another limitation is that the clinical interviews did not include informant reports, which improve detection of individuals with ADHD who lack insight into their conditions.[24] Our screening scale should consequently be considered optimal only for people with sufficient insight to recognize their symptoms, although even specialty treatment centers typically base diagnoses on patient reports rather than also requiring informant reports. An additional limitation is that clinical standard interviews did not evaluate important exclusionary comorbid diagnoses (eg, bipolar disorder). Another limitation is that, although the ACDS scale has been used previously in both academic research and US Food and Drug Administration registration trials for ADHD medications, no formal reliability and validity studies have been performed. Although having such studies would be ideal, it is unlikely that low reliability or validity would have led to the findings reported in this and other studies using the ACDS scale.

Conclusions

We found that a new machine-learning algorithm was able to build a screening scale for *DSM-5* adult ADHD from responses to 6 questions in the WHO ASRS screening scale using optimal integer scoring rules that had excellent cross-validated concordance with blinded clinical diagnoses both in the general population and among individuals seeking help. The new scale is short, easily scored, can detect nearly all of adult ADHD cases in the general population with high sensitivity and specificity, and also discriminates well among patients presenting for evaluation and specialty treatment.

ARTICLE INFORMATION

Accepted for Publication: February 8, 2017.

Correction: This article was corrected on October 25, 2017, to fix an error in the Results section; and on October 2, 2019, to fix errors in the footnote to Table 1..

Published Online: April 5, 2017. doi:10.1001/jamapsychiatry.2017.0298

Author Affiliations: Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge (Ustun); Department of Psychiatry, New York University Langone School of Medicine, New York (Adler); Department of Child and Adolescent Psychiatry, New York University Langone School of Medicine (Adler); Department of Computer Science, Duke University, Durham, North Carolina (Rudin); Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina (Rudin); Department of Psychiatry, State University of New York Upstate Medical University, Syracuse (Faraone); Department of Neuroscience and Physiology, State University of New York Upstate Medical University, Syracuse (Faraone); Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston (Spencer); Institute for Social Research, University of Michigan, Ann Arbor (Berglund); Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts (Gruber, Kessler).

Author Contributions: Dr Kessler had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Ustun, Adler, Spencer, Kessler.
Acquisition, analysis, or interpretation of data: Adler, Rudin, Faraone, Spencer, Berglund, Gruber, Kessler.
Drafting of the manuscript: Ustun, Adler, Berglund, Kessler.
Critical revision of the manuscript for important intellectual content: Adler, Rudin, Faraone, Spencer, Gruber, Kessler.
Statistical analysis: Ustun, Adler, Rudin, Faraone, Berglund, Gruber.
Obtained funding: Kessler.

Administrative, technical, or material support: Adler, Kessler.
Supervision: Kessler.

Conflict of Interest Disclosures: Dr Adler has received grant and research support from Sunovion Pharmaceuticals, Purdue Pharmaceuticals, Enzymotec, Shire Pharmaceuticals, and Lundbeck; served as a consultant for Enzymotec, Alcobra Pharmaceuticals, Rhodes Pharmaceuticals, Shire Pharmaceuticals, National Football League, and Major League Baseball; and has received royalty payments (as inventor) from New York University for license of adult attention-deficit/hyperactivity disorder (ADHD) scales and training materials since 2004. With his institution, Dr Faraone has US patent US20130217707 A1 for the use of sodium-hydrogen exchange inhibitors in the treatment of ADHD and has received income, potential income, travel expenses, and/or research support from Rhodes Pharmaceuticals, Arbor Pharmaceuticals, Pfizer, Ironshore, Shire, Akili Interactive Labs, CogCubed, Alcobra, VAYA Pharma, NeuroLifeSciences, and NACE. Dr Spencer receives research support from or is a consultant for Alcobra Pharma, Avekshan, Heptares Therapeutics, Impax Laboratories, Ironshore, Lundbeck, Shire Pharmaceuticals, Sunovion Pharmaceuticals, VAYA Pharma, the US Food and Drug Administration, and the US Department of Defense; serves on an advisory board for Alcobra Pharmaceuticals; receives research support from Royalties and Licensing fees on copyrighted ADHD scales through Massachusetts General Hospital (MGH) Corporate Sponsored Research and Licensing; and has a US Patent Application pending (provisional number 61/233,686), through MGH corporate licensing on a method to prevent stimulant abuse; consultant fees are paid to the MGH Clinical Trials Network and not directly to Dr Spencer. Dr Kessler has received support for his epidemiologic studies from Sanofi; was a consultant for Johnson & Johnson Wellness and Prevention, Shire Pharmaceuticals, and Takeda Pharmaceuticals; served on an advisory board for the Johnson & Johnson Services Inc Lake Nona Life Project; and is a co-owner of DataStat, Inc, a market research firm that carries out healthcare research. No other conflicts were reported.

Funding/Support: The Adult Attention-Deficit/Hyperactivity Disorder Self-Report (ASRS) scale was developed in collaboration with the World Health Organization (WHO) World Mental Health (WMH) Survey Initiative in conjunction with the revision of the WHO Composite International Diagnostic Interview (CIDI)[25] with support from Eli Lilly and Company. The ASRS Screener cross-validation study was supported by Eli Lilly and Company and HealthCore Inc. The larger series of WMH methodological studies in which the ASRS development work was embedded were also supported National Institute of Mental Health grant R01 MH070884; the John D. and Catherine T. MacArthur Foundation; the Pfizer Foundation; US Public Health Service grants R13-MH066849, R01-MH069864, and R01 DA016558; the Fogarty International Center grant R01-TW006481; the Pan American Health Organization; Eli Lilly and Company; Ortho-McNeil Pharmaceutical Inc; GlaxoSmithKline; and Bristol-Myers Squibb. The *DSM-IV* ASRS Screening Scale was developed with support from an unrestricted educational grant from the Eli Lilly Company. The *DSM-5* ASRS Screening Scale was developed with support from Shire Pharmaceuticals.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Information: The WMH CIDI Advisory Group that participated in developing the ASRS included Lenard A. Adler, MD (New York University Langone School of Medicine), Russell Barkley, MD (Medical University of South Carolina), Joseph Biederman, MD (Massachusetts General Hospital and Harvard Medical School), Keith Conners, PhD (Duke University School of Medicine), Stephen V. Faraone, PhD (State University of New York Upstate Medical University), Laurence Greenhill, MD (New York State Psychiatric Institute), Molly Howes, PhD (Harvard Medical School), Ronald C. Kessler, PhD (Harvard Medical School), Thomas J. Spencer, MD (Massachusetts General Hospital and Harvard Medical School), and T. Bedirhan Ustun, MD (World Health Organization). None of the members of the Advisory Group were compensated for the current analysis.

REFERENCES

1. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. *Psychol Med*. 2006;36(2):159-165.

2. Sobanski E. Psychiatric comorbidity in adults with attention-deficit/hyperactivity disorder (ADHD). *Eur Arch Psychiatry Clin Neurosci*. 2006;256(suppl 1):i26-i31.

3. de Graaf R, Kessler RC, Fayyad J, et al. The prevalence and effects of adult attention-deficit/hyperactivity disorder (ADHD) on the performance of workers: results from the WHO World Mental Health Survey Initiative. *Occup Environ Med*. 2008;65(12):835-842.

4. Küpper T, Haavik J, Drexler H, et al. The negative impact of attention-deficit/hyperactivity disorder on occupational health in adults and adolescents. *Int Arch Occup Environ Health*. 2012;85(8):837-847.

5. London AS, Landes SD. Attention deficit hyperactivity disorder and adult mortality. *Prev Med*. 2016;90:8-10.

6. Fayyad J, Sampson NA, Hwang I, et al. The descriptive epidemiology of DSM-IV adult ADHD in the World Health Organization World Mental Health Surveys. *Atten Defic Hyperact Disord*. 2017;9(1):47-65.

7. Surman CB, Hammerness PG, Pion K, Faraone SV. Do stimulants improve functioning in adults with ADHD? a review of the literature. *Eur Neuropsychopharmacol*. 2013;23(6):528-533.

8. Minkoff NB. ADHD in managed care: an assessment of the burden of illness and proposed initiatives to improve outcomes. *Am J Manag Care*. 2009;15(5)(suppl):S151-S159.

9. Kessler RC, Lane M, Stang PE, Van Brunt DL. The prevalence and workplace costs of adult attention deficit hyperactivity disorder in a large manufacturing firm. *Psychol Med*. 2009;39(1):137-147.

10. Adler LA, Shaw DM, Alperin S. ADHD diagnostic and symptom assessment scales for adults. In: Adler LA, Spencer TJ, Wilens T, eds. *Attention-Deficit Hyperactivity Disorder in Adults and Children*. New York, NY: Cambridge University Press; 2014:224-232.

11. Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychol Med*. 2005;35(2):245-256.

12. Kessler RC, Adler LA, Gruber MJ, Sarawate CA, Spencer T, Van Brunt DL. Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) screener in a representative sample of health plan members. *Int J Methods Psychiatr Res*. 2007;16(2):52-65.

13. Kessler RC, Berglund P, Chiu WT, et al. The US National Comorbidity Survey Replication (NCS-R): design and field procedures. *Int J Methods Psychiatr Res*. 2004;13(2):69-92.

14. Brod M, Johnston J, Able S, Swindle R. Validation of the Adult Attention-Deficit/Hyperactivity Disorder Quality-of-Life Scale (AAQoL): a disease-specific quality-of-life measure. *Qual Life Res*. 2006;15(1):117-129.

15. Kessler RC, Green JG, Adler LA, et al. Structure and diagnosis of adult attention-deficit/hyperactivity disorder: analysis of expanded symptom criteria from the Adult ADHD Clinical Diagnostic Scale. *Arch Gen Psychiatry*. 2010;67(11):1168-1178.

16. Wender P. *Attention-Deficit Hyperactivity Disorder in Adults*. New York, NY: Oxford University Press; 1998.

17. Adler L, Cohen J. Diagnosis and evaluation of adults with attention-deficit/hyperactivity disorder. *Psychiatr Clin North Am*. 2004;27(2):187-201.

18. Adler LA, Shaw DM, Kovacs K, Alperin S. Diagnosing ADHD in children and adults. In: Adler LA, Spencer TJ, Wilens T, eds. *Attention-Deficit Hyperactivity Disorder in Adults and Children*. New York, NY: Cambridge University Press; 2014:16-23.

19. Ustun B, Rudin C. Learning optimized risk scores on large-scaled datasets. arXiv:1610.00168. http://arxiv.org/abs/1610.00168. Updated February 7, 2017. Accessed October 1, 2016.

20. Ustun B, Rudin C. Supersparse linear integer models for optimized medical scoring systems. *Mach Learn*. 2016;102(3):349-391.

21. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36.

22. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence. Burlington, MA: Morgan Kaufmann; 1995:1137-1143.

23. Wynants L, Bouwmeester W, Moons KG, et al. A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. *J Clin Epidemiol*. 2015;68(12):1406-1414.

24. Sibley MH, Mitchell JT, Becker SP. Method of adult diagnosis influences estimated persistence of childhood ADHD: a systematic review of longitudinal studies. *Lancet Psychiatry*. 2016;3(12):1157-1165.

25. Kessler RC, Ustün TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res*. 2004;13(2):93-121.