Title

Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity

Journal

Journal of Child Psychology and Psychiatry, 58(6)

ISSN

0021-9630

Authors

Sibley, Margaret H
Swanson, James M
Arnold, L Eugene
et al.

Publication Date

2017-06-01

DOI

10.1111/jcpp.12620

Peer reviewed

HHS Public Access

Author manuscript

J Child Psychol Psychiatry. Author manuscript; available in PMC 2018 June 01.

Published in final edited form as:

J Child Psychol Psychiatry. 2017 June ; 58(6): 655–662. doi:10.1111/jcpp.12620.

Defining ADHD Symptom Persistence in Adulthood: Optimizing Sensitivity and Specificity

Margaret H. Sibley, Ph.D.[1], **James M. Swanson, Ph.D.**[2], **L. Eugene Arnold, M.D.**[3], **Lily T. Hechtman, M.D.**[4], **L. Elizabeth Owens, Ph.D.**[5], **Annamarie Stehli, M.P.H.**[6], **Howard Abikoff, Ph.D.**[7], **Stephen P. Hinshaw, Ph.D.**[8], **Brooke S.G. Molina, Ph.D.**[9], **John T. Mitchell, Ph.D.**[10], **Peter S. Jensen, M.D.**[11], **Andrea Howard, Ph.D.**[12], **Kimberley D. Lakes, Ph.D.**[13], **William E. Pelham, Ph.D.**[14], and **for the MTA Cooperative Group**

[1]Department of Psychiatry, Florida International University

[2]Child Development Center, School of Medicine, University of California, Irvine

[3]Department of Psychiatry, Ohio State University, Nisonger Center, Columbus Ohio

[4]Division of Child Psychiatry, McGill University, Montreal Children's Hospital, Montreal, Quebec, Canada

[5]Institute of Human Development, University of California, Berkeley

[6]Department of Pediatrics, University of California, Irvine

[7]Child Study Center at New York University Langone Medical Center

[8]Department of Psychology, University of California, Berkeley

[9]Departments of Psychiatry and Psychology, University of Pittsburgh School of Medicine

[10]Department of Psychiatry & Behavioral Sciences, Duke University Medical Center

[11]The REACH Institute

[12]Department of Psychology, Carleton University

[13]Department of Pediatrics, University of California, Irvine

[14]Deparatment of Psychology, Florida International University

Abstract

Objective—Longitudinal studies of children diagnosed with ADHD report widely ranging ADHD persistence rates in adulthood (5-75%). This study documents how information source (parent vs. self report), method (rating scale vs. interview), and symptom threshold (DSM vs. norm-based) influence reported ADHD persistence rates in adulthood.

Correspondence: Margaret H. Sibley, Department of Psychiatry, Florida International University, 11200 SW 8th Street, AHC1 Room 146, Miami, FL 33133, USA; msibley@fiu.edu.

Conflict of interest statement: See Acknowledgements for full disclosures.

The remaining authors have no conflicts to disclose.

The remaining authors have declared that they have no competing or potential conflicts of interest to declare.

Method—579 children were diagnosed with DSM-IV ADHD-Combined Type at baseline (ages 7.0-9.9 years) and 289 classmates served as a local normative comparison group (LNCG), 476 and 241 of whom respectively were evaluated in adulthood (Mean Age= 24.7). Parent and self reports of symptoms and impairment on rating scales and structured interviews were used to investigate ADHD persistence in adulthood.

Results—Persistence rates were higher when using parent rather than self reports, structured interviews rather than rating scales (for self report but not parent report), and a norm-based (NB) threshold of 4 symptoms rather than DSM criteria. Receiver-Operating Characteristics (ROC) analyses revealed that sensitivity and specificity were optimized by combining parent and self reports on a rating scale and applying a NB threshold.

Conclusion—The interview format optimizes young adult self-reporting when parent reports are not available. However, the combination of parent and self reports from rating scales, using an "or" rule and a NB threshold optimized the balance between sensitivity and specificity. With this definition, 60% of the ADHD group demonstrated symptom persistence and 41% met both symptom and impairment criteria in adulthood.

Introduction

Longitudinal studies of children diagnosed with attention-deficit hyperactivity disorder (ADHD) and followed into adulthood present an extraordinarily wide range of ADHD persistence rates (5% to 75%; Barkley, Murphy, & Fischer, 2008; Biederman, Petty, Clarke, Lomedico, & Faraone, 2011; Biederman, Petty, O'Connor, Hyder, & Faraone, 2012; Halperin et al., 2008; Hinshaw et al., 2012; Klein et al., 2012; Mannuzza, Klein, & Moulton, 2002; Sibley et al., 2012; Weiss & Hechtman, 1993). Sample heterogeneity across these studies is probably not the sole explanation. Close examination of persistence definitions reveals considerable variability in sources, methods, and symptom thresholds used to define persistence. Source of information may include only self (Barkley et al., 2008; Klein et al., 2012; Sibley et al., 2012; Weiss & Hechtman, 1993) or parent report (Barkley et al., 2008; Sibley et al., 2012) or their combination (Biederman et al., 2011, 2012; Halperin et al., 2008; Hinshaw et al., 2012; Mannuzza et al., 2002; Sibley et al., 2012). Method of data collection may involve rating scales (Sibley et al., 2012; Weiss & Hechtman, 1993), structured interviews (Barkley et al., 2002; Hinshaw et al., 2012), or semi-structured interviews (Barkley et al., 2008; Biederman et al., 2011, 2012; Halperin et al., 2008; Klein et al., 2012; Mannuzza et al., 2002). Symptom threshold may reflect DSM criteria (Barkley et al., 2002; Biederman et al., 2011, 2012; Halperin et al., 2008; Hinshaw et al., 2021; Klein et al., 2012; Mannuzza et al., 2002; Sibley et al., 2012) or a developmentally adjusted, norm-based (NB) threshold (Barkley et al., 2008; Sibley et al., 2012; Weiss & Hechtman, 1993). Some investigators report symptom persistence that merely exceeds this threshold (Barkley et al., 2008; Biederman et al., 2011, 2012; Sibley et al., 2012), but others require impairment and rule out other mental health disorders as the source of symptoms (Biederman et al., 2011, 2012; Halperin et al., 2008; Klein et al., 2012; Mannuzza et al., 2002; Weiss & Hechtman, 1993). Determining the optimal method of defining ADHD symptom persistence in adults has important implications for clinical evaluation and treatment of adult ADHD.

Herein, we examine persistence of ADHD symptoms in the Multimodal Treatment Study of Children with ADHD (MTA; MTA Cooperative Group, 1999) in young adulthood, when the mean age of the sample was 24.7 years (range 19 to 28). Our interest in defining persistence began in the 8-year follow-up, when only 30% of the adolescent cases in the MTA met DSM-IV criteria for diagnosis of ADHD, even though many more cases displayed elevated symptoms and impairment (Molina et al., 2009). We suggested that this rate was an underestimate due to "…. symptom-count thresholds developed for the diagnosis of ADHD in children that may be overly stringent for adolescents and adults" (p.497).

In a companion investigation, we showed persistence of ADHD symptom severity through early adulthood to age 24.7 (Swanson et al., under review). Here we evaluate factors that affect observed categorical adult ADHD persistence rates, to generate a definition of persistence that optimizes sensitivity and specificity. The MTA is well-positioned to examine this question with prospective collection of data on symptoms and impairment from multiple sources and methods in adolescence and early adulthood. Our aims were (1) to identify factors that contribute to the wide range of persistence rates reported in longitudinal studies, aiding in interpretation of disparate results reported in past studies and (2) to determine the optimal method for ascertaining symptom persistence of ADHD into adulthood.

Methods

The MTA (MTA Cooperative Group, 1999) was originally designed to compare 2-year effects of pharmacological and psychosocial treatments for children (7.0 to 9.9 years old) with ADHD-Combined Type. Two years after baseline, 289 classmates were recruited as a local normative comparison group (LNCG). Due to our recruitment strategies, the ADHD and LNCG groups were not significantly different in childhood on sex, age, and minority status (White, non-Hispanic versus other). However, the LNCG had slightly higher socioeconomic advantage than the ADHD group. The MTA continued with prospective follow-up approximately biennially until 16 years after baseline (Jensen et al., 2007; Molina et al., 2007, 2009, 2013). Informed consent was obtained for all participants in adulthood.

Participants

The current subsample includes participants with at least one adult assessment (12, 14, or 16 year follow-up). Overall retention rate in adulthood was 82% for the ADHD group (N=476 out of 579) and 94% for the LNCG (N=272 out of 289); however, the current subsample excluded an additional 31 LNCG participants with a baseline diagnosis of ADHD and 23 ADHD participants with only one reporter at adult follow-up when that reporter indicated less than four ADHD symptoms. The latter decision was based on concerns that underreporting by adults with ADHD (Sibley et al., 2012; Swanson et al., under review) might lead to false negative diagnoses in the absence of supplemental informant report. The resulting subsample represented 78% of the original ADHD group and 83% of the original LNCG. Average age at provision of adult report was 24.8 (SD= 1.31) for the ADHD group and 24.4 (SD= 1.36) for the LNCG.

Procedures

The MTA childhood assessment protocol (MTA Cooperative Group, 1999) was adapted to be age-appropriate for participants≥18 years during later follow up assessments. These assessments were administered by trained bachelor's level assessment staff who were trained to be objective, but were not blind to participant group.

Measures

ADHD Symptoms—ADHD symptoms were measured by the Conners Adult ADHD Rating Scale (CAARS; Conners, Erhardt, & Sparrow, 1999) and the young adult and parent versions of the Diagnostic Interview Schedule for Children (DISC; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000) structured interview, completed by two sources (self, parent). The CAARS is a 72-item scale, which includes the 18 DSM-IV-TR symptoms of ADHD. ADHD symptoms were deemed present on the 0-3 CAARS scale if the respondent endorsed “2=Pretty Much” or “3=Very Much.” The DISC is a structured interview that queries the presence of each ADHD symptom (0=No, 1=Yes). Though supplemental probes for symptom-specific impairment are included (Shaffer et al., 2000), these probes were not considered when deeming a symptom as present.

Impairment—Impairment was evaluated via the Impairment Rating Scale (IRS; Fabiano et al., 2006) an individually administered paper and pencil 7-point scale (0=no problem, 6=extreme problem) that measures severity of adult impairment both globally and across separate functional domains. The IRS demonstrates strong psychometric properties for identifying impairment in adults with ADHD (Fabiano et al., 2006; Sibley et al., 2012). Based on the work of Fabiano and colleagues (2006) and Sibley and colleagues (2012) we used a cutoff of ≥3 on the IRS overall impairment item as endorsed by either informant (self, parent) to designate clinically significant impairment.

Analyses

Analysis 1 investigated how persistence rates vary when applying different methods, sources, and symptom thresholds. We anticipated a NB threshold of four symptoms of Inattention (IN) or Hyperactivity/Impulsivity (HI) based on past research (Barkley et al., 2008; Sibley et al., 2012) and verified this threshold by calculating the LNCG symptom count mean+2 standard deviations across measures and sources. Analysis 1 compared 36 candidate persistence definitions, generated by crossing two instruments (CAARS vs. DISC), three symptom thresholds (6 symptoms: DSM-IV; 5 symptoms: DSM-5; and 4 symptoms: NB), and six informant sources (parent and young adult reports examined alone and in four combinations). Four combination rules were derived by crossing two criteria for symptom presence (“and” rule: symptom counted present only if endorsed by both sources; “or” rule: symptom counted present if endorsed by either source) with two combination procedures (combining at the item level vs. at the symptom count level). Item level combination occurred by first combining reports at the item level and then counting symptoms. Count level combination occurred by first counting symptoms separately for each source and then using the higher estimate to represent symptom level. McNemar's chi-square tests of marginal probability were used to test specific hypotheses about persistence

rates using an SPSS Macro (Newcombe, 1998). Bonferroni corrections were enforced to correct for multiple comparisons within set.

Receiver Operating Curve (ROC) analyses (Analysis 2) compared six diagnostic methods to compare sensitivity and specificity when detecting cases who met four basic criteria for ADHD in adulthood (elevated symptoms, current impairment, and childhood onset/chronicity; Hanley & McNeil, 1982). Cases who met the symptom threshold but did not meet impairment or age of onset/chronicity criteria were considered to be false positives. For each evaluated diagnostic method, ROC curves plot true positives on the vertical axis and false positives on the horizontal access, creating a probability curve that indicates the balance between sensitivity and specificity as a function of the plot's area under the curve (AUC). In the ROC analysis, we examined the AUC values and their confidence intervals to identify the statistically optimal solution. Fifteen paired comparisons between curves were conducted to assess statistically significant differences between AUCs using a method developed by DeLong and colleagues (1988). Bonferroni correction was enforced to correct for multiple comparison ($p<.003$).

Results

Analysis 1

Table 1 reveals that across methods and sources, the average of the LNCG M+2SD thresholds was 4.18 for IN and 2.94 for HI, which conservatively supports the previously suggested NB threshold of four symptoms of either IN or HI (Barkley et al., 2008). This threshold of four DSM symptoms henceforth served as the NB symptom threshold for all analyses. Table 2 displays the percentages of cases deemed persistent for 36 combinations of method, source, and threshold. McNemar's tests (see Table 2) confirmed that, as expected: higher rates of symptom persistence were obtained when using an “or” vs. “and” rule (6 contrasts); for the “or” rule, item-level combination led to significantly higher persistence rates than count-level combination (6 contrasts); parent reports led to significantly higher persistence rates than self reports (6 contrasts); combined report (using an “or” item-level rule) led to significantly higher persistence rates than parent report (6 contrasts); for self-report, ratings on the DISC led to significantly higher persistence rates than the CAARS (3 contrasts); for parent report, ratings on the DISC and CAARS were not significantly different (3 contrasts); and the NB threshold led to significantly higher persistence rates than the DSM-5 criterion (2 contrasts), which resulted in higher persistence rates than the DSM-IV (2 contrasts).

Analysis 2

Given evident incremental information when combining parent and self reports (Table 2), we specified the ROC analysis to compare six different diagnostic methods that varied on instrument (CAARS vs. DISC) and threshold (DSM-IV, DSM-5, NB), but not source (all six final candidate definitions used combined report and an item-level “or” rule). All six methods discriminated true positive and negative cases at a level that was significantly greater than chance (AUC>.5, $p<.001$; Figure 1 & Table 3). A stepwise series of curves indicated an increasing ratio of sensitivity to specificity when moving from the DSM-IV to

NB symptom count threshold for both the DISC and CAARS. After correcting for multiple tests, paired z-score comparisons indicated that these differences were statistically significant for comparison of the NB threshold on the CAARS vs. the DSM-IV threshold on the DISC and comparison of the DSM-IV vs. DSM-5 thresholds on the CAARS. These results indicate that balance between sensitivity and specificity was maximized using combined parent and self reports from the CAARS and applying a NB threshold. Table 4 displays classification rates for each candidate persistence definition separated by childhood diagnostic group. Under this optimal method, 41.1% of the ADHD group met both symptom and impairment criteria in adulthood and 7.1% of the LNCG appeared to experience onset of above threshold ADHD symptoms after childhood.

Discussion

We evaluated 36 candidate definitions for symptom persistence of ADHD into adulthood based on combinations of sources, methods, and symptom thresholds. Across these definitions, symptom persistence rates varied dramatically (from 1.9% to 61.4%--see Table 2). Prevalence analyses suggested that parent and self reports offered unique diagnostic information. Findings suggest that when using combined parent and young adult reports the balance between diagnostic sensitivity and specificity was optimized by a rating scale method (e.g., the CAARS) and a NB symptom threshold (i.e., 4 symptoms of either IN or HI). Using this definition, symptom persistence of ADHD in young adulthood was approximately 60% (Table 2) and 41.1% of the ADHD group met both the optimized persistence criteria and presence of impairment (see Table 4).

Persistence rates reported in other studies were also substantially higher when using combined vs. parent-only or self-only reports and NB vs. strict DSM criteria (Barkley et al., 2002; Biederman et al., 2011, 2012; Halperin et al., 2008; Hinshaw et al., 2012; Klein et al., 2012; Mannuzza et al., 2002; Sibley et al., 2012; Weiss & Hechtman, 1993). These estimates are elevated by virtue of casting a wider net on symptoms; however, doing so appears key to preserving diagnostic sensitivity since symptoms of ADHD in adults may be subjective, often unrecognized by the patient, and difficult for informants to observe. The incremental benefit of combined report was particularly expected given noted symptom underreporting by adults with ADHD—a tendency attributed to a self-perception bias or inattentiveness during assessment (Molina & Sibley, 2014). There is additional concern that parents of adults have only intermittent interactions with their offspring, limiting familiarity with the individual's daily functioning. Thus, use of parent or self reports alone may falsely deflate persistence rates in adulthood. Obtaining parent report for young adults may be challenging in settings where a parent is not immediately accessible during the assessment. Though it may require extra efforts to obtain these reports, doing so not only enhances the collection of accurate information about the adult's current functioning, but also may clarify symptom chronicity during childhood and adolescence. This requirement is particularly critical in young adulthood, when false positive diagnoses may be common and duplicitously sought by young adults without ADHD who seek diagnoses to obtain stimulant medication prescriptions or educational accommodations (Molina & Sibley, 2014). In cases where it is impossible to obtain a parent report, the report of another informant (e.g., spouse, sibling,

roommate, coworker, supervisor) can be useful given the important limitations to self-report (see Table 2).

The optimal trade-off between sensitivity (.86) and specificity (.73) was obtained using a NB criterion (see Table 4). When defining persistence in research or epidemiological settings, optimizing sensitivity and specificity refines estimation of disorder prevalence. Overall, false negative diagnoses were substantially higher (19.4% vs. 6.6%) under the strict DSM threshold compared to a NB one (see Table 4). Perhaps the gravest implication of false negative diagnoses is failure to provide treatment to individuals with impairing symptoms. On the other hand, false positive diagnoses among unimpaired individuals with a childhood history of ADHD were higher when using a NB threshold (19.9%) vs. strict DSM criteria (9.9%). These individuals likely represented childhood cases with well-managed symptoms that are no longer impairing. Importantly, these false positive diagnoses do not likely pose a true threat to false diagnosis in clinical settings because they do not meet the DSM impairment criterion.

Most longitudinal studies solely used structured interviews to ascertain symptom presence; thus, our report offers a first glimpse of the relative utility of rating scales vs. structured interview. These two methods led to similar persistence rates when using parent reports alone; however, the interview produced a significantly higher persistence estimate than the rating scale when considering self reports alone. A face-to-face format may enhance accurate reporting among individuals with ADHD, who may rush through or carelessly complete rating scales. However, ROC curves indicated that combined report on a rating scale possessed stronger sensitivity and specificity than the DISC. Thus, this method is typically advised; however, the interview may be particularly useful for symptom detection when parent or facsimile report is unavailable. Although we did not test the utility of semi-structured interviews, this format may be particularly helpful for detecting symptoms in young adults when an informant report is unavailable. Unlike structured interviews, semi-structured interviews allow clinicians to probe the presence of symptoms when the interviewee provides ambiguous or incomplete information upon initial query. Since clinical diagnosis requires interview to assess the fifth DSM criterion—that symptoms are not attributable to another disorder—interviews should continue to play an important role in clinical assessment, despite the superior symptom detection properties of the CAARS.

Using our optimal definition of persistence, there was a 7.1% rate of de novo adult ADHD cases in the LNCG. Many de novo cases meet symptom and impairment criteria for the disorder. Some of these may be true positive cases of adult-onset ADHD if the symptoms are not due to another disorder. These individuals may have experienced subclinical ADHD symptoms prior to adulthood that became significant with the increase of environmental demands in adulthood. On the other hand, it is possible that these cases possess a qualitatively different type of attention problems that represents a distinct adult-onset disorder. Other de novo cases may be false positives, impaired individuals with mood, anxiety, or other disorders that share symptoms with ADHD; these are clinically ruled out by the fifth DSM criterion. LNCG individuals who are unimpaired but met criteria for symptom presence (7.4%) are unlikely to be diagnosed in clinical settings because the DSM requires impairment for diagnosis.

The most important limitation to this study is that the symptoms of ADHD are subjective. The field lacks a gold standard objective test of this disorder. Our ROC analyses employ three important markers of adult ADHD (current symptoms, current impairment, and childhood ADHD history) as stand-ins for a true gold standard criterion. Additionally, because of the nature of the IRS—a single item, reported by the same informants who provided symptom information—we were unable to fully determine the source of impairments. Approximately 6% of the ADHD group reported significant impairment in the absence of clinically elevated symptoms (by any definition; see Table 4) and some of these individuals may have been misclassified if their impairments were due to a mental health disorder other than ADHD. Though we do not believe that the small number of individuals affected by this limitation would meaningfully influence analyses, approximately 2% of the desistent ADHD group had a mood disorder and 5% had an anxiety disorder at adult follow-up, while a drug use disorders were present in 12% of desistent cases (Hechtman et al., under review). In addition, to protect against false negative classification due to underreporting by young adults with ADHD (Barkley et al., 2008; Sibley et al., 2012), we required two sources to verify full absence of symptoms in a case. This led to the omission of 16 cases from analyses, and some may have been true negative cases. It is also important to note that the ADHD symptom count distribution in a normative sample is positively skewed; thus, using a mean + 2 SD approach to defining age-adjusted thresholds (Barkley et al., 2008) may not be optimal. In addition, our data suggested that an asymmetrical NB threshold of either four symptoms of IN or three symptoms of HI may warrant further investigation, but we did not pursue this. Finally, all study assessments were conducted by research staff who were not blind to initial group assignment.

Despite these limitations, our findings suggest that a thorough and optimal assessment of ADHD in adulthood should include: reports by other informants (e.g., parents)-- especially in cases when available self reports do not indicate clinically significant symptoms; rating scales to determine symptom count as a supplement to clinical or semi-structured interviews that assess presence of impairment, chronicity, and explanation of symptoms by other disorders; and a developmentally referenced threshold for symptom persistence. These strategies may lead to a narrower range of symptom persistence estimates and improved methods for evaluating prevalence in adulthood.

Acknowledgments

The work reported was supported by cooperative agreement grants and contracts from NIMH and the National Institute on Drug Abuse (NIDA) to the following: University of California–Berkeley: U01 MH50461, N01MH12009, and HHSN271200800005-C; DA-8-5550; Duke University: U01 MH50477, N01MH12012, and HHSN271200800009-C; DA-8-5554; University of California– Irvine: U01MH50440, N01MH12011, and HHSN271200800006- C; DA-8-5551; Research Foundation for Mental Hygiene (New York State Psychiatric Institute/Columbia University): U01 MH50467, N01 MH12007, and HHSN271200800007-C; DA-8-5552; Long Island–Jewish Medical Center U01 MH50453; New York University: N01MH 12004, and HHSN271200800004-C; DA-8-5549; University of Pittsburgh: U01 MH50467, N01 MH 12010, DA039881, and HHSN271200800008-C; DA-8-5553; and McGill University N01MH12008, and HHSN271200800003-C; DA-8-5548.

The Multimodal Treatment Study of Children with ADHD (MTA) was a National Institute of Mental Health (NIMH) cooperative agreement randomized clinical trial, continued under an NIMH contract as a follow-up study and finally under a National Institute on Drug Abuse (NIDA) contract. Collaborators from NIMH: Benedetto Vitiello (Child & Adolescent Treatment and Preventive Interventions Research Branch), Joanne B. Severe (Clinical Trials Operations and Biostatistics Unit, Division of Services and Intervention Research), Peter S. Jensen (currently at REACH Institute and Mayo Clinic), L. Eugene Arnold (currently at Ohio State University), Kimberly Hoagwood

(currently at Columbia); previous contributors from NIMH to the early phases: John Richters (currently at National Institute of Nursing Research); Donald Vereen (currently at NIDA). Principal investigators and co-investigators from the sites are: University of California, Berkeley/San Francisco: Stephen P. Hinshaw (Berkeley), Glen R. Elliott (San Francisco); Duke University: Karen C. Wells, Jeffery N. Epstein (currently at Cincinnati Children's Hospital Medical Center), Desiree W. Murray; previous Duke contributors to early phases: C. Keith Conners (former PI); John March; University of California, Irvine: James Swanson, Timothy Wigal, Ph.D.; previous contributor from UCLA to the early phases: Dennis P. Cantwell (deceased); New York University: Howard B. Abikoff; Montreal Children's Hospital/McGill University: Lily Hechtman; New York State Psychiatric Institute/Columbia University/ Mount Sinai Medical Center: Laurence L. Greenhill (Columbia), Jeffrey H. Newcorn (Mount Sinai School of Medicine). University of Pittsburgh: Brooke Molina, Ph.D., Betsy Hoza, Ph.D. (currently at University of Vermont), William E. Pelham, Ph.D. (PI for early phases, currently at Florida International University). Follow-up phase statistical collaborators: Robert D. Gibbons, Ph.D. (University of Illinois, Chicago); Sue Marcus (Mt. Sinai College of Medicine); Kwan Hur (University of Illinois, Chicago). Original study statistical and design consultant: Helena C. Kraemer (Stanford University). Collaborator from the Office of Special Education Programs/US Department of Education: Thomas Hanley. Collaborator from Office of Juvenile Justice and Delinquency Prevention/Department of Justice: Karen Stern.

Dr. Hechtman has received research support, served on advisory boards and has been a speaker for Ely Lilly, GlaxoSmithKline, Ortho Janssen, Purdue and Shire. Dr. Arnold has received research funding from Curemark, Forest, Lilly, Neuropharm, Novartis, Noven, Shire, and YoungLiving (as well as NIH and Autism Speaks) and has consulted with or been on advisory boards for Gowlings, Neuropharm, Novartis, Noven, Organon, Otsuka, Pfizer, Roche, Seaside Therapeutics, Sigma Tau, Shire, and Tris Pharma and received travel support from Noven. Dr. Swanson acknowledges research support, advisory board membership, speaker's bureau membership, and/or consulting for Alza, Richwood, Shire, Celgene, Novartis, Celltech, Gliatech, Cephalon, Watson, CIBA, UCB, Janssen, McNeil and Lilly.

L.T.H. has received research support, served on advisory boards and has been a speaker for Ely Lilly, GlaxoSmithKline, Ortho Janssen, Purdue and Shire. L.E.A. has received research funding from Curemark, Forest, Lilly, Neuropharm, Novartis, Noven, Shire, and YoungLiving (as well as NIH and Autism Speaks) and has consulted with or been on advisory boards for Gowlings, Neuropharm, Novartis, Noven, Organon, Otsuka, Pfizer, Roche, Seaside Therapeutics, Sigma Tau, Shire, and Tris Pharma and received travel support from Noven. Dr. Swanson acknowledges research support, advisory board membership, speaker's bureau membership, and/or consulting for Alza, Richwood, Shire, Celgene, Novartis, Celltech, Gliatech, Cephalon, Watson, CIBA, UCB, Janssen, McNeil and Lilly.

References

Barkley, RA., Murphy, KR., Fischer, M. ADHD in adults: What the science says. Guilford Press; 2008.

Barkley RA, Fischer M, Smallish L, Fletcher K. The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. Journal of abnormal psychology. 2002; 111:279. [PubMed: 12003449]

Biederman J, Petty CR, Clarke A, Lomedico A, Faraone SV. Predictors of persistent ADHD: an 11-year follow-up study. Journal of psychiatric research. 2011; 45:150–155. [PubMed: 20656298]

Biederman J, Petty CR, O'Connor KB, Hyder LL, Faraone SV. Predictors of persistence in girls with attention deficit hyperactivity disorder: results from an 11- year controlled follow- up study. Acta Psychiatrica Scandinavica. 2012; 125:147–156. [PubMed: 22097933]

Conners, CK., Erhardt, D., Sparrow, E. Conner's Adult ADHD Rating Scales: CAARS Toronto. MHS; 1999.

DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 44:837–845. [PubMed: 3203132]

Fabiano GA, Pelham WE Jr, Waschbusch DA, Gnagy EM, Lahey BB, Chronis AM, et al. Burrows-MacLean L. A practical measure of impairment: psychometric properties of the impairment rating scale in samples of children with attention deficit hyperactivity disorder and two school-based samples. Journal of Clinical Child and Adolescent Psychology. 2006; 35:369–385. [PubMed: 16836475]

Halperin JM, Trampush JW, Miller CJ, Marks DJ, Newcorn JH. Neuropsychological outcome in adolescents/young adults with childhood ADHD: profiles of persisters, remitters and controls. Journal of Child Psychology and Psychiatry. 2008; 49:958–966. [PubMed: 18573145]

Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982; 143:29–36. [PubMed: 7063747]

Hinshaw SP, Owens EB, Zalecki C, Huggins SP, Montenegro-Nevado AJ, Schrodek E, Swanson EN. Prospective follow-up of girls with attention-deficit/hyperactivity disorder into early adulthood: continuing impairment includes elevated risk for suicide attempts and self-injury. Journal of consulting and clinical psychology. 2012; 80:1041. [PubMed: 22889337]

Jensen PS, Arnold LE, Swanson JM, Vitiello B, Abikoff HB, Greenhill LL, et al. Hur K. 3-year follow-up of the NIMH MTA study. Journal of the American Academy of Child & Adolescent Psychiatry. 2007; 46:989–1002. [PubMed: 17667478]

Klein RG, Mannuzza S, Olazagasti MAR, Roizen E, Hutchison JA, Lashua EC, Castellanos FX. Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder 33 years later. Archives of general psychiatry. 2012; 69:1295–1303. [PubMed: 23070149]

Mannuzza S, Klein RG, Moulton JL Iii. Young adult outcome of children with “situational” hyperactivity: a prospective, controlled follow-up study. Journal of abnormal child psychology. 2002; 30:191–198. [PubMed: 12008657]

Molina BSG, Sibley MH. The Case for Including Informant Reports in the Assessment of Adulthood ADHD. The ADHD Report. 2014; 22(8):1–7. DOI: 10.1521/adhd.2014.22.8.1

Molina BS, Flory K, Hinshaw SP, Greiner AR, Arnold LE, Swanson JM, et al. Wigal T. Delinquent behavior and emerging substance use in the MTA at 36 months: prevalence, course, and treatment effects. Journal of the American Academy of Child & Adolescent Psychiatry. 2007; 46:1028–1040. [PubMed: 17667481]

Molina BS, Hinshaw SP, Arnold LE, Swanson JM, Pelham WE, Hechtman L, et al. MTA Cooperative Group. Adolescent substance use in the multimodal treatment study of attention-deficit/hyperactivity disorder (ADHD)(MTA) as a function of childhood ADHD, random assignment to childhood treatments, and subsequent medication. Journal of the American Academy of Child & Adolescent Psychiatry. 2013; 52:250–263. [PubMed: 23452682]

Molina BS, Hinshaw SP, Swanson JM, Arnold LE, Vitiello B, Jensen PS, et al. MTA Cooperative Group. The MTA at 8 years: prospective follow-up of children treated for combined-type ADHD in a multisite study. Journal of the American Academy of Child & Adolescent Psychiatry. 2009; 48:484–500. [PubMed: 19318991]

MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. Archives of General Psychiatry. 1999; 56:1073. [PubMed: 10591283]

Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. Statistics in medicine. 1998; 17:2635–2650. [PubMed: 9839354]

Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. Journal of the American Academy of Child & Adolescent Psychiatry. 2000; 39:28–38. [PubMed: 10638065]

Sibley MH, Pelham WE Jr, Molina BS, Gnagy EM, Waxmonsky JG, Waschbusch DA, et al. Kuriyan AB. When diagnosing ADHD in young adults emphasize informant reports, DSM items, and impairment. Journal of consulting and clinical psychology. 2012; 80:1052. [PubMed: 22774792]

Swanson JM, Arnold LE, Molina BSG, Sibley MH, Hechtman LT, et al. the MTA Cooperative Group. Prospective Follow-up of the Multimodal Treatment Study of ADHD: Effects of Medication, Maturation, and Source on Symptom Severity in Adulthood. Paper Submitted for Publication as a companion article to this one. under review.

Weiss, G., Hechtman, LT. Hyperactive children grown up: ADHD in children, adolescents, and adults. Guilford Press; 1993.

Key Points

- Longitudinal studies of children diagnosed with ADHD report widely ranging ADHD persistence rates in adulthood (5-75%).
- This study indicated that rates of ADHD persistence into adulthood vary greatly, depending on how investigators collect and analyze information: structured interviews vs. rating scales, self-reported vs. parent/other-reported information, and selection of symptom threshold.
- This study indicated that parent reports yielded higher persistence rates than self-reports.
- The combination of parent and self reports using an “or” rule and then applying an age-adjusted norm-based DSM threshold was considered optimal in providing a balance between sensitivity and specificity.
- With this definition, 60% of the ADHD group demonstrated symptom persistence and 41% met both symptom and impairment criteria in adulthood.

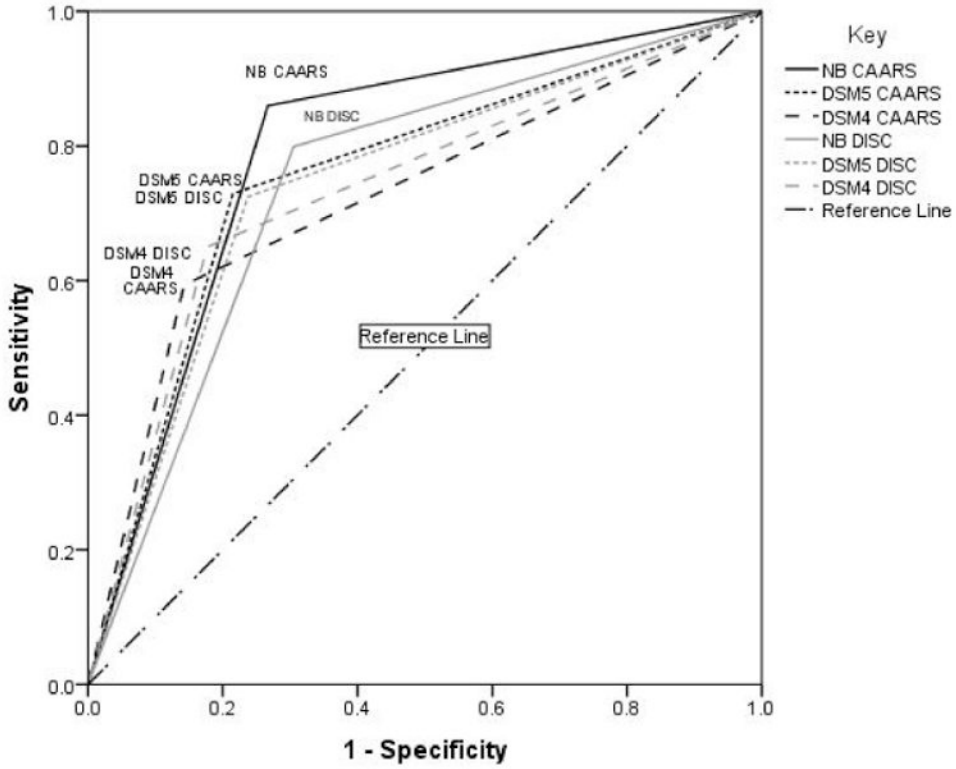

Figure 1. Receiver Operating Characteristics for Method and Threshold

Table 1

LNCG Symptom Count Means and Standard Deviations used to Derive Norm-based (Mean+2 SD) Thresholds

	CAARS Rating Scale						DISC Interview					
	IN			HI			IN			HI		
	M	SD	M+2 SD	M	SD	M+2 SD	M	SD	M+2 SD	M	SD	M+2 SD
Self report	.64	1.56	**3.76**	.83	1.29	**3.41**	.82	1.58	**3.98**	.64	1.27	**3.18**
Parent report	.84	2.01	**4.85**	.65	1.33	**3.31**	.79	1.67	**4.13**	.29	.78	**1.85**

[1] NB thresholds presented in bold.

Table 2

Variability in Symptom Persistence by Source, Method, and Threshold

	DISC Interview (%)			CAARS Rating Scale (%)		
	DSM-IV (6)	**DSM-5 (5)**	**NB (4)**	**DSM-IV (6)**	**DSM-5 (5)**	**NB (4)**
Self Report	17.7	25.3	32.2	12.1	15.9	24.5
Parent Report	28.4	35.8	45.1	28.7	35.8	44.6
Combined Sources						
And Rule/Count Level	5.8	10.6	16.2	6.2	8.9	14.8
And Rule/Item Level	**1.9**	2.9	4.4	3.8	5.3	7.5
Or Rule/Count Level	38.4	48.9	59.2	35.5	44.7	57.4
Or Rule/Item Level	44.0	52.5	**61.4**	38.2	49.9	60.1

Note. (6)= six symptom threshold, (5)=five symptom threshold, (4)=four symptom threshold. No DISC was available for fifteen individuals with CAARS rating scales. The four combination rules were derived by crossing two criteria for symptom presence ("and" rule symptom counted present only if endorsed by both sources; "or" rule: symptom present counted present if endorsed by either source) with two sequences to combine sources (item level: first combining reports at the item level and then counting symptoms vs. count level: first counting symptoms separately for each source and then using the higher estimate to represent symptom level). Values in bold represent the lowest and highest estimates provided using systematically varied diagnostic methods. The following comparisons were made for the persistence rates presented in Table 2. And rule vs. or rule comparisons: CAARS DSM-IV: 3.8 % vs. 38.2%, $\chi^2(1)=147.00$, $p<.001$, CAARS DSM-5: 5.3% vs. 49.9%, $\chi^2(1)=154.00$, $p<.001$, CAARS NB: 7.5% vs. 60.1%, $\chi^2(1)=216.00$, $p<.001$, DISC DSM-IV: 1.9% vs. 44.0%, $\chi^2(1)=111.00$, $p<.001$, DISC DSM-5: 2.9% vs. 52.5%, $\chi^2(1)=135.00$, $p<.001$, DISC NB: 4.4% vs. 61.4%, $\chi^2(1)=159.00$, $p<.001$); Item vs. symptom level comparisons: NB And rule: 4.4% vs. 16.2%, $\chi^2(1)=22.00$, $p<.001$, DSM-5 And rule: 10.6% vs. 2.9% $\chi^2(1)=48.00$, $p<.001$, DSM-IV And rule: 5.8% vs. 1.9%, $\chi^2(1)=9.00$, $p<.003$, NB Or rule: 57.4% vs. 60.1% $\chi^2(1)=20.00$, $p<.001$, DSM-5 Or rule: 44.7% vs. 49.9%, $\chi^2(1)=14.00$, $p<.001$, DSM-IV Or rule: 35.5% vs. 38.2%, $\chi^2(1)=28.00$, $p<.001$; self report vs. parent report comparisons: DISC DSM-IV: 17.7% vs. 28.4%, $\chi^2(1)=11.97$, $p<.001$, DISC DSM-5: 25.3% vs. 35.8%, $\chi^2(1)=9.47$, $p=.002$, DISC NB: 32.2% vs. 45.1% $\chi^2(1)=10.46$, $p<.001$, CAARS DSM-IV: 12.1% vs. 28.7%, $\chi^2(1)=50.00$, $p<.001$, CAARS DSM-5: 15.9% vs. 35.8%, $\chi^2(1)=59.08$, $p<.001$, CAARS NB: 24.5% vs. 44.6%, $\chi^2(1)=48.78$, $p<.001$; parent vs. combined report comparisons: DISC DSM-IV: 28.4% vs. 44.0%, $\chi^2(1)=41.00$, $p<.001$, DISC DSM-5: 35.8% vs. 52.5%, $\chi^2(1)=51.00$, $p<.001$, DISC NB: 45.1% vs. 61.4%, $\chi^2(1)=61.00$, $p<.001$, CAARS DSM-IV: 28.7% vs. 38.2%, $\chi^2(1)=38.00$, $p<.001$, CAARS DSM-5: 35.8% vs. 49.9%, $\chi^2(1)=58.00$, $p<.001$, CAARS NB: 44.6% vs. 60.1%, $\chi^2(1)=6500$, $p<.001$); DISC vs. CAARS comparisons: self DSM-IV: 17.7% vs. 15.9%, $\chi^2(1)=6.72$, $p<.009$, self DSM-5: 25.3% vs. 15.9%, $\chi^2(1)=22.77$, $p<.001$, self NB: 32.2% vs. 24.5%, $\chi^2(1)=15.80$, $p<.001$, parent DSM-IV: 28.4% vs. 28.7%, $\chi^2(1)=.29$, $p=.588$, parent DSM-5: 35.8% vs. 35.8%, $\chi^2(1)=.16$, $p<.686$, parent NB: 45.1% vs. 44.6%, $\chi^2(1)=.04$, $p<.840$); NB vs DSM-5: DISC: 61.4% vs. 52.5%, $\chi^2(1)=46.00$, $p<.001$, CAARS: 60.1% vs. 49.9%, $\chi^2(1)=38.00$, $p<.001$, DSM-5 vs DSM-IV: DISC: 52.5% vs. 44.0%, $\chi^2(1)=53.00$, $p<.001$, CAARS: 49.9% vs. 38.2%, $\chi^2(1)=31.00$, $p<.001$).

Table 3

Receiver Operating Characteristic Curves

	AUC	*SE*	*p*	95% CI	*z-score for paired comparison*				
DISC					vs. 1	vs. 2	vs. 3	vs. 4	vs. 5
(1) DSM-IV	.736	.022	<.001	.694 - .779	---				
(2) DSM-5	.744	.021	<.001	.703 - .784	.64	---			
(3) NB	.747	.020	<.001	.708 - .786	.71	.33	---		
CAARS									
(4) DSM-IV	.725	.022	<.001	.682 - .769	.59	.92	1.09	---	
(5) DSM-5	.757	.021	<.001	.717 - .797	1.10	.78	.58	2.60	---
(6) NB	.796	.018	<.001	.762 - .832	3.16*	2.92	2.85	4.32*	3.08*

Note. AUC=Area Under the Curve; SE= standard error; p= statistical significance; CI= Confidence Interval.

*Indicates statistically significant paired comparison $p<.003$.

Table 4

Absolute Classification Rates within Childhood ADHD and LNCG Groups

	DISC Interview			CAARS Rating Scale		
	DSM-IV (6)	**DSM-5 (5)**	**NB (4)**	**DSM-IV (6)**	**DSM-5 (5)**	**NB (4)**
Childhood ADHD						
Persistence Rate (%)	44.0	52.5	61.4	38.2	49.9	**60.1**
Total Correct Classification (%)	70.3	68.9	67.2	70.7	71.8	74.4
True Positives: Persistent with impairment	30.5	33.9	37.5	28.3	34.7	41.1
True Negatives: Desistent without impairment	39.8	35.0	29.7	42.4	37.1	33.3
Total Incorrect Classification (%)	29.7	31.1	32.8	29.3	28.2	25.6
False Positives: Persistent without impairment	13.1	18.0	23.3	9.9	15.2	19.0
False Negatives: Desistent with impairment	16.6	13.1	9.5	19.4	13.0	6.6
No Childhood ADHD (LNCG)						
Symptomatic Rate (%)	7.9	10.9	14.2	7.1	11.2	14.5
De Novo ADHD or other disorder: Symptomatic with impairment	4.2	4.6	5.0	4.6	6.2	7.1
False Positives: Symptomatic without impairment	3.7	6.3	9.2	2.5	5.0	7.4
Full Sample (ADHD + LNCG)						
Sensitivity	64.9	72.4	79.9	59.3	72.7	86.1
Specificity	82.3	76.4	69.5	85.9	78.4	73.3

Note. Value in bold represents the persistence rate with an optimized balance between sensitivity and specificity according to ROC analyses.