



Published in final edited form as:

J Child Psychol Psychiatry. 2017 June ; 58(6): 663–678. doi:10.1111/jcpp.12684.

Young adult outcomes in the follow-up of the Multimodal Treatment Study of Attention-Deficit/Hyperactivity Disorder: symptom persistence, source discrepancy and height Suppression

James M. Swanson, Ph.D.¹, L. Eugene Arnold, M.D.², Brooke S. G. Molina, Ph.D.³, Margaret H. Sibley, Ph.D.⁴, Lily T. Hechtman, M.D.⁵, Stephen P. Hinshaw, Ph.D.⁶, Howard Abikoff, Ph.D.⁷, Annamarie Stehli, M.P.H.⁸, Elizabeth B. Owens, Ph.D.⁹, John T. Mitchell, Ph.D.¹⁰, Quyen Nichols, B.A.¹¹, Andrea Howard, Ph.D.¹², Laurence L. Greenhill, M.D.¹³, Betsy Hoza, Ph.D.¹⁴, Jeffrey Newcorn, M.D.¹⁵, Peter Jensen, M.D.¹⁶, Benedetto Vitiello, M.D.¹⁷, Timothy Wigal, Ph.D.¹⁸, Jeffery Epstein, Ph.D.¹⁹, Leanne Tamm, Ph.D.²⁰, Kimberly D. Lakes, Ph.D.²¹, James Waxmonsky, M.D.²², Marc Lerner, M.D.²³, Joy Etcovitch, M.A.²⁴, Desiree Murray, Ph.D.²⁵, Maximilian Muenke, M.D.²⁶, Maria Acosta, M.D.²⁷, Mauricio Arcos-Burgos, Ph.D.²⁸, William E. Pelham, Ph.D.²⁹, and Helena Kraemer³⁰ >MTA Cooperative Group

¹Department of Pediatrics, School of Medicine, University of California, Irvine, CA

²Department of Psychiatry, Ohio State University, Nisonger Center, Columbus, OH

³Departments of Psychiatry and Psychology, University of Pittsburgh School of Medicine, PA

⁴Department of Psychiatry, Florida International University, Miami, FL

⁵Division of Child Psychiatry, McGill University, Montreal Children's Hospital, Montreal, Quebec, Canada,

⁶Department of Psychology, University of California, Berkeley, CA

⁷Child Study Center at New York University Langone Medical Center, NY

⁸Department of Pediatrics, School of Medicine, University of California, Irvine, CA

⁹Institute of Human Development, University of California, Berkeley, CA

¹⁰Department of Psychiatry & Behavioral Sciences, Duke University Medical Center, Durham, NC

¹¹Department of Psychological Science, University of Vermont, Burlington, VT

¹²Department of Psychology, Carleton University, Ottawa, ON, Canada

¹³Department of Psychiatry, Columbia University, New York, NY

¹⁴Department of Psychological Science, University of Vermont, Burlington, VT

Correspondence James M. Swanson, Department of Pediatrics, University of California, Irvine, 3 Harvey Court Irvine, CA 92617, USA; jmswanso@uci.edu.

Ethical Guidelines: This research meets the ethical guidelines of the US National Institutes of Health and of the multiple universities that participated in the study, and it adheres to the legal requirements for research of the United States and Canada.

Conflict of interest statement: See Acknowledgements for disclosures.

¹⁵Department of Psychiatry, Mount Sinai Medical School, New York, NY

¹⁶Department of Psychiatry, University of Arkansas for Medical Sciences, Little Rock, AR

¹⁷National Institute of Mental Health, Bethesda, MD

¹⁸Avida, Inc., Newport Beach, CA

¹⁹Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

²⁰Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

²¹Department of Pediatrics, School of Medicine, University of California, Irvine, CA

²²Pennsylvania State College of Medicine, Hershey, PA

²³Orange County Department of Education, Costa Mesa, CA

²⁴Montreal Children's Hospital, Montreal, Quebec, Canada,

²⁵Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill, NC

²⁶National Human Genome Research Institute, Bethesda, MD

²⁷National Human Genome Research Institute, Bethesda, MD

²⁸National Human Genome Research Institute, Bethesda, MD

²⁹Department of Psychology, Florida International University, Miami, FL

³⁰Department of Psychiatry, Stanford University, CA

Abstract

BACKGROUND: The Multimodal Treatment Study (MTA) began as a 14-month randomized clinical trial of behavioral and pharmacological treatments of 579 children (7–10 years of age) diagnosed with ADHD-combined type. It transitioned into an observational long-term follow-up of 515 cases consented for continuation and 289 classmates (258 without ADHD) added as a local normative comparison group (LNCG), with assessments 2 to 16 years after baseline.

METHODS: Primary (symptom severity) and secondary (adult height) outcomes in adulthood were specified. Treatment was monitored to age 18, and naturalistic subgroups were formed based on three patterns of long-term use of stimulant medication (Consistent, Inconsistent, and Negligible). In the follow-up, hypothesis-generating analyses were performed on outcomes in early adulthood (at 25 years of age). Planned comparisons were used to estimate ADHD-LNCG differences reflecting persistence of symptoms and naturalistic subgroup differences reflecting benefit (symptom reduction) and cost (height suppression) associated with extended use of medication.

RESULTS: For ratings of symptom severity, the ADHD-LNCG comparison was statistically significant for the parent/self-report average (0.51 ± 0.04 , $p < 0.0001$, $d = 1.11$), documenting symptom persistence, and for the parent/self-report difference (0.21 ± 0.04 , $p < 0.0001$, $d = 0.60$), documenting source discrepancy, but the comparisons of naturalistic subgroups reflecting medication effects were not significant. For adult height, the ADHD group was 1.29 ± 0.55 cm shorter than the LNCG ($p < 0.01$, $d = 0.21$), and the comparisons of the naturalistic subgroups were

significant: the treated group with the Consistent or Inconsistent pattern was 2.55 ± 0.73 cm shorter than the subgroup with the Negligible pattern ($p < 0.0005$, $d = 0.42$), and within the treated group, the subgroup with the Consistent pattern was 2.35 ± 1.13 cm shorter than the subgroup with the Inconsistent pattern ($p < 0.04$, $d = 0.38$).

CONCLUSIONS: In the MTA follow-up into early adulthood, ADHD group showed symptom persistence compared to local norms from the LNCG. Within naturalistic subgroups of ADHD cases, extended use of medication was associated with suppression of adult height but not with reduction of symptom-severity.

Keywords

Attention-deficit/hyperactivity disorder; follow-up studies; growth; longitudinal studies; treatment trials; medication effects

INTRODUCTION

The Multimodal Treatment Study of Children with Attention-Deficit/Hyperactivity Disorder (ADHD), known as the MTA, was initiated in 1994 (just before release of DSM-IV) and continued until 2013 (just before release of DSM-5), spanning the entire DSM-IV era. The background, design, methods, and findings from childhood to adolescence have been presented in over 100 publications (Appendix S1). This report describes outcomes of the MTA follow-up into adulthood (at an average age of 25 years).

The MTA began as a 14-month randomized clinical trial (RCT) to test hypotheses about four treatment strategies: medication management (Med), behavior modification (Beh), their combination (Comb), or treatment as-usual in a community comparison (CC) group. After diagnosis with ADHD-Combined Type, 579 children (7.0–9.9 years old) were randomly assigned to treatment conditions and showed high compliance and adherence to the MTA protocol. Intent-to-treat (ITT) analyses revealed a significant relative benefit (greater decline in ratings of symptom severity) as well as relative cost (reduced height gain) in the groups with (Med and Comb) compared to without (Beh and CC) stimulant medication provided by-protocol for 14 months.

After the RCT, the MTA transitioned into an observational long-term follow-up (LTF), and analyses were performed to generate hypotheses for future studies (see Appendix S2). At 2, 3, and 8 years after baseline, ITT analyses of the randomly-assigned groups were continued and indicated the initial symptom-related benefit and growth-related cost dissipated after the 14-month treatment-by-protocol phase. During the LTF phase, medication use was monitored prospectively. This documented different patterns of starting and stopping medication resulting in convergence in the rates of medication use across the assigned treatment groups, generating the hypothesis that continued relative benefit and cost may be associated with continued use of medication. However, comparison of naturalistic subgroups (see Swanson, Elliott, Greenhill et al, 2007) with different patterns (i.e., no use, new use, inconsistent use, and consistent use of medication in childhood) showed that symptom severity was not significantly different for these naturalistic subgroups, generating the hypothesis that symptom-related benefit may dissipate even when medication is continued

and that childhood height was reduced in the naturalistic subgroup with consistent treatment, generating the hypothesis that growth-related costs may not dissipate when medication is continued.

Here, adult outcomes will be evaluated based on the final assessment of the MTA. Methods developed by Swanson, Kraemer, Hinshaw et al (2001) will be used: to focus on a primary outcome, adult symptom-severity averaged across ADHD domains and information sources, instead of many outcomes (to limit the number of statistical tests); to evaluate outcome at the end-point in adulthood instead of the trajectory into adulthood (to simplify interpretation of findings); and to apply planned comparisons instead of pairwise comparisons (to address efficiently the main questions): ‘Do symptoms persist into adulthood?’, ‘Is there a significant long-term symptom-related benefit (e.g., reduced symptom severity) of extended use of medication compared to no use?’, and ‘Does continuous use of medication result in greater benefit than interrupted use?’. Following the Consolidated Standard of Reporting Trials (CONSORT) revised guidelines (Ioannidis, Evans, Gotzsche et al, 2004), parallel analyses were performed to address the same questions about adult physical size and a possible growth-related cost (i.e., suppressed adult height).

METHODS:

MTA Sample

The CONSORT diagram of the MTA is presented in Table 1. Entry into the LTF phase occurred 2 years after baseline, when 540 of the ADHD participants from the RCT phase were assessed and 515 re-consented for the follow-up. At the same time, 289 LNCG participants were recruited from the same schools as the ADHD cases and assessed with the same protocol. Following the precedent set by Molina et al (2009), 31 LNCG participants with a diagnosis of ADHD at recruitment were excluded, so only 258 were included in the analyses reported here. Participants were assessed eight times from 2 to 16 years after baseline. As shown in Table 2, retention was high in childhood, adolescence, and adulthood. Previous analyses of adult outcomes (Howard, Strickland, Murray et al, 2016) indicated MTA participants with and without complete data were not significantly different on most baseline demographic variables and ‘missing at random’ criteria were met.

Use of Medication

Starting at the 14-month assessment and continuing through the 10-year assessment (or until participants were 18 years of age), the Services for Children and Adolescents Parent Interview (SCAPI: Jensen, Hoagwood, Roper et al, 2004) was administered. Part of this interview documented the number of days treated with stimulant medication and the daily doses administered since the previous assessment. In most cases the stimulant medication was immediate-release (IR) methylphenidate (Ritalin® or Metadate®, 50:50 racemic mixtures of the d- and l-isomers). Doses of other approved and available stimulants were transformed to d,l-methylphenidate equivalent (ME) doses. For the d-methylphenidate isomer (Focalin®), dose was multiplied by 2; for amphetamine (Dexedrine®, a pure d-isomer, or Adderall®, a 75:25 racemic mixture), the dose was multiplied by 2; for pemoline (Cylert®), the dose was multiplied by 6. Sustained-release (SR) formulations of d,l-

methylphenidate (Ritalin SR®, Concerta®, Ritalin LA®, and Metadate CD®) were included at the prescribed doses, and prescribed doses of SR formulations of d-methylphenidate (Focalin XR®) and amphetamine (Dexedrine Spansules® or Adderall XR®) were multiplied by 2. If an IR and an SR formulation were administered on the same day, the sum of daily ME doses was used. Non-stimulant medications were excluded. The total ME dose for each interval was estimated as the product of the days treated times the daily dose. The totals were summed for the 6 intervals from Baseline to the 10-year assessment to estimate the cumulative ME dose from childhood through adolescence.

Also, a minimum ME regimen (at least 10 mg/day for at least 50% of days since the previous assessment) was used to classify treatment during each assessment intervals as ‘minimal’ or ‘< minimal’ (see Jensen, Arnold, Swanson, et al, 2007). Rather than representing adequate or optimal treatment for all cases, this was intended to avoid exclusion of regimens based on low doses that might be effective for a few cases and regimens with medication administered only on school days (i.e., with planned drug holidays). Extending a previous method (Swanson et al, 2007), sequences of intervals above or below the cutoff were used to define three long-term patterns of prospective treatment with medication from childhood through adolescence: Consistent (minimal in all intervals), Inconsistent (minimal in some but not all intervals), and Negligible (< minimal in all intervals). Prior use of medication before entering the MTA was documented in the baseline assessment, but the SCAPI was not administered then to provide the same detailed information on prior as for prospective treatment. Prior Status was specified as a binary variable based on the presence or absence of a history of prior treatment with stimulant medication.

Outcomes

In adulthood (i.e., at assessments 12-, 14-, and 16-years after baseline), the primary outcome was symptom-severity measured by the Conners Adult ADHD Rating Scale (CAARS: Conners, Erhardt, and Sparrow, 1999), which was completed by two informants (or sources) to provide parent-ratings and self-ratings. The CAARS includes items for the 18 DSM-IV ADHD symptoms, each rated on a 4-point severity scale (0=Not at All to 3=Very Much). For both sources (and within the ADHD group and the LNCG), average rating-per-item for the 9 Inattention and 9 Hyperactive/Impulsive items were highly correlated with each other ($r \sim 0.70$ to 0.80), as well as with the overall average for the 18 ADHD items ($r \sim 0.90$ to 0.95). After decreasing from baseline, the symptom severity ratings were stable across the three assessment points in adulthood, indicating an asymptote in early adulthood. Correlations of adjacent points in adulthood were high ($r \sim 0.61$ to 0.76), indicating adequate test-retest reliability in adulthood. At each assessment, correlations of ratings from the two sources were significant but low for the ADHD group ($r \sim 0.22$ to 0.29) and LNCG ($r \sim 0.26$ to 0.47).

A composite measure (CAARS-avg) was created to reflect symptom persistence in adulthood by averaging parent- and self-ratings to capture source convergence (Kraemer, Measelle, Ablow, et al, 2003) and to increase precision of measurement (Swanson, Kraemer, Hinshaw et al, 2001). Typically, self-ratings by children with ADHD are lower than parent-ratings (see Owens, Goldfine, Evangelista et al, 2007), but it is unclear whether the

difference continues (see Sibley, Pelham, Molina et al, 2012) or dissipates (see Barkley, Murphy, and Fisher, 2008) in adulthood. To evaluate this, a second composite measure (CAARS-diff) was created to reflect source divergence by taking the difference of parent- and self-ratings of symptom severity.

At clinic visits, MTA staff measured height (in cm) with stadiometers, as well as weight (in kg) with digital scales. For comparison to critical studies in the literature (i.e., Spencer et al, 1998 and Swanson et al, 2007), norms for the United States (Kuczmarski, Ogden, Grummer-Steawn et al, 2009) were used to transform the absolute (raw score) to a relative (z-score) measure adjusted for age and sex.

Analyses

Using the Statistical Analysis System (SAS) General Linear Model (GLM) program, simple one-way analyses of the primary outcome (CAARS-avg) and secondary outcome (Height-cm) were performed, using a between-group factor with 4 levels: Consistent, Inconsistent, Negligible, and LNCG. Sex and Age were included to evaluate possible gender and developmental effects. For Height-cm, Mid-Parent height (average z-score) was included to account for expected genetic effects. Comparisons of the groups were made using the SAS-GLM estimates of least square means (LSMs) adjusted for the other factors in the model.

To decompose a significant Group effect, 3 planned comparisons (see Appendix S3) were used instead of 6 (all possible) paired comparisons. This limited the number of comparisons and addressed critical questions specified in the literature. As recommended by Barkley, Murphy, and Fischer (2008), a norm-based comparison (i.e., ADHD-vs-LNCG) was used to provide an estimate of symptom persistence in adulthood. As recommended by Spencer et al (1998, p 503), a comparison of ‘... treated children with ADHD with untreated children, and not with unaffected control subjects’ was used to differentiate treatment and disorder effects, with the treated group represented by the pooled naturalistic subgroups with Consistent or Inconsistent treatment and the untreated group represented by the naturalistic subgroup with Negligible treatment. As recommended by Faraone, Biederman, Morley, et al (2008, p 994), to evaluate ‘... continuous treatment from childhood to adulthood’ the group with treatment through age 18 (the naturalistic subgroup with Consistent treatment) was compared to a group with interrupted treatment (the naturalistic subgroup with Inconsistent treatment). Each comparison represents a difference between two groups, which was standardized by dividing by the pooled standard deviation (RMSE, the square root of the mean square error in the model defined for the SAS-GLM analysis of the outcome measure) to estimate effect size (Cohen’s *d*).

Within the ADHD group, supplementary regression analyses were performed to address effects of cumulative ME dose on adult outcomes. This approach was suggested by Charach, Figueroa, Chen, et al (2006, p 419), who proposed that height suppression would not ‘... become statistically significant until the dose of MPH is 2.5 mg/kg/day for 4 years’. In the MTA, cumulative ME dose was specified as absolute instead of relative dose, and escalation over time with increasing size was not observed, as implied by the longitudinal mg/kg dose. For a typical case in the MTA (e.g., an 8.4-year old with an average weight of 30.5 kg at baseline), the high 2.5 mg/kg dose would be 76.5 mg/day. For 4 years, this would

result in a cumulative ME dose of 111,325 mg, but over 10 years the cumulative ME dose would be much higher (i.e., 278,313 mg). To accommodate such a wide range (e.g., 0–300,000 mg), cumulative ME dose was transformed from mg to grams by dividing by 10,000, and the natural logarithm (after adding 1 to avoid 0) was used as the predictor in the regression analysis. As in the analyses of Group (see above), Age and Sex were included as factors in the regression analyses of CAARS-avg and Height-cm, and also Mid-Parent Height in the analysis of Height-cm.

RESULTS:

By the 12-year assessment point, all participants in the MTA were beyond age 18 (and thus in adulthood), so retention was based on the percentage having at least one observation from this point forward. This was higher than the percentage observed at any one assessment point, since some participants returned to the follow-up after missing an earlier assessment. As shown in Table 2a, observations in adulthood were available for 92.4% (476/515) of the ADHD and 93.4% (241/258) of the non-ADHD LNCG participants consented for the LTF, representing 82.2% (476/579 for the ADHD group) and 83.4% (241/289 for the LNCG) of those recruited for the study. The final assessments in adulthood were used for all participants, which were primarily from the 16-year assessment (88.0% for the ADHD group and 92.1% for the LNCG), with progressively fewer from the 14-year (8.6% and 5.8%) and 12-year (3.4% and 2.8%) assessment points.

Prospectively gathered information from the SCAPI was used to estimate patterns of extended use of medication for the 476 ADHD cases. As shown in Table 2b, in adolescence there was a 4-fold decrease in the overall percentage of the ADHD cases with >minimal medication use at successive assessment points (i.e., from 57.2% at the 3-year assessment to 14.3% at the 10-year assessment). For the individual patterns based on sequences of use from childhood (baseline) through adolescence, 23.5% of the ADHD cases had the Negligible pattern (<minimal in all intervals), 69.1% had the Inconsistent pattern (minimal in some but not all intervals), and 7.4% had the Consistent pattern (minimal in all intervals). For these naturalistic subgroups, the prospective average cumulative ME doses were 2,153 mg, 60,567 mg, and 117,102 mg.

As shown in Table 2b, at the end of the 14-month treatment-by-protocol RCT, the rates of medication use were dramatically different for the 4 assigned treatment groups (e.g., 92.5% for MedMgt to 13.9% for Beh), but then converged in adolescence, resulting in 10-year averages for cumulative ME dose that were similar but not identical (e.g., 58,986 mg for MedMgt and 41,140 mg for Beh). Before performing the analyses of naturalistic subgroups, ITT analyses of the initially assigned treatment groups were performed to evaluate possible ' sleeper effects '. However, there were no re-emergent significant assigned treatment group differences: the 4 assigned treatment groups did not differ significantly on average symptom-severity ratings (Med=0.93, Comb= 0.91, Beh=0.94, and CC=0.90) or adult z-height (Med=0.21, Comb=0.10, Beh=0.26, and CC=0.18). The critical MTA Medication Algorithm (Med+Comb-vs-Beh+CC) contrast remained non-significant, despite a non-trivial enduring difference (~17,000 mg) in cumulative ME dose (see Table 2b).

As shown in Table 2c, the naturalistic subgroups use did not differ significantly on most demographic variables assessed in the children at baseline, including age at entry into the MTA, percentage male, symptom severity (except for slightly lower parent ratings for the Negligible subgroup), or birthweight. However, the naturalistic subgroups differed ($p < 0.05$) on some of the demographic variables: the Consistent naturalistic subgroup had significantly higher family income, greater household advantages, and higher percentage of white participants. The LNCG differed from the overall ADHD group and was similar to the Consistent naturalistic subgroup with higher income and greater household advantages. Since these differences were not predicted or considered during the design phase of the MTA to set the sample size and statistical power (see Kraemer, 2015), these variables were not included initially as covariates, but subsequently they were included separately as possible moderators of findings in the initial analyses of the outcome variables.

Analysis of Primary and Secondary Outcomes

The analyses of symptom severity (CAARS-avg and CAARS-diff) and adult height (Height-cm and Height-z) outcomes are shown in Table 3. The 3 planned comparisons (ADHD-vs-LNCG; Consistent or Inconsistent-vs-Negligible; Consistent-vs-Inconsistent) are shown in Table 3 (for more detail, see Appendix S3). The observed means and standard deviations for the 4 levels of the between-group factor (as well as for the overall ADHD group based on the pooled naturalistic subgroups) are shown in Figure 1.

In the analysis of CAARS-avg, the Group factor was significant ($p < 0.0001$), but effects of Age and Sex were not significant. As shown in Table 3a, the comparison of diagnostic groups (ADHD-vs-LNCG) was significant ($p < 0.0001$), due to a higher value for the parent/self-report average rating in the ADHD group than the LNCG (by 0.51 ± 0.04), with a large effect size ($d=1.11$). Within the ADHD group, the medication-related comparisons of the naturalistic subgroups were not significant. In analyses of CAARS-diff, the Group factor was significant, and the ADHD-LNCG comparison revealed significant ($p < 0.0001$) parent/self-report discrepancy (by 0.21 ± 0.04), with a medium effect size ($d=0.61$), due to higher parent- than self-ratings in the ADHD group but not in the LNCG (see Figure 1). Within the overall ADHD group, the two medication-related comparisons of the naturalistic subgroups were not significant.

In analysis of Height-cm, the effect of the Group factor was significant ($p < 0.0001$). Age was not significant, as expected in adulthood (age 25) when final height has been achieved, but Sex and Parent Height were significant. As shown in Table 3a, the ADHD group was significantly ($P < 0.01$) shorter than the LNCG (by 1.29 ± 0.55 cm or about 0.51 of an inch) with a small effect size ($d=0.21$). Within the overall ADHD group, the two medication-related comparisons of the naturalistic subgroups were significant with medium effect sizes for Height-cm: (a) the average was significantly ($p < 0.0005$, $d=0.42$) lower for the treated ADHD group (the pooled Consistent and Inconsistent subgroups) compared to the Negligible subgroup (by 2.55 ± 0.73 cm or about 1 inch) and (b) the average was significantly ($p < 0.0378$, $d=0.38$) lower for the Consistent compared to Inconsistent subgroup (by 2.36 ± 1.13 cm or about 0.93 inch). To supplement the 3 planned comparisons, the 6 paired-

comparisons are provided in Appendix S3, which show a 4.7 cm difference between the extreme subgroups (Consistent-vs-Negligible).

Analyses of Height-z showed similar significance levels and effect sizes for relative height as for absolute height adjusted for Sex and Age (see Table 3b). The Negligible subgroup had an average z-score of 0.467, which indicated the 'untreated' clinical control group was taller than the population average (z-score = 0). In the previous report (Swanson et al, 2007), a speculative hypothesis was proposed that the larger-than-average size of the untreated ADHD cases may be associated with a theoretical biological factor (a dopamine deficit), which is expected to increase growth. However, also the LNCG was taller than the population average (with an average z-score of 0.271), confirming a previous finding in childhood (Swanson et al, 2007) and suggesting secular factors (e.g., the recruitment methods, the site locations, etc.) contributed to increased height compared to the population average. The z-score difference was only 0.196 (a small effect size).

All analyses were repeated with covariates added for Household Advantage, Income, and Race/Ethnicity. For symptom severity, Income was significant ($p < 0.03$), but Household Advantage and Race/Ethnicity were not significant. For adult height, Income, Household Advantage, and Race/Ethnicity were not significant. For symptom severity and adult height, analyses were repeated with covariates added (whether they were significant or not), and significance levels and effect sizes for outcome variables were essentially unchanged.

Weight-kg was not a primary outcome, but exploratory analyses revealed the ADHD group was significantly ($p < 0.0001$) heavier (by 4.6 ± 1.78 kg) than the LNCG (see Figure 1b). Unlike for Height-cm, the medication-related comparisons of naturalistic subgroups were not significant ($p > 0.05$). Analyses of BMI-m/kg² paralleled those for Weight-kg, with significantly higher body mass for the ADHD group than the LNCG, but neither comparison of the naturalistic subgroups was significant. Analyses of relative measures showed similar significance levels and effect sizes as analyses of absolute measures adjusted for Sex and Age.

The naturalistic subgroups were divided further based on history of prior use of medication before entry into the MTA (Prior Status). The naturalistic subgroups differed in the percentage of participants with prior treatment with medication (28% for Negligible, 41% for Inconsistent, and 67% for Consistent). Prior Status was added as a factor in analyses with 3 levels of Group (without the LNCG). Neither Prior Status nor its interaction with Group was significant in any of the analyses of adult outcome (see Appendix S4).

Regression analyses:

Regression analyses of cumulative ME dose in the ADHD group are shown in Table 3c for symptom severity and adult height. Scatter plots of the observed outcomes are presented for CAARS-avg (Figure 1a), Height-cm (Figure 1b), and Height-z (Figure 1c). For CAARS-avg, the slope was positive but not significant ($+0.04 \pm 0.03$ cm/ln-unit, $p = 0.14$), suggesting cumulative ME dose was not related to symptom severity in adulthood (see Figure 1a). Neither Age nor Sex was significant (see Table 3)

For Height-cm, the slope was negative and significant (-1.02 ± 0.34 cm/ln-unit, $p < 0.003$), suggesting an increase in cumulative dose was associated with lower adult height (see Figure 1b). For a cumulative ME dose of 100,000 mg, the regression equation predicts height suppression in adulthood of 2.4 ± 0.33 cm or almost an inch. Age was not significant, but Sex was significant (with female participants 14.2 cm shorter on-average than male participants -- see Table 3c). Regression analysis of the relative measure adjusted for age and sex (Height-z) showed a similar effect of cumulative ME dose (i.e., a significant slope) as in the analysis of the absolute measure (Height-cm) with Age and Sex included as factors.

Analyses with covariates added for Income, Household Advantage, and Race/Ethnicity were performed, but none of these were significant or changed the findings. Analyses were repeated with Prior Status (and its interaction with cumulative ME dose) added as predictors, but neither was significant (see Appendix S4).

In additional analyses for weight and body mass, the slopes were not significant for raw scores (Weight-kg and BMI-m/kg²) or for standard scores (Weight-z and BMI-z), suggesting cumulative ME doses was not related to these outcome measures in adulthood. These findings are consistent with the non-significant medication-related comparisons of the ADHD subgroups for these outcomes.

DISCUSSION:

The findings from LTF phase of the MTA address the 3 critical questions specified in the introduction about long-term symptom-related and growth-related outcomes of children with ADHD and about possible long-term benefits (adult symptom reduction) and costs (adult height suppression) associated with the treatment of ADHD with stimulant medication. First, comparisons of ADHD group to contemporary (as well as prospective) norms from unaffected classmates (the LNCG) suggest significant persistence of the disorder due to higher levels of symptom severity in adulthood but only a small elevation of relative height. Second, the comparisons of ADHD-treated and ADHD-untreated groups suggest that in the long-term, symptom-related benefit of treatment with medication may dissipate and not remain significant but growth-related cost may remain statistically significant in adulthood. Third, the comparisons of groups with continuous and interrupted use of medication from childhood through adolescence suggest greater treatment may not result in greater symptom-related benefit but may result in greater growth-related cost. These findings add to the literature and to some controversies about long-term outcomes and long-term effects of treatment of ADHD.

The significant symptom persistence effect described here was based on a dimensional outcome (a scaled outcome based on effect size) instead of a traditional categorical outcome (a binary outcome based on a cutoff). Kraemer and Kupfer (2006) discussed issues about when each of these alternative approaches should be used, and described the decision as a ‘... struggle between statistical considerations favoring dimensional outcomes and clinical consideration favoring categorical outcomes’. One reasons for using the dimensional approach was to maximize statistical power for detecting possible long-term benefits and costs of medication. However, a categorical approach (based on symptom count rather than

symptom severity) was presented in companion article to enhance clinical relevance related to symptom persistence. Sibley, Swanson, Arnold et al (2016) described persistence rates for combinations of three factors: cutoff criteria (6 symptoms for DSM-IV, 5 for DSM-5, and 4 for statistical norms), source (parent-report or self-report), and method (interview or rating scale). A broad range (from 1.9% to 61.4%) highlighted the controversy about which combination to use. Hechtman, Swanson, Sibley et al (2016) evaluated the validity of one combination that is currently clinically relevant (DSM-5 criteria, rating scale method, and combined sources, with a persistence rate of 49.9%) and showed significant differences between Persistent and Desistent subgroups on functional adult outcomes assessed in the MTA (educational, occupational, sexual, emotional, and substance use). These large dimensional and categorical effects of symptom persistence are consistent with the notion that ADHD is a chronic condition. This adds to the controversy about the magnitude of persistence of ADHD in adulthood, which depends on how it is measured (see Faraone, Biederman, and Mick, 2008 and Barkley et al, 2002).

The significant source discrepancy effect described here adds to the literature on another controversial issue. For example, Barkley et al (2008) followed (to age 27) a group of 158 cases (diagnosed with Hyperactivity in 1979–1980 before the criteria were established for diagnosing ADHD) and 81 non-affected controls, and concluded ‘... self-reported information [converges with] parent/other reported information’ (p 66), but Sibley, Pelham, Molina et al (2012) followed (to age 20) a group of 200 cases (diagnosed with ADHD from 1986 to 1997) and 121 non-affected controls, and concluded ‘... young adults with ADHD tended to underreport symptoms’ (p 1). If source convergence occurs, when is this manifested? Using a larger sample, a dimensional outcome with statistical advantages, and follow-up to age 25, source discrepancy was still significant in adulthood. For categorical outcomes, Sibley et al (2016) also found symptom persistence to be 40% to 137% higher for parent-report than self-report depending on the combination of factors used to classify cases. These findings from the MTA are consistent with the hypothesis that self-report of symptoms remains far below parent-report in early adulthood (at least to age 25).

The significant effect of height suppression effect described here adds to the controversy about a possible growth-related cost of long-term use of medication. The findings from the MTA are not consistent with many LTF studies in the literature (see below), but they are consistent with a recent study that was based on similar methods. Poulton, Melzer, Taft et al (2013) described a small clinical sample (n=22 ADHD cases) selected for consistent use of medication from childhood to adolescence (i.e., average ME dose starting at 46 mg/day with continuous treatment for an average of 7 years, suggesting a cumulative ME dose of 117,530 mg). For this naturalistic subgroup, average adolescent height (assessed between 14 and 16 years of age) was significantly ($p < 0.01$) lower (by 2.7 cm or more than an inch) than for a non-ADHD group (n=132). For an equivalent comparison in the MTA, the height of Consistent naturalistic subgroup (n=33) -- with about the same cumulative ME dose (117,102 mg) -- was compared to a non-ADHD group, the LNCG (n=221). Average height in adulthood was significantly ($p < 0,001$, $d=0.65$) lower (by 4.0 cm or about 1.5 inch). This replicates and extends the findings of Poulton et al (2013) with a larger sample and longer follow-up.

The discrepancy with other LTF studies in the literature may be related to changes in the clinical use of medication over the past 50 years and differences in the average cumulative ME dose for treatment-as-usual. For example, most studies of ADHD and growth report average ME daily dose (mg/day) and average duration (years) for treated cases, and for some often-cited studies the product provides an estimate of cumulative ME dose. For a clinical sample recruited in the 1960s (Kramer, Loney, Porto et al, 2000), the product (31.2 mg/day x 3.02 years) was 34,350 mg. For a clinical sample recruited in the 1970s (Mannuzza and Klein, 1988), the product (44.9 mg/day x 2.24 years) was 36,710 mg, indicating a slight increase in cumulative ME dose. For a population-based sample of ADHD cases identified from an epidemiological cohort in the 1980s (Harstad, Weaver, Katusic et al, 2014), the product (26.2 mg x 4.42 years) was 42,268 mg, indicating a substantial increase in cumulative ME dose over the decade from 1970 to 1980. However, none of these studies reported a significant difference in adult height between the treatment-as-usual ADHD group and a non-ADHD control group. To match the methods of these historical studies, the pooled group of treated cases with the Inconsistent or Consistent pattern (i.e., with treatment-as-usual in the LTF during the 1990s and 2000s) was considered. The cumulative ME dose was 66,003 mg, representing a 92% increase from the 1960s, an 80% increase from the 1970s, and a 56% increase from the 1980s. Compared to the LNCG, adult height was significantly lower (by 2.0 cm, $p < 0.01$), suggesting height suppression with a medium effect size ($d=0.32$). The cumulative exposure hypothesis (see Charach et al, 2006) and this post hoc finding from the MTA generate a hypothesis about the discrepancy with previous studies: a secular trend of increasing cumulative ME dose occurred from the 1960s to the 1990s, and in the MTA (initiated in the 1990s) the average exposure was sufficient to produce significant height suppression in cases treated with medication.

The MTA findings generate hypotheses about other prominent studies that do not report cumulative ME dose. For example, based on a 10-year follow-up in a study initiated in the 1980s, Biederman, Spencer, Monuteaux et al (2010) reported that duration of treatment (dose was not available) was not significantly related to adult height (although linear growth curve analysis suggested the males were 2.4 cm shorter than non-affected controls). However, in a previous report on substance use outcomes in the male subjects, Biederman, Monuteaux, Spencer et al (2008) reported that 27% of the ADHD cases were untreated, but this naturalistic subgroup was not compared to the treated subgroup as recommended by Spencer et al (1998). Also, 50% were treated for less than 2 years, suggesting 23% of the sample had extended use of medication for 2 years or more. The MTA findings generate the hypothesis that a post hoc analysis the naturalistic subgroup of males with extended use of medication would show height suppression compared to the clinical control group of untreated cases. Also, consider the example provided by Peyre, Hoertel, Cortese et al (2013). In a large representative sample of over 35,000 adults identified in 2004–2005, adult ADHD cases with ($n=216$) and without ($n=591$) use of stimulant medication were identified. Peyre et al (2013) reported that these subgroups did not differ in adult height. However, they also reported the average age of treatment initiation was 15.9 years, which likely resulted in a high percentage of cases with treatment only after most growth was attained. The MTA findings generate the hypothesis that a subgroup of cases with treatment initiated in childhood and continued through adolescence would show height suppression. Thus,

findings from the MTA suggest alternative analyses that may be warranted before accepting the null hypothesis that extended use of medication does not suppress adult height.

Implications for Guidelines:

The most recently published guidelines (American Academy of Pediatrics, 2011) recommend expanding the diagnosis and treatment beyond school-aged children and using stimulant medication as first-line treatment for adolescents as well as school-aged children. Since this would increase the average duration of treatment and cumulative ME dose of medication in some individuals, the MTA findings suggest growth-related costs may increase. This possibility may need additional emphasis, similar to the contemporary European guidelines for managing adverse side effects of medication for ADHD (Graham, Banaschewski, Buitelaar et al, 2011), which presciently highlight this relevant issue: 'It is possible that the effects of stimulants on growth are dose-dependent. Significant effects on weight and height may require average doses of methylphenidate exceeding 1.5 mg/kg per day which are given continuously' (p 24). The regression analyses of the MTA support this speculation. Revised guidelines could address strategies for reducing cumulative ME dose. In the RCT phase of the MTA, combining behavioral intervention with the use of medication reduced average daily dose by 20% without a reduction in efficacy (see Vitiello et al, 2001), and recently, Pelham, Fabiano, Waxmonsky et al (2016) showed that initiating treatment with low-intensity psychosocial treatment before adding medication reduced the daily dose on school days by 25% without a reduction in efficacy. The findings of the MTA suggest that these strategies – which appear to achieve full short-term symptom-related benefits in childhood -- may reduce long-term growth-related costs in adulthood.

Limitations:

First, as for any observational study, the findings from the LTF phase of the MTA should be used to generate hypotheses and not be interpreted as tests of hypotheses. Without protection by randomization, the suggestions based on comparisons of naturalistic subgroups are speculative, both for significant effects and for non-significant effects. It remains unclear how much selection or confounding factors contribute to or account for the findings reported here. For example, the cases receiving consistent treatment with medication may have been the most severe cases based on factors not evaluated in the MTA, and if so, this could have masked benefit. Therefore, the possibility of long-term symptom-related benefits should not be excluded by the hypothesis-generating analyses of the LTF phase of the MTA. Instead, the findings should be used to design future studies to tests refined hypotheses about alternative treatments and to identify alternative follow-up methods that may reveal some long-term benefits of medication in some cases in some conditions.

Second, with alternative designs the findings of the MTA may have been different. For example, treatment provided as-usual in the follow-up likely resulted in lower doses and fewer adjustments than treatment by-protocol would have provided. If treatment according to the MTA medication algorithm had continued (including monthly 30-minute physician visits, with adjustments to the medication regimen based on regular review of current status based on parent and teacher ratings), the significant relative benefit of medication observed in the RCT phase (relative reduction of symptom severity) may have persisted into early

adulthood. Also, the acute effect of medication in adulthood was not evaluated. For example, a placebo substitution manipulation was not used to evaluate the cases that continued to use medication in adulthood. If this design had been incorporated into the LTF phase of the MTA, support for continued benefit of medication may have emerged due to increased symptom severity on placebo, even though when on medication, overall symptom severity may not have been reduced compared to other subgroups.

Third, the non-significant effect of medication on the primary outcome reported here (symptom-severity) may not hold for other outcomes. Other studies have reported long-term benefits of stimulant medication on non-symptom outcomes, including reduction in substance use in adolescence (e.g., see Wilens et al, 2006 for a review of historical studies and Groenman, Oosterlaan, Rommelse et al, 2013 for a contemporary study). However, apparent protective effects in adolescence may not be manifested later in adulthood (see Biederman, Monuteaux, Spencer et al, 2008). Mannuzza, Klein, Truong et al (2008) suggested that a protective effect on substance use in adulthood may depend on starting stimulant medication early, and this hypothesis was supported by recent epidemiological studies of medication records in Norway (Dalsgaard, Mortensen, Frydenberg et al, 2014) and a high-school survey in the USA (McCabe, Dickinson, West et al 2016). Analyses of MTA substance use outcomes in adulthood are in progress.

Fourth, the large symptom persistence effect and the medium source discrepancy effect reported here were based on the CAARS rating scale, which in the non-clinical population (i.e., the LNCG) is not normally distributed. The impact of this on the significance levels and effect sizes reported here is unclear. Further evaluation of this is planned for future reports. Also, these effects were observed in early adulthood (at an average age of 25 years), and if the MTA LTF had been continued into mid-adulthood it is possible that smaller or non-significant effects of symptom persistence and source discrepancy may have been observed.

Fifth, the significant association of extended use of medication with reduced adult height may not be causal. The formation of naturalistic subgroups using SCAP1-based sequences of extended use of medication may have resulted in selection biases or confounding related to physical size as well as to severity (see above). For unknown reasons, selective treatment of ADHD cases that are constitutionally short may have occurred. If so, then the height suppression reported here would be due to characteristics of the treated cases rather than to their treatment with medication. An analysis was performed to evaluate the possible association of future use of medication with the baseline height of the naturalistic subgroups described here (i.e., before pattern of medication use could have caused height suppression). This did not show a significant effect of prospective medication use on baseline height, but the sample size of the MTA was not sufficient to evaluate fully multiple factors and their possible interactions (age, sex, maturation, and prior use of medication). A previous analysis (see Swanson et al, 2007) noted that among the ADHD cases without a history of prior use of medication before entry into the MTA, those remaining untreated prospectively during childhood were already about 1 cm taller at baseline than those destined to receive consistent prospective treatment. To compensate for this, Swanson et al (2007) suggested a reduction in estimates of medication-related height suppression in late childhood by 1 cm (i.e., from

about 3 cm to 2 cm). A similar adjustment may be prudent for estimates of height suppression in adulthood.

Sixth, the lack of complete information on prior medication use limited the analysis of this important factor (see Poulton and Nanan, 2008). In analyses of the naturalistic subgroups, neither the main effect of Prior Status nor the interaction of Prior and Prospective use of medication in the MTA (defined by different patterns of extended use by the naturalistic subgroups or by cumulative ME dose) had a significant effect on adult height. However, the direction of small effects (lower intercept and reduced slope in cases with a history of prior use of medication) suggest with a larger sample the effect of Prior Status may reach significance.

Seventh, timing of treatment may affect growth mediated by different underlying biological processes (e.g., growth hormone and sex hormones). Differential exposure to medication during childhood and adolescence were not evaluated here. Also, analyses using mathematical modeling to identify effects of medication on milestones of growth (i.e., height, height velocity, and age of the ‘take-off’ or ‘peak’ points of the adolescent growth spurt) were not reported here. Analyses of trajectory of height from childhood to adulthood (rather than analyses of the adult end-point) may provide additional information on mechanisms involved in suppression of final adult height. These possibilities will be addressed in future reports.

Eighth, outcomes related to weight and body mass were considered secondary here. In the limited analyses of these outcomes, the findings from the planned comparisons are consistent with the findings from some studies in the literature about disorder-related effects, such as those suggesting adults with ADHD are at risk for obesity (see Cortese, Angriman, Maffei et al, 2008), overweight (see Hanc, Stopien, Wolanczyk et al, 2014), and increased body mass (Schwartz, Bailey-Davis, Banded-Roche et al, 2014). However, the findings from the MTA generate the hypothesis that use of medication may increase BMI by decreasing adult.

Ninth, the SCAPI was used to establish the naturalistic subgroups, but the use of this instrument has several limitations. The SCAPI was a new measure developed for the MTA, and information was not available on reliability and validity of parent reports of medication use. The parent-based interview was not administered after the age of 18 years, so information on use of medication in adulthood was not obtained. The SCAPI-based patterns of extended medication use were based on logical analysis of all possible sequences across the assessment intervals from childhood through adolescence, and recently developed data-driven classification methods to establish subgroups with different trajectories of medication use (see Schveren, Groenman, von Rhein et al, 2016) were not evaluated here.

Conclusions:

The observational LTF phase of the MTA generated hypotheses that could be tested in future studies. The findings of the MTA suggest the following: childhood-onset ADHD is a chronic disorder with persistence of symptoms into adulthood; source discrepancy due to lower self-

ratings than parent-ratings is manifested in early adulthood; and extended use of stimulant medication from childhood through adolescence is associated with suppression of adult height but is not associated with reduced symptom-severity in adulthood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The initial RCT phase of the MTA 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, and HHSN271200800008-C; DA-8–5553; DA039881; and McGill University N01MH12008, and HHSN271200800003-C; DA-8–5548. Continuation support was provided by NIDA. Funding support for Dr. Mitchell was provided by NIDA K23 DA032577.

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. 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; Betsy Hoza (currently at University of Vermont); William E. Pelham (PI for early phases, currently at Florida International University). Follow-up phase statistical collaborators: Robert D. Gibbons (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.

The National Institute of Mental Health (NIMH) was involved in the initial study design and treatment phase and early follow-up but not in the adult follow-up phase. National Institute on Drug Abuse negotiated contract deliverables for the final follow-up. J.M.S. 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.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. L.H. has received research support, served on advisory boards and has been a speaker for Eli Lilly, GlaxoSmithKline, Ortho Janssen, Purdue and Shire. The remaining authors declare that they have no competing or potential conflicts of interest in relation to this article.

REFERENCES

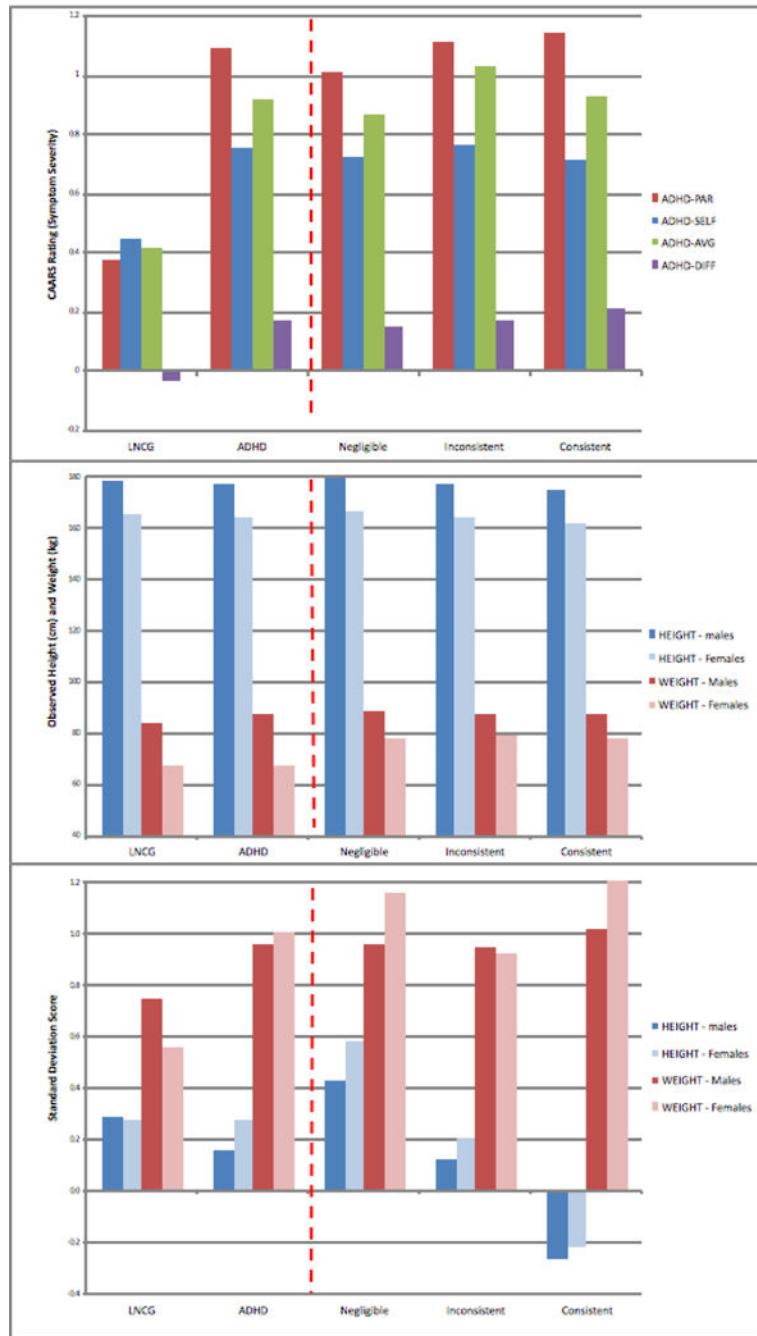
- American Academy of Pediatrics, Subcommittee on Attention-Deficit/Hyperactivity Disorder, Steering Committee on Quality Improvement and Management. *Pediatrics*. 2011; 128: 1–14. [PubMed: 21646265]
- Barkley R, Murphy K, Fischer M. *ADHD in Adults: What Science Says*. NY: Guilford; 2008.

- Biederman J, Spencer TJ, Monuteaux MV, et al. A naturalistic 10-year prospective study of height and weight in children with attention-deficit hyperactivity disorder grown up: sex and treatment effects. *J Pediatrics*. 2010; 157: 635–640.
- Biederman J, Petty CR, Clarke A, et al. Predictors of persistent ADHD: An 11-year follow-up study. *Journal of Psychiatric Research*. 2011; 45: 150–155. [PubMed: 20656298]
- Biederman J, Monuteaux MC, Spencer T, et al. Stimulant therapy and risk for subsequent substance use disorders in male adults with ADHD: A naturalistic controlled 10-year follow-up study. *Am J Psychiatry*. 2008; 165: 597–603. [PubMed: 18316421]
- Charach A, Figueroa M, Chen S et al. Stimulant treatment over 5 years: effects on growth. *J Am Acad Child Adolesc Psychiatry*. 2006; 45: 415–421. [PubMed: 16601646]
- Conners C, Erhardt D, Sparrow E Conner's Adult ADHD Rating Scale. Toronto: MHS; 1999.
- Cortese S, Angriman M, Maffei C, et al. Attention-deficit/hyperactivity disorder (ADHD) and obesity: A systematic review of the literature. *Critical Reviews Food Sci Nutr*. 2008; 48: 524–537.
- Dalsgaard S, Mortensen PB, Frydenberg M et al. ADHD, stimulant treatment in childhood and subsequent substance abuse in adulthood – a naturalistic long-term follow-up study. *Addictive Beh*. 2014; 39: 325–328.
- Faraone SV, Biederman J, and Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. *Psychological Medicine*. 2006; 36: 159–165. [PubMed: 16420712]
- Faraone SV, Biederman J, Morley CP, et al. Effect of stimulants on height and weight: a review of the literature. *J Am Acad Child Adolesc Psychiatry*. 2008; 47: 994–1009. [PubMed: 18580502]
- Graham J, Banaschewski T, Buitelaar J, et al. European guidelines on managing adverse effects of medications for ADHD. *Eur Child Adolesc Psychiatry*. 2011; 20: 17–37. [PubMed: 21042924]
- Groenman AP, Oosterlaan J, Rommelse NNJ, et al. Stimulant treatment for attention-deficit hyperactivity disorder and risk of developing substance use disorder. *Br J Psychiatry*. 2013; 203: 112–119 [PubMed: 23846996]
- Hanc T, Slopian A, Wolanczyk T et al. ADHD and overweight in boys: cross-sectional study with birth weight as a controlled factor. *Eur Child Adolesc Psychiatry*. 2014;
- Harstad EB, Weaver AL, Katusic SK, et al. ADHD, stimulant treatment, and growth: a longitudinal study. *Pediatrics*. 2014; 134: e935–e944. [PubMed: 25180281]
- Hechtman L, Swanson JM, Sibley MH, et al. Functional Adult Outcomes 16 years after Childhood Diagnosis of Attention-Deficit/Hyperactivity Disorder: MTA Results. Submitted.
- Hernan MA, Hernandez-Diaz S, and Robins JM. A structural approach to selection bias. *Epidemiology*. 2004; 15: 615–625. [PubMed: 15308962]
- Al Howard, Strickland NJ Murray DW, et al. Progression of impairment in adolescent with attention-deficit/hyperactivity disorder through transition out of high school: contribution of parent involvement and college attendance. *J Abnormal Psychology*. 2016; 125: 233–247.
- Ioannidis JPA, Evans JW, Gotzche PC et al. Better reporting of harms in randomized trials: An extension of the CONSORT statement. *Ann Intern Med*. 2004; 141: 718–788. [PubMed: 15520429]
- Jensen PS, Hoagwood K, Roper M, et al. The services for children and adolescents-parent interview: development and performance characteristics. *J Clin Child Adolesc Psychology*. 2004; 43:1334–44.
- Jensen PM, Arnold LE, Swanson JM, et al. 3-year follow-up of the NIMH MTA study. *J Amer Acad Child Adolesc Psychiatry*; 2007; 48: 988–1001.
- Klein RG and Mannuzza S Hyperactive boys almost grown up: III. Methylphenidate effects on ultimate height. *Arch Gen Psychiatry*. 1988; 45: 1131–1134. [PubMed: 3058089]
- Klein RG, Mannuzza S, Olazasgasti MAR, et al. Clinical and functional outcome of childhood ADHD 33 years later. *Arch Gen Psychiatry*. 2012; 69: 1295–1303. [PubMed: 23070149]
- Kraemer HC. A source of false findings in published research studies: Adjusting for covariates. *JAMA Psychiatry*. 2015; 72: 961–962. [PubMed: 26244634]
- Kraemer HC and Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry*. 2005; 59: 990–996. [PubMed: 16368078]

- Kraemer HC, Measelle JR, Ablow JC et al. A new approach to integrating data from multiple informants in psychiatric assessment and research: mixing and matching contexts and perspectives. *Am J Psychiatry* 2003 160: 1566–1577. [PubMed: 12944328]
- Kramer JR, Loney J et al. Predictors of adult height and weight in boys treated with methylphenidate for childhood behavior problems. *J Amer Acad Child Adolesc Psychiatry*. 2000; 39: 517–524. [PubMed: 10761355]
- Kuczumski RJ, Ogden CL, Grummer-Steawn LM et al. (2000), Advance data 314,1Y27. Available at: <http://www.cdc.gov/growthcharts>.
- Mannuzza S, Klein RG, Truong NL, et al. Age of methylphenidate treatment initiation in children with ADHD and later substance abuse. *Am J Psychiatry*. 2008; 165: 604–609. [PubMed: 18381904]
- McCabe SE, Brady K, West BT, et al. Age of onset, duration, and type of medication therapy for attention-deficit/hyperactivity disorder and substance use during adolescence: a multi-cohort national study. *J Amer Acad Child Adolesc Psychiatry*. 2016; 55: 479–486. [PubMed: 27238066]
- Molina BSG Pelham WE, Cheong JW, et al. Childhood ADHD and growth in adolescent alcohol use: The roles of functional impairments, ADHD symptom persistence, and parental knowledge. *J Abnormal Psychology*. 2012; 121: 922–935.
- Owens JS, Goldfine ME, Evangelista NM et al. A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical, Child, & Family Psychological Rev*. 2007; 10: 335–351.
- Peyre H, Hoertel N, Cortese S, et al. Long-term effects of ADHD medication on adult height: results from the NESARC. *J Clin Psychiatry*. 2013 74; 1123–1125.
- Poulton AS and Nanan R. Prior treatment with stimulant medication: a much neglected confounder of studies of growth in children with ADHD. *J Child Adolesc Psychopharm*, 2008; 18: 385–387.
- Poulton AS, Melzer E, Talt PR et al. Growth and pubertal development in adolescent boys on stimulant medications for attention deficit hyperactivity disorder. *MJA*. 2013; 198: 29–32. [PubMed: 23330767]
- Schwartz BS, Bailey-Davis L, Bandeen-Roche K et al. Attention deficit disorder, stimulant use, and childhood body mass index trajectory. *Pediatrics*. 2014; 133: 1–8. [PubMed: 24344106]
- Schweren L, Groenman A, von Rhein D et al. Stimulant treatment trajectories are associated with neural reward processing in ADHD. *European Neuropsychopharmacology*. 2016; 26.
- Sibley MH, Pelham WE, Molina BS, et al. When diagnosing ADHD in young adults emphasize informant reports, DSM items, and impairment. *J Consult Clin Psychology*. 2012; 80: 1052–1061.
- Sibley MH, Swanson JM, Arnold LE, et al. Defining ADHD Symptom Persistence in Adulthood: Optimizing Sensitivity and Specificity. *Journal of Child Psychology and Psychiatry*. In Press.
- Spencer T, Biederman J and Wilens T. Growth deficits in children with attention deficit hyperactivity disorder. *Pediatrics*; 1998; 102: 501–506. [PubMed: 9685453]
- Swanson JM, Elliott GR, Greenhill LL et al. Effects of stimulant medication on growth rates across 3 years in the MTA follow-up. *J Amer Acad Child Adolesc Psychiatry*. 2007; 46: 1015–1027. [PubMed: 17667480]
- Swanson J, Kraemer H, Hinshaw S, et al. Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *J Amer Acad Child Adolesc Psychiatry*. 2001; 40: 168–179. [PubMed: 11211365]
- Weiss G, Kruger E, Danielson U, et al. Effects of long-term treatment of hyperactive children with methylphenidate. *CMA Journal*. 1975; 25: 159–165.

Key points

- ADHD is a chronic disorder that persists into adulthood in some cases, but the magnitude of persistence has been unclear. Also, when initiated in childhood stimulant medication has clear benefits (e.g., reduced symptom-severity) that may outweigh costs (e.g., reduced height gain), but the long-term effects of extended use of medication on these outcomes have been unclear.
- The MTA follow-up provides an opportunity to explore long-term outcome of ADHD and effects of long-term treatment with stimulant medication. The findings reported here indicate:
 - In comparison to local norms, a group of cases with childhood-onset ADHD-Combined Type showed high levels of symptom severity that persisted into adulthood.
 - Treatment-as-usual with stimulants in community settings did not alter symptom severity (or symptom persistence), but it was associated with slight height suppression in adulthood.
 - Extended use of medication consistently from childhood to adulthood occurred in less than 10% of ADHD cases, and this pattern of adherence to medication was associated with greater suppression of adult height but still no relative reduction in symptom severity in adulthood.
- Short-term treatment with stimulant medication seems to be well justified by benefits that outweigh costs, but long-term treatment may be associated with residual growth-related costs that may not be balanced by residual symptom-related benefits in adulthood.



	ADHD-SELF			ADHD-PAR	
	n	mean	sd	mean	sd
LNCG	233	0.446	0.419	0.379	0.437
ADHD	439	0.752	0.578	1.09	0.7
Negligible	102	0.72	0.609	1.01	0.727
Inconsistent	307	0.765	0.581	1.11	0.693
Consistent	30	0.718	0.428	1.14	0.68

	ADHD-AVG		ADHD-DIFF		
	n	mean	sd	mean	sd
LNCG	233	0.413	0.363	-0.034	0.227
ADHD	439	0.921	0.503	0.17	0.398
Negligible	102	0.867	0.372	0.147	0.411
Inconsistent	307	1.028	0.569	0.174	0.392
Consistent	30	0.929	0.53	0.211	0.429

	Males		HTcm		WT kg	
	n	mean	sd	mean	sd	sd
LNCG	176	178.92	6.65	83.94	19.13	
ADHD	314	177.99	7.51	87.05	21.37	
Negligible	71	179.89	7.84	88.440	24.89	
Inconsistent	216	177.75	7.3	87.2	20.33	
Consistent	27	174.9	7.19	87.84	20.3	

	Females		HTcm		WT kg	
	n	mean	sd	mean	sd	sd
LNCG	45	165.14	7.61	67.55	14.6	
ADHD	88	165.05	7.61	67.55	14.6	
Negligible	21	167.14	5.67	78.520	19.5	
Inconsistent	61	164.63	8.1	79.02	28.48	
Consistent	6	161.93	7.25	78.31	10.2	

	Height Z-score			Weight z-score	
	n	mean	sd	mean	sd
LNCG	221	0.2913	0.984	0.7128	1.11
ADHD	402	0.1867	1.08	0.9705	1.16
Negligible	92	0.4674	1.05	1.010	1.05
Inconsistent	277	0.1466	1.08	0.9435	1.08
Consistent	33	-0.2588	1.01	1.087	1.01

	Males		HT zscore		WT zscore	
	n	mean	sd	mean	sd	sd
LNCG	176	0.2937	0.93	0.7515	1.15	
ADHD	314	0.1640	1.05	0.9591	1.14	
Negligible	71	0.4312	1.10	0.964	1.23	
Inconsistent	216	0.1304	1.02	0.949	1.14	
Consistent	27	-0.2685	1.00	1.023	0.99	

	Females		HT zscore		WT zscore	
	n	mean	sd	mean	sd	sd
LNCG	45	0.2817	1.18	0.5614	0.924	
ADHD	88	0.2817	1.17	1.011	1.24	
Negligible	21	0.5900	0.878	1.165	0.991	
Inconsistent	61	0.2038	1.25	0.9229	1.36	
Consistent	6	-0.2152	1.12	1.362	0.531	

Figure 1: Orthogonal Comparisons of Diagnostic Groups and Naturalistic Subgroups^a

^a Observed means and standard deviations from SAS GLM analyses for the four groups (LNCG, Negligible, Inconsistent, and Consistent). Least Square Means (LSM), adjusted for other factors in the models, were used for the 3 planned comparisons (ADHD-vs-LNCG; Consistent or Inconsistent-vs-Negligible; Consistent-vs-Inconsistent). The LSM estimates for the four groups are presented in footnotes of Table 3.

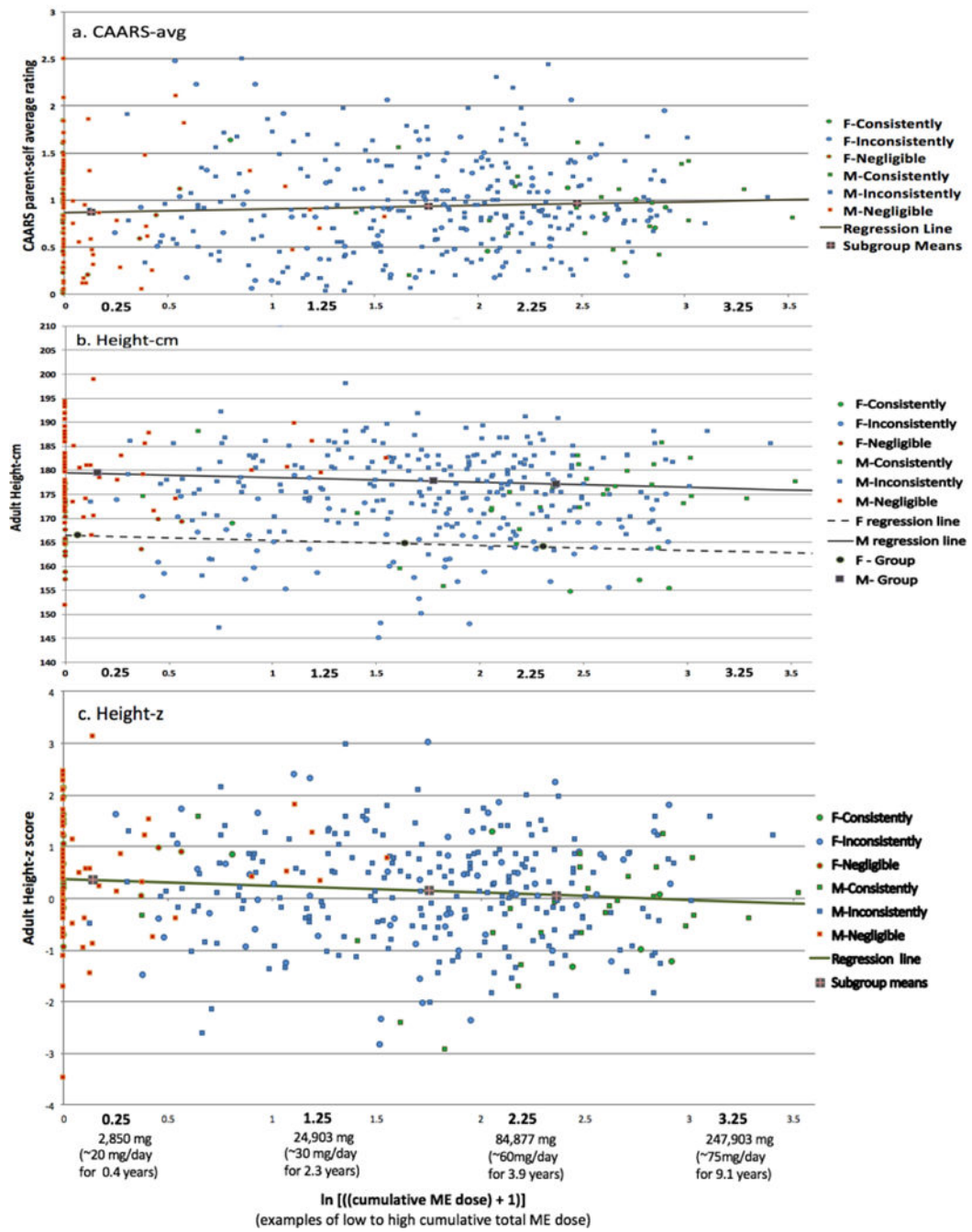


Figure 2:
Regression Analyses of Symptom Severity and Height

Table 1:

The CONSORT Chart for the MTA

Group	Identified	Eligible	Enrolled	Allocated	Followed up	Analyzed
MTA	1000	1000	1000	1000	1000	1000
Control	500	500	500	500	500	500
MTA	500	500	500	500	500	500
Control	500	500	500	500	500	500

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2:

Retention, Medication Use, and Demographic Characteristics of the MTAS Sample

2a: Retention of the Sample in Childhood, Adolescence, and Adulthood

2b: Rates of Medication Use from SCAPIs Baseline to 10 yrs after baseline

2c: Demographic Characteristic of the Diagnostic Groups and Naturalistic Subgroups

	0	14 mo.	2 yr.	3 yr.	6 yr.	8 yr.	10 yr.	12 yr.	14 yr.	16 yr.
	Childhood				Adolescence			Adulthood		
Randomized sample for entire ADHD group and recruited LNCG	ADHD, n = 579			LNCG, n = 289						
Re-consented ADHD group and LNCG without ADHD				ADHD, n = 515 LNCG, n = 258						
Adolescent ADHD group and LNCG with self-report added					ADHD, n = 498 LNCG, n = 249					
Adult ADHD group and LNCG with parent- and self-report								ADHD, n = 476 LNCG, n = 241		

		% Adequate Medication									
	Time Since Baseline	0	14 mo.	2 yr.	3 yr.	6 yr.	8 yr.	10 yr.	12 yr.	14 yr.	16 yr.
	Developmental Stage	Childhood			Adolescence			Adulthood			
Overall for ADHD group	Overall % of ADHD cases	38.3	61.5	57.2	59.2	42.2	29.6	14.3	7.1 (self-report)		
		Average 10-yr total									
For Assigned Treatment Groups	Med, n = 114	42.9	90.2	66.1	66.0	42.1	31.96	11.25	58,986 mg		
	Comb, n = 125	38.5	85.6	69.2	69.8	44.7	28.7	15.38	59,023 mg		
	CC, n = 113	42.1	56.9	58.9	56.7	43.9	30.6	13.0	44,632 mg		
	Beh, n = 124	36.9	13.9	31.2	42.5	39.5	28.9	17.2	41,140 mg		
For Subgroups with Extended Use of Med. ^{a,b}	Negligible, n = 112	28.2	0.0	0.0	0.0	0.0	0.0	0.0	2,153 mg		
	Inconsistent, n = 329	41.3	77.4	68.9	72.8	50.2	32.0	13.3	60,567 mg		
	Consistent, n = 35	66.7	100.0	100.0	100.0	100.0	100.0	100.0	117,102 mg		

GROUPS:	LNCG	ADHD	Negligible	Inconsistent	Consistent
Assessed in adulthood (n) ^a	241	476	112	329	35
Age at assessment (yrs)	24.4	24.8	24.9	24.8	24.9
Age at baseline (yrs) ^b	10.4	8.4	8.6	8.4	8.4
Baseline SNAP (parent)	0.38	2	1.85	2.03	2.08
Baseline SNAP (teacher)	0.54	2.08	2.05	2.09	2.18
Sex (% male)	80%	78%	80%	77%	83%
Birth Weight (kg)	3.44	3.35	3.36	3.35	3.31
Race/Ethnicity (% White)	66.40%	62.70%	52.70%	64.90%	74.3
(% Black)	11.2%	19.4%	27.7%	17.1%	14.3%
(% Hispanic)	12.9%	7.6%	9.8%	7.6%	0.0%
(% Other)	9.5%	10.3%	9.8%	10.4%	11.4%
Intelligence (IQ):	110	102	103	101	105

GROUPS:	LNCG	ADHD	Negligible	Inconsistent	Consistent
Household income (\$10K)	5.9	4.9	4.4	4.9	6.6
Household Advantage ^c 1	13%	20%	27%	18%	9%
2	39%	40%	38%	43%	20%
3	48%	40%	35%	39%	71%

^aNumber of participants with at least one observation in adulthood (at 12, 14, or 16 years after baseline)

^bBaseline for the LNCG obtained 2 years after the baseline for the ADHD group

^cComposite household advantage measures developed by Molina, Pelham, Cheong et al, (2012), with 3 levels (1: one-parent household and no college-educated parent; 2: either two-parent household or at least one college-educated parent; 3: two-parent family and at least one college-educated parent).

^dIQ was higher in the LNCG than in the ADHD group, but based on precedent (see Barkley et al ,2008 and Sibley et al, 2012) it was not included as a covariate.

Table 3:

Analyses of Groups with Planned Comparisons and Regression Analyses of Medication Use

A. Group Analyses:						
Symptom Severity ^a	CAARS-avg			CAARS-diff		
	df	F	p > F	df	F	p > F
Group ^b (Consistent: C, Inconsistent: I, Negligible: N, LNCG: L)	3	60.83	<0.0001	3	18.26	<0.0001
Sex (Male, Female)	1	0.01	0.9429	1	2.24	0.1354
Age (Centered for the MTA Sample)	1	0.04	0.8326	1	0.60	0.4380
Error	656			656		
Planned Comparisons ^{c,d}						
	Contrast	t	p > t	Contrast	t	p > t
ADHD-vs-LNCG (C+I+N _{weighted} -vs-L)	0.514	13.45	<0.0001	-0.212	-7.35	<0.0001
Consistent+Inconsistent-vs-Negligible (C+I _{weighted} -vs-N)	0.063	1.19	0.2331	-0.025	-0.63	0.5282
Consistent-vs-Inconsistent (C-vs-I)	-0.011	-0.13	0.9003	-0.039	-0.58	0.5610
Adult Height ^e						
	Height-cm			Height-z		
	df	F	p > F	df	F	p > F
Group (Consistent: C, Inconsistent: I, Negligible: N, LNCG: L)	3	7.60	<0.0001	3	7.78	<0.0001
Sex (Male, Female)	1	536.74	<0.0001			
Age (Centered for the MTA Sample)	1	1.31	0.2527			
Mid-parent z-Ht	1	235.52	<0.0001	1	240.75	<0.0001
Error	616			622		
Planned Comparisons ^{d,f}						
	Contrast	t	p > t	Contrast	t	p > t
ADHD-vs-LNCG (C+I+N _{weighted} -vs-L)	-1.293	-2.47	<0.0316	-0.197	-2.66	0.0079
Consistent+Inconsistent-vs-Negligible (C+I _{weighted} -vs-N)	-2.550	-3.49	0.0005	-0.363	-3.47	0.0006
Consistent-vs-Inconsistent (C-vs-I)	-2.358	-2.08	0.0378	-0.332	-2.05	0.0411
B. Regression Analyses ^g						
Parameter	Height (cm)			Height (z)		
	Estimate	t	p > t	Estimate	t	p > t
Intercept	178.44	282.05	<0.0001	0.187	2.18	0.0296
Slope (change per unit {ln[((10-year total ME dose)/10,000)+1]})	-1.02	-2.95	0.0034	-0.136	-2.78	0.0058
Sex (Male, Female)	-14.26	-18.05	<0.0001			
Age (Centered for the MTA Sample)	0.33	-1.66	0.0987			
Mid-parent Ht.	4.84	11.54	<0.0001	0.692	11.60	<0.0001
Error	401			399		

^aRatings of symptom severity by both sources were required for the composite scores (CAARS parent/self-report-average and parent/self-report difference). Ratings from both sources were available in adulthood for n=233 participants in the LNCG (90.5%) and for n=439 participants in the ADHD group (85.2%), with n=102 (N), n=307 (I), and n=30 (C) in the subgroups.

^bAnalyses were performed with SAS GLM program, using the end-point observations in adulthood for each participant. For the Group factor, 4 levels were used: LNCG (L), Negligible (N), Inconsistent (I), and Consistent (C).

^cThe GLM Least Square Mean (LSM) estimates of symptom severity (adjusted for other factors in the model) for the 4 groups were C=0.929, I=0.940, N=0.876, L=0.420 (for CAARS-avg) and (C=-0.198, I=-0.159, N=-0.138, L=0.055 (for CAARS-diff).

^dFor the 3 comparisons, contrast coding was used to combine the 4 LSM estimates, with the LSMs either equally ('unweighted') or proportionally weighted by the subgroup sample size/total sample size of the subgroups being combined (see Appendix S2).

^eSome participants were assessed by telephone instead of at a clinic visit. Clinic visits occurred and height was measured for n=221 of the LNCG (85.7%) and n=402 of the ADHD group (78.1%), with n=92 (N), n=276 (I), and n=33 (C) in the subgroups.

^fThe LSM estimates of adult height for the four groups were C=168.3, I=170.7, N=173.0, L=172.3 (for Height-cm) and C=-0.226, I=0.106, N=0.434, and L=0.351 (for Height-z).

^gThese analyses included the ADHD participants with information on the SCAPI on the use of medication. As described in the text, the natural logarithm of cumulative ME dose of medication (after transforming from mg to g by dividing by 10,000 and adding 1).