The Journal of Clinical Psychiatry

Supplementary Material

Article Title: Characteristics and Predictors of Fluctuating ADHD in the Multimodal Treatment of ADHD (MTA) Study

Authors: Margaret H. Sibley, PhD; Traci M. Kennedy, PhD; James M. Swanson, PhD; L. Eugene Arnold, MD; Peter S. Jensen, MD; Lily T. Hechtman, MD; Brooke S.G. Molina, PhD; Andrea Howard, PhD; Laurence Greenhill, MD; Andrea Chronis-Tuscano, PhD; John T. Mitchell, PhD; Jeffrey H. Newcorn, MD; Luis A. Rohde MD, PhD; and Stephen P. Hinshaw, PhD; for the MTA Cooperative Group

DOI Number: 10.4088/JCP.24m15395

LIST OF SUPPLEMENTARY MATERIAL FOR THE ARTICLE

DISCLAIMER

This Supplementary Material has been provided by the authors as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Supplement

Supplementary Table 1. Baseline Characteristics of the MTA Sample

Variable	Total Across All Treatment Groups
Age *M* (*SD*)	8.5 (0.8)
Male n (%)	465 (80.3)
Ethnicity n (%)	
White	351 (60.6)
African-American	115 (19.9)
Hispanic	48 (8.3)
Full Scale IQ *M* (*SD*)	100.9 (14.8)
Comorbidity (DISC) n (%)	
Anxiety Disorder	194 (33.5)
Conduct Disorder	83 (14.3)
Oppositional-Defiant Disorder	231 (39.9)
Affective Disorder	22 (3.8)

Appendix 1: Measurement of Clinical Variables

Impairment. Based on normative analyses in the MTA's non-ADHD group (Sibley et al., 2022), absence of impairment was optimally defined as a "1" or lower on all CIS items. For the IRS, absence of impairment was optimally defined as a "2" or lower on all items (combining parent- and self-reports using an "OR rule").

Comorbidity. The DISC interview assessed mood disorders (major depression, dysthymia, mania), anxiety disorders (agoraphobia, generalized anxiety disorder, obsessive compulsive disorder, panic disorder, separation anxiety disorder, social phobia, selective mutism, post-traumatic stress disorder), disruptive behavior disorders (oppositional defiant disorder, conduct disorder), substance use disorders (abuse and dependence), and eating disorders (anorexia nervosa, bulimia nervosa).

Appendix 2: Additional information about childhood prediction measures.

Parents reported the participant's age, sex, and race/ethnicity at baseline. Parent and teacher ADHD symptom severity was measured on the SNAP. A six-point biological risk score reflecting pre and peri-natal risks (e.g., maternal smoking during pregnancy, birth prior to 37 weeks) was calculated based on the work of Leffa et al., (2023). Based on the work of Rutter et al., (1975) we adapted a psychosocial risk index. For details about calculation of these scores, see Supplement 3. Based on Roy et al., (2016), we measured parental psychopathology based on the total number of parental mental health diagnoses (out of 28 lifetime disorders; from biological mother or father, whichever was higher) assessed with the Structured Clinical Interview for DSM Disorders–Non Patient (SCID) at baseline. Alcohol use disorder was examined separately as a measure of problematic parental drinking. Baseline maternal depression was measured dimensionally on the Beck Depression Inventory; BDI).

For childhood comorbidities, 13 physical health comorbidities were assessed via parent report at baseline. A physical health score aggregated one point for each health condition endorsed (e.g., diabetes, thyroid problems, asthma, allergies). The DISC parent interview administered at baseline assessed 23 mental health disorders comorbid to ADHD (see Supplement 2). A mental health score aggregated one point for each condition endorsed. Presence of ODD/CD, anxiety disorder, and mood disorder were also calculated. To assess severity of anxiety and depression the Multidimensional Anxiety Scale for Children and the Children's Depression Inventory were administered to the child at baseline.

At baseline, the presence of 33 negative life events in the past 12 months were reported on the Coddington Life Event Scale, parent report. Total event score was calculated as a count of the endorsed items. The Wechsler Intelligence Scale for Children (WISC)-3rd Edition was administered to participants at baseline. Full scale IQ was computed for each participant. A continuous performance test (CPT) presented twelve letters on a video monitor in quasi-random sequence until a total of 400 letters were presented. The entire task lasted

approximately 12 minutes. The number of omission errors, commission errors, reaction time, and reaction time variability were calculated. For detailed information about this task see Halperin et al., 1988.

Initial randomized treatment group as well as response to initial randomized treatment (regardless of group) served as predictors. Treatment response by 36 months was measured by membership in one of three latent classes described by Swanson et al., 2007 (see Supplement 3). Pre-study medication, psychosocial treatment, and educational interventions were also examined.

Count of extracurricular activities was calculated from the Child Behavior Checklist. Parents could list involvement in up to three extracurricular activities for their child. Two self-report parenting variables (Negative/Ineffective Discipline and Positive Parenting) were examined as computed by Hinshaw et al., 2000 (see Supplement 3). Parents reported on the number of close friends the child had at baseline using the Child Behavior Checklist. Response options were 0=none, 1=one, 2=two or three, 3=four or more.

Biological Risk Score (Leffa et al., 2023): One point was contributed for each of the following variables that were present: (1) low maternal age at birth (lowest quartile), (2) maternal smoking during pregnancy, (3) maternal hypertension during pregnancy, (4) cesarean section birth, (5) birth prior to 37 weeks of pregnancy, and (6) postnatal smoke exposure in the home up to 5 years of age.

Psychosocial Risk Score (Rutter et al., 1975): We contributed one point each for the following variables: (1) both parents without a college degree, (2) single parent household, and (3) three or more children in the household.

Treatment Response Latent Classes (Swanson et al., 2007): Class 1 (n = 199, 34% of the sample) manifested a linearly decreasing (improving) symptom trend over time; class 2 (n = 299, 52% of the sample) manifested a large initial symptom decrease that was maintained over time; class 3 (n = 81, 14%) manifested a quadratic trend, with an initial decrease followed by a return to baseline (Fig. 2).

Parenting (Hinshaw et al., 2000): Hinshaw and colleagues created at baseline from items on the Alabama Parenting Questionnaire (APQ) and the Parent Child Relationship Questionnaire (PCRQ). First, each questionnaire was separately submitted to a principal components analysis (see Hinshaw et al. 2000 for summary), and then the first-order factors derived from those analyses were factor analyzed.

References for Measures not Cited in Main Document

Spitzer RL, Williams JBW, Gibbon M: Instruction Manual for the Structured Clinical Interview for DSM-III-R (SCID). New York, New York State Psychiatric Institute, Biometrics Research, 1989

Beck, A. T., & Steer, R. A. (1984). Internal consistencies of the original and revised Beck Depression Inventory. Journal of Clinical Psychology, 40(6), 1365–1367.

March JS, Parker JD, Sullivan K, Stallings P, Conners CK. The Multidimensional Anxiety Scale for Children (MASC): factor structure, reliability, and validity. Journal of the American academy of child & adolescent psychiatry. 1997 Apr 1;36(4):554-65.

Kovacs M. Child depression inventory. Personality and Individual Differences. 1992.

Coddington, R.D. (1984). Measuring the stressfulness of a child's environment. In J.H. Humphrey (Ed.), Stress in Childhood (97-126). New York: AMC Press.

Wechsler, D. (1991). Wechsler Intelligence Scale for Children—Third Edition manual. San Antonio: The Psychological Corporation.

Achenbach TM, Edelbrock C. Child behavior checklist. Burlington (Vt). 1991;7:371-92.

Appendix 3: Details of Sibley et al., remission classification system

Full remission required symptoms to fall below the full remission threshold (3 symptoms of inattention—IN-- and hyperactivity/impulsivity--HI) according to all informants, absence of clinically significant impairment, and discontinuation of all ADHD intervention for at least a month prior to assessment. For persistent, we utilized a previously validated definition of persistence, which applied the DSM-5 symptom threshold (5 or 6 symptoms of either Inattention or Hyperactivity/Impulsivity, depending on age) using the CAARS (or SNAP) and impairment threshold of "3 or higher" on the IRS (or CIS). Partially remitted cases met criteria for neither persistence nor full remission, typically because they had low symptoms but continued impairment, high symptoms but insufficient impairment, or met symptom and impairment criteria for full remission, but were currently treated.

Appendix 4: Multilevel model Sensitivity Analysis. As a sensitivity analysis, we also reconducted the analysis with both a between-person comorbidity index and a time-varying comorbidity index as covariates in the model,

as well as a time-varying demands x time-varying comorbidity interaction term to understand whether the association between demands and ADHD status remains after considering comorbidity. Because of increased rates of missing DISC data over time, participants in this secondary analysis contributed an average of 4.27 of 6 possible data points (70.5% complete data).

Appendix 5: Proposed Future directions related to the time course of remission/recurrence: a commentary provided by Dr. Swanson.

ADHD is considered to be a chronic condition (e.g., like substance use disorder) or an extreme of a trait (e.g., like extraversion), but in the MTA follow-up it was not a stable condition (see Sibley et al., 2022). The current article (Sibley et al., 2024) characterized the fluctuations between two clinical states, remission and recurrence of ADHD, based on rigorously defined categorical cutoffs (e.g., counts of symptoms and impairments). A clear and important finding is that majority of cases met criteria for fluctuating status (63.8%), which is described in detail and discussed extensively. One research direction outside the scope of the current paper is investigating factors related to temporal course of outcome or stable status defined by states of remission and recurrence.In a future investigation, the MTA will address this limitation by applying methods for survival analysis to characterize time-to-remission (which occurred at some point of the MTA follow-up in 92% of the cases) and duration of remission or time-to-recurrence (which occurred in 82% of the cases). This could be accomplished by applying the method described by Snappin (2005), "the extended Kaplan-Meier method with Cox regression", which could provide an estimate of the condition probability of occurrence up-to the time of occurrence of an event (defined either as "remission" or "recurrence" of ADHD) and a comparison of the subgroups defined by Sibley et al. (2022) and characterized by Sibley et al. (2024) on the average time of onset and average duration of these stable components of these binary measures of outcome. This alternative approach would supplement the current set of analyses by building off of the specific aims specified in Sibley et al. (2024).

Supplementary Table 2: Between-group comparisons for multinomial categorical childhood predictors.

	1 vs. 2	1 vs. 3	1 vs. 4	2 vs. 3	2 vs. 4	3 vs. 4
Assigned Treatment Group						
Med vs. Beh	1.33	1.70	1.20	.78	1.11	1.41
Med vs. Comb	1.20	1.57	1.20	.76	1.00	1.30
CC vs. Med	.79	1.05	.91	.76	.87	1.15
Comb vs. Beh	1.11	1.08	1.00	1.02	1.11	1.09
CC vs. Beh	1.05	1.77	1.09	.59	.96	1.64
CC vs. Comb	.95	1.64	1.09	.58	.87	1.49
36 months Tx Response						
Class 1 vs. Class 2	2.72	1.48	.549	.55	**4.95**	2.70
Class 2 vs. Class 3	**.23**	.77	1.98	3.38	**.11**	.39
Class 1 vs. Class 3	.62	1.13	1.08	1.84	.57	1.04

Note. Statistically significant effects noted in bold. Effects are represented by odds ratios. 1=stable persistence; 2=stable partial remission; 3=recovery; 4=fluctuating; Med=medication; Beh=Behavioral Treatment; Comb=combined medication and behavioral treatment; CC=community comparison; Class 1=gradual improvement; class 2=large initial improvement with maintenance; class 3= large initial improvement with return to baseline

Supplementary Table 3: Relationship between Demands and ADHD Fluctuations with comorbidity as a covariate

	Persistence vs. Full Remission				Persistence vs. Partial Remission			
	b	*SE*	*p*	*OR*	*b*	*SE*	*p*	*OR*
Age	.098	.029	<.001	1.103	-.037	.018	.040	.964
Demands: Person-Centered Mean	.106	.217	.625	1.112	.210	.138	.127	1.234
Demands: Time-Varying	.322	.160	.044	1.380	.101	.089	.253	1.107
Comorbidity: Person-Centered Mean	-1.203	.189	<.001	.300	-.631	.079	<.001	.532
Comorbidity: Time-Varying	-.611	.159	<.001	.543	-.261	.058	<.001	.770
Demands: Time-Varying x Age	-.045	.040	.268	.956	-.027	.022	.220	.973
Demands: Time-Varying x Comorbidity	.095	.145	.514	1.099	.069	.062	.268	1.071

Supplementary Table 4. Sensitivity analyses with restricted sample (six or more follow-up assessments)

	1. Fluctuating *M*(*SD*) **N=279**	2. Stable Persistence *M*(*SD*) **N=24**	3. Stable Partial Remission *M*(*SD*) **N=29**	4. Recovery *M*(*SD*) **N=47**	1 vs. 2 *p*	1 vs. 3 *p*	1 vs. 4 *p*	2 vs. 3 *p*	2 vs. 4 *p*	3 vs. 4 *p*
Total Fluctuations	3.73(1.39)	.00(.00)	1.00(.00)	3.26 (1.13)	**<.001**	**<.001**	**.018**	**.004**	**<.001**	**<.001**
IN Count Peak	8.60(1.03)	8.96 (.20)	8.55(1.09)	7.06(2.49)	.191	.840[a]	**<.001**	.249[a]	**<.001**	**.027**
H/I Count Peak	7.19 (2.05)	8.33(1.09)	7.10(2.14)	5.19(2.63)	**.011**	.834[a]	**<.001**	**.034**	**<.001**	**<.001**
IN Count Trough	1.39(1.95)	5.50(1.93)	1.03(1.30)	.06(.32)	**<.001**	.307	**<.001**	**<.001**	**<.001**	**.022**
H/I Count Trough	.99(1.36)	2.92(2.59)	.83(1.14)	.13(.41)	**<.001**	.549	**<.001**	**<.001**	**<.001**	**.029**
Age at First Remission Episode	12.32(3.37)	---	19.44(5.57)	11.50(2.30)	---	**<.001**	.135	---	---	**<.001**
Proportion of Assessments Impaired	83.92(17.94)	100.00(.00)	90.95(17.13)	45.35(20.43)	**<.001**	.042[a]	**<.001**	.063	**<.001**	**<.001**
Proportion of Assessments with Comorbidity										
Anxiety	17.06(16.91)	28.68(22.72)	24.91(22.9)	11.29(14.77)	**.002**	**.023**	**.038**[b]	.438	**<.001**	**.001**
Mood	4.27(8.54)	14.45(19.23)	6.70(12.70)	1.33(4.69)	**<.001**	.193	.052[a]	**.004**	**<.001**	**.018**
Substance Use[c]	26.21(28.88)	27.15(25.40)	17.24(22.85)	12.23(20.81)	.872	.094	**.001**	.191	.031[a]	.439[a]
Proportion of Assessments Medicated	28.93(23.63)	29.02(26.79)	34.48(27.42)	20.85(19.20)	.891	.298	**.017**	.400[a]	.166	**.014**
Proportion of Assessments with Psychosocial Tx	19.82(20.82)	36.28(25.89)	38.15(26.99)	9.27(12.98)	**<.001**	**<.001**	**.001**	.736	**<.001**	**<.001**
DSM-5 Symptom Persistence at Adult Endpoint (%)	44.4	100.0	20.7	0.00	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**	**<.001**
Number of Assessments	7.53(.70)	7.58(.72)	7.31(.85)	7.51(.78)	.748[a]	.113[a]	.837	.171[a]	.688[a]	.240[a]
Age at Final Assessment	24.84(1.21)	24.81(1.37)	24.61(1.09)	24.99 (1.23)	.896	.327	.446	.554	.555	.188

[a]Significance lost when using six or more assesssments vs. adult data as the inclusion criterion. [b]Significance gained when using six or more assessments vs. adult data as the inclusion criterion. [c]Substance use disorder was only collected during the 6 through 16 year assessment