

The Treatment of Severe Childhood Aggression Study: 12 Weeks of Extended, Blinded Treatment in Clinical Responders

Robert L. Findling, MD, MBA,^{1,2} Lisa Townsend, PhD,^{1,3} Nicole V. Brown, MS,⁴
L. Eugene Arnold, MD, MEd,⁵ Kenneth D. Gadow, PhD,⁶ David J. Kolko, PhD,⁷
Nora K. McNamara, MD,⁸ Devin S. Gary, PhD,² Dana B. Kaplin, MPH,¹ Cristan A. Farmer, PhD,⁹
Heidi Kipp, MD, LPC,⁷ Craig Williams, MD,¹⁰ Eric M. Butter, PhD,¹¹ Oscar G. Bukstein, MD, MPH,¹²
Robert Rice, Jr., PhD,¹⁰ Kristin Buchan-Page, BS,¹⁰ Brooke S.G. Molina, PhD,⁷ and Michael G. Aman, PhD¹⁰

Abstract

Objectives: Previous “Treatment of Severe Childhood Aggression” (TOSCA) reports demonstrated that many children with severe physical aggression and attention-deficit/hyperactivity disorder (ADHD) responded well to two randomized treatments (parent training [PT]+stimulant+placebo=Basic vs. PT+stimulant+risperidone=Augmented) for 9 weeks. An important clinical question is whether these favorable outcomes are maintained over longer times.

Methods: Clinical responders to the 9-week trial ($n=103/168$), defined as Clinical Global Impressions (CGI)-Improvement of much/very much improved plus substantial reduction in parent ratings of disruptiveness, were followed another 12 weeks (21 weeks total) while remaining on blinded treatment. Outcome measures included Clinical Global Impressions scale, Nisonger Child Behavior Rating Form (NCBRF), other parent/teacher-rated scales, laboratory tests, clinician ratings of abnormal movement, and other adverse events (AEs).

Results: Parent ratings of problem behavior showed minimal worsening of behavior from end of the 9-week acute trial (expected from regression to the mean after selecting best responders), but outcomes at Extension endpoint were meaningfully improved compared with acute study baseline. As expected, outcomes for Basic and Augmented treatment did not differ among these children selected for good clinical response. During Extension, more Augmented subjects had elevated prolactin; there were no clinically confirmed cases of tardive dyskinesia. Delayed sleep onset was the most frequent Basic AE. We also conducted a last-observation-carried-forward analysis, which included both nonresponders and responders. We found that, at the end of Extension, Augmented subjects had more improvement than Basic subjects on the NCBRF Positive Social subscale ($p=0.005$; $d=0.44$), the Antisocial Behavior Scale Reactive Aggression subscale ($p=0.03$; $d=0.36$), and marginally so on the Disruptive Behavior Total subscale ($p=0.058$; $d=0.29$, the primary outcome).

¹Division of Child and Adolescent Psychiatry, Johns Hopkins University, Baltimore, Maryland.

²Department of Psychiatry, Kennedy Krieger Institute, Baltimore, Maryland.

³Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

⁴Center for Biostatistics, Ohio State University Medical Center, Columbus, Ohio.

⁵Department of Psychiatry, Ohio State University, Columbus, Ohio.

⁶Department of Psychiatry, Stony Brook University, Stony Brook, New York.

⁷Department of Psychiatry and Psychology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania.

⁸Department of Psychiatry, Case Western Reserve University, Cleveland, Ohio.

⁹Pediatrics & Developmental Neuroscience Branch, National Institute of Mental Health, Bethesda, Maryland.

¹⁰The Nisonger Center (UCEDD), Ohio State University, Columbus, Ohio.

¹¹Division of Pediatric Psychology and Neuropsychology, Nationwide Children’s Hospital, Columbus, Ohio.

¹²Harvard University Children’s Hospital, Boston, Massachusetts.

Funding: This study was supported by grants from National Institute of Mental Health (NIMH) to The Ohio State University (R01 MH077907), Case Western Reserve University (R01 MH077750), the University of Pittsburgh (R01 MH077676), and Stony Brook University (R01 MH 077997). The project was supported by a National Institutes of Health (NIH) General Clinical Research Center grant M01RR10710 (State University of New York Stony Brook) and Clinical and Translational Science Awards from the National Center for Advancing Translational Sciences grants 8UL1TR000090-05 (The Ohio State University) and UL1 RR024153 and UL1TR000005 (University of Pittsburgh).

Conclusions: The medium-term outcomes were good for the participants in both treatment groups, perhaps because they were selected for good response. When nonresponders were included in ITT analyses, there was some indication that Augmented surpassed Basic treatment.

Keywords: disruptive behavior disorders, stimulant, clinical trial, aggression

Introduction

SERIOUS DYSFUNCTIONAL AGGRESSION is common in children referred for psychiatric evaluation (Stattin and Magnusson 1989). Early disruptive behavior disorders (DBDs) in children result in significant impairment and may be predictors of negative consequences later in life. Such consequences include delinquency, risky sexual behavior, substance abuse, and perpetrating serious crimes (Stattin and Magnusson 1989; Brody et al. 2003; Timmermans et al. 2008). Given the persistence of impairments related to DBDs, the durability of treatments is important to consider. Therefore, studies that extend beyond brief acute phases are needed.

The Treatment of Severe Childhood Aggression (TOSCA) study started with an initial 9-week, four-site, double-blind, placebo-controlled clinical trial in children with severe physical aggression and attention-deficit/hyperactivity disorder (ADHD) (Aman et al. 2014). The primary results showed that risperidone (RIS), when added to psychostimulant pharmacotherapy (STIM) and parent training (PT), was moderately more effective than placebo added to STIM and PT in treating disruptive behavior measured by the Nisonger Child Behavior Rating Form (NCBRF) disruptive total score (D-total) in the short term (Aman et al. 2014). Secondary NCBRF subscales and a complementary report (Gadow et al. 2014) showed that the children receiving risperidone also had modestly to moderately improved parent-rated oppositional-defiant disorder (ODD) symptoms, aggression, home impairment, prosocial behavior, and teacher-rated ADHD symptoms.

The purpose of this report is to describe a 12-week Extension phase (Extension) of additional blinded treatment in participants who responded clinically to randomized interventions during the initial 9-week acute trial.

Study questions

We had two main interests in conducting this extension of our earlier randomized, controlled clinical trial. First, we wanted to determine if therapeutic effect was maintained over time. We expected that because Extension participants were selected for being good responders, they would undergo some regression to the mean, but, for the most part, would maintain their improved state. If symptomatic deterioration should occur, we expected it might occur more frequently in the group that received only STIM and PT. Second, we wanted to assess longer term tolerability and safety of STIM alone versus STIM+RIS. The combined condition may lead to either additive adverse events (AEs) or neutralization of AEs (Farmer et al. 2011). For instance, stimulants and antipsychotics may have opposite effects on weight, arousal, and prolactin. Therefore, we did not formulate a directional hypothesis in terms of what condition might show greater AEs. An exploratory comparison included all 154 participants from the double-blind phase of the acute trial, irrespective of clinical response, to endpoint of their participation (either double-blind or Extension phase). This was done to compare their outcomes with those of the responders alone and to gauge, in an unbiased manner, any decline of *Augmented* superiority in the undifferentiated sample. As we already knew that the treatment groups

diverged in the acute trial, we predicted continued separation of outcomes in this last-observation-carried-forward analysis. Complementary to this issue, we also examined participant attrition from the acute trial to the endpoint of the Extension and clinical response status at end of the Extension. As the number of clinical responders did not differ significantly at end of the acute trial, we were agnostic about any differences at Extension endpoint.

Methods

Study design

Before study procedures, parents/guardians of participants consented and children assented, using forms and procedures approved by each site's institutional review board and by the TOSCA data safety and monitoring board. The initial phase of TOSCA was a 9-week, randomized, double-blind, placebo-controlled study conducted across four sites—Case Western Reserve University, Ohio State University, University of Pittsburgh, and Stony Brook University. The *a priori* primary outcome measure for this study was the NCBRF–Typical IQ *Disruptive Total score* (NCBRF D-total) (Aman et al. 2008). Participants were randomized to two treatment groups: *Basic* (PT+STIM+placebo) or *Augmented* (PT+STIM+RIS). Randomization occurred at baseline in a 1:1 ratio, was stratified by site, and balanced for diagnosis of comorbid conduct disorder (CD) versus ODD.

During weeks 1–3 of the acute trial, participants in both groups received PT and STIM only. Subsequently, in weeks 4–8, participants with a less-than-optimal clinical response received a second double-blind medication added to PT and STIM: those in the *Basic* group received placebo (PBO), whereas those in the *Augmented* group received RIS.

PT was administered via an empirically established program for children (*Community Parent Education Program* [COPE]), nine sessions focusing on strategies for management of impulsive behavior—including reactive aggression. COPE sessions ran concurrently from baseline through week 9 of the acute phase, with two booster COPE sessions at 1 and 2 months into the Extension (Cunningham 2005; Cunningham et al. 2009).

Further details regarding the background, design, methods, and variables of the acute phase of the study can be obtained from Aman et al. (2014) and Farmer et al. (2011).

Participants

Inclusion criteria for children in the acute trial were ages 6–12 years, inclusive; evidence of serious physical aggression as determined by the Modified Overt Aggression Scale (Coccaro et al. 1991) (score ≥ 3 involving assaults against self, objects, or other people); a *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition (DSM-IV) (American Psychiatric Association 1994) diagnosis of CD or ODD; a DSM-IV diagnosis of ADHD (any type); evidence of seriously disruptive behavior as indicated by a parent/guardian rating of ≥ 27 (90th percentile) on the NCBRF D-Total; and a score of at least 4 (moderately ill or worse) for

aggression on the Clinical Global Impressions-Severity (CGI-S) subscale (Guy 1976). Participants entering the acute trial were free of psychotropic medications before randomization.

Exclusion criteria were as follows: full-scale IQ <71; pregnancy; a history of seizures or a neurologic or general medical disorder that could present undue risk in combination with the study drugs; a pervasive developmental disorder; schizophrenia, schizoaffective disorder, and other psychoses; eating disorders; evidence of current or previous major depressive disorder; diagnosis of bipolar disorder; evidence of recent or current child abuse/neglect; history of a suicide attempt in the past year or current suicidal ideation; a family history of type 2 diabetes in two or more first-degree relatives; current use of psychotropic medication from which a discontinuation may present a health risk; and an active substance use disorder.

At the conclusion of the acute trial (week 9), participants' assessments were reviewed and clinical responders (NCBRF D-Total decrease of >25% and a CGI-Improvement [CGI-I] of 1 or 2 compared to acute trial baseline scores) were eligible for transition into the 12-week blinded Extension. Eligibility also required adherence to study-related procedures during the acute trial and the ability to tolerate prescribed medications, in the treating physician's judgment. Participants in the Extension were seen every 4 weeks, for a total of three additional visits during the 12-week period. During the Extension, participants continued to receive the double-blind treatment to which they were initially randomized and which had appeared successful.

Study medication

During the acute phase, participants started on a once-daily dose of OROS methylphenidate (MPH) at 18 mg, which was titrated over the first 2 weeks to optimal dose (up to 72 mg). Dosing was based on participant weight, side effects, and clinical response to MPH. Participants who were initially unable to swallow or tolerate MPH were able to receive alternative long-acting STIMs. Participants generally remained on the established dose of STIM during the Extension; however, dosing could be adjusted to enhance efficacy or tolerability.

Participants randomized to RIS (*Augmented* group) or PBO (*Basic* group) received the following dosage schedule: 0.5 to 2.5 mg/day of RIS (or PBO) for children <25 kg and 0.5 to 3.5 mg/day RIS (or PBO) for children ≥25 kg. As with STIM, RIS/PBO doses could be titrated by an individual patient's primary clinician as needed to maximize benefit and minimize side effects.

Efficacy assessments and schedule

The NCBRF includes one prosocial subscale (Positive/Social) and six problem behavior subscales (Conduct Problem, Oppositional, Hyperactive, Inattentive, Overly Sensitive, and Withdrawn/Dysphoric) (Aman et al. 2008). The Conduct Problem and Oppositional subscales are summed to give the D-Total score (the primary outcome), and Hyperactive and Inattentive subscales are totaled to provide an ADHD composite.

The NCBRF D-Total has high internal consistency, can differentiate between controls and patients with disruptive behavioral disorders, and is known to be highly treatment sensitive (Aman et al. 2008). The parent-completed NCBRF was obtained at key points in the acute study and at each Extension visit. The following secondary measures were obtained at the final Extension visit in addition to the last acute trial visit: (1) Antisocial Behavior Scale (ABS) (Brown et al. 1996), comprising the Proactive Aggression and Reactive Aggression subscales; (2) CGI-I; and (3) the Child-

Adolescent Symptom Inventory, fourth edition revised (CASI-4R; parent-completed version) (Gadow and Sprafkin 2005).

Safety assessments and schedule

Safety measures at baseline of the Extension (acute study week 9) and all three Extension visits included (1) the Abnormal Involuntary Movement Scale (AIMS; clinician-rated tremor, dyskinesia, and other neuromotor side effects) (Guy 1976), (2) the Barnes Akathisia Scale (BAS; clinician-completed scale integrating clinical and objective judgment and patient-reported experience of restlessness) (Barnes 1989), and (3) the Simpson-Angus Scale (SAS; detecting extrapyramidal symptoms) (Simpson and Angus 1970). Vital signs (including height and weight), a review of side effects, and open-ended elicitation of AEs and concomitant medications were collected at each study visit.

Clinical laboratory values (lipids, glucose, and prolactin concentration), physical examinations (including hip-waist ratio), and electrocardiogram recordings were taken at the baseline for the acute study and final Extension visits. Prolactin concentrations ≥100 ng/mL, even in the absence of any clinical signs or symptoms, were recorded as AEs. A final medical history was taken on the last Extension visit.

Statistical analysis

For the primary outcome, the NCBRF D-Total score, a mixed-effects model was used to assess within and between treatment changes for the group of participants who entered the Extension ($n = 103$). Fixed effects included those for time, treatment, treatment-by-time interaction, site, and disorder type. An unstructured variance covariance matrix was assumed for the correlated measurements within each participant and empirical-based sandwich estimators were obtained to assess the group differences at week 9 and 21.

Secondary longitudinal outcome variables, such as the remaining NCBRF subscales, ABS, CGI-I, and CASI-4R were modeled in a similar manner to the D-Total, with appropriate adjustments made for quantitative versus categorical responses. Although mixed models can be used in the presence of missing data, the missing-at-random assumption is not directly testable, and model results for behavioral measures reflect estimations made via the restricted maximum likelihood method. The CASI-4R score was calculated as a mean symptom severity score and an impairment rating (defined as parent ratings of *Often* or *Very Often* vs. *Never* or *Sometimes*). The details regarding these measures are provided in Gadow et al. (2016). Model assumptions were assessed by examination of residuals. Some outcomes were transformed by square-root to accommodate the assumption of normality. Data for AEs, BAS, SAS, laboratory, and vital signs were reported descriptively.

Secondary analyses. As a sensitivity analysis, we examined the effect of randomization at week 21 on the NCBRF D-Total scale using week 3 as "baseline." These analyses were not reported here as the conclusions did not differ from the primary analyses. In addition, we performed exploratory analyses for three select outcomes by using last-observation-carried-forward (LOCF) for the NCBRF D-Total, NCBRF Positive Social, and the ABS Reactive scales. Observations at week 4 or later were carried forward through week 21 for the NCBRF and ABS measures. These LOCF analyses used mixed models to assess differences in treatment groups as described for the primary outcome measurements. No adjustments for multiple comparisons were made to reported outcomes. Cohen's d was used to estimate effect size

(ES) with ES calculated using LOCF after week 9 for primary outcomes. Pooled standard deviations (SDs) were used to compute *d*, with the relevant baseline and endpoint SDs used for each analysis. All analyses were conducted in SAS/STAT, version 9.3 (SAS Institute, Cary, NC).

Results

Study participants

One hundred sixty-eight participants began the acute trial; 137 of them (81.5%) completed it. Most (*n*=103, 75.2%) of the 137 participants who completed the acute trial entered the 12-week Extension. Most participants responding to the acute treatment went on to participate in the Extension. There were, however, seven responders who declined the Extension. The *Basic* group continued with PT+STIM+PBO (*n*=49) and the *Augmented* group continued with PT+STIM+RIS (*n*=54). Eighty-eight (85.4%) of the 103 children enrolled in the Extension completed it. Of note, one nonresponder in the *Basic* group was inadvertently enrolled in the Extension. Disposition of youth in the Extension is detailed in Figure 1. That participant was included in the analyses.

The demographic characteristics of the 103 Extension participants are summarized in Table 1. Compared to the other 65 participants not enrolled in the Extension, mothers of the Extension participants had higher levels of education, defined as some college experience or more (*p*=0.02). Other differences between participants and nonparticipants were not statistically significant.

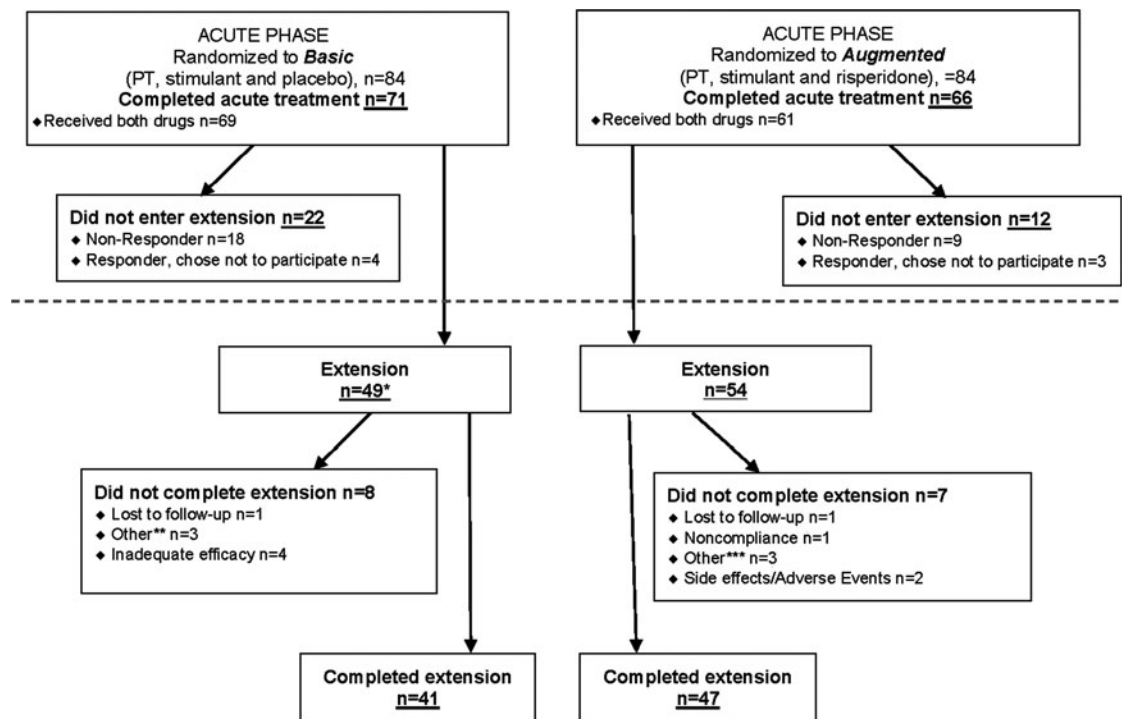
Treatment in the extension phase

The mean length of total study participation (acute+Extension) for the 103 responders was 19.8 weeks (±2.9) for the *Basic* group and 20.2 weeks (±2.3) for the *Augmented* group out of a possible 21 weeks. On entering the Extension, the *Basic* participants were receiving a mean of 44.9±14.9 mg/day MPH; and the *Augmented* participants were receiving a mean of 46.8±16.6 mg/day MPH plus 1.58±0.74 mg/day RIS. During the Extension, the mean modal dose (the average of the most frequently occurring dose for each participant) of MPH for the *Basic* treatment participants was 44.8±14.6 mg/day. The *Augmented* participants' mean modal Extension doses were 46.8±16.9 mg/day of MPH and 1.56±0.73 mg/day of RIS. At end of the Extension, *Basic* participants were receiving a mean dose of 48.0±13.1 mg/day of MPH and *Augmented* participants a mean dose of 48.4±16.4 mg/day of MPH, plus a mean of 1.55±0.72 mg/day of RIS.

Numbers of COPE sessions attended during the Extension were compared using the Wilcoxon rank-sum test, due to non-normal distribution. Thirty (61%) of the *Basic* participants and 42 (78%) of the *Augmented* participants received at least one Extension COPE session. The mean number of COPE sessions received by *Basic* participants was 1.18±1.11 and by *Augmented* participants 1.39±1.02 (*p*=0.33; Mann-Whitney test).

NCBRF outcome measure results

Tables 2A, 2B, and 3A, 3B below report the outcomes for the NCBRF and ABS, respectively. Part A of the tables gives the



* One participant was a Non-Responder that was entered into this Extension Phase in error.

**Other includes mother becoming ill, inconvenience of study visits, and continued travel being too demanding.

***Other includes participants no longer living in the area, family too busy, mother/father withdrew consent.

FIG. 1. CONSORT diagram accounting for participants during the TOSCA Extension phase (weeks 9 through 21). TOSCA, Treatment of Severe Childhood Aggression.

TABLE 1. DEMOGRAPHIC FEATURES OF PARTICIPANTS IN THE TREATMENT OF SEVERE CHILDHOOD AGGRESSION EXTENSION (WEEKS 9 THROUGH 21)

Characteristic	Basic (n=49)	Augmented (n=54)	Overall (n=103)
Gender, n (%)			
Male	37 (75.5)	44 (81.5)	81 (78.6)
Female	12 (24.5)	10 (18.5)	22 (21.4)
Disorder, n (%)			
CD	13 (26.5)	14 (25.9)	27 (26.2)
ODD	36 (73.5)	40 (74.1)	76 (73.8)
Age (years) at week 9, mean (SD)	9.1 (1.9)	9.4 (2.1)	9.2 (2.0)
IQ at screening, mean (SD)	97.6 (14.8)	97.6 (12.6)	97.6 (13.6)
Race, n (%)			
White	24 (49.0)	35 (64.8)	59 (57.3)
Black	18 (36.7)	14 (25.9)	32 (31.1)
Multiracial	7 (14.3)	5 (9.3)	12 (11.6)
Child's type of school, n (%)			
Regular public (or parochial)	41 (83.7)	48 (88.9)	89 (86.4)
Other	8 (16.3)	6 (11.1)	14 (13.6)
Mother's employment, n (%)			
Keeping house	6 (12.2)	6 (11.1)	12 (11.6)
Working full/part time	25 (51.0)	31 (57.4)	56 (54.4)
Other	18 (36.7)	17 (31.5)	35 (34.0)
Father's employment, n (%)			
Keeping house	1 (2.0)	0 (0.0)	1 (1.0)
Working full/part time	23 (46.9)	31 (57.4)	54 (52.4)
Unknown	2 (4.1)	0 (0.0)	2 (1.9)
Other	23 (46.9)	23 (42.6)	46 (44.7)
Mother's education, n (%)			
Some high school or less	0 (0.0)	5 (9.3)	5 (4.9)
High school graduate or GED	8 (16.3)	17 (31.5)	25 (24.3)
Some college or more	41 (83.7)	32 (59.3)	73 (70.9)
Father's education, n (%)			
Some high school or less	2 (4.1)	3 (5.6)	5 (4.8)
High school graduate or GED	16 (32.7)	16 (29.6)	32 (31.1)
Some college or more	14 (28.6)	20 (37.0)	34 (33.0)
Not in household or unknown	17 (34.7)	15 (27.8)	32 (31.1)
Income, n (%)			
Less than \$20,000	22 (44.9)	16 (29.6)	38 (36.9)
\$20,001–\$40,000	9 (18.4)	11 (20.4)	20 (19.4)
\$40,001–\$60,000	7 (14.3)	11 (20.4)	18 (17.5)
\$60,001–\$90,000	5 (10.2)	6 (11.1)	11 (10.7)
More than \$90,000	5 (10.2)	8 (14.8)	13 (12.6)
Unknown	1 (2.0)	2 (3.7)	3 (2.9)

CD, conduct disorder; GED, General Education Diploma; ODD, oppositional-defiant disorder; SD, standard deviation; TOSCA, The Treatment of Serious Childhood Aggression Study.

respective means and probability levels for group-by-time interactions, whereas Part B gives within-group comparisons over time and between-group comparisons from baseline to week 9 and from week 9 to 21. NCBRF D-Total scores for the 103 Extension study participants at various time points are detailed in Table 2A and Figure 2. Baseline (week 0) acute-phase NCBRF D-Total scores did not differ significantly between those who participated in the Extension and those who did not ($p=0.94$; Table 2A). Likewise,

Extension baseline (week 9) NCBRF D-Total scores did not differ significantly between *Basic* participants and *Augmented* participants who entered the Extension ($p=0.80$). The remaining analyses concern only Extension participants.

Compared to the acute-phase baseline, both randomized groups were significantly improved at week 9 (main effect of visit, $p<0.0001$) and week 21 (main effect of visit, $p<0.0001$ for both *Basic* and *Augmented*; Table 2B). Despite this overall improvement from the acute-trial baseline, both groups were slightly worse at week 21 than they had been at week 9 (the end of acute phase/beginning of Extension; Table 2B) (main effect of visit: *Basic*, $p<0.0005$, $d=0.51$; *Augmented*, $p<0.0006$, $d=0.50$).

Although the mean scores in both groups of Extension participants were significantly improved at week 21 from baseline, there was no evidence of differences between *Augmented* and *Basic* groups at week 9 ($p=0.48$) or at week 21 ($p=0.88$; Table 2A). This differs from our acute 9-week study (Aman et al. 2014), which included both clinical responders and nonresponders.

Antisocial Behavior Scale

Results for the ABS variables are listed in Tables 3A and 3B. These showed the same pattern as NCBRF D-Total scores; significant improvements on Proactive and Reactive Aggression Subscales during the acute phase ($p<0.0001$ – 0.009 for both groups; Table 3B) followed by nominal worsening during the Extension. The modest worsening of behavior for the Proactive Aggression subscale was significant in the *Basic* group ($p=0.0004$, $d=0.50$; Table 3B) but missed significance in the *Augmented* group ($p=0.08$, $d=0.29$). However, the visit-by-treatment interactions between groups for the ABS scores during the Extension were not significant (Table 3A).

CASI-4R measures

Changes in the CASI-4R symptom and impairment subscales for those who completed the Extension generally had a similar temporal pattern as the changes in the primary outcome measure. There were mostly improvements during the acute trial without between-group differences in scores. Subsequently, in the Extension, there were generally no differences observed between the *Basic* and *Augmented* groups either at baseline or at the end of study participation. During the Extension, within-participants comparisons showed either no change or a modest deterioration, relative to improvements seen during the acute trial in symptoms/impairment in both treatment groups (data not shown).

CGI-I scores

Similarly, between-group differences in CGI-I scores either at the end of the acute trial ($p=0.78$) or at the end of the Extension ($p=0.30$) were not significant when analyzed using the original ordinal responses (1–7) of the CGI-I. At the end of the Extension, 83% ($n=34$) of *Basic* treatment participants and 89% ($n=42$) of the *Augmented* treatment participants had a CGI-I score of 1 or 2. The difference was not statistically significant ($p=0.23$).

Responder status

As noted above, with the exception of one individual in the *Basic* group who was inadvertently enrolled, all participants in the *Extension* were classified as clinical responders at the conclusion of the acute trial. For those who completed the Extension, at week 21, the majority of participants in both groups remained clinical responders (*Basic*: $n=33$; 80%, *Augmented*, $n=41$; 87%), although

TABLE 2A. COMPARISON OF BASIC VERSUS AUGMENTED ON NISONGER CHILD BEHAVIOR RATING FORM: TYPICAL IQ VERSION FOR THE 103 IN THE EXTENSION

NCBRF measure	BL	Week 3	Week 9	Week 21	p-Values		
					Trt by visit interaction	Week 9 ^a	Week 21 ^b
Basic, <i>n</i>	49	49	49	41			
Augmented, <i>n</i>	54	54	54	47			
D-Total							
Basic	43.0 (9.9)	22.3 (14.4)	11.2 ^c (9.8)	15.1 (11.5)	0.3143 ^d	0.4827 ^d	0.8780 ^d
Augmented	42.6 (10.9)	26.1 (15.0)	9.8 ^c (8.4)	14.4 (12.4)			
Positive social							
Basic	9.3 (4.6)	14.4 (4.4)	16.9 ^e (6.7)	16.9 (6.8)	0.2772	0.0787	0.0608
Augmented	9.6 (3.8)	13.8 (5.2)	19.0 ^e (6.9)	19.3 (6.4)			
Overly sensitive							
Basic	6.6 (3.5)	4.5 (3.3)	1.8 (2.5)	2.4 (2.5)	0.4056	0.6727	0.4902
Augmented	6.0 (3.2)	3.6 (2.8)	1.7 (2.2)	2.0 (2.3)			
ADHD							
Basic	26.4 (5.4)	13.7 (9.1)	7.3 (7.1)	9.0 (7.0)	0.5921	0.4682	0.5536
Augmented	25.2 (5.8)	13.9 (7.4)	6.5 (5.6)	7.8 (6.1)			
W/D-dysphoric							
Basic	13.7 (7.9)	7.4 (5.6)	3.0 (3.6)	4.2 (4.3)	0.4616 ^d	0.6834 ^d	0.8600 ^d
Augmented	13.6 (8.2)	6.9 (4.9)	3.5 (4.6)	4.3 (5.9)			

Unlike the acute trial (Aman et al. 2014), superiority of Augmented versus Basic is no longer present because all nonresponders have been removed from the comparisons. Observed cases only were analyzed.

^aBetween-group differences at week 9.

^bBetween-group differences at week 21.

^cFor all subjects, including nonresponders, week 9 scores in the acute study (Aman et al. 2014) were 17.8 (Basic) and 10.7 (Augmented).

^dSquare root transformation.

^eFor all subjects, including nonresponders, week 9 scores in the acute study (Aman et al. 2014) were 15.5 (Basic) and 18.7 (Augmented). ADHD, attention-deficit/hyperactivity disorder; BL, baseline; NCBRF, Nisonger Child Behavior Rating Form; W/D, withdrawn.

the rate did not differ significantly between groups ($p=0.56$; two-tailed Fisher’s exact test).

As previously mentioned, one patient in the *Basic* group who did not meet acute-trial responder criteria participated in the Extension in error. However, when this participant withdrew from the study in week 13 due to lack of efficacy, the participant’s NCBRF D-Total had decreased by just over 27%, with a CGI-I of 2. The participant therefore met responder criteria at his/her first (and last) Extension visit.

Secondary analyses: LOCF—initially significant outcomes

The original TOSCA article (Aman et al. 2014) reported statistically significant improvements and moderate advantage of *Augmented* versus *Basic* for the following variables: (1) NCBRF D-

Total, (2) NCBRF Positive Social subscale, and (3) ABS, Reactive Aggression subscale with *Augmented* surpassing *Basic*. The 103 Extension participants, biased by being selected for good response, showed no significant treatment-group differences on those outcomes. To extend those three outcomes to 21 weeks in an intention-to-treat approach, we conducted an LOCF analysis by carrying forward observations through week 21 starting at week 4 (when *Augmented* treatment began) for NCBRF and ABS measures. The goal of this analysis was to attempt to un-bias the Extension analyses by including all participants, regardless of response during the acute phase. Thus, this analysis differs from those above, because it retained nonresponders from the acute trial and children whose families opted out of the Extension ($n=154$). The results from the NCBRF are plotted in Figure 3 and summarized in Table 4. NCBRF D-Total scores were marginally better for *Augmented* than *Basic* ($p=0.06$) and Positive Social scores were

TABLE 2B. COMPARISONS FOR TIME AND GROUP BY TIME ON NISONGER CHILD BEHAVIOR RATING FORM: TYPICAL IQ VERSION FOR THE 103 IN THE EXTENSION

NCBRF measure	Basic		Augmented		Between-group comparison	
	BL to week 9	Week 9–21	BL to week 9	Week 9–21	BL to week 9	Week 9–21
D-Total	<0.0001	0.0006	<0.0001	0.0005	0.5709	0.6029
Positive social	<0.0001	0.6848	<0.0001	0.8064	0.2207	0.8807
Overly sensitive	<0.0001	0.0328	<0.0001	0.0692	0.5877	0.7369
ADHD	<0.0001	0.0273	<0.0001	0.0346	0.8225	0.9270
W/D-dysphoric	<0.0001	0.0461	<0.0001	0.1896	0.6145	0.5605

ADHD, attention-deficit/hyperactivity disorder; BL, baseline; NCBRF, Nisonger Child Behavior Rating Form; W/D, withdrawn.

TABLE 3A. COMPARISONS OF BASIC VERSUS AUGMENTED ON ANTI SOCIAL BEHAVIOR SCALE FOR THE 103 IN THE EXTENSION

ABS measure	BL	Week 9	Week 21	p-Values		
				Trt by visit interaction	Week 9 ^a	Week 21 ^b
Parent						
Basic, <i>n</i>	49	49	41			
Augmented, <i>n</i>	54	54	47			
Proactive						
Basic	20.0 (4.6)	13.8 (3.4)	15.0 (4.0)	0.4437	0.6517	0.2064
Augmented	19.7 (4.6)	13.6 (3.4)	14.1 (3.7)			
Reactive						
Basic	15.8 (1.9)	11.2 (2.8)	11.8 (3.3)	0.7581	0.6111	0.3165
Augmented	15.4 (2.6)	10.6 (2.6)	11.4 (2.8)			
Teacher						
Basic, <i>n</i>	28	26	19			
Augmented, <i>n</i>	24	24	17			
Proactive						
Basic	15.3 (4.1)	12.9 (3.1)	12.1 (3.2)	0.1555	0.2264	0.3793
Augmented	16.8 (5.9)	11.6 (2.8)	12.0 (2.2)			
Reactive						
Basic	11.3 (3.6)	8.3 (2.9)	8.3 (2.8)	0.1095	0.6095	0.6994
Augmented	12.4 (4.3)	7.9 (2.8)	8.9 (3.4)			

^aBetween-group differences at week 9.

^bBetween-group differences at week 21.

Observed cases only were analyzed.

ABS, Antisocial Behavior Scale; BL, baseline.

significantly better ($p=0.005$). Reactive aggression scores on the ABS also differed between *Basic* and *Augmented* ($p=0.03$). Thus, there was separation of *Basic* and *Augmented* participants on NCBRF Positive Social and ABS Reactive scores and marginal separation on D-Total when nonresponder data were reintroduced into the analyses.

Secondary analysis: attrition from acute trial to endpoint of extension and CGI-I status

Another relevant issue concerns the clinical status at Extension endpoint relative to the time when drug augmentation began. We examined attrition from end of week 3 of the acute trial [when 154 were left; see Fig. 1 from Aman et al. (2014)] to the end of the Extension (Fig. 1). Three *Basic* participants and 11 *Augmented* participants dropped out before week 3. As the administered treatments were identical up to that point, those participants were not included in further analyses (resulting sample sizes were $n=81$ in *Basic* and $n=73$ in *Augmented*).

Dropouts before week 21. Forty of 81 *Basic* participants dropped out before completion of the Extension, and 26 of 73 within *Augmented* dropped out ($\chi^2=2.97$, $p=0.085$).

CGI-I = 1 or 2 classification at endpoint of extension. Within *Basic*, 47 children (58%) dropped out before week 21 or were rated with CGI-I ≥ 3 , leaving 34 (42%) at week 21 with CGI-I scores of 1 or 2. Within *Augmented*, 31 (42%) dropped out before week 21 or were rated ≥ 3 at week 21, leaving 42 (58%) at week 21 rated as 1 or 2 on the CGI [$\chi^2(1)=3.72$; $p=0.054$].

Adverse events

Serious AEs did not occur during the Extension. There were no deaths and no suicidal ideation. However, two participants, both in the *Augmented* group, had side effects that led to their discontinuation from the study. One had excessive weight gain, the other a 7-point increase in AIMS score at week 13. The participant described this as “arm flailing”; although this resulted in a higher AIMS score, the clinical team did not regard this as a manifestation of dyskinesia or other nonvoluntary movements. There were no other reports of dyskinesia. Table 5 summarizes AEs that occurred in two or more participants in either treatment group at each postbaseline visit of the Extension.

Two participants in the *Augmented* group, including the one who discontinued due to “arm flailing,” had a 2-point increase in the BAS from week 9 baseline to the end of their study participation at

TABLE 3B. COMPARISONS FOR TIME AND GROUP BY TIME ON ANTI SOCIAL BEHAVIOR SCALE FOR THE 103 IN THE EXTENSION

ABS	<i>Basic</i>		<i>Augmented</i>		<i>Between-group comparison</i>	
	BL to week 9	Week 9–21	BL to week 9	Week 9–21	BL to week 9	Week 9–21
Parent–Proactive	<0.0001	0.0004	<0.0001	0.0830	0.9947	0.2175
Parent–Reactive	<0.0001	0.0591	<0.0001	0.0035	0.5879	0.5276
Teacher–Proactive	0.0088	0.6360	<0.0001	0.8408	0.0546	0.7990
Teacher–Reactive	<0.0001	0.9107	<0.0001	0.8252	0.0458	0.9527

ABS, Antisocial Behavior Scale; BL, baseline.

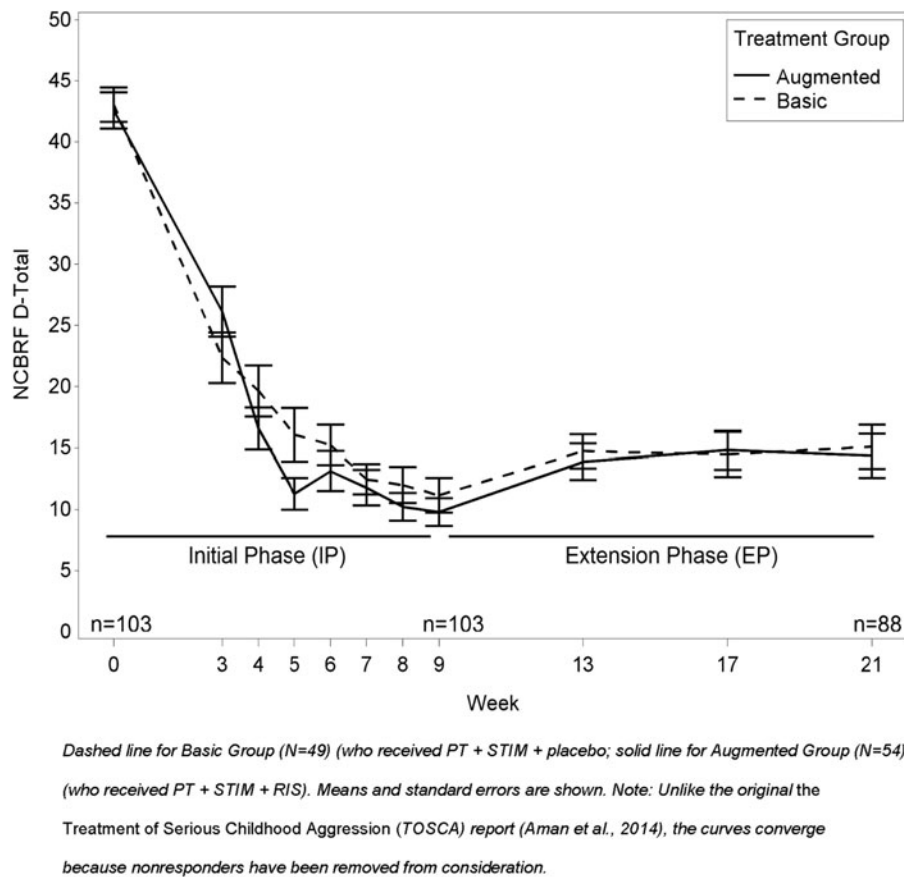


FIG. 2. Nisonger Child Behavior Rating Form Disruptive Behavior Total (D-Total) by Treatment Group (Augmented vs. Basic) over time for those in Extension.

weeks 13 and 21. Of note, akathisia was not reported as an AE during the Extension.

On the SAS, seven participants (six from the *Augmented* group and one from the *Basic*) had a 2-point or more increase from Extension baseline to endpoint. Of these, three had a 2-point increase, three had a 3-point increase, and one had a 4-point increase. In the *Augmented* group, one participant had a dystonic reaction at week 13.

The weight of participants in the *Basic* group increased from a mean of 31.5 ± 9.2 kg to 31.6 ± 9.4 kg during the Extension ($p = 0.58$). In the *Augmented* group, mean weight increased from 37.7 ± 14.9 kg to 37.9 ± 14.9 kg during the Extension (model-adjusted week 9 weight = 37.7 kg; week 21 weight = 39.8 kg; $p = 0.0001$). Both the *Basic* and *Augmented* groups, on average, neither lost nor gained substantial weight during the Extension, but as noted above, one participant in *Augmented* reported weight gain as an AE and discontinued study participation. No participants reported weight loss as an AE.

No participants had systolic or diastolic blood pressure measurements that were considered AEs. During the Extension, one participant had a “tachycardia” event reported at week 13 and another at week 17; their heart rates at the visit when they reported the event were 82 and 98 bpm, respectively; both patients were from *Basic*.

Two participants, both in the *Basic* group, were found to have an ectopic atrial rhythm on their electrocardiograms—one at week 13 and the other at week 17.

We also examined the reasons for study participant discontinuation that occurred after the initiation of double-blind treatment

(after week 3 of the acute study). There were 154 patients who participated in TOSCA beyond week 3 of the acute trial. Sixty-five participants (42%) discontinued for a variety of reasons. Only five participants over both studies (Aman et al. 2014 and this study) withdrew due to intolerable side effects; three (3.7%) from the *Basic* group and two (2.7%) from the *Augmented* group. Other reasons for discontinuance are detailed below. Participants who withdrew from the study as nonresponders by week 9 made up a large percentage of each group. Participants who withdrew due to “other” reasons reported logistical challenges to remaining in the study, including illness of a parent, moving to another state, or the study requiring too many family resources.

Laboratory tests

One laboratory report met the criteria for an AE. A *Basic* participant had thrombocytopenia (platelet count of 126 K) at week 21. At week 21, other potentially clinically meaningful laboratory results included one participant from *Augmented* with an LDL of 136 mg/dl and another from *Augmented* with a urine protein of 30 mg/dL. No other clinically significant laboratory findings occurred during the Extension.

During the Extension, the mean fasting glucose for patients in the *Basic* group decreased from 81.4 ± 12.3 mg/dL to 81.0 ± 11.5 mg/dL. The mean fasting glucose for the *Augmented* group increased from 82.8 ± 8.2 mg/dL to 83.1 ± 9.1 mg/dL. Surprisingly, mean cholesterol, triglycerides, and LDL concentrations decreased modestly for *Augmented*.

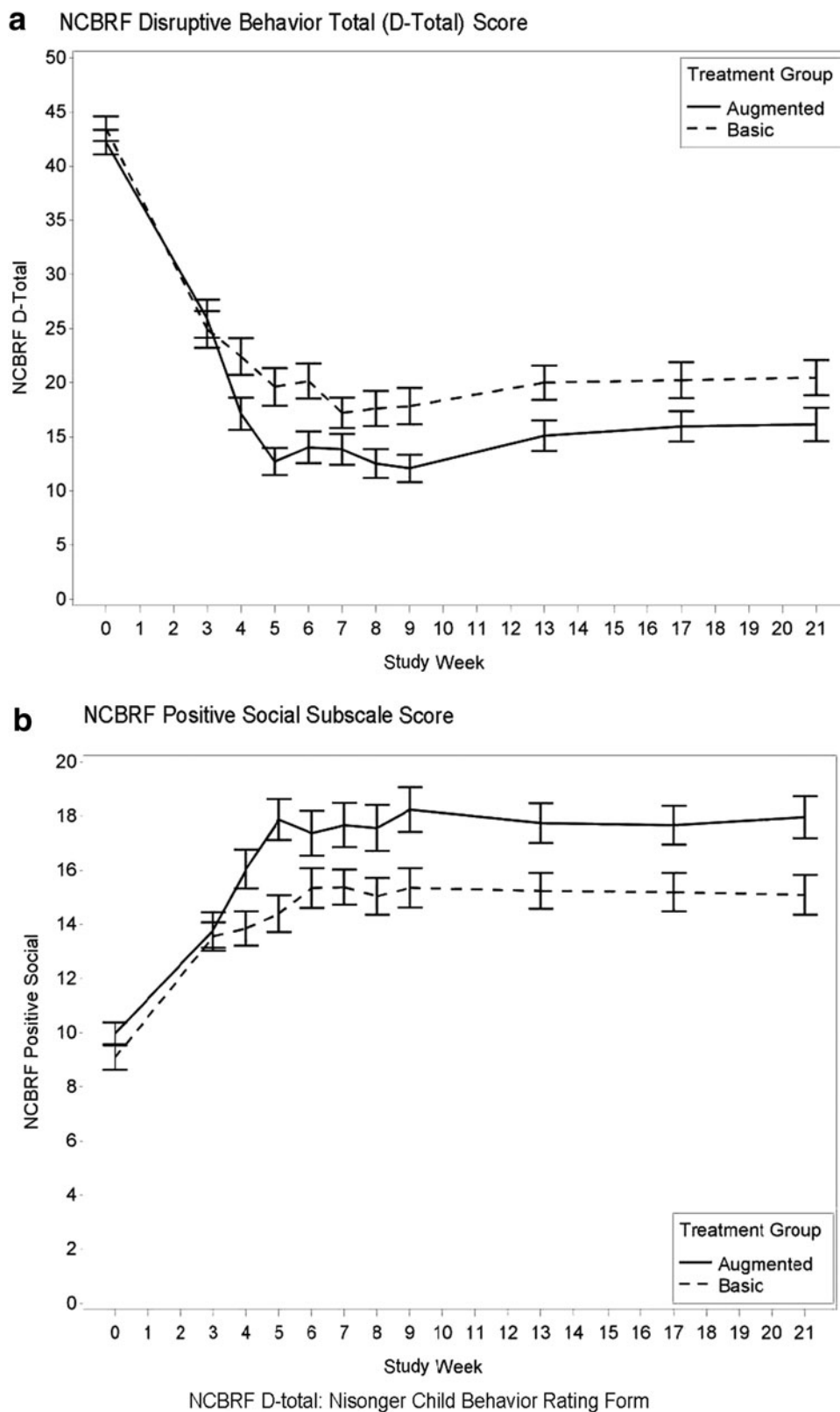


FIG. 3. LOCF at week 4 or later for all randomized subjects present at week 3 ($n=81$ for *Basic* and 73 for *Augmented*). **(a)** NCBRF Disruptive Behavior Total (D-Total) Score as a function of treatment group (*Augmented* vs. *Basic*) over time. **(b)** NCBRF Positive Social Subscale Score as a function of treatment group (*Augmented* vs. *Basic*) over time. LOCF, last-observation-carried forward. Note: means and standard errors are shown.

TABLE 4. LAST-OBSERVATION-CARRIED-FORWARD COMPARISON OF BASIC VERSUS AUGMENTED ON NISONGER CHILD BEHAVIOR RATING FORM: TYPICAL IQ VERSION AND ANTISOCIAL BEHAVIOR SCALE

LOCF measures	BL	Week 3	Week 9	Week 21	p-Values		
					Trt by visit interaction	Week 9 ^a	Week 21 ^b
NCBRF							
Basic, <i>n</i>	84	82	80	80			
Augmented, <i>n</i>	84	75	73	73			
D-Total							
Basic	43.5 (10.3)	24.9 (15.3)	17.8 (15.0)	20.5 (14.6)	0.0487	0.0102 ^c	0.058 ^c
Augmented	42.2 (10.3)	25.9 (15.2)	12.1 (10.9)	16.0 (13.1)			
Positive social							
Basic	9.1 (4.2)	13.6 (4.7)	15.4 (6.6)	15.1 (6.6)	0.0961	0.0058	0.005
Augmented	10.0 (3.7)	13.8 (5.6)	18.2 (7.0)	18.0 (6.7)			
ABS							
Basic, <i>n</i>	84		75	75			
Augmented, <i>n</i>	84		71	71			
Reactive							
Basic	15.9 (1.8)		12.3 (3.1)	12.7 (3.2)	0.1887	0.0051	0.0314
Augmented	15.5 (2.4)		10.9 (2.7)	11.7 (2.7)			

^aBetween-group differences at week 9.

^bBetween-group differences at week 21.

^cSquare root transformation.

ABS, Antisocial Behavior Scale; BL, baseline; LOCF, last-observation-carried-forward; NCBRF, Nisonger Child Behavior Rating Form.

The mean prolactin concentration decreased from 33.4 ± 23.5 ng/mL (week 9) to 26.9 ± 23.0 ng/mL (week 21) in the *Augmented* group (*p* = 0.0054). For *Basic*, the respective mean values were 7.1 ± 8.1 to 7.9 ± 12.9 ng/mL. Prolactin concentrations higher than 18.0 ng/mL for boys and 30.0 ng/mL for girls designated a threshold for elevated prolactin. At week 21, prolactin elevation was significantly more common in *Augmented* (46%) than in *Basic* (3%). The number of patients with a prolactin concentration above 50 ng/mL at week 21 was eight (17%) in *Augmented* and one (3%) in *Basic*. No partici-

pants had a prolactin concentration >100 ng/mL, gynecomastia, or breast enlargement during the Extension.

Discussion

Summary of extension-only observations

Our findings are somewhat complex, so we summarize them here for clarity. To recap the behavioral changes that occurred only during the Extension for these clinical responders, no between-group

TABLE 5. ADVERSE EVENTS IN TWO OR MORE PARTICIPANTS IN ANY SINGLE TREATMENT GROUP AT ANY STUDY VISIT (OBSERVED CASES)

	Week 13		Week 17		Week 21	
	Basic n=49 (%)	Augmented n=54 (%)	Basic n=42 (%)	Augmented n=50 (%)	Basic n=41 (%)	Augmented n=47 (%)
Headache	5 (10)	6 (11)				
Difficulty initiating sleep	7 (14)	2 (4)	4 (10)	2 (4)	2 (5)	
Cough	4 (8)	2 (4)	3 (7)	3 (6)	2 (5)	2 (4)
Appetite increase		6 (11)	2 (5)			
Anxiety	2 (4)	3 (6)				
Enuresis		3 (6)	2 (5)	2 (4)		
Musculoskeletal injury	2 (4)					
Behavioral, NOS	2 (4)					
Rhinorrhea		3 (6)				
Sedation		3 (6)	4 (10)		2 (5)	
Appetite decrease			3 (7)			
Constipation	2 (4)					
Fever	2 (4)			3 (6)		
Nasal congestion		2 (4)				
Otitis media	2 (4)					
Pharyngitis (strep)	2 (4)					
Oral/dental injury			2 (5)			
Bronchopulmonary congestion				2 (4)		

NOS, not otherwise specified.

comparisons showed significant deterioration favoring *Basic* or *Augmented* during the Extension. Regardless of how the participants improved by end of the acute trial, they stayed improved, at similar percentages, in the Extension. The marginal Positive Social subscale difference between treatments approached significance ($p=0.06$) favoring *Augmented*, which is of interest as this variable was one of two significant NCBRF variables in the acute trial (the other being D-Total). Statistically significant deterioration was seen over time within *Basic* on the NCBRF D-Total, Overly Sensitive, ADHD, and Withdrawn/Dysphoric subscales; within *Augmented*, deterioration occurred for D-Total and ADHD. Nevertheless, on a group basis, these changes were modest and of little clinical importance. Group-by-time analyses failed to show any significant changes (Table 2B) indicating that the minor changes associated with time did not indicate differences between treatment groups.

For the ABS (Table 3B), parent ratings of the *Basic* group indicated mild within-participant worsening in Proactive Aggression from week 9 to 21, whereas within *Augmented* treatment, no significant worsening occurred (Table 3B). Once again, the mean changes were minor and all group-by-time analyses showed no significant differences.

All in all, these outcomes are in line with our hypotheses. Essentially, the findings indicate that children who were responders at week 9 generally continued to be responders over the ensuing 12 weeks. The slight clinical slippage (as reflected by increases in mean scores) could be explained as possible regression to the mean with time, because entry to the Extension required selection for substantial improvement. To place the changes over the Extension in more clinical terms, where deterioration occurred, it is possible that dosing needed to be increased slightly (whereas doses were basically stable throughout), or that some tolerance to treatment developed, or that the booster parent management sessions were too infrequent or poorly attended to avoid any clinical deterioration. Nevertheless, clinical improvement was mostly stable for both groups.

These positive results for clinical responders out to week 21 were substantially better than our follow-up results derived from week 52 assessments (i.e., one year after the baseline) (Gadow et al. 2016). In that follow-up study, two-thirds of participants were rated as symptomatic (i.e., CGI-S scores ≥ 3). By parent rating, 45% of children were impaired from ADHD, noncompliant, or aggressive behavior. Important differences between the circumstances leading to week 21 and the week 52 follow-up are that all treatments were free to week 21, PT was still available to participants, and treatment remained structured with regular monthly contact with the clinical teams. Following week 21, medication was no longer provided free of cost, and services had to be obtained through the usual community resources. In fact, 42% of *Augmented* patients stopped one or both drugs by 52 weeks (although 77% of them were receiving psychotropic medication at follow-up). Clearly, both *Basic* and *Augmented* treatments were largely successful for children who responded in the short term, and suggest that there is a need, somehow, to continue these essential services to such severely afflicted children and families. This difference between well-monitored study-directed treatment and subsequent treatment found in the community is consistent with findings of the Multimodal Treatment Study of ADHD (the MTA). Although the MTA sample was not selected for severe aggression as this one was, 54% did have a comorbid DBD (MTA Cooperative Group 1999a). In the MTA, the treatment groups assigned study-supplied medication fared significantly better for 14 months than the group who received treatment as usual in the community (MTA Cooperative Group 1999b; Arnold et al. 2004). At the 24-month follow-up, the ran-

domly assigned MTA treatment groups still differed significantly in both symptoms and medication use, with the medication difference mediating the symptom difference (MTA Cooperative Group 2004). However, symptoms converged across treatment groups by 36-month follow-up (Jensen et al. 2007), and medication use declined significantly by 8 years (Molina et al. 2009).

Findings from acute trial to end of extension

Taking a different perspective, we also examined changes in functioning from end of week 3 (when RIS was introduced for *Augmented*) to week 21. Hence, the LOCF analyses mapped attrition for 154 participants, including nonresponders and those declining participation in the Extension, from week 3 of the acute trial to endpoint of the Extension. For these comparisons, we only analyzed the three variables that were significant in the first report of the acute study (Aman et al. 2014). Interestingly, these showed marginally better NCBRF scores for disruptive behavior (D-Total, $p=0.058$, $d=0.29$), significantly better Positive Social behavior ($p=0.005$, $d=0.44$), and better Reactive Aggression scores ($p=0.03$, $d=0.36$) for *Augmented* treatment. Figure 3 and Table 4 graphically demonstrate what “signal” was lost when clinical nonresponders were removed from analyses. We do not know what clinical significance to attach to enhanced social behavior in the context of DBDs, although logically it would seem to be beneficial. The interplay of positive social behavior and disruptive behavior appears to be a potentially productive although challenging avenue for future research.

The attrition analysis also included nonresponders and dropouts from end of week 3 through endpoint of the acute trial, and participants whose parents declined to join the Extension. This was a secondary analysis and it did not reflect large differences between groups and treatments. Nevertheless, the treatment failures were likely important to the families and children affected, so we certainly cannot discount them, secondary or not. If there are future studies of this type, we feel that it would be informative to follow *all* study participants into any treatment extension. The reality in this trial was that all initial responders, regardless of treatment, tended to fare quite well. Any differences appear to reside more among the children who did not respond well acutely to their assigned treatments or dropped out. Although the comparison of *Basic* and *Augmented* group outcomes in the attrition analysis was not statistically significant ($p=0.054$), some clinicians may consider the difference in survival (42% for *Basic*; 58% for *Augmented*) as clinically significant.

We were only able to find one other study that randomized typical-IQ children with DBDs to ongoing antipsychotic treatment. Reyes et al. (2006) assigned 335 clinical responders to continued risperidone ($n=167$) or placebo replacement ($n=159$) for a period of 24 weeks. Of those assigned to placebo, 47% had a symptom recurrence, compared with 28% maintained on risperidone (derived from CONSORT diagram; $p=0.002$). By comparison, in this study, 16% of children assigned to *Basic* and 13% of those assigned to *Augmented* had symptom recurrence. Thus, the outcomes were remarkably good in the current study. A likely explanation for the difference in results relates to differences in study design. In the report by Reyes et al., all patients initially received the same open-label treatment (i.e., risperidone), and only responders were subsequently randomized on initiating ongoing treatment. In our study, there was no randomization on entering the ongoing treatment phase—patients entered the maintenance phase (Extension) on two different treatment regimens and were continued on the same blinded pharmacotherapy. Other possible explanations for between-study differences include (1) the shorter Extension period (12 weeks here;

24 weeks in Reyes et al.), (2) medication differences (all participants received at least one medication in TOSCA), (3) the fact that all participating families initially received PT, and (d) that all were offered booster parent management sessions in the Extension. Hence, the fact that we followed established responders who continued to receive their assigned drug treatments and whose parents had the potential benefit of PT seemed to auger well for the intermediate-term outcome of clinical responders who initially presented with very severe symptoms.

Adverse events

There were few surprises in terms of AEs. A large proportion of participants were overweight when they enrolled. Therefore, the slight weight loss observed for *Basic* was a net benefit for most children affected, and the maintenance of constant weight for children within *Augmented* implies sustained overweight status for many children in this treatment. Difficulty initiating sleep was common even before the trial commenced, with ~27% of participants (evenly distributed within *Basic* and *Augmented*) experiencing some difficulty during baseline of the acute study. To be recorded as an AE during the study, sleep difficulty either had to abate and then worsen or simply worsen relative to baseline. Difficulty initiating sleep was problematic both in the acute trial and the Extension for several children, especially within *Basic*. This may have been the result of stimulant monotherapy.

By comparison, Findling et al. (2004) reported a 48-week study of open-label risperidone in 107 children with DBDs. Their IQs were below 85, age ranged from 5 to 12 years (mean=9), mean dose was 1.51 mg/day, and NCBRF Conduct Problem subscale cutoff (≥ 24) similar to that in our acute study (Aman et al. 2014). Findling et al. reported the following as the most common AEs: somnolence (33%), headache (33%), rhinitis (28%), and unacceptable weight gain (21%). Treatment was terminated because of weight increase ($n=4$), depression ($n=3$), suicide attempt ($n=2$), dystonia ($n=1$), and emotional lability ($n=1$). On average, it seems that AEs normally attributed to risperidone were lower in the present study. It is possible that some of these AEs (e.g., somnolence, weight gain) may have been partially offset by the stimulants that were begun 3 weeks before the risperidone. This possibility will be one issue addressed in a planned subsequent article to be devoted to AEs alone.

Table 6 summarizes changes in prolactin concentration and weight of participants over the entire 21 weeks—from the acute trial to the end of the Extension.

TABLE 6. LAST-OBSERVATION-CARRIED-FORWARD ANALYSIS

Prolactin			Weight		
<i>Basic</i>	Mean	SD	<i>Basic</i>	Mean	SD
Week -1 ^a (screen)	15.33	5.25	Week 0 (BL)	32.43	9.62
Week 9	16.09	7.15	Week 9	31.11	9.16
Week 21	17.02	22.64	Week 21	31.30	9.32
<i>Augmented</i>			<i>Augmented</i>		
Week -1 ^a (screen)	15.27	5.09	Week 0 (BL)	36.79	14.58
Week 9	39.09	21.77	Week 9	37.72	14.92
Week 21	34.99	21.54	Week 21	39.40	15.35

Mean and SD for participant prolactin concentration and weight by study visit.

^aWeek -1 refers to the screen that preceded study BL.

BL, baseline; SD, standard deviation.

Aman et al. (2004) assessed AEs and behavioral effects in 155 children who took part in two 6-week randomized clinical trials (RCTs) of risperidone and placebo. About half of the participants were previously prescribed psychostimulants, and these were continued unchanged during the RCT. Statistical analyses compared the effects of previously prescribed stimulants (or not) with and without concomitant risperidone. The presence of stimulant did not have any significant effect on the appearance of new AEs, including appetite increase, somnolence, and insomnia. Thus, the presence of previously prescribed stimulants did not appear to alter the emergence of AEs in the study by Aman et al.

In summary, both study treatments were generally well-tolerated. Rates of reported AEs decreased over time. Only one patient discontinued due to AEs. Changes in weight, laboratory, and other physical parameters (including extrapyramidal side effects) appeared clinically unremarkable.

Limitations

Some limitations of the study include the following. First, the sample size at 21 weeks was modest. Less than two-thirds of those in the acute trial participated in the Extension. The study was not powered definitively to test similarities or differences between *Basic* and *Augmented* interventions out to 21 weeks of extended treatment. Also, the LOCF and attrition comparisons were exploratory, rather than primary questions. They were intended to clarify how the acute advantage from *Augmented* treatment, clearly helpful in the acute study, projected forward to week 21 status. Excluded from the LOCF analyses were the 3 *Basic* and 11 *Augmented* participants who dropped out before week 3. The main purposes of the Extension were to check durability of good responses and accumulation of side effects. We did not intend to follow treatment failures from the acute trial or the responders who declined to join the Extension. Ethically, it would have been difficult to justify retaining the treatment failures on their suboptimal treatments. We did not correct for multiplicity of analyses, and therefore, some of the significant outcomes reported here may well have occurred by chance. We do know from our demographic comparisons that mothers of the Extension participants had higher levels of education ($p=0.02$) than responders who declined. That being said, it would be helpful to know more about the clinical and environmental circumstances of the children whom we could not follow. In a sense then, the representativeness of the Extension sample is likely diminished and slightly slanted more toward educationally advantaged families, which would lead to improved outcomes. Nevertheless, the sample in this and other TOSCA reports is informative because of the overall severity of aggressiveness, careful evaluation, exposure to common treatment combinations, and geographic dispersion across multiple sites.

Conclusion

To summarize, our comparison of clinical responders showed no group differences from week 9 to 21 on numerous clinical outcome measures. However, exploratory analyses indicated that when we included nonresponders (acute study) from week 3 forward and responders who were nonparticipants in the Extension, there was evidence suggesting that modestly better outcomes occurred in those assigned to *Augmented*. To some extent, this further supports several outcomes observed in our acute study; much of the difference between the two treatments seems to have occurred in terms of different response rates in the acute trial and possibly differential attrition during the first 3 weeks. Irrespective of these observations,

children who initially received and responded to STIM+PT (i.e., both groups) generally did quite well over time.

Clinical Implications and Clinical Significance

Practitioners treating children with severe physical aggression and ADHD can hope for reasonably good success in the short and medium term once these children have shown a positive response to PT+STIM and/or PT+STIM+RIS. In general, there may be modest deterioration over a period of weeks to a few months, providing they actually receive the kind of care described herein. The presence of stimulant, past exposure to PT, and continued availability of PT appear to auger for a relatively good medium-term outcome. When analyzed from the point of adding risperidone (i.e., end of week 3 forward), children receiving *Augmented* (PT+STIM+RIS) displayed less reactive aggression and more positive social behavior. There was a suggestion that fewer children receiving *Augmented* (PT+STIM+RIS) relapsed from week 0 through 21 than with Basic treatment.

Acknowledgments

The authors acknowledge Aditi Kantipuli, BS, MS, for her thoughtful contributions to an earlier iteration of this article and also thank Melissa Miller for technical assistance on this article.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Centers for Advancing Translational Sciences or the NIH.

Disclosures

Dr. Findling receives or has received research support, acted as a consultant, and/or served on a speaker's bureau for Alcobra, American Academy of Child & Adolescent Psychiatry, American Physician Institute, American Psychiatric Press, Bracket, CogCubed, Cognition Group, Coronado Biosciences, Dana Foundation, Elsevier, Forest, Guilford Press, Ironshore, Johns Hopkins University Press, Jubilant Clinsys, KemPharm, Lundbeck, Merck, NIH, Neurim, Novartis, Otsuka, Oxford University Press, Pfizer, Physicians Postgraduate Press, Purdue, Rhodes Pharmaceuticals, Roche, Sage, Shire, Sunovion, Supernus Pharmaceuticals, Transept Pharmaceuticals, Tris, Validus, and WebMD. Dr. Arnold has received research funding from Curemark, Forest, Lilly, Neuropharm, Novartis, Noven, Shire, and Young Living (as well as NIH and Autism Speaks) and has consulted with or been on advisory boards for Arbor, Gowlings, Neuropharm, Novartis, Noven, Organon, Otsuka, Pfizer, Roche, Seaside Therapeutics, Sigma Tau, Shire, Tris Pharma, and Waypoint and received travel support from Noven. Dr. Gadow is shareholder in Checkmate Plus, publisher of the *Child and Adolescent Symptom Inventory-4R*. Dr. Kolko has received research funds from the NIMH and SAMHSA. Dr. Butter has received funding from NIMH, MCHB/HRSA, Autism Speaks, and the Simons Foundation and is board member of the RUBI Foundation, a 501(C)3 organization supporting autism treatment research. Dr. Bukstein has acted as a consultant and/or served on a speaker's bureau or received royalties from Cephalon, Forest, Lilly, Novartis, Shire, Johnson and Johnson, and Routledge Press. Dr. Aman has received research contracts, consulted with, served on advisory boards, or done investigator training for BioMarin Pharmaceutical, Bristol-Myers Squibb, Cogstate, Inc., Confluence Pharmaceuticals, Cogstate Clinical Trials, Ltd., Coronado Biosciences, Forest Research, Hoffman

LaRoche, Johnson and Johnson, MedAvante, Inc., Novartis, Pfizer, ProPhase LLC, and Supernus Pharmaceuticals. Lisa Townsend, Nicole Brown, Nora McNamara, Devin Gary, Dana Kaplin, Cristan Farmer, Heidi Kipp, Craig Williams, Robert Rice, Kristin Buchan-Page, and Brooke Molina do not have any disclosures.

References

- Aman MG, Binder C, Turgay A: Risperidone effects in the presence of psychostimulant medicine in children with ADHD and disruptive behavior disorders. *J Child Adolesc Psychopharmacol* 14:243–254, 2004.
- Aman MG, Bukstein OG, Gadow KD, Arnold L, Molina S, McNamara N, Rundberg-Rivera E, Li X, Kipp H, Schneider J, Butter E, Baker J, Sprafkin J, Rice R, Bangalore S, Farmer C, Austin A, Buchan-Page K, Brown N, Hurt E, Grondhuis S, Findling R: What does risperidone add to parent training and stimulant for severe aggression in child attention-deficit/hyperactivity disorder? *J Am Acad Child Adolesc Psychiatry* 53:47–60 e41, 2014.
- Aman M, Leone S, Lecavalier L, Park L, Buican B, Coury D: The Nisonger Child Behavior Rating Form: Typical IQ version. *Int Clin Psychopharmacol* 23:232–242, 2008.
- American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. Washington, DC: American Psychiatric Association; 1994.
- Arnold LE, Chuang S, Davies M, Abikoff HB, Conners CK, Elliott GR, Greenhill LL, Hechtman L, Hinshaw SP, Hoza B, Jensen PS, Kraemer HC, Langworthy-Lam KS, March JS, Newcorn JH, Pelham WE, Severe JB, Swanson JM, Vitiello B, Wells KC, Wigal T: Nine months of multicomponent behavioral treatment for ADHD and effectiveness of MTA fading procedures. *J Abnorm Child Psychol* 32:39–51, 2004.
- Barnes T: A rating scale for drug-induced akathisia. *Br J Psychiatry* 154:672–676, 1989.
- Broidy LM, Nagin DS, Tremblay RE, Bates JE, Brame B, Dodge KA, Fergusson D, Horwood JL, Loeber R, Laird R, Lynam DR, Moffitt TE, Pettit GS, Vitaro F: Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Dev Psychol* 39:222–245, 2003.
- Brown K, Atkins MS, Osborne ML, Milnamow M: A revised teacher rating scale for reactive and proactive aggression. *J Abnorm Child Psychol* 24:473–480, 1996.
- Coccaro EF, Harvey PH, Kupshaw-Lawrence E, Herbert JL, Bernstein DP: Development of neuropharmacologically based behavioral assessments of impulsive aggressive behavior. *J Neuropsychiatry Clin Neurosci* 3:S44–S51, 1991.
- Cunningham CE: COPE: large group, community based, family-centered parent training. In: *Attention Deficit Hyperactivity: A Handbook for Diagnosis and Treatment*, 3rd edition. Edited by Barkley RA. New York, NY: The Guilford Press, 2005, pp. 480–498.
- Cunningham CE, Bremner, Secord M, Harrison R: COPE, The Community Parent Education Program: Large Group Community Based Workshops For Parents of 3 to 18 Year Olds. Hamilton, ON, Canada: COPE Works, 2009.
- Farmer CA, Arnold LE, Bukstein OG, Findling R, Gadow K, Li X, Butter E, Aman M: The treatment of severe child aggression (TOSCA) study: Design challenges. *Child Adolesc Psychiatry Ment Health* 5:36, 2011.
- Findling RL, Aman MG, De Smedt G, Derivan A, Lyons B, the Risperidone Disruptive Behavior Study Group: A long-term open-label study of risperidone in children with severe disruptive behaviors and subaverage IQ. *Am J Psychiatry* 161:677–684, 2004.
- Gadow KD, Arnold LE, Molina S, Findling R, Bukstein O, Brown N, McNamara N, Rundberg-Rivera E, Li X, Kipp H, Schneider J, Farmer C, Baker J, Sprafkin J, Rice R, Bangalore S, Butter E, Buchan-Page K, Hurt E, Austin A, Grondhuis S: Risperidone added to parent training and stimulant medication: Effects on attention-deficit/hyperactivity

- disorder, oppositional defiant disorder, conduct disorder, and peer aggression. *J Am Acad Child Adolesc Psychiatry* 53:948–959, 2014.
- Gadow KD, Brown NV, Arnold LE, Buchan-Page KA, Bukstein OG, Butter E, Farmer CA, Findling RL, Kolko DJ, Molina BS, Rice RR Jr., Schneider J, Aman MG: Severely aggressive children receiving stimulant medication versus stimulant and risperidone: 12-month follow-up of the TOSCA trial. *J Am Acad Child Adolesc Psychiatry* 55:469–478, 2016.
- Gadow KD, Sprafkin J: *Child and Adolescent Symptom Inventory-4R*. Stony Brook, NY: Checkmate Plus, 2005.
- Guy W: Abnormal Involuntary Movement Scale (AIMS). In: *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: National Institute of Mental Health, 1976, pp. 534–537.
- Jensen PS, Arnold LE, Swanson JM, Vitiello B, Abikoff HB, Greenhill LL, Hechtman L, Hinshaw SP, Pelham WE, Wells KC, Conners CK, Elliott GR, Epstein JN, Hoza B, March JS, Molina BS, Newcorn JH, Severe JB, Wigal T, Gibbons RD, Hur K: 3-year follow-up of the NIMH MTA study. *J Am Acad Child Adolesc Psychiatry* 46:989–1002, 2007.
- Molina BS, Hinshaw SP, Swanson JM, Arnold LE, Vitiello B, Jensen PS, Epstein JN, Hoza B, Hechtman L, Abikoff HB, Elliott GR, Greenhill LL, Newcorn JH, Wells KC, Wigal T, Gibbons RD, Hur K, Houck PR; MTA Cooperative Group. The MTA at 8 years: Prospective follow-up of children treated for combined-type ADHD in a multisite study. *J Am Acad Child Adolesc Psychiatry* 48:484–500, 2009.
- MTA Cooperative Group: A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 56:1073–1086, 1999a.
- MTA Cooperative Group: Moderators and mediators of treatment response for children with attention-deficit/hyperactivity disorder: The multimodal treatment study of children with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 56:1088–1096, 1999b.
- MTA Cooperative Group: The NIMH MTA follow-up: 24-month outcomes of treatment strategies for attention-deficit/hyperactivity disorder (ADHD). *Pediatrics* 113:754–761, 2004.
- Reyes M, Buitelaar J, Toren P, Augustyns I, Eerdekens M: A randomized, double-blind, placebo-controlled study of risperidone maintenance treatment in children and adolescents with disruptive behavior disorders. *Am J Psychiatry* 163:402–410, 2006.
- Simpson GM, Angus JW: A rating scale for extrapyramidal side effects. *Acta Psychiatr Scand* 212:11–19, 1970.
- Stattin H, Magnusson D: The role of early aggressive behavior in the frequency, seriousness, and types of later crime. *J Consult Clin Psychol* 57:710–718, 1989.
- Timmermans M, van Lier PA, Koot HM: Which forms of child/adolescent externalizing behaviors account for late adolescent risky sexual behavior and substance use? *J Child Psychol Psychiatry* 49:386–394, 2008.

Address correspondence to:

Robert L. Findling, MD, MBA
Division of Child and Adolescent Psychiatry
1800 Orleans Street
12344-A
Baltimore, MD 21287

E-mail: rfindli1@jhmi.edu