



Gender Effects in the Efficacy of Racemic Amphetamine Sulfate in Children with Attention-Deficit/Hyperactivity Disorder

Ann C. Childress · Jeffrey H. Newcorn · Andrew J. Cutler

Received: October 2, 2018 / Published online: April 10, 2019
© The Author(s) 2019

ABSTRACT

Introduction: A laboratory classroom study in children aged 6–12 years with attention-deficit/hyperactivity disorder (ADHD) found that racemic amphetamine sulfate (RA-AMPH) significantly improved performance versus placebo from 45 min through 10 h post-dose (NCT01986062). A secondary analysis assessed gender as a potential moderator of response to treatment comparing the ADHD Rating Scale-IV (ADHD-RS-IV) and Swanson, Kotkin, Agler, M-Flynn and Pelham (SKAMP) rating scales.

Methods: After 8 weeks of open-label RA-AMPH dose optimization using improvement in ADHD-RS-IV symptoms as a guide, 97 subjects (38 females and 59 males) were randomized to the sequence of 2 weeks of double-blind treatment with the optimized dose of RA-AMPH followed by placebo or vice versa during a laboratory classroom day. Efficacy measures

included the SKAMP and the Permanent Product Measure of Performance (PERMP). The average difference for RA-AMPH versus placebo was estimated using least-square (LS) means. Treatment interaction by gender was analyzed using a cross-sectional fixed-effects model.

Results: ADHD-RS-IV scores were comparable for males and females at study entry and at the end of open-label treatment. During double-blind treatment, LS mean scores significantly improved for both genders versus placebo on the SKAMP scale and the PERMP (average $p < 0.0001$ for all post-dose time points). Beginning at baseline, males had significantly higher (worse) SKAMP scores than females but not worse ADHD-RS-IV or PERMP scores.

Conclusion: Both genders responded well to treatment with RA-AMPH, with comparable onset and duration of effect. The ADHD-RS-IV and SKAMP scales both measure changes in attention and hyperactive-impulsive behavior, but the SKAMP scale also measures associated disruptive behaviors, such as frustration, lying, and interpersonal conflict, that are more characteristic of oppositional and conduct disorders and more prevalent in boys with ADHD. Therefore, the SKAMP may be more sensitive for measuring the range of symptoms of boys with ADHD than the ADHD-RS-IV.

Funding: Arbor Pharmaceuticals, LLC.

Enhanced digital features To view enhanced digital features for this article go to <https://doi.org/10.6084/m9.figshare.7880102>.

A. C. Childress (✉)
Center for Psychiatry and Behavioral Medicine, Las Vegas, NV, USA
e-mail: drann87@aol.com

J. H. Newcorn
Mount Sinai Medical Center, New York, NY, USA

A. J. Cutler
Meridien Research, Bradenton, FL, USA

Keywords: Attention-deficit/hyperactivity disorder; Gender; Laboratory classroom study; Racemic amphetamine; Rating scales

INTRODUCTION

First-line medication therapy for children diagnosed with attention-deficit/hyperactive disorder (ADHD) is with one of the two classes of psychostimulant: methylphenidate or amphetamine [1, 2]. Although the use of stimulant medication leads to clinically significant reduction of ADHD symptoms in approximately 70% of cases, not all children respond to any particular stimulant medication [1]. Factors that contribute to an inadequate response include poor adherence, the severity and/or complexity of the ADHD, the adequacy of dosing, and/or the occurrence of dose-limiting adverse events. It is also recognized that some children respond preferentially to particular types of stimulant medication [3]. Guidelines of the American Academy of Child and Adolescent Psychiatry and the American Academy of Pediatrics have therefore recommended switching to another stimulant medication, particularly of a different class, if the expected reduction in ADHD symptoms does not occur or if adverse events are intolerable [1, 2, 4].

Because response to any particular stimulant medication is not certain, understanding possible predictors of clinical response can be useful for clinicians. Gender is a potential moderator of stimulant response in ADHD. Females differ from males in terms of ADHD prevalence, clinical presentation, and symptom profiles [5]. In the classroom, males exhibit more rule breaking and externalizing behaviors [6]. Gender differences in the density of dopamine transporters and clearance of extracellular dopamine [7, 8] have been suggested as possible explanations for why ADHD prevalence rates are higher in males than in females [9]. These factors could also be relevant to medication response.

Varying gender effects have been observed with the application of the ADHD rating scales often used for diagnosis and monitoring response to treatment in clinical trials and

clinical practice, but these effects have rarely been explored or compared on a measure-by-measure basis. Male subjects score higher (greater symptom severity) than female subjects on the commonly used ADHD Rating Scale IV (ADHD-RS-IV) [10] when age and ADHD subtype are considered [11]. Although no consistent pattern of gender effects has been found with the commonly used Swanson, Kotkin, Agler, M-Flynn and Pelham (SKAMP) Rating Scale [12–14], male subjects did have significantly higher SKAMP Combined (SKAMP-C) and SKAMP Department (SKAMP-D) scores (higher scores signifying greater impairment) in a laboratory classroom study of lisdexamfetamine in 6–12 year olds with ADHD [13]. In that same study, however, no gender differences were seen with the Permanent Product Measure of Performance (PERMP), an individualized 400-problem math test used in laboratory classroom studies of ADHD drugs [13, 15].

Amphetamine contains a single chiral center that gives rise to distinct enantiomeric active forms, dextro (D)-amphetamine and levo (L)-amphetamine isomers. The two isomers have different pharmacokinetic properties (mean elimination half-life ranging from 9.77–11 h for D-amphetamine versus 11.5–13.8 h for L-amphetamine) and different neuropharmacological properties [16]. In animal models, D-amphetamine has been found to reduce overactivity and impulsiveness more efficiently, while L-amphetamine appears to be relatively more specific for sustaining attention [17]. Several amphetamine products with different proportions of each isomer have been developed for the treatment of ADHD. While some of these consist entirely of D-amphetamine, one currently available enantiomeric formulation consists of a ratio of 3:1 D-amphetamine to L-amphetamine.

Evekeo (RA-AMPH) (Arbor Pharmaceuticals, Atlanta, GA, USA) is an immediate-release racemic formulation of amphetamine that by definition consists of a 1:1 ratio of the D-amphetamine isomer and the L-amphetamine isomer. The efficacy and safety of RA-AMPH have been evaluated in a multicenter, dose-optimized, double-blind, randomized, placebo-controlled crossover laboratory classroom study

[18]. The current paper reports a preliminary secondary analysis of that study to estimate effect sizes and facilitate power calculation for future studies regarding the possible presence of gender effects. A particular focus of the analysis is the possibility of differences as a function of gender between the findings with the ADHD-RS-IV-RS and SKAMP scales, and the relationship of any such differences to the findings with the laboratory-classroom-specific PERMP test.

METHODS

Study Design

Full details of the methodological design and conduct of the study (NCT01986062) have been previously published [18]. Briefly, a laboratory school protocol [15] was used to measure efficacy, including onset and duration of effect. The study was conducted at 7 sites in the United States, in accordance with Good Practice guidelines, and the study protocol was approved by an appropriate institutional review board at each study site. The parent/guardian of each subject provided voluntary written informed consent for the subject to participate in this study, and the subject gave written assent.

Study inclusion criteria were a severity score of at least 3 (mildly ill) on the clinician-administered 7-point Clinical Global Impressions–Severity (CGI-S) scale and a score on the ADHD-RS-IV scale [19] at screening or baseline \geq 90th percentile normative values for gender and age in at least one of the following categories: hyperactive-impulsive subscale, inattention subscale, or total score. A primary psychiatric diagnosis other than ADHD excluded a subject from participating in the study, as did secondary or comorbid diagnoses with the exception of simple phobias, oppositional defiant disorder (ODD), motor skill disorders, communication disorders, learning disorders, adjustment disorders, and sleep disorders. Subjects could be treatment-naïve or previously treated with an ADHD medication.

The study design consisted of a screening period, an 8-week open-label dose optimization

phase, and a 2-week randomized double-blind crossover phase. Subjects took RA-AMPH twice daily (once in the morning and again 4–6 h later). The starting dose of RA-AMPH was 10 mg per day (5 mg bid), which was titrated weekly in 5-mg increments until an optimal dose based on clinical response and tolerability was achieved. A single down-titration for tolerability reasons was also permitted. To assess global changes in the severity of ADHD over the course of the open-label phase of the study, clinicians used the ADHD-RS-IV, the CGI-S, and the Clinical Global Impressions–Improvement (CGI-I) scale, which is scored from 1 (very much improved) to 7 (very much worse). Safety and tolerability were assessed via weekly monitoring of blood pressure and pulse and the tracking of spontaneously reported adverse events. The Columbia-Suicide Severity Rating Scale was also used to assess changes from baseline in suicidal ideation and behavior.

At the end of the open-label phase, subjects able to tolerate a stable dose of at least 10 mg per day were randomized on a 1:1 basis to a treatment sequence of 1-week RA-AMPH/1-week placebo or 1-week placebo/1-week RA-AMPH. No adjustment of RA-AMPH dose was permitted during the double-blind phase. The subjects continued to take the medication twice daily until the 7th day of each crossover week, when a final single dose was administered at the laboratory school site by study staff on the morning of each full-length laboratory classroom day. On the two laboratory classroom days, only the single morning dose was administered in order to establish duration of effect of the racemic formulation. At the end of the first laboratory classroom day, the subjects were dispensed their double-blind crossover medication for the following week leading up to the second laboratory classroom day. One week following the second laboratory classroom day, a final safety assessment was conducted during a post-withdrawal follow-up phone call or visit to the study site.

Efficacy Measures

The primary objective of the study was to establish that an optimal dose of RA-AMPH would result in a significant reduction in ADHD signs and symptoms compared to placebo at 2 h post-dose using the SKAMP-C scale. The ADHD-RS-IV was used during the open-label dose-optimization period to guide RA-AMPH dose changes. The ADHD-RS has been updated to be consistent with the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), but this newer version (ADHD-RS-5) was not available at the time this study was conducted. The secondary efficacy objective was to determine the onset and duration of clinical effect for RA-AMPH after a single morning dose. The secondary efficacy variables included SKAMP-C scores, SKAMP Attention (SKAMP-A) subscale scores, SKAMP-D subscale scores, PERMP scores, each measured before dosing and at 0.75, 2, 4, 6, 8, and 10 h post-dose on the two full-length laboratory classroom days. For the PERMP, an objective individualized 400-problem mathematics exam, subjects were instructed by site staff to work at their seats and complete as many problems as possible in 10 min.

The safety and tolerability of RA-AMPH compared with placebo were assessed and reported in the primary outcomes paper [18] and will not be repeated here.

Description of the ADHD Scales

The ADHD-RS-IV includes all 18 symptoms described in the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, Text Revision (DSM-IV-TR) [20] and measures the presence and severity of these symptoms. The 18 symptoms assessed by the ADHD-RS-IV are each rated on a 4-point scale; total scores range from 0 to 54, with higher scores indicating greater ADHD symptom severity. Normative data for age and gender have been published for the ADHD-RS-IV. Scores \geq 90th percentile for age and gender are often used for clinical trial entry [10]. The ADHD-RS-IV was originally designed to be completed by a teacher or parent. However, an investigator interview version

of the ADHD-RS-IV was later developed specifically for use in ADHD clinical trials and regulatory documents [21]. In the current trial (as in most clinical trials), the ADHD-RS-IV was completed by an investigator while interviewing a parent or caregiver at each visit. Since parents are not with subjects during school hours, the ratings often describe behaviors that are seen outside of school, but they may also reflect information that teachers have reported to caregivers. Completion of the ADHD-RS-IV is dependent on subjective recall and reporting of symptoms over the previous week, which can be inconsistent and imprecise.

The SKAMP Rating Scale differs from the ADHD-RS-IV in that ratings are based on direct observation in real time. The SKAMP scale is completed by trained raters observing subject behaviors during several 20-min sessions over the course of a day in a simulated classroom. The SKAMP measures core ADHD behaviors as well as associated features such as oppositional symptoms and conduct disturbances (e.g., frustration, lying, and relationships with staff and peers). The SKAMP scale consists of 13 items (grouped under the subcategories of attention, department, quality of work, and compliance) on each of which subjects are rated according to a 7-point scale (0 = normal to 6 = maximal impairment) by trained study personnel [15]. The SKAMP Combined score is obtained by summing the rating values for all of the 13 items. Higher SKAMP scores signify greater impairment.

Multiple behaviors are coded to each SKAMP item (Table 1) [22]. Two of the four items on the Department subscale and one of the two items on the Cooperation subscale do not describe core ADHD symptoms, while all four of the items on the Attention subscale and all three of the items on the Quality of Work (SKAMP-QOW) subscale describe core ADHD symptoms (although the items on the SKAMP-QOW subscale are scored in relation to the number of problems attempted and completed correctly during a programmed academic exercise and neatness of handwriting).

The SKAMP scale was originally completed by teachers after a classroom period and validated using the Conners, Loney and Milich

Table 1 Items in the SKAMP Combined scale and in the SKAMP subscales

SKAMP subscales	SKAMP Combined items	Coded behaviors
Attention	1. Getting started on assignments for classroom periods	Not initiating, not following directions
	2. Sticking with tasks or activities for the allotted time	Poor eye contact, staring, talking, playing (avoids attending to activity or tasks by playing with materials or peers)
	3. Participating in an activity or classroom discussion	Poor eye contact, staring
	4. Stopping and making transition to the next period	Not initiating, not following directions
Department	5. Interacting with other children (e.g., other students)	Mean, aggressive voice/body
		Passive (sulking), whiney voice/body
		Touching without permission (peers or materials)
		Teasing (verbal or nonverbal), provoking peers
	6. Interacting with adults (e.g., teacher or aide)	Aggression (hitting, kicking, biting, spitting, pushing, restraining)
		Not accepting (frustration tolerance)
		Lying (deception)
		Lying (deception)
7. Remaining quiet according to classroom rules	Mean, aggressive voice/body	
	Passive (sulking), whiney voice/body	
	Aggression (hitting, kicking, biting, spitting, pushing, restraining)	
	Not accepting (frustration tolerance)	
8. Staying seated according to classroom rules	Lying (deception)	
	Rudeness to staff	
	Interruption (verbal)	
Quality of work	9. Completing assigned work	Noise making (nonverbal)
	10. Performing work accurately	Calling out to teacher (unsolicited blurting out)
	11. Being careful and neat while writing or drawing	Not sitting properly (fidgeting)
		Out of seat
		Attention
		Careless mistakes
		Messy

Table 1 continued

SKAMP subscales	SKAMP Combined items	Coded behaviors
Cooperation	12. Complying with teacher's usual requests or directions	Refuses to comply with verbal or nonverbal requests of staff, including requests to begin or end an activity
	13. Following the rules established for the classroom	Interruption, not sitting properly, out of seat, disruptive body language, noise making, inappropriate language, calling out, destructive behavior

SKAMP Swanson, Kotkin, Agler, M-Flynn, and Pelham

scale (CLAM) and Overactivity with Aggression (IOWA) Conners Rating Scales, which measure both ADHD and oppositional defiant symptoms [12, 23, 24]. ODD is more common in males than females, with a prevalence ratio estimated at 1.59:1 [25]. Since ODD is present in 40–70% of subjects who have ADHD, oppositional behaviors are common in the laboratory classroom even in patients who do not meet full criteria for ODD.

The 10-min, 400-question PERMP test was developed to be administered throughout a laboratory school day and has become a secondary outcome measure in most laboratory school studies [15]. The appropriate level of math difficulty for each student was determined based on results from an initial administration of the PERMP at the baseline visit at the beginning of the open-label phase. Performance on the PERMP test is evaluated using two scores, indicating the number of problems attempted (PERMP-A) and the number of problems correct (PERMP-C). PERMP scores at each laboratory classroom time point were also translated into a score of 0–6 and were entered as items 9 and 10 under Quality of Work on the SKAMP (Table 1).

Statistical Analyses

Primary and secondary efficacy analyses were performed on the intent-to-treat (ITT) population, consisting of all randomized subjects who received at least one dose of double-blind study medication and had at least one post-baseline assessment of the primary efficacy variable.

Descriptive statistics for the SKAMP-C, SKAMP-A, SKAMP-D, and PERMP problems attempted, and PERMP problems correct, scores were calculated for each time point on the laboratory classroom days. Parameter estimates and effect sizes were generated from a cross-sectional fixed-effects model. The model included terms for treatment sequence (RA-AMPH/placebo and placebo/RA-AMPH), period (class effect: 1 and 2), treatment, gender, and treatment-by-gender interaction. All available data were used in the model, and there was no imputation for missing data. No corrections for multiple testing were applied on time points and subgroup statistical comparisons. Secondary analyses evaluated ADHD-RS-IV, SKAMP, and PERMP subscale values for males and females separately, and differences between them. The treatment comparison was conducted as a 2-sided test at the 5% level of significance. Standard errors (SEs) and 95% confidence intervals (CIs) were calculated. Statistical analyses were performed by Rho (Chapel Hill, NC, USA).

RESULTS

Subgroup Demographics

Of 107 initially enrolled subjects, 10 withdrew prior to randomization at the end of the open-label phase, leaving 97 subjects for the ITT population: 47 subjects randomized to the treatment sequence of RA-AMPH followed by

placebo and 50 randomized to the sequence of placebo followed by RA-AMPH.

Of the 97 subjects in the ITT population, 59 (60.8%) were male and 38 (39.2%) female. The mean (SD) subject age was 9.5 (1.87) years for males and 9.7 (1.88) years for females, each group with a range of 6–12 years. There were no significant differences between the male and female subjects in demographic and baseline characteristics, including ADHD type, except for weight: female subjects were on average heavier than male subjects—93.12 (33.54) pounds versus 79.12 (28.32) pounds ($p = 0.0295$)—a not unexpected difference between preadolescent boys and girls (Table 2).

The mean (SD) daily dose of RA-AMPH was 17.2 (4.55) mg; the mean (SD) and median final doses were 23.4 (8.18) mg and 20.0 mg, respectively. The mean final dose did not differ for males and females (23.5 mg for males and 23.3 mg for females).

Eight-Week Open-Label Dose-Optimization Phase

ADHD-RS-IV Scores by Gender

For the entire ITT population, during the open-label dose-optimization phase, mean (SE) ADHD-RS-IV total scores decreased (improved)

Table 2 Demographic and baseline characteristics by gender

Variables	Male ($n = 59$)	Female ($n = 38$)	p value
Age, years			0.716
Mean \pm SD	9.5 \pm 1.87	9.7 \pm 1.88	
Median (IQR)	10.0 (8, 11)	10.0 (8, 11)	
Range (min, max)	6, 12	6, 12	
Race, n (%)			0.739
White	36 (61.0%)	22 (57.9%)	
Black/African American	20 (33.9%)	13 (34.2%)	
Other	3 (5.1%)	3 (7.9%)	
ADHD type, n (%)			0.117
Inattentive	8 (13.6%)	10 (26.3%)	
Hyperactive/impulsive	0	0	
Combined	51 (86.4%)	28 (73.7%)	
Height, in			0.137
Mean \pm SD	54.5 \pm 5.45	56.2 \pm 5.23	
Median (IQR)	54.0 (51.2, 58.0)	57.5 (52.0, 59.5)	
Range	44.5, 67.0	46.8, 66.0	
Weight, lb			0.029
Mean \pm SD	79.12 \pm 28.32	93.12 \pm 33.54	
Median (IQR)	70.8 (59.0, 97.2)	85.5 (66.0, 121.8)	
Range	39.8, 165.0	50.2, 166.0	

IQR Interquartile range

by 27.8 points from 40.8 (7.64) at baseline to 13.0 (7.32) at the end of the 8-week study period [18]. All subjects were either drug-naïve or washed off of their ADHD medication prior to baseline ratings. Figure 1 shows the ADHD-RS-IV total scores and the inattention and hyperactivity/impulsivity subscale scores by gender during the open-label dose-optimization phase. At baseline, male and female subjects were comparable in ADHD-RS-IV total scores ($p = 0.065$) and inattention subscale scores ($p = 0.523$), but male subjects did have higher hyperactivity/impulsivity subscale scores than female subjects—20.8 (5.16) versus 18.4 (5.91) ($p = 0.038$)—a difference that disappeared at the end of the optimization period. For males and females, the improvement of ADHD symptoms was consistent and comparable at each visit through at least week 7 of the open-label phase. ADHD-RS-IV total scores changed by 25.9 points from 39.0 (7.97) at baseline to 13.1 (6.81) at week 8 for females, and by 29 points from 41.9 (7.25) to 12.9 (7.69) for males. There were no significant differences between the genders

in terms of the optimized week-8 total scores ($p = 0.876$) or in terms of total score improvement at week 8 compared to baseline ($p = 0.097$), but the level of improvement showed a trend favoring boys—a trend related to the original difference between the genders in the hyperactive/impulsivity subscale scores. At week 8 of the open-label dose-optimization phase, the proportion of responders to treatment ($\geq 50\%$ improvement in the ADHD-RS-IV total score) was 86.8% (33/38) for females and 88.1% (52/59) for males. For males and females, improvements on the ADHD-RS-IV hyperactivity/impulsivity subscale and on the ADHD-RS-IV inattention subscale were also consistent and comparable. ADHD-RS-IV hyperactivity/impulsivity scores changed from 18.4 (1.0) at baseline to 5.6 (0.7) at week 8 for females and from 20.8 (0.7) at baseline to 5.9 (0.5) at week 8 for males. On this subscale, there was no significant difference between the genders in terms of optimized week-8 scores ($p = 0.710$), but the level of improvement at week 8 showed a trend favoring boys ($p = 0.069$). ADHD-RS-IV inattention

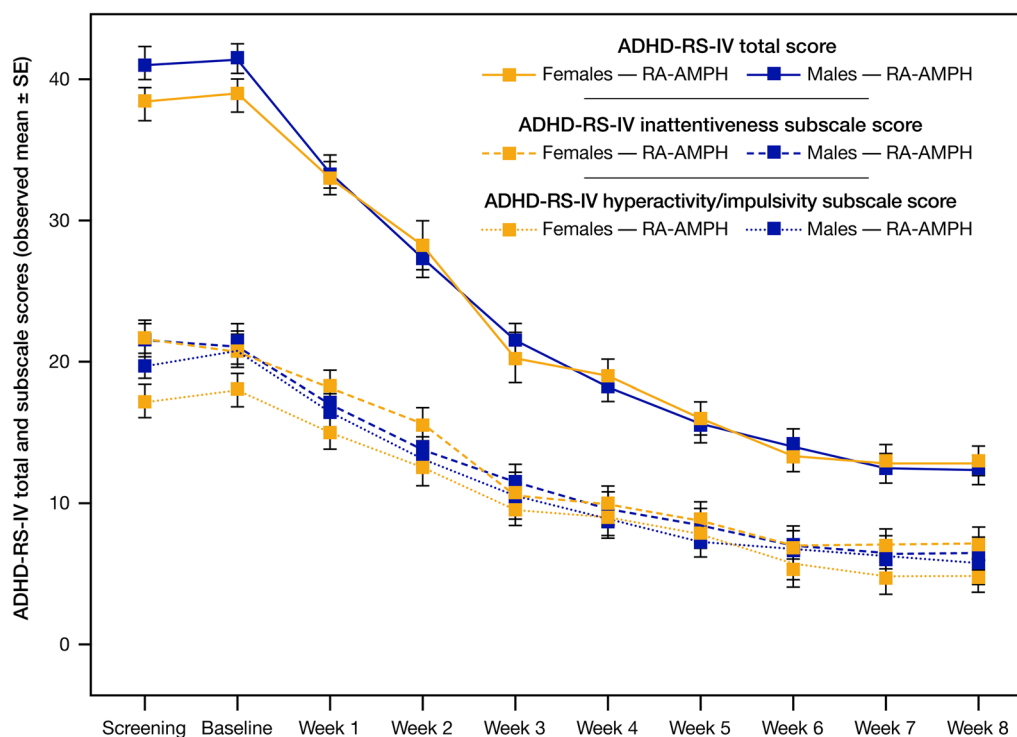


Fig. 1 ADHD-RS-IV total scores, ADHD-RS-IV inattention subscale scores, and ADHD-RS-IV hyperactivity/impulsivity subscale scores by gender during the 8-week open-label dose-optimization phase

scores changed from 20.7 (0.7) at baseline to 7.5 (0.6) at week 8 for females and from 21.2 (0.5) at baseline to 7.1 (0.6) at week 8 for males. On this subscale, there were no significant differences between the genders in terms of optimized week-8 scores ($p = 0.664$) or level of improvement at week 8 ($p = 0.342$).

Two-Week Double-Blind Randomized Crossover Phase

Analysis of Post-dose SKAMP Scores by Gender

On the SKAMP-C score at 2 h post-dose on the 2 full-length laboratory classroom days, RA-AMPH was superior to placebo: LS mean (SE) scores were 10.3 (1.09) for RA-AMPH versus 18.1 (1.09) for placebo ($p < 0.0001$) [18]. When SKAMP-C scores at 2 h post-dose for the ITT population were examined by gender, the treatment effect was greater for the male subjects, with statistically significant differences in LS means favoring RA-AMPH treatment over placebo of -9.9 points for male subjects ($n = 57$, $p < 0.0001$) and -4.8 points for female subjects ($n = 38$, $p = 0.0002$).

For the ITT population, significant separation from placebo was observed for RA-AMPH on the SKAMP-C scale and the SKAMP-D and SKAMP-A subscales at each time point tested (0.75, 2, 4, 6, 8, and 10 h) over the course of the 2 laboratory classroom days, during which each subject received RA-AMPH and placebo [18]. There were no significant or consistent differences in effect size (treatment difference divided by standard deviation) between the genders, and the magnitude of the effect sizes appeared typical for laboratory-classroom studies. Table 3 presents the cross-sectional fixed-effects analysis by treatment, gender, and treatment by gender for predose and post-dose time points for the SKAMP-C, SKAMP-A, and SKAMP-D scales. A significant treatment-by-gender interaction occurred only at the 2-h time point on the SKAMP-C scale ($f = 5.86$, $p = 0.0174$) and on the SKAMP-A subscale ($f = 4.71$, $p = 0.0326$). Subject scores on the SKAMP-C scale and the SKAMP-D and SKAMP-A subscales over the course of the two laboratory classroom days are broken out by gender in Fig. 2. Male subjects

started at the predose time point with higher scores on all three of the SKAMP scales, and these higher scores continued throughout the day. Significant separation from placebo was observed for both genders for RA-AMPH on all three SKAMP scales at each time point tested (average $p < 0.0001$ over all post-dose time points), with the exception of the 10-h time point on the SKAMP-A subscale for female subjects (LS mean treatment effect at that time point was -0.6 points, $p = 0.2233$).

Analysis of Post-Dose PERMP Scores by Gender

For the PERMP-A and PERMP-C tests, there were no significant and consistent differences in effect size between the genders, and the magnitude of change appeared typical for this type of study. In the cross-sectional fixed-effects analysis, there were no significant treatment-by-gender interactions on the PERMP-A and PERMP-C tests (Table 3). Subject scores on the PERMP test over the course of the two laboratory classroom days are broken out by gender in Fig. 3. Significant separation from placebo was observed for both genders for RA-AMPH in terms of both problems attempted and problems correct at each time point tested, with the greatest RA-AMPH effect seen at 2 h post-dose for the male subjects and at 4 h post-dose for the female subjects, although effect levels were similar throughout the laboratory classroom days for both genders. Scores for females receiving placebo were consistently numerically higher than those for males receiving placebo, and scores for females receiving RA-AMPH were consistently higher than those for males receiving RA-AMPH, but no statistically significant differences were seen between males and females at any time point (Table 3).

DISCUSSION

In the first classroom study of 1:1 racemic amphetamine, an optimal dose of RA-AMPH resulted in significant reduction in symptoms of ADHD and related problems in children aged 6–12 years as measured by the SKAMP. Scores on all three SKAMP scales were significantly improved for all children combined from the

Table 3 Cross-sectional fixed-effects analysis by treatment, gender, and treatment by gender for predose and post-dose time points

Time (h)	Mixed model	SKAMP-C		SKAMP-A		SKAMP-D		PERMP-A		PERMP-C	
		f value	p value	f value	p value	f value	p value	f value	p value	f value	p value
0	Treatment	10.14	0.0020	0.77	0.3811	0.10	0.7557	5.68	0.0191	5.13	0.0258
	Gender	7.28	0.0083	2.68	0.1049	6.94	0.0099	0.43	0.5154	0.45	0.5023
	Treatment by gender	0.18	0.6683	0.07	0.7982	0.31	0.5795	0.04	0.8515	0.00	0.9519
0.75	Treatment	28.99	< 0.0001	16.77	< 0.0001	19.67	< 0.0001	24.03	< 0.0001	31.44	< 0.0001
	Gender	7.00	0.0096	2.74	0.1012	7.40	0.0078	0.85	0.3587	1.30	0.2579
	Treatment by gender	0.60	0.4410	0.21	0.6445	2.22	0.1395	0.86	0.3572	0.22	0.6415
2	Treatment	42.78	< 0.0001	28.88	< 0.0001	24.61	< 0.0001	36.57	< 0.0001	46.67	< 0.0001
	Gender	8.37	0.0047	3.00	0.0866	12.17	0.0007	0.95	0.3318	0.88	0.3510
	Treatment by gender	5.86	0.0174	4.71	0.0326	2.65	0.1069	0.68	0.4122	2.89	0.0925
4	Treatment	56.55	< 0.0001	37.93	< 0.0001	35.08	< 0.0001	53.12	< 0.0001	54.51	< 0.0001
	Gender	10.34	0.0018	2.00	0.1604	8.34	0.0048	1.49	0.2253	1.73	0.1911
	Treatment by gender	0.02	0.8978	0.22	0.6364	0.53	0.4690	0.05	0.8314	0.02	0.8906
6	Treatment	33.69	< 0.0001	26.58	< 0.0001	9.44	0.0028	49.86	< 0.0001	45.79	< 0.0001
	Gender	6.71	0.0111	0.83	0.3644	7.22	0.0085	0.79	0.3752	1.23	0.2703
	Treatment by gender	0.77	0.3812	2.38	0.1259	0.02	0.8859	1.36	0.2460	1.41	0.2372
8	Treatment	40.70	< 0.0001	31.80	< 0.0001	18.15	< 0.0001	31.06	< 0.0001	34.63	< 0.0001
	Gender	5.32	0.0233	0.39	0.5344	3.58	0.0614	1.78	0.1856	2.83	0.0960
	Treatment by gender	0.13	0.7219	0.09	0.7605	0.09	0.7632	0.51	0.4762	0.11	0.7402
10	Treatment	13.50	0.0004	4.63	0.0339	14.82	0.0002	27.73	< 0.0001	27.86	< 0.0001
	Gender	1.81	0.1814	0.00	0.9953	3.68	0.0581	0.83	0.3634	1.13	0.2912
	Treatment by gender	0.51	0.4787	0.28	0.5968	1.20	0.2767	0.00	0.9627	0.03	0.8734

The treatment comparison was conducted as a 2-sided test at the 5% level of significance. Significant values are in bold type

PERMP Permanent Product Measure of Performance, PERMP-A problems attempted, PERMP-C problems correct, SKAMP Swanson, Kotkin, Agler, M-Flynn, and Pelham, SKAMP-C SKAMP Combined scale, SKAMP-A SKAMP Attention subscale, SKAMP D SKAMP Dependent subscale

first time point tested (0.75 h post-dose) through the end of the day at 10 h [18].

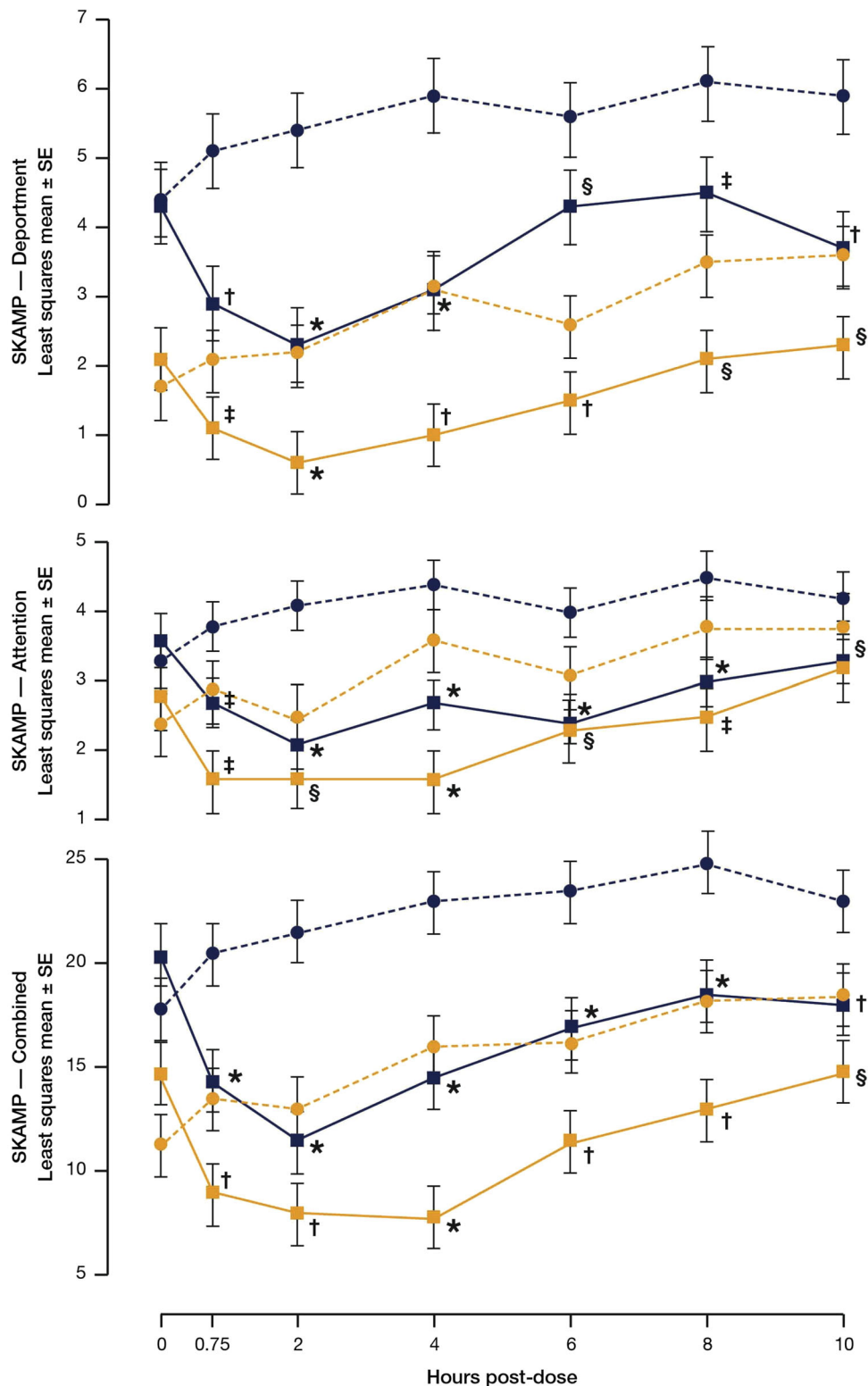
Although the SKAMP-C scores were significantly different from placebo for males and females at all time points except for 10 h post-dose, the treatment-by-gender interaction was not significant for any time point except at 2 h post-dose ($p = 0.0174$). Since both genders entered the study with similar baseline ADHD-RS-IV scores and optimized ADHD-RS-IV total scores were also comparable, one might expect SKAMP ratings to also be similar. However, the SKAMP-C scores were significantly higher (worse) for males than for females at all time points except 10 h post-dose (Table 3). These differences were driven primarily by the SKAMP-D scores, which were significantly different between the genders at all time points except at 8 h ($p = 0.0614$) and at 10 h ($p = 0.0581$) post-dose (Table 3). Since there were no significant gender differences on the SKAMP-A subscale or on PERMP, ADHD-RS-IV subscales were evaluated to determine whether there were gender differences between the inattention and the hyperactivity/impulsivity subscales. There were no significant differences in either subscale by gender at the end of open-label dose optimization.

So why did males have significantly higher (worse) SKAMP scores but not significantly higher (worse) ADHD-RS-IV or significantly lower (worse) PERMP scores than females in the trial? The answer might be related to the following fact, which is not frequently highlighted in studies of ADHD medication effects: namely, in addition to ADHD symptoms, the SKAMP scale also measures associated behaviors such as frustration, lying, and relationships with staff and peers, which are more characteristic of ODD and conduct disorder than of ADHD (Table 1). Two of the four items on the SKAMP Department subscale (items 5 and 6) and one of the two items on the SKAMP Cooperation subscale (item 13) do not describe core ADHD symptoms, while all seven of the items on the SKAMP-A and SKAMP-QOW subscales assess inattention directly or by evaluating functional outcomes (PERMP scores and handwriting neatness).

In the current study, an ODD diagnosis was not an exclusion criterion, except when ODD symptoms were severe enough to cause problems with study participation. In the actual enrolled population, however, only 7 (7.2%) of the randomized subjects met DSM-IV-TR criteria for ODD, a percentage far lower than the range of 39–84% meeting ODD criteria in other studies [26, 27]. Even in the Multimodal Treatment Study of Children with ADHD (MTA), 40% of the subjects with ADHD met ODD criteria [28]. Consequently, the presence of ODD cannot account for the gender difference in the current study. Since stimulant treatment is not only effective for treatment of core ADHD symptoms but also improves behavioral noncompliance and aggression [29], which are characteristic of youth with ODD and high levels of impulsivity, it is not surprising to find a lack of difference between genders in treatment effects, as all SKAMP symptoms tend to improve. However, the reason for improvement is less clear. Is it a direct effect of medication on associated symptoms or secondary to reduced ADHD symptoms? If core ADHD symptoms were only evaluated using the SKAMP, it is possible that the effect sizes of stimulants would be lower.

Another difference between the SKAMP and the ADHD-RS scales is that the SKAMP scale measures behaviors using direct observation by skilled raters in real time in a classroom setting, where subjects interact with peers, while the ADHD-RS scale is based on historical information about attention and hyperactive-impulsive behavior (typically, behavior during the past week in most trials, including this one). Although the ADHD-RS ratings are completed by a trained investigator, symptoms are based primarily on parent/caregiver reports, which are imprecise and most heavily influenced by behavior witnessed at home for only parts of most days. In addition, the SKAMP scale is useful in determining the onset, duration, and peak effects of stimulant medications because it measures real-time effects, which is not possible with the ADHD-RS scale. Thus, it is clear that the two scales are not identical, although both scales are accepted as study outcome measures by the US Food and Drug Administration for drug registration studies in ADHD.

■ Females — RA-AMPH ■ Males — RA-AMPH * $p < 0.0001$ vs placebo † $p < 0.001$ vs placebo
● Females — Placebo ● Males — Placebo ‡ $p < 0.01$ vs placebo § $p < 0.05$ vs placebo



◀ **Fig. 2** Analysis of laboratory-classroom SKAMP Department, SKAMP Attention, and SKAMP Combined scores by treatment and gender. * $p < 0.0001$ vs. placebo. † $p < 0.001$ vs. placebo. ‡ $p < 0.01$ vs. placebo. § $p < 0.05$ vs. placebo

Differences between the SKAMP and ADHD-RS scales are important for practitioners to understand. The ADHD-RS scale can be used to evaluate treatment effects during clinic visits, but even when there is much improvement in ADHD symptoms, patients may still have sig-

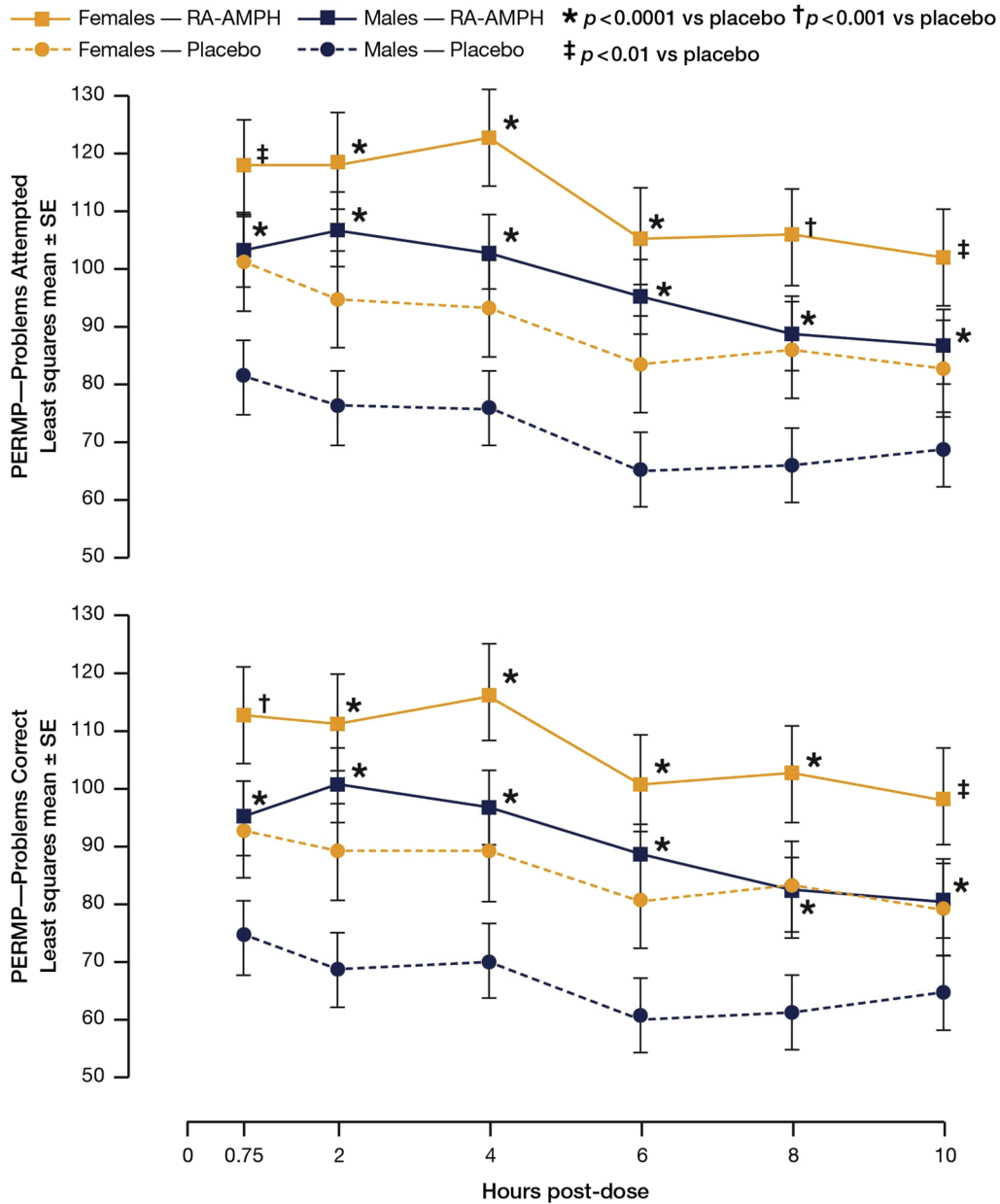


Fig. 3 Analysis of laboratory-classroom PERMP number of problems attempted and number of problems correct by treatment and gender. * $p < 0.0001$ vs. placebo. † $p < 0.001$ vs. placebo. ‡ $p < 0.01$ vs. placebo

nificant impairment due to associated behaviors not measured by the ADHD-RS scale. The SKAMP is better for assessing a range of ADHD and non-ADHD symptoms; however, the SKAMP scale is not practical to administer in a typical classroom or clinic and only useful when assessments are obtained in a controlled environment by trained raters.

Differences for females versus males are reported in the literature in terms of ADHD prevalence and type of presentation [30], patterns of referral/diagnosis and prescription [5, 31], neurobiology and pharmacokinetics [7, 8], and the timing or rhythm of response. Although most of the studies on the moderating effects of gender are based on populations with relatively limited female enrollment—contrasting with the female enrollment of nearly 40% in the randomized cohort in the study on which the current subgroup analysis is based—the studies converge in asserting that gender does not moderate overall outcomes and effect sizes [5, 13, 32–35]. In an analysis of results from the MTA study, investigators found that gender did not affect treatment outcomes, although 80% of that cohort were male [32, 36, 37]. In a reanalysis of data from the Comparison of Methylphenidates in an Analog Classroom Setting study, Sonuga-Barke et al. compared 136 boys (74%) and 48 girls (26%) aged 6–12 years in a double-blind crossover trial and found no main effect of gender on measures of overall response to treatment [33]. In the same trial, the authors also compared laboratory classroom measures to *parent* Swanson, Nolan, and Pelham IV (SNAP-IV) ratings [38] and found modest correlation at certain time points during the laboratory classroom day. Interestingly, after correction for multiple measures, there were no significant correlations at the end of the day (12 h after dosing), and the only significant correlation with the SKAMP-A subscale was prior to dosing. More significant correlations were noted during the time periods of 1.5–7.5 h after dosing than before dosing and at the end of the day. When the measures were analyzed by gender, only the SKAMP-C scale was significantly correlated with the SNAP-IV overall rating for females 1.5–4.5 h after dosing (-0.18 , $p < 0.001$).

To summarize, males and females both have significant improvement in core ADHD symptoms and a range of other behavior problems when treated with stimulants. However, as supported by data from this RA-AMPH study, associated behavioral features and oppositional symptoms are more prevalent in boys (though clearly not exclusively so), are often a presenting complaint, and should be a target of treatment. It is therefore important to evaluate associated behavioral symptoms in youth with ADHD during the initial assessment and over the course of treatment. Since the ADHD-RS rating scale is not sufficient to evaluate associated behavioral symptoms, clinicians may use other rating scales to assess disruptive behaviors such as the IOWA Conners scale [23], the NICHQ Vanderbilt Assessment Scale [39], or the Conners 3rd Edition rating scale (Conners 3) [40].

The study results reported here should be considered in the context of several limitations. Because this secondary analysis was applied to the results of a laboratory classroom study of RA-AMPH that was not specifically designed to assess gender differences in treatment response, it was underpowered. The current paper presents the results of an exploratory post hoc analysis. As stated, there were no comparisons for the multiple testing, and the rate of false detection was not determined, which could for example have affected the finding of non-significance for the differences between males and females on the PERMP scales. Children with significant psychiatric and medical comorbidities were excluded from the laboratory classroom study, which may have limited generalizability. This included subjects with conduct disorder and severe ODD. Subjects who were randomized and completed the double-blind phase of the laboratory classroom study, including the classroom sessions, also had to both respond to and tolerate RA-AMPH, which is not true of all patients with ADHD. Further, dosing in this study did not fully replicate what would be done in the clinical setting. To illustrate, during the 8-week open-label dose-optimization phase, RA-AMPH was given twice daily. This was also the case during the double-blind period, except for the two laboratory

classroom days, when only the single morning dose was administered. This was done to test the duration of action of RA-AMPH. Consequently, notwithstanding the possibility of amphetamine tachyphylaxis, the SKAMP results reported for the laboratory classroom days (and especially the results for the afternoon sessions)—in showing what might possibly be considered a “fatigue effect”—likely underrepresent the level of symptom control that can be achieved in clinical settings. Also, data from the SKAMP and ADHD-RS-IV were not obtained using the same dosing, as the former represented only QD and not BID dosing. The recently published ADHD-RS-5 reflects the revised ADHD criteria appearing in the Diagnostic and Statistical Manual of Mental Disorders, 5th edition, but that scale was not available at the time the RA-AMPH laboratory classroom study described in this manuscript was conducted.

CONCLUSION

In this secondary analysis of a laboratory classroom study of children with ADHD treated with R-AMPH, both genders responded equally well to treatment with RA-AMPH, with comparable onset and duration of effect in the laboratory classroom, as evidenced by the ADHD-RS-IV and SKAMP scales and by the laboratory classroom-specific PERMP. However, in the laboratory classroom setting, behavioral differences were found between boys and girls with ADHD. Boys demonstrated significantly more oppositional and conduct disordered behaviors than girls before and after treatment with RA-AMPH.

Although the ADHD-RS-IV and SKAMP scales measure changes in attention and behavior with drug treatment, the SKAMP scale also measures associated disruptive behaviors, such as frustration, lying, and interpersonal conflict, that are more characteristic of oppositional and conduct disorders and that are more prevalent in boys than in girls with ADHD. The SKAMP scale might thus be more sensitive for measuring the range of symptoms of boys with ADHD than the ADHD-RS-IV.

ACKNOWLEDGEMENTS

The authors thank all participants in the study.

Funding. Sponsorship for this study, the article-processing charges, and the open-access fee were funded by Arbor Pharmaceuticals, LLC. All authors had full access to all of the data in this study and take complete responsibility for the integrity of the data and accuracy of the data analysis.

Medical Writing and/or Editorial Assistance. Medical writing support was provided by representatives of Galen Press, Inc., and was funded by Arbor Pharmaceuticals, LLC.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this manuscript, take responsibility for the integrity of the work as a whole, and have given final approval of the version to be published.

Disclosures. Ann C. Childress has been a consultant for, and has received research support from, Arbor Pharmaceuticals, KemPharm, Ironshore Pharmaceuticals, Neos Therapeutics, Neurovance, Noven Pharmaceuticals, Otsuka Pharmaceutical, Purdue Pharma, Rhodes Pharmaceuticals, Shire Pharmaceuticals, Sunovion Pharmaceuticals, Supernus Pharmaceuticals, and Tris Pharma; she has been a speaker for Arbor Pharmaceuticals, Neos Therapeutics, Tris Pharma, and Shire Pharmaceuticals; and she has received research support from Akili Interactive, Eli Lilly, Forest Research Institute, and Johnson & Johnson. She has also been an advisor for Akili Interactive and NLS Pharma. Jeffrey H. Newcorn has been a consultant for, and has received research support from, Lundbeck and Shire Pharmaceuticals; he has received research support from Enzymotec; and he has been a consultant for Akili Interactive, Alcobra Pharmaceuticals, Arbor Pharmaceuticals, Cingulate Therapeutics, KemPharm, Ironshore Pharmaceuticals, Medice, Neos Therapeutics, NLS Pharma, Sunovion Pharmaceuticals, and Supernus Pharmaceuticals. Andrew J. Cutler has been a consultant and speaker for, and has received

research support from, AbbVie, Aevi Genomics, Allergan, Arbor Pharmaceuticals, Astra Zeneca, Janssen Pharmaceuticals, Eli Lilly, Lundbeck, Novartis Pharmaceuticals, Neos Therapeutics, NextWave Pharmaceuticals, Otsuka, Shionogi Pharma, Shire Pharmaceuticals, and Sunovion Pharmaceuticals; and he has been a consultant for, and has received research support from, Akili Interactive, Forest Research Institute, Ironshore Pharmaceuticals, Johnson & Johnson, KemPharm, Neurovance, Noven Pharmaceuticals, Purdue Pharma, Rhodes Pharmaceuticals, Supernus Pharmaceuticals, Targacept, and Therovance Biopharma; he has been a consultant for Cingulate Therapeutics and NLS Pharmaceuticals; and he serves on the board for the Neuroscience Education Institute.

Compliance with Ethics Guidelines. The protocol of the RA-AMPH laboratory classroom study was approved by an appropriate institutional review board at each study site. The parent/guardian of each subject provided voluntary written informed consent for the subject to participate in this study, and the subject gave written assent.

Data Availability. The datasets analyzed for the current study are available from the corresponding author upon reasonable request.

Open Access. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

1. Pliszka S. Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry.* 2007;46:894–921.
2. Wolraich M, Brown L, Brown RT, DuPaul G, Earls M, Feldman HM, et al. ADHD: clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics.* 2011;128:1007–22.
3. Childress AC, Sallee FR. Attention-deficit/hyperactivity disorder with inadequate response to stimulants: approaches to management. *CNS Drugs.* 2014;28:121–9.
4. Wolraich M, Brown L, Brown RT, DuPaul G, Earls M, Feldman HM, et al. Process of care supplemental appendix. ADHD: clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics.* 2011;128:S11–21.
5. Rucklidge JJ. Gender differences in attention-deficit/hyperactivity disorder. *Psychiatr Clin North Am.* 2010;33:357–73.
6. Abikoff HB, Jensen PS, Arnold LL, Hoza B, Hechtman L, Pollack S, et al. Observed classroom behavior of children with ADHD: relationship to gender and comorbidity. *J Abnorm Child Psychol.* 2002;30:349–59.
7. Mozley LH, Gur RC, Mozley PD, Gur RE. Striatal dopamine transporters and cognitive functioning in healthy men and women. *Am J Psychiatry.* 2001;158:1492–9.
8. Bhatt SD, Dluzen DE. Dopamine transporter function differences between male and female CD-1 mice. *Brain Res.* 2005;1035:188–95.
9. Willcutt EG. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics.* 2012;9:490–9.
10. DuPaul GJ. ADHD rate scale IV for children and adolescents. New York: Guildford; 1998.
11. Casat CD, Norton HJ, Boyle-Whitesel M. Identification of elementary school children at risk for disruptive behavioral disturbance: validation of a combined screening method. *J Am Acad Child Adolesc Psychiatry.* 1999;38:1246–53.
12. Wigal SB, Gupta S, Guinta D, Swanson JM. Reliability and validity of the SKAMP rating scale in a laboratory school setting. *Psychopharmacol Bull.* 1998;34:47–53.
13. Wigal SB, Kollins SH, Childress AC, Adeyi B. Efficacy and tolerability of lisdexamfetamine dimesylate in children with attention-deficit/hyperactivity disorder: sex and age effects and effect size across the day. *Child Adolesc Psychiatry Ment Health.* 2010;4:32.

14. Sonuga-Barke EJ, Coghill D, DeBacker M, Swanson J. Measuring methylphenidate response in attention-deficit/hyperactivity disorder: how are laboratory classroom-based measures related to parent ratings? *J Child Adolesc Psychopharmacol*. 2009;19:691–8.
15. Wigal SB, Wigal TL. The laboratory school protocol: its origin, use, and new applications. *J Atten Disord*. 2006;10:92–111.
16. Heal DJ, Smith SL, Gosden J, Nutt DJ. Amphetamine, past and present—a pharmacological and clinical perspective. *J Psychopharmacol*. 2013;27:479–96.
17. Sagvolden T, Xu T. L-Amphetamine improves poor sustained attention while D-amphetamine reduces overactivity and impulsiveness as well as improves sustained attention in an animal model of Attention-Deficit/Hyperactivity Disorder (ADHD). *Behav Brain Funct*. 2008;4:3.
18. Childress AC, Brams M, Cutler AJ, Kollins SH, Northcutt J, Padilla A, et al. The efficacy and safety of Evekeo, racemic amphetamine sulfate, for treatment of attention-deficit/hyperactivity disorder symptoms: a multicenter, dose-optimized, double-blind, randomized, placebo-controlled crossover laboratory classroom study. *J Child Adolesc Psychopharmacol*. 2015;25:402–14.
19. DuPaul G, Power T, Anastopoulos A, Reid R. ADHD rating scale IV: checklists, norms, and clinical interpretations. New York: Guilford; 1998.
20. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed., text revision (DSM-IV-TR). Washington, DC: American Psychiatric Association; 2000.
21. Zhang S, Faries DE, Vowles M, Michelson D. ADHD Rating Scale IV: psychometric properties from a multinational study as a clinician-administered instrument. *Int J Methods Psychiatr Res*. 2005;14:186–201.
22. Swanson JM, Agler D, Fineberg E, Wigal S, Flynn D, Fineberg E, et al. University of California, Irvine, laboratory school protocol for pharmacokinetic and pharmacodynamic studies. In: Greenhill LL, Osman BB, editors. Ritalin: theory and practice. 2nd ed. Larchmont: Mary Ann Liebert; 2000.
23. Loney J, Milich R. Hyperactivity, inattention, and aggression in clinical practice. In: Wolraich M, Milich R, editors. Advances in developmental and behavioral pediatrics. Greenwich: JAI; 1982.
24. Pelham WE Jr, Gnagy EM, Greenslade KE, Milich R. Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *J Am Acad Child Adolesc Psychiatry*. 1992;31:210–8.
25. Demmer DH, Hooley M, Sheen J, McGillivray JA, Lum JA. Sex differences in the prevalence of oppositional defiant disorder during middle childhood: a meta-analysis. *J Abnorm Child Psychol*. 2017;45:313–25.
26. Bauermeister JJ, Shrout PE, Ramirez R, Bravo M, Alegria M, Martinez-Taboas A, et al. ADHD correlates, comorbidity, and impairment in community and treated samples of children and adolescents. *J Abnorm Child Psychol*. 2007;35:883–98.
27. Pliszka S. Comorbid psychiatric disorder in children with ADHD. In: Barkley RA, editor. Attention-deficit hyperactivity disorder: a handbook for diagnosis & treatment. 4th ed. New York: Guilford; 2015. p. 140–68.
28. Swanson J, Arnold LE, Kraemer H, Hechtman L, Molina B, Hinshaw S, et al. Evidence, interpretation, and qualification from multiple reports of long-term outcomes in the Multimodal Treatment study of Children With ADHD (MTA): part I: executive summary. *J Atten Disord*. 2008;12:4–14.
29. Spencer TJ, Abikoff HB, Connor DF, Biederman J, Pliszka SR, Boellner S, et al. Efficacy and safety of mixed amphetamine salts extended release (adderall XR) in the management of oppositional defiant disorder with or without comorbid attention-deficit/hyperactivity disorder in school-aged children and adolescents: a 4-week, multicenter, randomized, double-blind, parallel-group, placebo-controlled, forced-dose-escalation study. *Clin Ther*. 2006;28:402–18.
30. Ohan JL, Visser TA. Why is there a gender gap in children presenting for attention deficit/hyperactivity disorder services? *J Clin Child Adolesc Psychol*. 2009;38:650–60.
31. Derks EM, Hudziak JJ, Boomsma DI. Why more boys than girls with ADHD receive treatment: a study of Dutch twins. *Twin Res Hum Genet*. 2007;10:765–70.
32. Owens EB, Hinshaw SP, Kraemer HC, Arnold LE, Abikoff HB, Cantwell DP, et al. Which treatment for whom for ADHD? Moderators of treatment response in the MTA. *J Consult Clin Psychol*. 2003;71:540–52.
33. Sonuga-Barke EJ, Coghill D, Markowitz JS, Swanson JM, Vandenberghe M, Hatch SJ. Sex differences in the response of children with ADHD to once-daily formulations of methylphenidate. *J Am Acad Child Adolesc Psychiatry*. 2007;46:701–10.

34. Efron D, Jarman F, Barker M. Methylphenidate versus dexamphetamine in children with attention deficit hyperactivity disorder: a double-blind, crossover trial. *Pediatrics*. 1997;100:E6.
35. Cornforth C, Sonuga-Barke E, Coghill D. Stimulant drug effects on attention deficit/hyperactivity disorder: a review of the effects of age and sex of patients. *Curr Pharm Des*. 2010;16:2424–33.
36. Jensen PS, Arnold LE, Swanson JM, Vitiello B, Abikoff HB, Greenhill LL, et al. 3-year follow-up of the NIMH MTA study. *J Am Acad Child Adolesc Psychiatry*. 2007;46:989–1002.
37. Newcorn JH, Halperin JM, Jensen PS, Abikoff HB, Arnold LE, Cantwell DP, et al. Symptom profiles in children with ADHD: effects of comorbidity and gender. *J Am Acad Child Adolesc Psychiatry*. 2001;40:137–46.
38. Swanson JM. School-based assessments and interventions for ADD students. Irvine: K.C. Publishing; 1992.
39. Wolraich ML, Lambert W, Doffing MA, Bickman L, Simmons T, Worley K. Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. *J Pediatr Psychol*. 2003;28:559–67.
40. Conners CK. Conners 3rd edition manual. North Tonawanda: Multi-Health Systems; 2008.