Physician perception of clinical improvement in children with attention-deficit/hyperactivity disorder: a post hoc comparison of lisdexamfetamine dimesylate and mixed amphetamine salts extended release in a crossover analog classroom study

Objective: To assess effects of lisdexamfetamine dimesylate (LDX) and mixed amphetamine salts extended release (MAS XR) on symptom improvement in children with attention-deficit/hyperactivity disorder (ADHD).

Methods: Post hoc analysis of a randomized, double-blind, placebo-controlled, crossover analog-classroom environment was conducted. The primary efficacy outcome was the deportment subscale of the Swanson, Kotkin, Agler, M-Flynn, and Pelham (SKAMP-D) rating scale. The secondary efficacy outcome was the investigator-rated Clinical Global Impressions-Improvement (CGI-I), a 7-point scale ranging from 1 (very much improved) to 7 (very much worse), which assesses improvement over time from baseline. McNemar test was used to compare participants' responses to LDX and MAS XR on CGI-I scores dichotomized into 1 (very much improved) vs all other response scores (2 to 7) in a 2 × 2 table.

Results: Fifty-two children (aged 6 to 12 years) were enrolled, titrated, and randomized; 50 completed the study. Investigators rated 74% of LDX participants as either very much improved or much improved on the CGI-I scale relative to 72% of MAS XR participants and 18% of placebo participants. Of the 50 children who completed the study, 32% of LDX participants were very much improved vs 16% of MAS XR, and 2% of placebo participants relative to baseline. McNemar test indicated that 10 participants were very much improved with LDX, but not MAS XR; 2 participants were very much improved with MAS XR, but not LDX; 6 participants were very much improved with both, while 32 were not very much improved with either. Analysis showed that LDX had a significantly higher number of children with a very much improved score on the CGI-I than MAS XR (P = 0.0386).

Conclusion: Treatment of children with LDX resulted in a higher number of participants with a very much improved score on the CGI-I than treatment with MAS XR or placebo.

Keywords: lisdexamfetamine dimesylate, LDX, Vyvanse, mixed amphetamine salts, extended release, ADHD, CGI-I

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is one of the most common psychiatric disorders of childhood and is estimated to affect approximately 5.3% of children and adolescents worldwide.¹ According to the Diagnostic and Statistical Manual of
Mental Disorders (DSM-IV-TR) criteria for ADHD, the disorder is generally characterized by inattention, difficulty in performing tasks, hyperactivity, and impulsivity to a degree that is inconsistent with the developmental level for that age. Stimulants are considered a first-line choice of treatment for patients with ADHD, and the short- and long-term efficacy of these medications in improving ADHD symptoms is well established. Since immediate-release stimulant formulations need multiple dosing regimens per day, long-acting forms have been developed.

Lisdexamfetamine dimesylate (LDX; Vyvanse®, Shire US Inc.) is the first long-acting prodrug stimulant, which is therapeutically inactive, and is indicated for the treatment of ADHD in children 6 to 12 years of age, adolescents 13 to 17 years, and in adults. After oral ingestion, LDX is converted to l-lysine and active d-amphetamine. While a small amount of LDX is hydrolyzed to d-amphetamine in the gastrointestinal tract, the conversion of LDX into active d-amphetamine occurs primarily in the blood. There was low interpatient variability of pharmacokinetic parameters in a clinical trial of pediatric participants with ADHD, possibly due to the biological rather than mechanical delivery of LDX.

LDX has demonstrated safety consistent with long-acting stimulant use and an extended duration of efficacy as assessed in various trials. In a randomized, controlled trial in children with ADHD, LDX was effective in controlling symptoms throughout the day, up to 6 PM, as measured by parent ratings using Conners’ Parent Rating Scale-Revised (CPRS-R) scores after a median dosing time between 7:30 AM and 8:00 AM. Efficacy was also demonstrated by clinician measures such as Clinical Global Impressions (CGI) scale scores and ADHD Rating Scale IV (ADHD-RS-IV) total scores. The most common significant adverse events (AEs) included weight loss, decreased appetite, dizziness, insomnia, and irritability, which are typical of stimulants and no serious AEs were observed. All AEs were mild or moderate in intensity and subsided with time.

A randomized, double-blind, placebo- and active-controlled study in an analog classroom demonstrated that LDX treatment in children was effective compared with placebo, beginning 2 hours after dosing, and continuing through the last time point assessed (12 hours postdose). This study examined treatment with LDX, mixed amphetamine salts extended release (MAS XR), and placebo, in a crossover fashion.

Efficacy was assessed using the primary efficacy measure of Swanson, Kotkin, Agler, M-Flynn, and Pelham Deportment (SKAMP-D) scores and secondary measures of SKAMP Attention (SKAMP-A), Permanent Product Measure of Performance-Attention (PERMP-A), PERMP-Correct (PERMP-C), and CGI-Improvement (CGI-I), in which LDX and MAS XR were significantly more effective relative to placebo, with no differences between the active treatments in the results of the primary efficacy analysis.

LDX was also effective from 1.5 to 13 hours postdose in a later randomized controlled trial in children in a laboratory classroom study. AEs were consistent with other pediatric studies of LDX.

Although general efficacy of stimulant medications including LDX and MAS XR has been well characterized, there was a desire to further understand the proportion of significant responders to medication treatment. Medication trials of ADHD frequently report a 30% improvement in ADHD symptoms, and/or a response designated as improved on the CGI-I scale, which is a score of 1 or 2. We are unaware of studies that have focused on the more stringent definition of very much improved, or CGI-I rating score of 1. A search of the available literature did not identify other studies that reported participant data for ratings of very much improved (CGI-I rating score of 1) alone as an improvement measure. Since an inspection of the relative proportions of participants by CGI-I category suggested differences between the active treatments, we conducted a post hoc analysis of the earlier analog classroom study by Biederman et al to compare efficacy differences in the extent of global symptom improvement of very much improved in participants with ADHD, who responded to LDX and MAS XR treatment vs placebo in accordance with the CGI-I assessment scale, a clinician-rated measure of improvement.

Methods
Study design
Children aged 6 to 12 years with a primary diagnosis of ADHD of the combined or predominantly hyperactive-impulsive subtypes, as defined by DSM-IV-TR criteria, were eligible for enrollment in this randomized, double-blind, placebo-controlled, analog classroom study. The study design as illustrated in Figure 1 comprised 3 treatment periods: 1-week screening; 3-week open-label titration to optimal MAS XR dose (10, 20, or 30 mg/day); 3-week randomized, double-blind (ie, participants and investigators) treatment crossover periods with optimum dose of MAS XR, LDX dose approximately equivalent to the optimized MAS XR dose, and placebo. Final selection of optimal MAS XR
dose and randomization to order of double-blind treatment sequence occurred during visit 5 along with practice analog classroom sessions. Blinding was ensured through use of matching capsule formulations and the delivery of prepackaged individual drug kits based on randomization numbers. No unblinding or breaking of treatment codes occurred during the study. Participants were evaluated in the laboratory school on the last day of each week of the double-blind period. All study activities were performed in accordance with the principles of the International Conference on Harmonisation Good Clinical Practice, 18th World Medical Assembly (Helsinki 1964), and amendments of the 29th (Tokyo 1975), the 35th (Venice 1983), the 41st (Hong Kong 1989), and the 48th (South Africa 1996) World Medical Assemblies.

The primary efficacy measure was SKAMP-D, which measures classroom symptoms of ADHD using an independent observer. It is scored on a 7-point impairment scale from 0 to 6, with higher scores indicative of more severe symptoms. SKAMP-D results were reported as the least squares (LS) mean per item score. The secondary efficacy measures included the SKAMP-A, PERMP-A and PERMP-C, and the CGI-I scores. PERMP is a validated 10-minute math test to evaluate response to stimulant medications.

The CGI scale is an investigator-rated evaluation of a subject’s improvement over time by rating change in symptoms. Investigators were thoroughly trained on administration and evaluation based on the CGI ratings. Each subject was assessed at baseline on the CGI-S (CGI-S) 7-point scale, with severity of ADHD symptoms rated from 1 (no symptoms) to 7 (very severe symptoms). At each crossover visit, subject improvement in ADHD symptoms relative to baseline was determined by the investigator on the CGI-I scale, a 7-point scale ranging from 1 (very much improved) to 7 (very much worse).

**Analog classroom environment**

Each classroom day (weeks 6, 7, 8) lasted 13 hours (approximately 6:30 AM to 7:45 PM). Participants took the randomized treatment each morning at home for the first 6 days and then the day-7 dose was given during the analog classroom visit. Thirty-minute classroom sessions, including the practice visit, were scheduled at approximately 1, 2, 3, 4.5, 6, 8, 10, and 12 hours after the morning study dose was administered.

**Safety assessments**

AEs were collected throughout the study and at each study visit. A 30-day telephone follow-up screened for serious AEs. Vital signs were obtained at all clinic visits. Laboratory parameters and physical examination information were collected on all participants at screening and at final study visit. Electrocardiogram (ECG) measurements were taken at screening, practice visit (week 5), double-blind treatment visits (weeks 6 to 8), and final study visit.

**Statistics**

A mixed-effects model of analysis of variance was conducted to assess mean CGI-I scores for the intention-to-treat (ITT) population. Given a significant overall treatment effect ($P < 0.05$), pairwise comparisons of LS means between individual treatments were conducted using a $t$ test. No multiplicity adjustments were made on subgroup statistical
comparisons. Frequencies of CGI-I were also assessed by McNemar test, a statistical method that can be used to analyze differences in response, when participants serve as their own controls or when participants are matched pairs. This test was used to compare participants’ responses to LDX and MAS XR on CGI-I scores dichotomized into 1 (very much improved) vs all other response scores (2 to 7) in a $2 \times 2$ table. A P-value can be determined based on the numbers of discordant pairs, in this case, the number of participants whose CGI-I scores were different after treatment with LDX or MAS XR. Data from the crossover phase were analyzed independently of the dose administered and reflect the dose to which each participant was optimized to MAS XR in the crossover phase of the study.

**Results**

Study demographics, baseline characteristics, participant disposition, and primary results have been previously reported. On the primary efficacy measure, SKAMP-D, participants receiving LDX and MAS XR had lower (better) LS mean (SE) SKAMP-D scores at endpoint of 0.8 (0.1) and 0.8 (0.1), respectively, in each comparison with placebo, 1.7 (0.1) ($P < 0.0001$ for each).

**CGI-I**

A total of 52 children (aged 6 to 12 years) were enrolled, titrated, and randomized; 50 completed the study postrandomization. Physician rating of the participant’s overall clinical improvement during the study using the CGI-I measure showed significant clinical improvement for participants who received LDX or MAS XR. The LS mean (SE) score was significantly better for the LDX-treated group (2.2 [0.2]) and for the MAS XR group (2.3 [0.2]) than for the placebo group (4.2 [0.2]) ($P < 0.0001$). Of the 50 ITT participants, 32% were considered very much improved and 42% much improved in the LDX-treated group, as demonstrated by investigator-rated CGI-I scores at endpoint (Table 1). The investigators rated 74% of participants on LDX as either very much improved or much improved (32% very much improved and 42% much improved) on the CGI-I scale, compared with 72% (16% very much improved and 56% much improved) of participants when they received MAS XR and 18% (2% very much improved and 16% much improved) of participants when they received placebo (Table 1). Percentages of children rated very much improved after treatment included 32% of the 50 children treated with LDX; 16% of these same 50 participants after treatment with MAS XR; and 2% of placebo-treated participants (Table 1). Specifically, 10 participants were rated very much improved with LDX, but not MAS XR, while 2 participants were rated very much improved with MAS XR, but not LDX (Table 2). Also, 6 participants were rated very much improved with both LDX and MAS XR (Table 2). The McNemar test on the $2 \times 2$ table showed that LDX resulted in a significantly higher proportion of participants with a very much improved score on the CGI-I than did MAS XR ($P = 0.0386$) (Table 2).

Overall, 42 participants in the crossover study were not very much improved when treated with MAS XR, of which 10 were very much improved with LDX, but not MAS XR. Of these 10 participants, 3 received LDX at an earlier visit than MAS XR; and 7 received MAS XR at an earlier visit than LDX. Also, of the 34 participants who were not very much improved with LDX treatment, 2 were very much improved when treated with MAS XR, but not LDX. Both participants were treated with LDX at an earlier visit than MAS XR.

**Safety assessment**

Safety results have previously been reported in detail. Briefly, safety analysis indicated that no serious AEs were reported throughout the study. During the titration period

<table>
<thead>
<tr>
<th>Table 1</th>
<th>CGI-I scores at endpoint, ITT population, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Improved (CGI-I score = 2)</td>
</tr>
<tr>
<td>LDX</td>
<td>37 (74)</td>
</tr>
<tr>
<td>MAS XR</td>
<td>36 (72)</td>
</tr>
<tr>
<td>Placebo</td>
<td>9 (18)</td>
</tr>
</tbody>
</table>

**Notes:** CGI-I score of 1 = very much improved; CGI-I score of 2 = much improved.

**Abbreviations:** CGI-I, Clinical Global Impressions-Improvement; ITT, intention-to-treat; LDX, lisdexamfetamine dimesylate; MAS XR, mixed amphetamine salts extended release.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Frequencies of CGI-I very much improved$^a$ vs other$^b$ scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LDX, n</td>
</tr>
<tr>
<td>MAS XR, n</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>34 (68.0%)</td>
</tr>
</tbody>
</table>

$^a$McNemar test$^a$ $P = 0.0386$

**Notes:** $^a$CGI-I very much improved score $= 1$; $^b$CGI-I all other scores $= 2, 3, 4, 5, 6$, or 7; $^c$McNemar test is a test of agreement of participants’ responses to MAS XR and LDX and is not a direct test on the percentages displayed in the table.

**Abbreviations:** CGI-I, Clinical Global Impressions-Improvement; LDX, lisdexamfetamine dimesylate; MAS XR, mixed amphetamine salts extended release.
with MAS XR, 46% of participants reported an AE. The most common AEs were headache (15%), decreased appetite (14%), insomnia (10%), abdominal pain (6%), and upper abdominal pain (6%). During the double-blind phase, 16% of participants treated with LDX reported AEs compared with 18% of those treated with MAS XR and 15% of those in the placebo group. The most common AEs associated with LDX treatment were insomnia (8%), decreased appetite (6%), and anorexia (4%); with MAS XR, upper abdominal pain (4%) and decreased appetite (4%); and placebo, vomiting (4%). No clinically meaningful vital sign or ECG findings were reported in patients treated with LDX or MAS XR.

Discussion
This clinical trial assessed the effect of LDX throughout the day in children with ADHD in a classroom setting compared with placebo and MAS XR. In general, the primary efficacy measure, SKAMP-D, indicated that treatment with either LDX or MAS XR is associated with improvement in ADHD symptoms and behavior, compared with placebo. This was demonstrated by a decrease in score, which is indicative of improvement in participant impairment in classroom behavior as assessed by an independent observer. Post hoc analysis of the data from this trial indicated a statistically significant difference between LDX and MAS XR treatment when examining the likelihood of being very much improved (CGI-I score = 1) when treated with LDX compared with MAS XR. Treatment with LDX resulted in a significantly higher number of participants who had a very much improved score on the CGI-I than with MAS XR, suggesting that the extent of improvement was greater with LDX. Thus, CGI-I assessment suggested that LDX treatment, given at doses containing approximately equivalent amounts of d-amphetamine as MAS XR, resulted in greater improvement in ADHD symptoms as indicated by a higher number of very much improved scores relative to comparators. Whereas many factors may have contributed to the differential findings of similar efficacy on the SKAMP-D between LDX and MAS XR, and although LDX showed greater likelihood of very much improved scores, differences in timing of the assessments (repeatedly over 12 hours vs once during the analog classroom day) and the scope of assessments (specific measures of deportment vs global assessments of efficacy and tolerability) may have played a role.

Differences between LDX and MAS XR were observed using other secondary measures in this trial. Participants at 12 hours postdose receiving LDX had significantly higher scores on the PERMP-A (math problems attempted) and PERMP-C (math problems correct) compared with children treated with MAS XR in a post hoc analysis, but not at the earlier time points.10,11 These data indicate overall improvement in symptoms of ADHD in children relative to placebo with various outcome measures. As previously reported, at 12 hours postdose, SKAMP-D and SKAMP-A assessments indicated significant improvement in scores of participants treated with LDX relative to placebo. In a more recent, double-blind, placebo-controlled, laboratory school study of children with ADHD, while no comparator arm of MAS XR was included, behavior and math test performance in participants receiving LDX showed significant improvement compared with placebo from 1.5 to 13 hours postdose (last time point assessed).8 Taken together these findings suggest that the beneficial effects of LDX extend into later hours of the day.8

It is possible that this observed difference in treatment response to LDX and MAS XR may be due to the differences in delivery systems. LDX is a prodrug that depends on enzymatic cleavage to yield active d-amphetamine and L-lysine.12 MAS XR uses a microbead technology to provide 2 pulses of MAS with a pharmacokinetic profile equivalent to 2 equal doses of MAS taken 4 hours apart.13,14 On the other hand, it is unclear if the open-label treatment phase with MAS XR could have affected subjective assessments of improvement related to switching medications.

We are not aware of previous studies that have reported on the comparative results of the more stringent CGI-I very much improved rating. The findings of this subanalysis may be important because a rating of very much improved on the CGI-I may be indicative of a more robust treatment response in many participants. In fact, recent work by Goodman et al has shown a correlation between scores on the ADHD-RS-IV and the CGI-I.15

These results should be interpreted carefully, due to limitations of the post hoc analysis. A number of design features contribute to the limited ability to extrapolate the findings of this analysis. The parent study was designed as a noninferiority study as well as not being prospectively designed or powered to detect differences between the products. Moreover, there were no differences in primary endpoints in this study. The CGI scale is widely used and easily scored; however, findings of test-retest reliability and validity are mixed.16 For this trial, though, the limited number of raters scoring this measure and limited number of repeated assessments (once for each crossover period) eliminated some concerns over the reliability of...
this measure. This study had a small sample size, and, additionally, treatment sequence in this crossover study could conceivably have affected the results. In relation to this, while no apparent treatment sequence effect is discernible for those participants who responded to one but not the other active treatment, the numbers of participants who met these criteria were too few to analyze this statistically. The results of this analysis should be viewed as an exploratory endeavor to identify interesting preliminary results that can be confirmed in other well-controlled clinical trials with preplanned endpoints controlling for all pairwise comparisons.

In conclusion, the results of this analysis showed that, overall, participants with global illness were improved by both active treatments and more participants taking LDX were very much improved compared with those receiving MAS XR and placebo. These findings should be considered preliminary in light of the limitations of the analysis.

Funding statement and acknowledgment
Clinical research was funded by the sponsor, Shire Development Inc. Under the direction of the authors, Huda Abdullah, PhD, and Michael Pucci, PhD, employees of Ogilvy CommonHealth Scientific Communications (OCHSC), provided writing assistance for this publication. Editorial assistance in formatting, proofreading, copy editing, and fact checking was also provided by OCHSC. Ben Adeyi, MS, Thomas Babcock, DO, and Bryan Dirks, MD, from Shire Development Inc. also reviewed and edited the manuscript for scientific accuracy. Shire Development Inc. provided funding to OCHSC for support in writing and editing this manuscript. Although the sponsor was involved in the design, collection, analysis, interpretation, and fact checking of information, the content of this manuscript, the ultimate interpretation, and the decision to submit it for publication in Neuropsychiatric Disease and Treatment were made by the authors independently.

Disclosure
Frank A López, MD, is a consultant for Bristol-Myers Squibb, Celltech, Cephalon, Eli Lilly, New River Pharmaceuticals, Novartis, Pfizer, Shire US and; has received grant/research support from Bristol-Myers Squibb, Cephalon, Celltech, Eli Lilly, New River Pharmaceuticals, Novartis, Pfizer, Shire US, and; is on the speakers bureau for Cephalon, Novartis, and Shire US; and is an advisory board member for Celltech, Cephalon, Eli Lilly, Novartis, and Shire US.

Brian Scheckner, PharmD, is an employee of Shire and holds stock and/or stock options in Shire.

Ann C Childress, MD, is a consultant for NextWave, Shire and Novartis; has received grant/research support from Sepracor, Shire, Novartis, Bristol-Myers Squibb, Somerset, NextWave, Abbott, Lilly USA, LLC, Ortho-McNeil Janssen Scientific Affairs, and Johnson & Johnson Pharmaceutical Research and Development, LLC; and is a speaker for GlaxoSmithKline, Shire, Novartis, and Bristol-Myers Squibb.

References