Symptom Improvement and Residual Symptoms during Acute Antidepressant Treatment in Pediatric Major Depressive Disorder

Rongrong Tao, M.D., Ph.D.,^{1,2} Graham J. Emslie, M.D.,^{1,2} Taryn L. Mayes, M.S.,^{1,2}
Paul A. Nakonezny, Ph.D.,¹ and Betsy D. Kennard, Psy.D.^{1,2}

Abstract

Objective: Knowing the timing of specific depressive symptom improvement will enable clinicians to prepare their patients well and improve treatment outcome, whereas recognizing which depressive symptoms may show delayed improvement will help clinicians to provide additional interventions early in treatment. In a prospective open-label fluoxetine study, we investigated the timing of depressive symptom improvement during acute treatment, and identified common remaining symptoms following 4, 8, and 12 weeks of acute treatment in depressed youths.

Method: A total of 168 children and adolescents, aged 7–18 years, with primary diagnoses of major depressive disorder (MDD) received 12 weeks of fluoxetine treatment. Youths were evaluated using the Kiddie Schedule for Affective Disorders and Schizophrenia. The outcome measure included the Children's Depression Rating Scale–Revised.

Results: All depressive symptoms improved, particularly during the first 4 weeks of acute treatment. Forty-seven percent of remitters reported at least one residual symptom following 12 weeks, with most common residual symptoms being impaired school performance, insomnia, and irritability.

Conclusions: Residual symptoms are common, even among remitters, at the end of 12 weeks of acute treatment. There is a need for clinicians to monitor symptom improvement and potentially provide additional interventions for the more resistant symptoms, such as insomnia and school performance.

Introduction

DURING THE PAST DECADE, research in the treatment of pediatric-onset major depressive disorder (MDD), particularly the acute phase of treatment, has flourished. Although controversy has occurred regarding the relative effectiveness and safety of medication in placebo-controlled trials, little attention has focused on the relatively low rate of remission in short-term trials. However, it has been clearly established that residual symptoms are common, even after adequate acute treatment. For example, the mean total score on the Children's Depression Rating Scale–Revised (CDRS-R) (Poznanski and Mokros 1996) of the three double-blind placebo-controlled fluoxetine studies (Emslie et al. 1997; Emslie et al. 2002; March et al. 2004) had decreased from 57 to 60 at baseline to 35 to 38 at the end of acute treatment, indicating that many youths have significant residual symptoms at the end of the acute treatment (CDRS-R ≤ 28 is generally consistent with remission). Kennard and colleagues examined the residual symptoms of youth in the

Treatment for Adolescents with Depression Study (TADS) (Kennard et al. 2006). On the basis of the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (K-SADS-PL) (Kaufman et al. 1997), 50% of responders ($n = 214$) and 20% of remitters ($n = 95$) continued to have residual symptoms at the end of 12 weeks of acute antidepressant treatment. This is particularly germane given the evidence that the presence of residual symptoms predicts future relapse (Emslie et al. 2008).

Studies have examined the timing of overall symptom improvement in adolescents in the TADS study (Kratochvil et al. 2006) and both children and adolescents in an open-label trial (Tao et al. 2009). No study, however, has investigated the timing of improvement of specific symptoms or symptom clusters or the differences between children and adolescents. Do all depressive symptoms improve at a similar pace, or do some symptoms resolve more quickly than others? It is clinically valuable to know which symptoms improve early and which symptoms are likely to be slow

¹University of Texas Southwestern Medical Center at Dallas and ²Children's Medical Center of Dallas.

The study was supported by a grant entitled Childhood Depression: Remission and Relapse (R01 MH39188) from the National Institute of Mental Health (Principal Investigator, Graham Emslie). Eli Lilly provided the medication for the study, but had no role in the study design or implementation, analysis of data, or authorship of this manuscript. In addition, support for Dr. Tao was provided in part by the Klingenstein Third Generation Foundation Fellowship in Child and Adolescent Psychiatry.

The statistical expert for this study was Paul A. Nakonezny, Ph.D.

to respond to treatment. This type of information will enable clinicians to prepare their patients well, and consequently may improve adherence by informing patients about what to expect from their treatment. In addition, adjunctive treatment may be warranted early on during the acute treatment phase if certain symptoms take longer to resolve as part of the depression treatment (e.g., sleep medication for severe sleep disturbance). Ultimately, this type of information will improve acute treatment outcome and reduce the likelihood of relapse.

The aims of this study were to examine the timing of symptom improvement over 12 weeks of open-label acute fluoxetine treatment in children and adolescents with nonpsychotic MDD and to examine which depressive symptoms are most commonly still problematic even after treatment response.

Methods

The data presented here are from the acute-phase treatment of a randomized continuation study (Emslie et al. 2008). The study was approved by Institutional Review Board at the University of Texas Southwestern Medical Center. All participants and their parents provided written informed consent and assent.

Participants

Details of the acute sample have been published previously (Tao et al. 2009). Briefly, 168 children and adolescents, age 7–18 years, with MDD based on the *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) (American Psychiatric Association 1994) for MDD for at least 4 weeks entered the acute treatment phase of the study. Although co-morbid psychiatric diagnoses were allowed, participants were excluded if they had a lifetime history of psychotic depression or bipolar I and II disorder (BP I and II) or past 6-month history of alcohol or substance abuse or dependence. Concurrent use of other psychotropic medications was not allowed except for stimulants for attention-deficit/hyperactivity disorder (ADHD), which could be initiated prior to or during acute treatment (sequentially after the antidepressant).

Procedures and measures

Study procedures have been published previously (Emslie et al. 2008; Tao et al. 2009). Children were evaluated by a trained research evaluator initially using the K-SADS-PL, CDRS-R, and the Clinical Global Impression–Severity Scale (CGI-S) (Guy 1976) by interviewing the child and parent separately. Diagnoses, co-morbid psychiatric disorders, and depression severity were confirmed 1 week later by a psychologist or psychiatrist using the same scales. At the third visit (baseline), children who continued to meet criteria for MDD (including CDRS-R score ≥ 40 and CGI-S ≥ 4) received fluoxetine treatment with 10 mg/day for 1 week, increasing to 20 mg/day a week later if the child was able to tolerate the medication. Fluoxetine dose could be increased to 30 mg or 40 mg after week 6 if the patient was not adequately improving. Children were followed up weekly for the first 4 weeks, then biweekly for the last 8 weeks (weeks 0, 1, 2, 3, 4, 6, 8, 10, and 12) of treatment by a child psychiatrist, who assessed depressive symptom severity, improvement and adverse events.

Depression Symptoms. Depressive symptoms and severity were assessed by the physician-rated CDRS-R at each visit. Depressive symptoms were considered present if the individual item was ≥ 3 . The CDRS-R is a 17-item clinician-rated pediatric depression scale. Except for three items rated on a scale of 1–5

(insomnia, appetite, speech), the items are rated on a Likert-type scale of 1–7, with higher scores indicating greater depression severity (range of the CDRS-R total score is 17–113). To examine symptom improvement over the 12-week study, we present data using symptom factors. Five depressive symptom factors of the CDRS-R scale identified recently are used as indicators of depressive symptom factors: Reported Depression, Anhedonia, Somatic Symptoms, Morbid Thoughts, and Observed Depression (Guo et al. 2006). As noted by Guo and colleagues, items 5 (Appetite), 8 (Irritability), and 9 (Guilt) had low loading with the above factors and were therefore not “assigned” to a specific factor. However, for a more complete presentation of the data, those items were included in our analyses, and they were grouped with the respective factor that they had the highest loading and or correlation, as recommended by Guo and colleagues (Guo et al. 2006) (see Table 1, below, for factor items). The score for each of the five symptom clusters (factors) was computed by averaging the scores of the individual items for each factor. Thus, the possible range for each factor score is 1–7 for all factors except for Somatic Symptoms and Observed Depression. The range is 1–6.2 for Somatic Symptoms and 1–6.3 for Observed Depression because some items use a 1- to 5-point scale instead of 1- to 7-point scale.

Descriptive data analyses on remaining symptoms (i.e., residual symptoms) at weeks 4 and 8 and at the end of treatment used each individual symptom rather than symptom factors. The CDRS-R item scores were used to calculate the presence of individual depressive symptoms. A symptom was present if a CDRS-R item had a rating equal or greater than 3 (a rating of 3 corresponds to mild severity of a symptom or difficulty) at baseline, weeks 4, 8, and 12. The number of subjects used to calculate percentage of depressed children and adolescents who had a particular depressive symptom at baseline, weeks 4, 8, and 12 was the number of subjects who remained in the study at those weeks. Similarly, a residual symptom was present if a CDRS-R item had a rating equal or greater than 3 at week 12.

Remission. Remission was defined as a CDRS-R total score of 28 or lower at week 12. Among the patients who completed 12 weeks of acute treatment, patients with a CDRS-R total score of greater than 28 were considered nonremitters.

Global Severity and Improvement. The clinician-rated CGI-S and the CGI-Improvement Scale (CGI-I) were also completed at each visit. The CGI-S is a 7-point Likert-type scale, where 1 indicates “Normal, not at all ill” and 7 indicates “Among the most extremely ill patients.” The CGI-I is also a 7-point Likert scale, where 1 indicates “Very much improved” and 7 indicates “Very much worse.” A score of 1 or 2 usually indicates response to treatment (Guy 1976).

Data analysis

A completely within-subjects mixed linear model analysis of repeated measures was used to examine the change in each of the five depressive symptom factor scores over the 12 weeks of treatment, with age included as a covariate in the model. A separate mixed-model analysis was conducted on each of the five factors. Restricted maximum likelihood estimation and Type III tests of fixed effects were used, with the Kenward–Roger correction (Kenward and Roger 1997) applied to the ante-dependence covariance model. Changes in each of the five depressive symptom mean factor scores over time periods during the acute treatment

(baseline to week 4, weeks 4–8, and weeks 8–12) were examined using least squares mean contrasts, and *p* values associated with the tests of the mean contrasts were adjusted for multiple comparisons using the SIMULATE method in SAS (SAS Institute, Inc., Cary, NC). The Hedges *g* (Hedges and Olkin 1985) was used to calculate the effect size of the change over each time period (mentioned above), which was adjusted for the correlated means across each respective repeated time period measure.

Next, in a separate model, age main effects were examined by a 2 age group (children vs. adolescents) $\times$ 8 time point (weeks 1, 2, 3, 4, 6, 8, 10, and 12) mixed linear model analysis of repeated measures. A separate mixed-model analysis was conducted on each of the five depressive symptom factors, controlling for each respective baseline depressive symptom rating. The main effect of age group and the age $\times$ time point (weeks) interaction effect were examined. A similar mixed model was also used to evaluate the age main effect and the age $\times$ time point interaction effect on CDRS-R total score over the 12 weeks of treatment.

Finally, descriptive statistics were used to evaluate the frequency of individual depressive symptoms at four time periods, baseline (*n* = 168), weeks 4 (*n* = 161), 8 (*n* = 149), and 12 (*n* = 138) and the frequency and number of residual symptoms among remitters (*n* = 105).

All analyses were intent-to-treat (ITT) with minor modification. One participant who did not have at least one postbaseline visit was excluded from the analyses. The mixed-model procedures of PROC

MIXED in SAS (SAS Institute, Inc., Cary, NC) were used for the mixed-model analyses. The maximum likelihood-based mixed-model approach is robust to the bias from random missing data (Mallinckrodt et al. 2003; Padilla and Algina 2007). Thus, missing observations were left as missing and the maximum likelihood-based method of estimation was used; we did not impute. The level of significance for all tests was set at $\alpha = 0.05$ and, to address multiple testing on the baseline symptom scores, *p* values were adjusted using the false discovery rate (Benjamini and Hochberg 1995).

Results

The sample was predominantly white (75%), with 48% children (ages 11 and under) and 42% females. The mean baseline CDRS-R total score was 57.6 [standard deviation (SD) 7.3], similar to the baseline depression severity of youths in previous randomized controlled trials (RCTs).

Baseline symptom factors

A total of 168 children and adolescents entered the acute phase of the treatment study, and the majority (82%; *n* = 138) completed the 12 weeks of acute treatment with fluoxetine. There were no significant differences in demographic and baseline characteristics including age, gender, ethnicity, baseline CDRS-R score, number of MDD episodes, length of illness and length of current MDD episode, suicidal behavior, number or type of co-morbid

TABLE 1. BASELINE SYMPTOM FACTORS AND CDRS-R ITEM SCORES

Symptom factors	Total (<i>n</i> = 168)	Children (<i>n</i> = 80)	Adolescents (<i>n</i> = 88)	<i>p</i> value ^a	Males (<i>n</i> = 97)	Females (<i>n</i> = 71)	<i>p</i> value ^b
Reported depression ^c	4.2 $\pm$ 0.6	4.2 $\pm$ 0.6	4.1 $\pm$ 0.7	0.17	4.1 $\pm$ 0.6	4.3 $\pm$ 0.6	0.04
8 Irritability	4.5 $\pm$ 1.0	4.6 $\pm$ 1.0	4.4 $\pm$ 1.0	0.30	4.3 $\pm$ 1.2	4.7 $\pm$ 0.7	0.07
10 Self-esteem	4.5 $\pm$ 1.1	4.7 $\pm$ 1.2	4.3 $\pm$ 1.0	0.13	4.6 $\pm$ 1.1	4.2 $\pm$ 1.1	0.07
11 Depressed mood	4.7 $\pm$ 0.8	4.5 $\pm$ 0.7	4.8 $\pm$ 0.9	0.09	4.6 $\pm$ 0.8	4.7 $\pm$ 0.8	0.49
14 Weeping	3.0 $\pm$ 1.6	3.2 $\pm$ 1.6	2.9 $\pm$ 1.5	0.30	2.7 $\pm$ 1.5	3.5 $\pm$ 1.5	0.01
Anhedonia ^c	4.1 $\pm$ 1.0	4.0 $\pm$ 1.0	4.2 $\pm$ 0.9	0.25	4.2 $\pm$ 0.9	4.0 $\pm$ 1.1	0.22
2 Social withdrawal	4.5 $\pm$ 1.0	4.5 $\pm$ 1.1	4.6 $\pm$ 1.0	0.66	4.6 $\pm$ 0.9	4.4 $\pm$ 1.1	0.31
3 Difficulty having fun	3.8 $\pm$ 1.2	3.6 $\pm$ 1.2	3.9 $\pm$ 1.2	0.28	3.9 $\pm$ 1.1	3.6 $\pm$ 1.2	0.29
Somatic symptoms ^c	3.7 $\pm$ 0.7	3.5 $\pm$ 0.7	3.8 $\pm$ 0.7	0.07	3.5 $\pm$ 0.7	3.9 $\pm$ 0.6	0.02
1 School work	4.2 $\pm$ 1.2	4.1 $\pm$ 1.1	4.3 $\pm$ 1.3	0.31	4.3 $\pm$ 1.2	4.0 $\pm$ 1.2	0.25
4 Sleep	3.6 $\pm$ 1.2	3.3 $\pm$ 1.2	3.9 $\pm$ 1.2	0.03	3.4 $\pm$ 1.3	3.8 $\pm$ 1.2	.13
5 Appetite	2.9 $\pm$ 1.2	2.8 $\pm$ 1.3	2.9 $\pm$ 1.2	0.66	3.7 $\pm$ 1.2	3.1 $\pm$ 1.2	.08
6 Fatigue	4.4 $\pm$ 1.3	4.1 $\pm$ 1.4	4.7 $\pm$ 1.2	0.03	4.2 $\pm$ 1.5	4.7 $\pm$ 1.1	.08
7 physical complaints	3.2 $\pm$ 1.4	3.3 $\pm$ 1.3	3.2 $\pm$ 1.5	0.66	2.9 $\pm$ 1.4	3.6 $\pm$ 1.4	0.02
Morbid thoughts ^c	2.3 $\pm$ 0.9	2.2 $\pm$ 0.9	2.4 $\pm$ 0.9	0.36	2.3 $\pm$ 0.9	2.3 $\pm$ 0.8	0.91
9 Guilt	2.4 $\pm$ 1.2	2.4 $\pm$ 1.2	2.5 $\pm$ 1.3	0.66	2.3 $\pm$ 1.2	2.5 $\pm$ 1.3	0.37
12 Morbid ideation	2.4 $\pm$ 1.2	2.4 $\pm$ 1.3	2.5 $\pm$ 1.2	0.83	2.4 $\pm$ 1.3	2.5 $\pm$ 1.2	0.47
13 Suicidal ideation	2.0 $\pm$ 1.2	1.9 $\pm$ 1.1	2.1 $\pm$ 1.3	0.31	2.1 $\pm$ 1.3	1.8 $\pm$ 1.1	0.20
Observed depression ^c	2.5 $\pm$ 0.8	2.4 $\pm$ 0.8	2.6 $\pm$ 0.9	0.17	2.5 $\pm$ 0.9	2.4 $\pm$ 0.8	0.57
15 Depressed affect	3.1 $\pm$ 1.0	3.0 $\pm$ 1.0	3.2 $\pm$ 1.1	0.28	3.2 $\pm$ 1.1	3.0 $\pm$ 1.0	0.33
16 Speech	2.1 $\pm$ 0.8	2.1 $\pm$ 0.8	2.2 $\pm$ 0.8	0.63	2.2 $\pm$ 0.8	2.0 $\pm$ 0.8	0.41
17 hypoactivity	2.3 $\pm$ 1.0	2.1 $\pm$ 0.9	2.5 $\pm$ 1.0	0.09	2.3 $\pm$ 1.0	2.3 $\pm$ 1.0	0.89

Note: Data are presented as mean $\pm$ standard deviation (SD).

^aTested for differences between children and adolescents on each factor and item score in a separate model, and the *p* values are adjusted using the false discovery rate described by Benjamini and Hochberg (1995).

^bTested for differences between males and females on each factor and item score in a separate model, and the *p* values are adjusted using the false discovery rate described by Benjamini and Hochberg (1995).

^cScore for each symptom factor is the mean score for the factor.

psychiatric illness, and family history of depression, between the 138 participants who completed the 12 weeks of acute treatment and the 30 participants who dropped out the study during the acute treatment. Table 1 presents the factor score for each symptom factor and mean scores of each of the individual item within those factors at baseline. Three symptom factors (Reported Depression, Anhedonia, and Somatic Complaints) were rated higher (means were 4.2 ± 0.6 , 4.1 ± 1.0 , and 3.7 ± 0.7 , respectively), whereas the severity of the other two symptom factors (Morbid Thoughts and Observed Depression) was substantially lower.

There were no significant age or gender differences on the CDRS-R total score (mean total scores were 56.4 ± 6.6 for children and 58.6 ± 7.8 for adolescents, $p = 0.052$). Similarly, there were no significant age differences on any factor total scores, although adolescents showed a trend for increased Somatic Symptoms (adjusted $p = 0.07$).

Although there were no significant gender differences on the CDRS-R total score at baseline, two symptom factors, Reported Depression (adjusted $p = 0.04$) and Somatic Symptoms (adjusted $p = 0.02$), were rated higher among females than males.

Because there were minimal age or gender differences on the symptom factors, we also examined individual symptoms and found that among individual symptoms, adolescents (≥ 12 years of age) had slightly more sleep difficulties (adjusted $p = 0.03$) and fatigue (adjusted $p = 0.03$). We also found that weeping (adjusted $p = 0.01$) and physical complaints (adjusted $p = 0.02$) were more common among females than males.

Time course of symptom improvement

To examine how symptom severity changes over the 12 weeks of acute treatment, least square means (adjusted for age) from the mixed-model analyses were used to show the decline of the mean ratings of the five symptom factors over the 12 weeks of treatment for the 168 depressed youth (Fig. 1). All depressive symptoms as indicated by the age-adjusted least squares means of each of the five CDRS-R symptom factors had significant improvement over the 12-week period.

To evaluate the timing (time period) of improvement of depressive symptom factors, The Hedges g was used to estimate the effect size of depression severity change over 4-week time intervals. Table 2 presents the least squares mean change and respective effect size estimates (adjusted for age) of the depressive severity change over time periods during the acute treatment (baseline to week 4, weeks 4–8, and weeks 8–12). The effect sizes of depression severity change were the largest for all five factors during the first 4 weeks of treatment, with Reported Depression having the largest effect size ($g = 2.16$). In other words, all depressive symptoms dropped significantly during the first 4 weeks of treatment, regardless of their severity at the baseline. From weeks 4 to 8, all symptoms continued to improve, but to a lesser degree than noted during the first 4 weeks; during this second 4-week treatment period (from weeks 4 to 8), Somatic Complaints had the largest improvement with an effect size of 0.64. From weeks 8 to 12, there was limited further improvement, particularly for

FIG. 1. Depression symptoms factor severity changes over the 12 weeks of acute treatment. Scores used in the plot are least square means adjusted for age. There was a significant difference in the slope of each of the five depressive symptom factors over the 12 weeks. Each p value < 0.0001 .

TABLE 2. EFFECT SIZE OF DEPRESSION SEVERITY CHANGE AT THREE TIME PERIODS DURING ACUTE TREATMENT

Factor/cluster	Weeks 0-4			Weeks 4-8			Weeks 8-12			Total (weeks 0-12)		
	Change	p	Effect size	Change	p	Effect size	Change	p	Effect size	Change	p	Effect size
Reported depression	1.98	0.0001	2.16	0.23	0.0007	0.37	0.19	0.04	0.25	2.40	0.0001	2.71
Anhedonia	1.99	0.0001	1.85	0.33	0.0001	0.46	0.25	0.001	0.35	2.57	0.0001	2.38
Somatic complaints	1.45	0.0001	1.95	0.35	0.0001	0.64	0.15	0.02	0.25	1.95	0.0001	2.39
Observed depression	0.86	0.0001	1.07	0.22	0.0001	0.40	0.10	0.24	0.20	1.17	0.0001	1.34
Morbid thoughts	0.94	0.0001	1.21	0.13	0.01	0.30	0.06	0.69	0.15	1.12	0.0001	1.31

Note: Change = Least square mean changes of depressive symptom cluster mean scores (adjusted for age) during the time period; effect size = the Hedges g , which was adjusted for the age-adjusted correlated means across each respective repeated time period; p values are from t -tests, in a within-subjects mixed linear model analysis of repeated measures, associated with the test of each contrast for change in symptom mean scores across each time period, while controlling for age; p values associated with the tests of the mean contrasts were adjusted for multiple comparisons using the SIMULATE method in SAS.

Observed Depression and Morbid thoughts, which were low at baseline.

To evaluate age effects on each factor mean score over the 12 weeks of treatment, the sample was grouped into children (≤ 11 years of age) and adolescents (≥ 12 years of age) and separate mixed models were conducted. The mixed linear model analysis of repeated measures analysis revealed that, after controlling for the respective baseline mean factor score, the pattern of each of the five adjusted mean factor scores was significantly different between children and adolescents (p values < 0.01) throughout acute treatment, with adolescents having higher omnibus least squares mean factor scores than children over the entire 12 weeks of acute treatment. To examine week-by-week differences between children and adolescents, simple group effects (adolescent vs. children) at weeks 1, 2, 4, 6, 8, 10, and 12 were examined on adjusted mean factor scores. Adolescents had higher adjusted mean factor scores than children during early weeks of acute treatment. However, by week 12, none of the adjusted mean factor scores was different between children and adolescents. In other words, depressive symptoms among children decreased their severity quicker than adolescents, but adolescents caught up at the end of the acute treatment.

Similarly, the mixed linear model analysis of repeated measures analysis revealed that, after controlling for baseline CDRS-R total score, the adjusted CDRS-R total scores were significantly different between children and adolescents throughout acute treatment, with adolescents having a higher omnibus least squares mean CDRS-R total score than children over the entire 12 weeks of acute treatment (36.07 vs. 31.95, $p = 0.0001$). However, no significant age group $\times$ time interaction effect was found ($p = 0.27$) and, by week 12, adolescents and children were similar on CDRS-R total scores (27.42 vs. 25.20, $p = 0.13$).

Similar (but separate) mixed models were used to evaluate gender effects on each mean factor score over the 12 weeks of treatment. However, no gender main effects or gender $\times$ time interaction effects were found for any of the five depression factors (p values > 0.30 ; results not reported).

Specific symptom changes over 12 weeks

The percent of depressed youths who had a particular CDRS-R depressive symptom at the four time periods (baseline, weeks 4, 8, and 12) is presented in Fig. 2. At baseline, the top five depressive symptoms of youths with MDD were depressed mood (99.4%; 167/168), irritability (97%; 163/168), difficulty having fun (96.4%; 162/168), decreased self-esteem (96.4%; 162/168), and impaired school performance (91.7%; 154/168). Rarer symptoms were listless speech (25.6%; 43/168), suicidal ideation (33.3%; 56/168), and hypoactivity (39.3%; 66/168). There was significant improvement during the first 4 weeks after initiating treatment. By week 4, the most frequent reported depressive symptoms were impaired school performance (50.3%; 81/161), decreased self-esteem (49.7%; 80/161), irritability (39.8%; 64/161), depressed mood (37.3%; 60/161), and difficulty having fun (34.2%; 55/161). At weeks 8 and 12, impaired school performance, decreased self-esteem, and irritability continued to be the three most frequently endorsed symptoms, although fewer patients continued to report these symptoms over time.

Because there were some differences in individual symptoms at baseline between children and adolescents, we examined the residual symptoms separately for children and adolescents. We found that the proportion of children versus adolescents who had

FIG. 2. Frequency of specific symptoms at baseline and weeks 4, 8, and 12. Symptoms were considered present if the Children's Depression Rating Scale–Revised item rating was ≥ 3 .

particular residual symptoms were similar at week 12, except for anhedonia (19.7%; 14/71 for adolescents and 4.5%; 3/67 for children) and fatigue (21.1%; 15/71 for adolescents and 7.5%; 5/67 for children). Even among remitters, residual symptoms were common. Of the 168 participants, 105 (62.5%; 105/168) were remitted (CDRS-R total score of ≤ 28) at week 12; 47% (49/105) reported at least one residual symptom, although few (8.6%; 9/105) had three or more residual symptoms. The two core depressive symptoms for MDD—depressed mood and difficulty having fun—were resolved among most remitters, with less than 5% (5/105) of remitters having depressed mood and only 1% (1/105) of remitters having difficulty having fun. Impaired school performance (20%; 21/105), insomnia (11.4%; 12/105), and irritability (10.5%; 11/105) were the most frequently rated residual symptoms among remitters. Low self-esteem (8.6%; 9/105) and appetite change (5.7%; 6/105) were the next most frequently reported residual symptoms among remitters.

Discussion

This is the first study to examine carefully the improvement of depressive symptoms over the course of 12 weeks of acute open-label fluoxetine treatment of pediatric MDD.

The current study indicates that all depressive symptoms improve over the 12 weeks of acute antidepressant treatment, with the largest improvement occurring during the first 4 weeks of treatment. Two symptom factors, Morbid Thoughts and Observed Depressed Mood, were rated low at baseline, but did show im-

provement over the course of treatment. The differences between children and adolescents on baseline symptoms were limited, although adolescents had slightly more severe depressive symptoms at baseline (i.e., sleep difficulties and fatigue). Adolescents had higher mean factor scores than children during the early weeks of acute treatment; however, by the end of treatment (at week 12), symptom severity was similar among both age groups. Females reported more weeping and physical complaints than males at baseline; although no significant differences between males and females in terms of time course of symptoms improvement emerged.

Despite the significant improvement in symptoms over 12 weeks, even among remitters, almost half of the group had at least one residual symptom. Impaired school performance, insomnia, irritability, low self-esteem, and appetite change were most common residual symptoms, which is consistent with findings from the TADS study (Kennard et al. 2006). One out of every 5 remitters continued to have impaired school performance, even after 3 months of treatment, and 1 out of 10 continued to report sleep disturbance. These findings are important clinically in helping patients and their family preparing for recovery from a depressive episode.

One important finding is that irritability was as common a presenting symptom for pediatric MDD as depressed mood at baseline, both in children and adolescents. Unlike depressed mood or difficulty having fun, which remitted rather quickly among most depressed pediatric patients, irritability remained one of the most

common continuing symptoms throughout the 12 weeks of acute treatment. Interestingly even though more than 95% of depressed youth reported the hallmark symptoms of MDD—depressed mood (99.5%) and difficulty having fun (96.4%) at baseline—very few continued to report these two symptoms at the end of acute treatment (13% and 12.3%, respectively). Along with irritability, impaired school performance and decreased self-esteem were the three most frequent and difficult-to-remit depressive symptoms.

In light of previous findings of the link between residual symptoms and relapse, clinicians may need to target those difficult to treat depressive symptoms early during the acute treatment. For example, working directly with schools (both teachers and counselors) may help youths to improve school performance. Behavioral interventions and/or pharmacotherapy to improve sleep during early weeks may also be important, as poor sleep potentially increases irritability and fatigue. Cognitive therapy targeting negative self-perception and irritability may be useful in improving outcomes.

The study presented does have several limitations. First, although the result of the study is more generalizable to clinical practice because it was an open-label study, the significant improvement seen during the early weeks may be partially attributed to clinicians' positive bias toward assessing improvement. The improvements, however, were maintained throughout the 12 weeks of acute treatment. Second, 18% (30/168) patients discontinued the study while still depressed, which may have lowered the frequency of depressive symptoms and inflated the improvement. Potentially, residual symptoms could be more frequent. Third, the pattern of improvement may be unique to fluoxetine treatment and may not generalize to other treatment modalities. Future studies with other treatment modalities are warranted to further examine symptom improvement during the acute treatment of pediatric MDD. Fourth, defining remission is often problematic, and there are numerous methods that are used, including cut-off scores, symptom counts, and absence of full disorder. A recent task force from the American College of Neuropsychopharmacology (ACNP) recommended a fairly broad definition of response to include individuals who were free of the two core symptoms (sadness and anhedonia) and had up to three of the remaining DSM-IV criterion symptoms (Rush et al. 2006). Unfortunately, there have been no studies conducted to identify the most useful definition of remission. For this study, remission was defined as a CDRS-R total score ≤ 28 , which is consistent to previously conducted antidepressant trials in this population. One issue with this cutoff score is that a child in remission could be up to 11 points away from being completely asymptomatic (CDRS-R total score = 17). However, this is likely consistent with the recommended criteria by the ACNP task force. Future studies to define remission may be worthwhile. Finally, the current study did not have a placebo control group, and we could not examine symptom changes due to placebo effect. Future studies with a placebo control group could provide valuable information of symptom change patterns to medication versus placebo.

Conclusions

Irritability is as common a presenting symptom as depressed mood and decreased interest among children and adolescents with MDD, but may take longer to remit than the two hallmark symptoms. There are few differences between children and adolescents or between males and females in symptom presentation during acute episodes, although adolescents may be more likely to present

with sleep disturbance and fatigue than younger children. The timing of specific symptom improvement is similar for the two age groups and both genders. Residual symptoms are common, even among remitters at the end of 12 weeks of acute treatment, with impaired school performance, decreased self-esteem, irritability, and insomnia being the most frequent residual symptoms. Because residual symptoms are associated with relapse (Fava et al. 1998; Emslie et al. 2008), these findings highlight the need for clinicians to monitor symptom improvement, and potentially provide additional interventions for the more resistant symptoms.

Disclosures

Dr. Emslie has received research support from Biobehavioral Diagnostics Inc., Bristol-Myers Squibb, Eli Lilly, Forest Laboratories, GlaxoSmithKline (formerly SmithKline Beecham), Novartis, Organon, Repligen, Shire Pharmaceuticals, Somerset, and Wyeth Pharmaceuticals; has served as a consultant for Biobehavioral Diagnostics Inc., Eli Lilly, Forest Laboratories, GlaxoSmithKline, Pfizer, Shire Pharmaceuticals, Validus Pharmaceuticals, and Wyeth Pharmaceuticals; and has been on the Speaker's Bureau for Forest Laboratories and McNeil. Drs. Tao, Nakonezny, Kennard, and Ms. Mayes have no conflicts or interest or financial ties to report.

References

- American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV). Washington (DC): American Psychiatric Association, 1994.
- Benjamini Y, Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Statist Soc* 57:289–300, 1995.
- Emslie GJ, Rush AJ, Weinberg WA, Kowatch RA, Hughes CW, Carmody T, Rintelmann J: A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Arch Gen Psychiatry* 54:1031–1037, 1997.
- Emslie GJ, Heiligenstein JH, Wagner KD, Hoog SL, Ernest DE, Brown E, Nilsson M, Jacobson JG: Fluoxetine for acute treatment of depression in children and adolescents: A placebo-controlled, randomized clinical trial. *J Am Acad Child Adolesc Psychiatry* 41:1205–1215, 2002.
- Emslie GJ, Kennard BD, Mayes TL, Nightingale-Teresi J, Carmody T, Hughes CW, Rush AJ, Tao R, Rintelmann JW: Fluoxetine versus placebo in preventing relapse of major depression in children and adolescents. *Am J Psychiatry* 165:459–467, 2008.
- Fava G, Rafanelli C, Grandi S, Conti S, Belluardo P: Prevention of recurrent depression with cognitive behavioral therapy: Preliminary findings. *Arch Gen Psychiatry* 55:816–820, 1998.
- Guo Y, Nilsson ME, Heiligenstein J, Wilson MG, Emslie G: An exploratory factor analysis of the Children's Depression Rating Scale-revised. *J Child Adolesc Psychopharmacol* 16:482–491, 2006.
- Guy W: ECDEU Assessment Manual for Psychopharmacology. Washington (DC): US Government Printing Office, 1976, pp 76–338.
- Hedges L, Olkin I: Statistical methods for meta-analysis. Orlando (Florida): Academic Press, 1985.
- Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N: Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 36:980–988, 1997.

- Kennard B, Silva S, Vitiello B, Curry J, Kratochvil C, Simons A, Hughes J, Feeny N, Weller E, Sweeney M, Reinecke M, Pathak S, Ginsburg G, Emslie G, March J: Remission and residual symptoms after short-term treatment in the Treatment of Adolescents with Depression Study (TADS). *J Am Acad Child Adolesc Psychiatry* 45:1404–1411, 2006.
- Kenward MG, Roger JH: Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* 53:983–997, 1997.
- Kratochvil C, Emslie G, Silva S, McNulty S, Walkup J, Curry J, Reinecke M, Vitiello B, Rohde P, Feeny N, Casat C, Pathak S, Weller E, May D, Mayes T, Robins M, March J: Acute time to response in the Treatment for Adolescents with Depression Study (TADS). *J Am Acad Child Adolesc Psychiatry* 45:1412–1418, 2006.
- Mallinckrodt CH, Sanger TM, Dube S, DeBrota DJ, Molenberghs G, Carroll RJ, Potter WZ, Tollefson GD: Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. *Biol Psychiatry* 53:754–760, 2003.
- March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, Burns B, Domino M, McNulty S, Vitiello B, Severe J: Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *JAMA* 292:807–820, 2004.
- Padilla M, Algina J: Type I error rates of the Kenward-Roger adjusted degree of freedom F-test for a split-plot design with missing values. *J Modern Appl Statist Methods* 6:66–80, 2007.
- Poznanski E, Mokros H: *Children's Depression Rating Scale-Revised (CDRS-R)*. Los Angeles: WPS, 1996.
- Rush AJ, Kraemer HC, Sackeim HA, Fava M, Trivedi MH, Frank E, Ninan PT, Thase ME, Gelenberg AJ, Kupfer DJ, Regier DA, Rosenbaum JF, Ray O, Schatzberg AF: Report by the ACNP Task Force on response and remission in major depressive disorder. *Neuropsychopharmacology* 31:1841–1853, 2006.
- Tao R, Emslie G, Mayes T, Nakonezny P, Kennard B and Hughes C: Early prediction of acute antidepressant treatment response and remission in pediatric major depressive disorder. *J Am Acad Child Adolesc Psychiatry* 48:71–78, 2009.

Address correspondence to:

Dr. Rongrong Tao

University of Texas Southwestern Medical Center

5323 Harry Hines Blvd.

Dallas, TX 75390-8589

E-mail: rongrong.tao@utsouthwestern.edu