Journal of Affective Disorders 143 (2012) 231–235

Contents lists available at SciVerse ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Preliminary communication

Respondent and item level patterns of response of aripiprazole in the acute treatment of pediatric bipolar I disorder

Robert L. Findling [a,*], Eric A. Youngstrom [b], Joan Zhao [c], Ron Marcus [d], Candace Andersson [e], Robert McQuade [c], Raymond Mankoski [e]

[a] University Hospitals Case Medical Center/Case Western Reserve University, Cleveland, OH, USA
[b] University of North Carolina, Chapel Hill, NC, USA
[c] Otsuka Pharmaceutical Development & Commercialization, Inc., Princeton, NJ, USA
[d] Bristol-Myers Squibb, Wallingford, CT, USA
[e] Bristol-Myers Squibb, Plainsboro, NJ, USA

ARTICLE INFO

Article history:
Received 14 March 2012
Accepted 17 April 2012
Available online 6 October 2012

Keywords:
Bipolar I disorder
Parent-rating
Clinician-rating
General Behavior Inventory
Clinical Global Impression
Young Mania Rating Scale

ABSTRACT

Background: Few studies have evaluated the value of a parent- and subject-rated scale in detecting symptom change in response to pharmacologic treatment.
Methods: This was a post-hoc analysis of data from a 4-week, randomized, double-blind, placebo-controlled study to evaluate which informants detect response to treatment with aripiprazole in pediatric subjects experiencing a mixed or manic episode associated with bipolar I disorder. Efficacy assessments included clinician-rated scales and the parent- and subject-rated 10-item General Behavior Inventory Mania (GBI-M10) and Depression (GBI-D10) scales. Cohen's d quantified effect sizes for total scale scores and individual line items.
Results: Parent-GBI-M10 total, clinician-rated Young Mania Rating Scale (YMRS) total, and Clinical Global Impression–Bipolar Disorder (CGI-BP) Mania scores produced similar effect sizes, suggesting that the parent-GBI-M10 is sensitive to treatment-related improvements in manic symptoms. Aripiprazole improved a broad spectrum of parent-rated mania symptoms; six parent-GBI-M10 line item effect sizes were moderate (>0.5) in at least one of the two aripiprazole treatment arms (10 or 30 mg/day). Subject-completed GBI-M10 line item effect sizes were consistently smaller, indicating that the subjects' experience of treatment effects were less pronounced.
Limitations: Study inclusion/exclusion criteria may limit generalizability of these findings.
Conclusions: Parent ratings of mania severity were in agreement with clinician ratings, indicating that parent-rated assessments can be valuable in detecting symptom change over the course of treatment. These data support the use of the parent-GBI-M10 as an outcome measure in research and clinical settings.

1. Introduction

Pediatric bipolar disorder is a serious and pernicious illness associated with significant functional impairment and marked reductions in quality of life compared with other pediatric illnesses (Freeman et al., 2009). There is a need for effective and well-tolerated treatments, and research in this area has expanded rapidly over the last decade. Recent research has provided important information about treatment options in this patient population, including the use of atypical antipsychotics.

Results from a multicenter, randomized, double-blind, placebo-controlled trial have demonstrated that the atypical antipsychotic aripiprazole is efficacious and generally safe and well tolerated in the treatment of pediatric subjects (aged 10–17 years) with a manic or mixed episode associated with bipolar I disorder (Findling et al., 2009). In this study, using the *a priori* defined primary endpoint of change from baseline to Week 4 in the clinician-rated Young Mania Rating Scale (YMRS) total score, aripiprazole resulted in significantly greater improvement in mania symptoms compared with placebo. Post-hoc analysis of the 11 items comprising the YMRS showed that aripiprazole improved a broad spectrum of discrete symptoms (Mankoski et al., 2011). In addition to evaluating symptoms of mania using the YMRS, this study also included parent- and subject-completed assessments of mania and depression using selected items from the General Behavior Inventory (GBI) scale and clinician-rated assessment of depression using the Children's Depression Rating Scale-Revised (CDRS-R).

* Corresponding author. Tel.: +1 216 844 1717; fax: +1 216 844 5883.
E-mail address: robert.findling@uhhospitals.org (R.L. Findling).

0165-0327/$ - see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jad.2012.04.033

Here, we report findings from a post-hoc analysis of data from these scales that was undertaken to evaluate more precisely: (1) which respondents (clinicians, parents, or the subjects themselves) demonstrated greater sensitivity in detecting drug versus placebo effects, and (2) to expand on prior work undertaken with the YMRS line items (specific to symptoms of mania) to determine which other specific symptoms may respond to treatment with aripiprazole.

2. Methods

2.1. Study design and efficacy assessments

This was a post-hoc analyses of data from a multicenter, randomized, double-blind, placebo-controlled, 4-week trial of aripiprazole (10 or 30 mg/day) versus placebo for the treatment of pediatric subjects (aged 10–17 years) with a Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) diagnosis of an acute manic or mixed episode associated with bipolar I disorder (Findling et al., 2009). Full details of study methodology have been described previously (Findling et al., 2009) and included the requirement for a YMRS (Young et al., 1978) total score ≥ 20 at baseline. After screening and appropriate medication washout, subjects were randomized to target doses of aripiprazole 10 or 30 mg/day, titrated over 5 and 13 days, respectively, or matching placebo for 4 weeks. Written informed consent/assent forms were obtained from all parents/legal guardians and subjects, respectively, prior to inclusion in the study.

The *a priori* primary efficacy outcome measure was the mean change from baseline to endpoint (Week 4) in the clinician-rated YMRS total score. Additional clinician-rated efficacy assessments included the Clinical Global Impression–Bipolar Disorder (CGI-BP) Overall Illness, Depression and Mania scales (Spearing et al., 1997) and the CDRS-R (Poznanski and Mokros, 1995). Additionally, parent- and subject-completed 10-item versions of the GBI-Mania (parent/subject-GBI-M10) (Youngstrom et al., 2008) and Depression (parent/subject-GBI-D10) scales (Danielson et al., 2003; Youngstrom et al., 2001; Youngstrom et al., 2005) assessed severity of mania and depression symptoms. Efficacy evaluations occurred weekly.

2.2. Statistical analysis

Post-hoc analyses included all patients randomized to treatment. Adjusted mean changes from baseline to endpoint (Week 4) for all efficacy scales were evaluated using the last observation carried observation carried forward data set, by analysis of covariance (ANCOVA), with baseline measurements as a covariate and treatment arms as main effects. ANCOVA also compared adjusted mean changes from baseline with endpoint in GBI-M10 and GBI-D10 individual item scores, and CDRS-R individual item scores in each aripiprazole treatment arm versus placebo. Cohen's d effect sizes at Week 4 were calculated as the difference (and its standard deviation) between the aripiprazole (10 or 30 mg/day) arm and placebo treatment arms using the formula $2*t/\sqrt{}$(d.f.), where t is the t statistic derived from a single degree of freedom contrast in an analysis of variance model with treatment as a factor. An effect size of ≥ 0.8 is considered large, an effect size of ≥ 0.5 is considered moderate, and an effect size of ≥ 0.2 is considered small (Cohen, 1988).

3. Results

3.1. Subject characteristics

A total of 296 subjects were randomized to treatment (aripiprazole 10 mg/day, $n=98$; aripiprazole 30 mg/day, $n=99$; placebo, $n=99$). Two hundred and thirty-seven (80.1%) subjects completed the 4-week study period. Full details of subject demographics and disposition have been reported previously (Findling et al., 2009). Baseline demographic characteristics were similar between the treatment groups. The current episode was classified as manic in 40.2% of subjects or mixed in 42.2% of subjects; for the remaining 17.6% of subjects, the most recent episode was classified as 'unknown'.

3.2. Total rating scale scores

Both doses of aripiprazole produced significantly greater improvements in the mean change from baseline to Week 4 on the CGI-BP Overall Illness score and on all mania rating scales (YMRS, CGI-BP Mania, and parent/subject-GBI-M10; all $p<0.05$ vs placebo) (Findling et al., 2009). Both aripiprazole doses (aripiprazole 10 mg/day and aripiprazole 30 mg/day, respectively) produced a moderate-to-large effect of treatment (effect size ≥ 0.6) on CGI-BP Overall Illness score (0.6; 0.9), YMRS (0.6; 0.8), CGI-BP Mania (0.6; 0.9), and parent-GBI-M10 (0.7; 0.6) ratings. Effect sizes were lower for subject-GBI-M10 ratings (0.2). Improvements in depression symptoms with aripiprazole were generally not statistically significantly different from placebo (Findling et al., 2009), and the effect sizes for all four depression rating scales (CDRS-R, parent/subject-GBI-D10 and CGI-BP Depression) were low (range 0.2 to -0.1).

3.3. Parent- and subject-GBI-M10 line item analyses

Six parent-rated GBI-M10 line item effect sizes were moderate (>0.5) for at least one aripiprazole treatment arm (Fig. 1a). For subject-GBI-M10 line items, the effect sizes were consistently smaller across all items (Fig. 1a).

3.4. Parent- and subject-GBI-D10 line item analyses

Effect sizes for both parent- and subject-GBI-D10 line items were consistently small, although parent-rated effect sizes were larger than those rated by the subjects themselves (Fig. 1b). Some line items on both the parent- and subject-rated scales demonstrated an effect size in the opposite direction of the total score.

3.5. CDRS-R line item analyses

The majority of CDRS-R line items showed a positive effect of treatment (Fig. 2). CDRS-R line items with the largest treatment effect were 'sleep disturbance', 'impaired schoolwork', and 'irritability'. As observed for some GBI-D10 line items described above, several CDRS-R line items also demonstrated a notable effect in the opposite direction to the CDRS-R total score.

4. Discussion

There are three major reasons to conduct item-level analyses of total rating scale scores: (a) the effects of an intervention may not be uniform across domains assessed, (b) psychometrically weak items may underestimate treatment effect, and (c) contradictory effects on specific symptoms may cancel each other out in the aggregate score. Each of these issues has the potential to be relevant in clinical trials. In this analysis, mania symptoms, when assessed using the total rating scale scores from a variety of instruments, were significantly improved with aripiprazole treatment, and the treatment effects were moderate-to-large.

However, despite the moderate effect of treatment on the parent-GBI-M10 total score and on other scales, there was

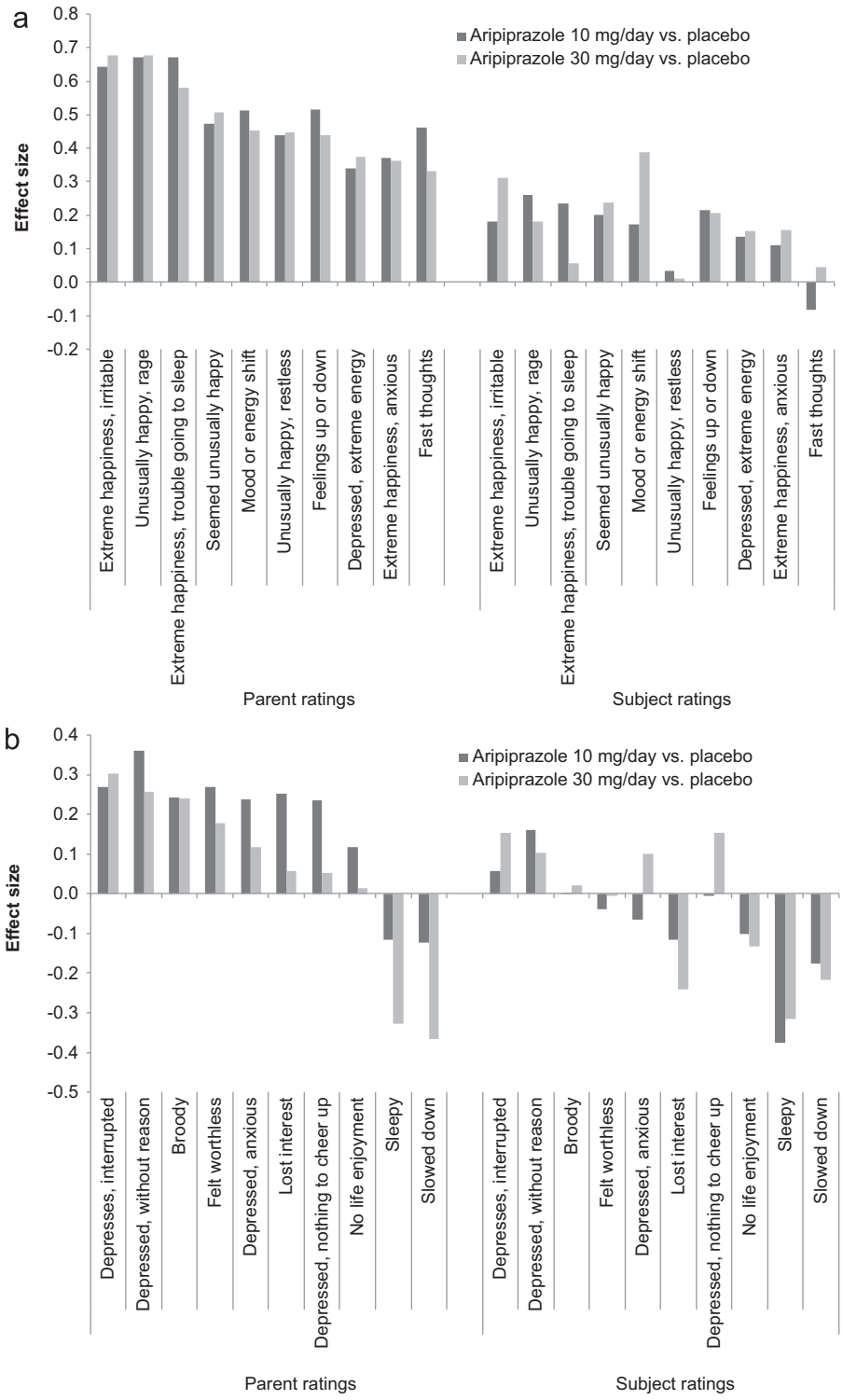

Fig. 1. Effect size (Cohen's d) of aripiprazole treatment on GBI Mania (a) and Depression (b) line items. GBI=General Behavior Inventory.

heterogeneity in symptom improvement detected on specific items of this scale. While items related to elated mood and irritability/aggression showed large treatment effects, the effect sizes for items related to depression and anxiety were generally smaller. This may partly be due to the fact that the subject population being studied had more severe mania symptoms than depressive symptoms at baseline.

With respect to the value of the assessments of treatment effects from different perspectives, the effect sizes observed on the parent-GBI-M10 total score were similar to the clinician-rated YMRS total score and the CGI-BP Mania score, suggesting that the parent-GBI-M10 scale is sensitive to improvements in mania symptoms resulting from pharmacologic treatment. However, compared with both parent- and clinician-rated improvements, the magnitude of treatment effect measured using the subject-GBI-M10 was substantially smaller. Furthermore, there was more heterogeneity in individual symptom improvement among subject ratings of mania symptoms than the corresponding parent ratings. Most notably, symptoms of 'unusually happy and restless' and 'fast thoughts' showed negligible improvement on the subject-rated scale compared with moderate

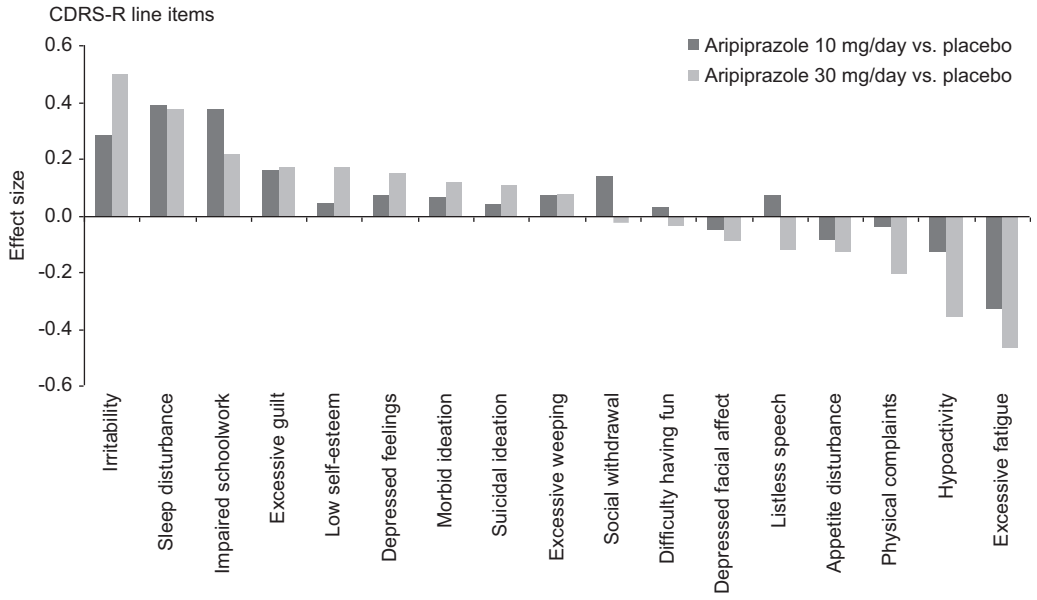

Fig. 2. Effect size (Cohen's d) of aripiprazole treatment on CDRS-R line items. CDRS-R=Children's Depression Rating Scale-Revised.

levels of improvement when rated by parents. The lack of improvement reported by subjects may reflect a lack of insight on the part of the subjects, and is in agreement with previous research that has shown that youths with a bipolar diagnosis tend to under-report manic symptoms compared with their parents (Youngstrom et al., 2004a). Thus, while subject-rated items are useful in the assessment of symptoms, they cannot replace clinician or parent evaluations.

As observed for the mania symptom ratings, there was a lack of agreement between improvements reported by parents and subjects on the GBI-D10. For example, there was a substantial difference between parents and subjects about changes in 'loss of interest'; on average, parent ratings indicated mild improvement, while subject ratings indicated mild worsening. This illustrates how line item analyses can expose interesting, and potentially clinically relevant, differences in perceived response. Also of note, the effect sizes for some items on both the GBI and CDRS-R related to energy levels (fatigue, sleep, and hypoactivity) indicated a negative effect of treatment. It is possible that these ratings were in part reflecting adverse events associated with treatment, as both fatigue and somnolence were commonly reported with short-term aripiprazole treatment (Findling et al., 2009).

To our knowledge, this is the first multicenter study to examine the value of the parent-rated GBI scale in detecting symptom change in response to pharmacologic treatment, and these data provide support for the use of this scale as an outcome measure in research and clinical settings. Given the difficulties with respect to inter-rater reliability when conducting trials in this patient population, parent-reported measures may provide a valuable alternative for assessment of symptom change. Of note, the Child-Mania Rating Scale-Parent version (CMRS-P) has recently been shown to be a valid measure of symptom change associated with pharmacotherapy (West et al., 2011), and parental assessment of symptoms has been shown to have greater validity for the diagnosis of bipolar disorder than subject or teacher reports (Hazell et al., 1999; Youngstrom et al., 2004b; Youngstrom et al., 2006). Thus, our findings extend this observation to indicate that parent assessments can also be valuable in detecting symptom change over the course of treatment.

Limitations include that these analyses were post hoc, and findings should be considered exploratory and descriptive. Inclusion criteria required subjects to have a YMRS total score ≥ 20 at baseline, and thus the population was enriched for having moderate-to-severe mania symptoms. Furthermore, symptoms of both mania and depression were not homogenous among subjects; some items had relatively higher baseline values than others, and differences in the symptom improvement may have in part resulted from psychometric artifacts that might bias effect size estimates to be smaller (i.e., floor effects) or larger (i.e., regression to the mean) depending on baseline symptom severity. Similarly, the findings of greater sensitivity to treatment change for some items over others should not be used to validate these items. It should be considered that there is variation in the psychometric properties of the individual items, with some items having relatively poor sensitivity, and the impact of this on the findings reported here was not evaluated. Ratings are also not completely independent of each other, as clinician ratings depend heavily on information obtained from parents and subjects. Finally, these were analyses of data from a short-term study, and long-term information on symptom improvement would be valuable—especially with regard to improvement in depressive symptoms, which might improve at a slower rate than manic symptoms.

Role of funding source

The study on which the analysis of data in this manuscript was based was supported by Bristol-Myers Squibb Company and Otsuka Pharmaceutical Company, Ltd.

Conflict of interest

Robert L. Findling receives or has received research support, acted as a consultant, received royalties from, and/or served on a speaker's bureau for Abbott, Addrenex, Alexza, American Psychiatric Press, AstraZeneca, Biovail, Bristol-Myers Squibb, Dainippon Sumitomo Pharma, Forest, GlaxoSmithKline, Guilford Press, Johns Hopkins University Press, Johnson & Johnson, KemPharm Lilly, Lundbeck, Merck, National Institutes of Health, Neuropharm, Novartis, Noven, Organon, Otsuka, Pfizer, Physicians' Post-Graduate Press, Rhodes Pharmaceuticals, Roche, Sage, Sanofi-Aventis, Schering-Plough, Seaside Therapeutics, Sepracore, Shionogi, Shire, Solvay, Stanley Medical Research Institute, Sunovion, Supernus Pharmaceuticals, Transcept Pharmaceuticals, Validus, WebMD and Wyeth.

Eric A. Youngstrom has received travel support from Bristol-Myers Squibb and consulted with Lundbeck.

Ron Marcus, Candace Andersson, Raymond Mankoski are employees of Bristol-Myers Squibb.

Joan Zhao and Robert McQuade are employees of Otsuka Pharmaceutical Development & Commercialization, Inc.

Acknowledgments

Editorial support for the preparation of this manuscript was provided by Ogilvy Healthworld Medical Education; funding was provided by Bristol-Myers Squibb. We gratefully acknowledge the patients and their caregivers for their participation in this study.

References

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Erlbaum, Hillsdale.

Danielson, C.K., Youngstrom, E.A., Findling, R.L., Calabrese, J.R., 2003. Discriminative validity of the general behavior inventory using youth report. Journal of Abnormal Child Psychology. 31, 29–39.

Findling, R.L., Nyilas, M., Forbes, R.A., McQuade, R.D., Jin, N., Iwamoto, T., Ivanova, S., Carson, W.H., Chang, K., 2009. Acute treatment of pediatric bipolar I disorder, manic or mixed episode, with aripiprazole: a randomized, double-blind, placebo-controlled study. Journal of Clinical Psychiatry 70, 1441–1451.

Freeman, A.J., Youngstrom, E.A., Michalak, E., Siegel, R., Meyers, O.I., Findling, R.L., 2009. Quality of life in pediatric bipolar disorder. Pediatrics 123, e446–e452.

Hazell, P.L., Lewin, T.J., Carr, V.J., 1999. Confirmation that child behavior checklist clinical scales discriminate juvenile mania from attention deficit hyperactivity disorder. Journal of Paediatrics and Child Health 35, 199–203.

Mankoski, R., Zhao, J., Carson, W.H., Mathew, S.J., Forbes, R.A., 2011. Young mania rating scale line item analysis in pediatric subjects with bipolar I disorder treated with aripiprazole in a short-term, double-blind, randomized study. Journal of Child and Adolescent Psychopharmacology 21, 359–364.

Poznanski, E.O., Mokros, H.B., 1995. Children's Depression Rating Scale, Revised (CDRS-R). Western Psychological Services, Los Angeles, CA.

Spearing, M.K., Post, R.M., Leverich, G.S., Brandt, D., Nolen, W., 1997. Modification of the clinical global impressions (CGI) Scale for use in bipolar illness (BP): the CGI-BP. Psychiatry Research 73, 159–171.

West, A.E., Celio, C.I., Henry, D.B., Pavuluri, M.N., 2011. Child mania rating scale-parent version: a valid measure of symptom change due to pharmacotherapy. Journal of Affective Disorders 128, 112–119.

Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A rating scale for mania: reliability, validity and sensitivity. British Journal of Psychiatry 133, 429–435.

Youngstrom, E.A., Findling, R., Calabrese, J.R., 2004a. Effects of adolescent manic symptoms on agreement between youth, parent, and teacher ratings of behavior problems. Journal of Affective Disorders 82, S5–S16.

Youngstrom, E.A., Findling, R.L., Calabrese, J.R., Gracious, B.L., Demeter, C., DelPorto Bedoya, D., Price, M., 2004b. Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. Journal of the American Academy of Child & Adolescent Psychiatry 43, 847–858.

Youngstrom, E.A., Findling, R.L., Danielson, C.K., Calabrese, J.R., 2001. Discriminative validity of parent report of hypomanic and depressive symptoms on the general behavior inventory. Psychological Assessment 13, 267–276.

Youngstrom, E.A., Frazier, T.W., Demeter, C., Calabrese, J.R., Findling, R.L., 2008. Developing a 10-item mania scale from the parent general behavior inventory for children and adolescents. Journal of Clinical Psychiatry 69, 831–839.

Youngstrom, E.A., Meyers, O., Youngstrom, J.K., Calabrese, J.R., Findling, R.L., 2006. Diagnostic and measurement issues in the assessment of pediatric bipolar disorder: Implications for understanding mood disorder across the life cycle. Development and Psychopathology 18, 989–1021.

Youngstrom, E.A., Meyers, O.I., Demeter, C., Kogos Youngstrom, J., Morello, L., Piiparinen, R., Feeny, N.C., Findling, R.L., Calabrese, J.R., 2005. Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. Bipolar Disorders 7, 507–517.