
BRIEF REPORT

Clinician, Parent, and Child Prediction of Medication or Placebo in Double-Blind Depression Study

Carroll W. Hughes, Ph.D., Graham Emslie, M.D., Robert Kowatch, M.D., Warren Weinberg, M.D., Jeanne Rintelmann, B.A., and A. John Rush, M.D.

To evaluate how well a blind is maintained in a double-blind study. Clinicians (n = 66), parents (n = 62), and depressed child/adolescent subjects (n = 62) predicted whether the patient had been on either placebo or active medication at the end of an eight-week double-blind placebo versus fluoxetine trial. Clinician, patient and parents' guesses as to which treatment they had received were at a chance level based on an overall analysis. However, when clinical response and condition assignment were controlled, all were correctly predicting placebo treatment but not medication treatment. The finding that subjects, parents

and clinicians predict at a chance level is important for double-blind study design integrity. However, clinicians, parents and subjects were accurately predicting placebo treatment when clinical response and the assigned condition were taken into account but not medication. Since they do not know condition however, all remain essentially blinded, and this is an important finding for design and analysis integrity for double-blind studies.

[Neuropsychopharmacology 23:591–594, 2000]

© 2000 American College of Neuropsychopharmacology.

Published by Elsevier Science Inc. All rights reserved

KEY WORDS: Placebo; Double-blind; Childhood depression; MDD; Fluoxetine; Prediction

In a randomized, double blind study of 96 patients, fluoxetine was found to be superior to placebo in the acute phase treatment of major depressive disorder in child and adolescent outpatients (Emslie et al. 1997). One issue of concern to double blind studies is how well the blind is maintained over the course of the study as the investigators become more familiar with the treatment outcomes. First, does the clinician accurately predict whether a patient received either drug or placebo treatment? Second, how well do the subjects or their parents predict which treatment they received following eight

weeks of treatment? And finally, what is the relationship of the clinical response to treatment and the prediction of whether they received medication or placebo? We report here the findings of a study to evaluate these predictions and to test the integrity of the double-blind study.

MATERIALS AND METHODS

Ninety-six child and adolescent outpatients (aged 7–17) with nonpsychotic major depressive disorder were randomized to 20 mg of fluoxetine or placebo (Emslie et al. 1997). All subjects provided written assent, and their parent(s) written consent to participate in the study. The research protocol was approved by the University IRB committee. They were seen weekly for eight consecutive weeks preceded by three evaluation visits and a one-week, single blind placebo run-in (Rintelmann et al. 1996). Subjects were administered a semistructured DSM-III-R-based diagnostic interview (Welner et al.

From the University of Texas Southwestern Medical Center, Dallas, TX.

Address correspondence to: Carroll W. Hughes, Ph.D., Children's Medical Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, Texas 75390-8589.

Received June 18, 1999; revised November 23, 1999; accepted February 1, 2000.

Table 1. Predictions of Medication or Placebo at the End of Eight Weeks of Double-Blind Treatment Were Made By the Clinicians, Subjects, and Their Parents

Individual's Guess	Clinician (n = 66) Both Groups Combined		Patient (n = 62) Both Groups Combined		Parent (n = 62) Both Groups Combined	
	Fluoxetine	Placebo	Fluoxetine	Placebo	Fluoxetine	Placebo
Guessed fluoxetine	19	12	16	10	17	9
Guessed placebo	15	20	18	18	17	19
Total Ns	34	32	34	28	34	28
Chi-square statistic	$\chi^2(1) = 2.24, p = ns$		$\chi^2(1) = 0.81, p = ns$		$\chi^2(1) = 2.01, p = ns$	

1987) to establish that the patient met DSM-III-R criteria for MDD and to identify other concurrent and lifetime psychiatric disorders (Chambers et al. 1985). At the end of the placebo run-in week, the patients were randomized to either fluoxetine treatment or placebo if they still met all of the enrollment criteria. Clinicians remained blind to assignment until the end of the study.

Outcome was measured weekly. Primary outcome measurements were the Clinical Global Improvement (CGI) improvement score ($CGI \leq 2$) and the Children's Depression Rating Scale, Revised (CDRS-R) (Poznanski et al. 1985). Following randomization, each patient was given one capsule of placebo or 20 mg of fluoxetine every morning and seen weekly for eight consecutive weeks. Compliance was monitored by counting returned pills and measuring serum levels. If patients were determined to be responders, they were encouraged to remain blind and on whatever treatment that they had been receiving. In no case was the blind broken for treating/rating clinicians until the study was completed. Non-responders were treated open label for an additional period or switched to an alternative medication.

At the end of eight weeks the clinician, parent(s), and child were asked to independently indicate whether they thought the subject had been on placebo or active medication. This resulted in a subsample ($n = 62$ for the patients, $n = 66$ for the psychiatrists, and $n = 62$ for the parents) of the original population. These independent ratings were implemented after the beginning of the study and a couple of patients and parents failed to

complete the forms at the eight-week evaluation. The prediction data were evaluated using Chi-square.

RESULTS

Clinicians

When subjects on both active treatment and placebo were combined, the clinicians were unable to predict who had been on either placebo or on medication (Table 1). Assuming that the clinical response at eight weeks should influence the prediction, a separate analysis controlled for clinical response and found that clinicians accurately predicted medication for the responders (27 out of 31) and placebo for nonresponders (26 out of 35) (Table 2). When the predictions for the separate groups of medication and placebo were tested, again controlling for clinical response, clinicians' predictions remained at a chance level for the medication group (Table 3) but they were able to predict correctly for placebo based on the designation of responder or nonresponder. However, since the clinician did not know which condition the subject was in, the blind remained essentially intact.

Parents

Similar to the clinicians, when subjects for both active treatment and placebo were combined, parents were not able to predict beyond a chance level what their child had been on (Table 1). However, when the prediction of

Table 2. Analyses by Those who Actually Responded ($CGI \leq 2$) Versus Those That Did Not and the Prediction of Medication or Placebo with Both Conditions Combined

Based on Clinical Response at the End of Eight Weeks	Clinician (n = 66) Both Groups Combined		Patient (62) Both Groups Combined		Parent (n = 62) Both Groups Combined	
	Predicted Active ^a	Predicted Placebo	Predicted Active	Predicted Placebo	Predicted Active	Predicted Placebo
Nonresponder	4	26	4	24	4	24
Responder ^b	27	9	22	12	22	12
Total Ns	31	35	26	36	26	36
Chi-square statistic	$\chi^2(1) = 24.98, p < .001$		$\chi^2(1) = 16.0, p < .001$		$\chi^2(1) = 16.0, p < .001$	

^aActive versus placebo refer to the individual prediction of what the subject was on.

^bResponders were defined *ex post facto* based on a CGI scale improvement rating of 1 or 2.

Table 3. Analyses by Those who Received Fluoxetine Versus Placebo Based on Who Actually Responded (CGI = 2) Versus Those That Did Not and Their Prediction of Medication or Placebo

CLINICIANS	Fluoxetine Group Only Actual ^a			Placebo Group Only Actual		
	Predicted Active ^b	Predicted Placebo	n = 34	Predicted Active	Predicted Placebo	n = 32
Nonresponder	4	8	12	0	18	18
Responder ^c	15	7	22	12	2	14
Chi-square statistic	$\chi^2(1) = 3.8, p^d = \text{ns}$			$\chi^2(1) = 24.69, p < .001$		
SUBJECTS	Fluoxetine Group Only Actual			Placebo Group Only Actual		
	Active Predicted	Placebo Predicted	n = 34	Active Predicted	Placebo Predicted	n = 28
Nonresponder	3	9	12	1	15	16
Responder	13	9	22	9	3	12
Chi-square statistic	$\chi^2(1) = 3.6, p < \text{ns}$			$\chi^2(1) = 14.1, p < .001$		
PARENTS	Fluoxetine Group Only Actual			Placebo Group Only Actual		
	Active Predicted	Placebo Predicted	n = 34	Active Predicted	Placebo Predicted	n = 28
Nonresponder	3	9	12	1	15	16
Responder	14	8	22	8	4	12
Chi-square statistic	$\chi^2(1) = 4.6, p = \text{ns}$			$\chi^2(1) = 11.5, p < .001$		

^aActual refers to the treatment condition (medication vs. placebo) that the subject received.^bActive versus placebo refer to the individual prediction of what the subject was on.^cResponders were defined *ex post facto* based on a CGI scale improvement rating of 1 or 2. ns = non-significant.^dAll analyses based on a $p < .001$ to control for multiple comparisons.

medication and placebo was tested controlling for clinical response, parents of responders predicted that the patient had been on active medication and nonresponders on placebo (Table 2). Separate analyses of actual medication and placebo indicated the above finding was accounted for by the placebo condition where responders were predicted to be on active medication and non-responders were predicted to have been on placebo (Table 3).

Child/Adolescent Subjects

The subjects' guesses were at a chance level similar to the clinicians and parents when subjects on both active treatment and placebo were combined (Table 1). As found for the clinicians and parents, when clinical response was taken into account, responders were predicted to be on medication and nonresponders were predicted to be on placebo (Table 2). When medication and placebo were tested separately, again it was only the placebo group that indicated responders had been predicted as on medication, whereas the non-responders predicted to be on placebo (Table 3).

DISCUSSION

First, prediction of which treatment condition a subject was assigned to, *without taking clinical response in to ac-*

count at eight weeks, remained at a chance level for the clinician, the child/adolescent patient, and the parent suggesting strong support that the blind had been maintained (Table 1). More importantly, when the clinical response was taken into account (Table 2), the clinician tended to attribute the responders to medication and the nonresponders to placebo. Further analysis (Table 3) which separated the subjects into those who were actually on medication versus placebo and then looking at response, found that the prediction based on clinical response for the medication group was at a chance level, whereas the placebo group was accurately identified based on response. However, since the clinicians did not know which group the subjects were in, the blind was clearly maintained.

Similar conclusions can be drawn for the child/adolescent subjects and the parents. Analyses that controlled for the clinical response at eight weeks (responder versus nonresponder) had the same findings for both parent and child subjects. *Based on the patient's clinical response*, subjects and parents were more likely to associate a positive clinical response outcome with receiving active medication (Tables 2 and 3). The opposite tendency was found for the non-responder child and parent who attributed the lack of improvement to placebo treatment. Subjective comments made by both the patients and parents in response to having been asked why they guessed the way they did were reviewed. These

responses clearly indicated that when they thought the individual was doing better it was due to the medication, and if no improvement (i.e., nonresponder) was seen, then it was due to having received placebo.

When the patient's clinical response is controlled for, the prediction by clinicians, parents and child/adolescent subjects remains at a chance level, which is critical for design integrity. Clearly the blind is maintained from the clinical researcher's point of view and suggests that studies that incorporate such procedures to control for experimenter bias are valid. The role that minimal side effects for the active medication played (Emslie et al. 1997), undoubtedly contributes to these findings. These data suggest that concern for the integrity of the blind being broken by clinicians with continued exposure of individuals receiving treatment, even when response to active treatment is positive, would appear to be not merited at least based on this study.

ACKNOWLEDGMENTS

This study was supported by grants MH-39188 (GJE) and MH-41115 (AJR, Dept of Psychiatry, UT Southwestern Medical Center at Dallas) from the NIMH. Dr. Weinberg's work was funded by contributions from Caleb C. and Julia W. Dula Education and Charitable Foundations, Mr. and Mrs. Woody

Hunt (Cimmaron Foundation), Mr. Morton Meyerson, and the Azoulay family.

REFERENCES

- Chambers WJ, Puig-Antich J, Hirsch M, Paez P, Ambrosini PJ, Tabrizi MA, Davies M (1985): The assessment of affective disorders in children and adolescents by semi-structured interview: Test-retest reliability of the K-SADS-P. *Arch Gen Psychiatry* 42:96–702
- Emslie GJ, Rush AJ, Weinberg WA, Kowatch RA, Hughes CW, Carmody T, Rintelmann J (1997): A double-blind, randomized placebo-controlled trial of fluoxetine in depressed children and adolescents. *Arch Gen Psychiatry* 54:1031–1037
- Poznanski EO, Freeman LN, Mokros HB (1985): Children's Depression Rating Scale—Revised (September 1984). *Psychopharmacol Bull* 21:979–989
- Rintelmann JW, Emslie, GJ, Rush AJ, Varghese T, Gullion CM, Kowatch RA, Hughes CW (1996): The effects of extended evaluation on depressive symptoms in children and adolescents. *J Affect Disord* 41:149–156
- Welner Z, Reich W, Herjanic B, Jung KG, Amado H (1987): Reliability, validity, and parent-child agreement studies of the Diagnostic Interview for Children and Adolescent (DICA). *J Am Acad Child Adolesc Psychiatry* 26:649–653