Association Between Patient Beliefs Regarding Assigned Treatment and Clinical Response: Reanalysis of Data From the *Hypericum* Depression Trial Study Group

Justin A. Chen, MD; George I. Papakostas, MD; Soo Jeong Youn, BA; Lee Baer, PhD; Alisabet J. Clain, MS; Maurizio Fava, MD; and David Mischoulon, MD, PhD

ABSTRACT

Objective: To reanalyze data from a 2002 study by the *Hypericum* Depression Trial Study Group to determine whether patients who believed they were receiving active therapy rather than placebo obtained greater improvement, independent of treatment.

Method: Three hundred forty adults with major depressive disorder (according to the Structured Clinical Interview for *DSM-IV*) and baseline scores of ≥ 20 on the 17-item Hamilton Depression Rating Scale (HDRS-17) were randomized to *Hypericum perforatum* 900–1,500 mg/d, sertraline 50–100 mg/d, or placebo and were asked to guess their assigned treatment after 8 weeks. This reanalysis of data was performed from October 1, 2009, to April 15, 2011. The intent-to-treat sample included 207 subjects (mean age = 44 years) who had (1) at least 1 postbaseline visit; (2) adherence data based on serum levels of hyperforin, sertraline, and desmethylsertraline; and (3) guess data. Univariate factorial analysis of variance was used to determine whether treatment assignment affected clinical improvement according to HDRS-17 score and whether this effect was moderated by patient guess of sertraline, *Hypericum*, or placebo. Analysis of covariance was used to determine whether side effects mediated improvement in the context of patient guess and assigned treatment. χ^2 analyses compared response rates (≥ 50% decrease in HDRS-17 score) between the guess groups and between the treatment groups within each guess group.

Results: Assigned treatment had no significant effect on clinical improvement ($P = .65$), but patient guess was significantly associated with improvement ($P < .001$), and treatment and guess interacted significantly ($P = .005$). Among subjects who guessed placebo, clinical improvement was small and did not differ significantly across treatments. Among subjects who guessed *Hypericum*, improvement was large and did not differ significantly across treatments. Among subjects who guessed sertraline, those who received placebo or sertraline had large improvements, but those who received *Hypericum* had significantly less improvement ($P < .001$). Similar findings were obtained for response rates.

Conclusions: Patient beliefs regarding treatment may have a stronger association with clinical outcome than the actual medication received, and the strength of this association may depend upon the particular combination of treatment guessed and treatment received.

J Clin Psychiatry 2011;72(12):1669–1676

Submitted: *July 28, 2010; accepted April 26, 2011.*
Online ahead of print: *October 4, 2011 (doi:10.4088/JCP.10m06453).*
Corresponding author: *David Mischoulon, MD, PhD, 1 Bowdoin Sq, 6th Floor, Massachusetts General Hospital, Boston, MA 02114 (dmischoulon@partners.org).*

Randomized placebo-controlled trials remain the gold standard for clinical research, given their ability to control for potential bias in factors such as investigator alliance and patient expectations.[1–3] The need for placebo controls to establish scientific validity suggests that the act of treatment administration itself can exert significant therapeutic effects that must be distinguished from any intrinsic efficacy of the intervention being tested.[2] While some researchers have argued that many of the effects that have been attributed to placebo could more convincingly be explained by various other factors,[4,5] others have tried to characterize some of the "nonspecific therapeutic effects" that might result from the act of placebo administration, including "attention; communication of concern; intense monitoring; diagnostic procedures; labeling of complaint; and alterations produced in a patient's expectancy, anxiety, and relationship to the illness."[1(p817)]

The beliefs of patients receiving treatment have been shown to exert a significant impact on clinical outcome in areas as varied as asthma, pain control, gastrointestinal motility, and hypertension.[1,6–10] For instance, patients who are told that an inert substance is a bronchodilator or a bronchoconstrictor demonstrate corresponding physiological responses.[6,7] In a landmark study exploring the intersection of cultural beliefs and mortality, Phillips and colleagues[11] found that Chinese Americans, but not whites, die 1.3–4.9 years earlier than normal if they have a combination of disease and a birth year that Chinese astrology and medicine consider ill-fated, and the loss of years is greater when individuals are more strongly attached to Chinese traditions,[11] suggesting that these differences in longevity are not due to genetics but to cultural beliefs.[12]

The impact of patient beliefs on clinical outcome has also been well documented in the psychiatric literature. In 1 study[13] by the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Program, depressed patients who had a greater degree of expectation of improvement at baseline demonstrated a greater probability of responding to treatment. A recent meta-regression analysis[14] of double-blind, randomized clinical trials in major depressive disorder (MDD) showed that a greater likelihood of receiving placebo predicted greater antidepressant-placebo separation at endpoint, presumably because of lower expectations of receiving active treatment. Conversely, studies in which there is a greater chance of receiving active treatment (eg, 3-armed studies comparing 2 active treatments with placebo) may in turn produce a greater placebo response.[14]

The power of patient belief has implications for study design. One investigation[15] surveying undergraduates in an introductory psychology course described 2 hypothetical clinical trials: 1 in which a medication would be compared with placebo, and another in which the same medication would be compared against another medication. Respondents reported a significantly higher expected likelihood and magnitude of improvement in the active comparator trial as opposed to the placebo-controlled trial.[15] The same investigators performed meta-analyses[16,17] of randomized controlled trials comparing antidepressants to a placebo or active comparator in geriatric outpatients with MDD and found that response and remission rates were higher in comparator as opposed to placebo-controlled trials. Even beliefs unrelated to a particular medication or intervention have been shown to exert an influence on clinical response; for instance, a recent study[18] concluded that those with strong beliefs in a personal and concerned God have an increased likelihood of response to treatment for depression.

Despite the influence that patients' beliefs can exert on clinical outcome, few studies have directly queried patients about their beliefs regarding their assigned treatments. In 1 clinical trial[19] of bupropion for smoking cessation, treatment arm guess was associated with quit rates, although that study found evidence that the blind integrity had been violated, most likely due to symptomatic improvement that may have influenced guess patterns.[19] Along these lines, in a controlled study of fluoxetine for treatment of depression, Hughes and colleagues[20] found that clinical response and assigned treatment impacted correct prediction of placebo treatment but not of medication.

We sought to further explore the impact of patients' beliefs by reanalyzing data from a large 2002 double-blind randomized trial[21] of St John's wort (*Hypericum perforatum* extract) for treatment of MDD. In this NIMH-funded investigation by the *Hypericum* Depression Trial Study Group,[21] no significant differences in improvement in the 17-item Hamilton Depression Rating Scale (HDRS-17)[22] total score or response rates were found between the 3 treatment arms: *Hypericum*, sertraline, and placebo. To ensure blinding, subjects were directly asked after 8 weeks of acute treatment which of the 3 possible interventions they believed they were receiving. This information presents an opportunity to examine whether a patient's belief that he or she is receiving an active treatment might be associated with a stronger response.

Given the availability of the original database, graciously provided by the *Hypericum* Depression Study Group investigators, we reanalyzed the data with a specific focus on whether patients who believed they were receiving sertraline or St John's wort showed greater improvement in symptoms compared to those who believed they were receiving placebo. We hypothesized that patient belief that he or she was receiving an active therapy would be associated with greater clinical improvement, regardless of actual treatment assignment. As an ancillary question, we examined whether age, gender, ethnicity, depression severity, and duration of index depressive episode predicted treatment guess.

Clinical Points

- In this reanalysis, patient beliefs regarding treatment were more strongly associated with clinical outcome than the actual medication received.
- The degree of influence of patient beliefs on clinical outcome may depend upon the particular combination of treatment guessed and treatment received.
- Results of this and future studies may have implications for how clinicians utilize the potentially therapeutic effects of patient beliefs in clinical practice.

METHOD

Detailed methods for the parent trial (clinicaltrials.gov Identifier: NCT00005013) have been described elsewhere.[21] In brief, the study was conducted at 12 academic or community psychiatry research clinics in the United States. Three hundred forty outpatients aged ≥ 18 years who met criteria for current major depressive episode according to the Structured Clinical Interview for *DSM-IV*[23] were recruited. Inclusion criteria included a screening and baseline HDRS-17 score ≥ 20 and a Global Assessment Scale (GAS)[24] score ≤ 60.

Subjects were randomized equally to receive double-blind treatment with *Hypericum*, sertraline, or placebo. After a 1-week placebo run-in, patients entered an 8-week acute phase. At baseline, patients received *Hypericum* (900 mg/d), sertraline (50 mg/d), or placebo. Daily doses could be increased to 1,200 mg/d of *Hypericum*, 75 mg/d of sertraline, or placebo equivalent after weeks 3 or 4 and to 1,500 mg/d of *Hypericum*, 100 mg/d of sertraline, or placebo equivalent at week 6 if the Clinical Global Impressions-Severity of Illness scale (CGI-S)[25] score was ≥ 4 at week 3 or ≥ 3 at weeks 4 or 6. At the end of the acute phase (week 8), responders (≥ 50% decrease in HDRS-17 score, HDRS-17 score ≤ 12, Clinical Global Impressions-Improvement scale [CGI-I][25] score ≤ 2) could enter an 18-week double-blind continuation phase. The HDRS-17 and other outcome measures, including the Beck Depression Inventory (BDI),[26] were assessed at all visits. At the week 8 visit (or at study exit in cases of early discontinuation), patients were asked to guess which of the 3 treatment assignments they had received.

This reanalysis, carried out from October 1, 2009, to April 15, 2011, focused on the acute treatment phase only (baseline to week 8). For the intent-to-treat (ITT) sample, we selected all subjects with at least 1 postbaseline visit for which guess data were available. In a previous examination of this database by Vitiello and colleagues,[27] they found that 1 in 6 patients taking placebo had measurable hyperforin and that 1 in 6 patients taking *Hypericum* had no measurable hyperforin, whereas all patients taking sertraline had measurable sertraline or desmethylsertraline. Since nonadherence would

© Copyright 2011 Physicians Postgraduate Press, Inc.

Table 1. Treatment Guess Patterns Among Patients Receiving Sertraline (n = 75), *Hypericum* (n = 66), or Placebo (n = 66) in the Intent-to-Treat Sample (N = 207)

Treatment Guessed[b]	Total Guesses (N = 207), n	Treatment Received[a] Sertraline, % (n) of Guesses	*Hypericum*, % (n) of Guesses	Placebo, % (n) of Guesses
Sertraline	71	37 (26)	30 (21)	34 (24)
Hypericum	90	46 (41)	28 (25)	27 (24)
Placebo	46	17 (8)	44 (20)	39 (18)

[a]Comparison of correct guess rates in each *treatment* group: sertraline vs *Hypericum*: $\chi^2_1 = 0.16$, $P = .73$; sertraline vs placebo: $\chi^2_1 = 0.89$, $P = .37$; *Hypericum* vs placebo: $\chi^2_1 = 1.69$, $P = .27$.
[b]Comparison of correct guess rates in each *guess* group: sertraline vs *Hypericum*: $\chi^2_1 = 1.43$, $P = .24$; sertraline vs placebo: $\chi^2_1 = 0.08$, $P = .85$; *Hypericum* vs placebo: $\chi^2_1 = 1.82$, $P = .24$.

most likely invalidate reported "beliefs," we excluded from our analysis any nonadherent patients and patients with no available adherence data.

Descriptive statistics were obtained to determine how many patients guessed each treatment, as well as which actual treatments were received. Significance was assessed by χ^2, comparing the different groups against each other (sertraline vs *Hypericum*, sertraline vs placebo, and *Hypericum* vs placebo).

Improvement in HDRS-17 score was determined for each guess group and assessed for significance by the paired *t* test. We similarly examined the other main outcome measures, including the CGI-S and CGI-I, the GAS, and the BDI. Comparisons between the 3 treatment groups within each guess group were made by univariate factorial analysis of variance and by 2-sample independent *t* tests.

Treatment response was defined as a 50% or greater improvement in HDRS-17 total score at study completion. Remission was defined as a final HDRS-17 score of < 8. The relationships between treatment guess and response and remission rates were calculated for each group by χ^2 analysis. Effect sizes were calculated by the odds ratio (OR) of response.

We calculated an adverse effect score on the basis of the number of adverse effects reported by each patient that were attributed to treatment, per data in the parent study database. To determine whether adverse effects had an impact on the association between guess and clinical improvement in the ITT sample as a whole, analysis of covariance was carried out with change in HDRS-17 score as the dependent variable, treatment guess and treatment assignment as the fixed factors, and medication-related adverse effect score as the covariate. We similarly examined the CGI-S, CGI-I, GAS, and BDI.

To examine the interaction between treatment guess and treatment assignment with regard to treatment response, χ^2 analyses were carried out comparing response rates between the 3 treatment groups within each guess group. Effect sizes were calculated using partial η-squared (η_p^2).

On the basis of the Explanatory Model for Depression Questionnaire by Bann and colleagues,[28] we also examined whether having internal or external locus of control (LOC) scores was predictive of improvement. Mean LOC scores between the different guess groups were compared by the 2-sample *t* test.

Logistic regression and χ^2 analysis were used to examine whether depression severity at baseline, duration of the index episode, age, gender, and ethnicity were predictors of treatment guessed.

For all analyses, 2-tailed statistical significance was set at $P < .05$. All calculations were performed with SPSS, version 17.0 (SPSS Inc, Chicago, Illinois).

RESULTS

The ITT sample consisted of 207 patients (mean ± SD age was 44 ± 13 years, 64% were female) with available adherence data and reported treatment guess at the conclusion of the acute treatment phase. The sample was composed of whites (77%; n = 159), blacks (11%; n = 22), Hispanics (6%; n = 12), Asians (5%; n = 11), Native Americans (0.5%; n = 1), and unknown ethnicity (1%; n = 2).

Treatment guess patterns are illustrated in Table 1. None of the comparisons of correct guess rates between guess groups or treatment groups reached significance ($P > .05$ for all). All of the 71 patients who guessed sertraline completed at least 7 weeks of treatment, and 67 completed the full 8 weeks; all 90 patients who guessed *Hypericum* completed at least 7 weeks of treatment, and 88 completed 8 weeks; all 46 patients who guessed placebo completed at least 6 weeks of treatment, 45 completed 7 weeks, and 41 completed 8 weeks.

As an initial analysis of the association between treatment guess and clinical improvement, we examined the change in HDRS-17 score in each treatment guess group (Table 2). All 3 treatment guess groups improved significantly with regard to HDRS-17 score ($P < .001$ for all), with the *Hypericum* guess group having the strongest improvement. Significant differences in clinical improvement were found when comparing the *Hypericum* guess group ($P < .001$) or the sertraline guess group ($P = .001$) with the placebo guess group. No significant differences were found in improvement between the *Hypericum* guess group and the sertraline guess group ($P = .25$).

To confirm the observed HDRS-17 improvement pattern, we also examined changes in the GAS, the BDI, and the CGI-S, as well as overall improvement (CGI-I), in each guess group. All these measures improved significantly in each guess group ($P < .001$ for all; data not shown). When we compared these 4 secondary outcome measures between the different guess groups, there was no significant difference in improvement between sertraline and *Hypericum* guessers ($P > .05$ for all outcome measures); both sertraline guessers ($P < .02$ for all measures) and *Hypericum* guessers ($P \leq .001$ for all measures) showed a significantly greater improvement compared to placebo guessers (data not shown).

To examine for any interaction between guess and treatment, while controlling for the impact of adverse effects, analysis of covariance was performed with change in HDRS-17

Table 2. 17-Item Hamilton Depression Rating Scale (HDRS-17) Score Improvement by Treatment Guess in the Intent-to-Treat Sample (N = 207)

Treatment Guessed	Baseline HDRS-17, Mean (SD)	Final HDRS-17, Mean (SD)	Change in HDRS-17, Mean (SD)	Significance[a]		
				t	*df*	*P*
Sertraline (n = 71)	22.77 (2.76)	11.23 (7.07)	11.55 (6.30)	15.44	70	< .001
Hypericum (n = 90)	22.74 (2.49)	10.12 (5.48)	12.62 (5.58)	21.46	89	< .001
Placebo (n = 46)	22.80 (2.77)	15.87 (7.78)	6.93 (7.60)	6.19	45	< .001

[a]One-way between-groups analysis of variance for all 3 choices: $F_2 = 12.74$, $P < .001$, $\eta_p^2 = 0.111$. For treatment guess of sertraline vs *Hypericum*: $t_{159} = -1.14$, $P = .25$, $\eta_p^2 = 0.008$; for sertraline vs placebo: $t_{115} = 3.56$, $P = .001$, $\eta_p^2 = 0.099$; for *Hypericum* vs placebo: $t_{70} = 4.49$, $P < .001$, $\eta_p^2 = 0.155$.

Table 3. 17-Item Hamilton Depression Rating Scale (HDRS-17) Score Improvement by Guess Group and Treatment Assignment in the Intent-to-Treat Sample (N = 207)

Treatment Guessed	Improvement by Treatment Received Within Guess Group			Significance and Effect Size[a,b]			
	Sertraline (n = 75), Mean (SD)	*Hypericum* (n = 66), Mean (SD)	Placebo (n = 66), Mean (SD)	*F*	*df*	*P*	η_p^2
Sertraline (n = 71)	12.12 (4.91)	7.48 (7.05)	14.50 (5.18)	8.69	2	< .001	0.204
Hypericum (n = 90)	12.07 (4.94)	14.72 (4.44)	11.37 (7.11)	2.66	2	.076	0.058
Placebo (n = 46)	9.13 (8.49)	7.15 (7.26)	5.72 (7.78)	0.56	2	.577	0.025

[a]Significance values are from univariate analysis of variance comparing improvement between the 3 treatment groups within each guess group.
[b]Analysis of covariance was performed with change in HDRS-17 as dependent variable, guess and treatment as fixed factors, and adverse effect load as covariate for the entire intent-to-treat sample: adverse effects: $F_1 = 0.65$, $P = .42$, $\eta_p^2 = 0.003$; guess: $F_2 = 12.39$, $P < .001$, $\eta_p^2 = 0.118$; treatment: $F_2 = 0.43$, $P = .65$, $\eta_p^2 = 0.005$; guess by treatment: $F_4 = 3.80$, $P = .005$, $\eta_p^2 = 0.076$.

score as the dependent variable, treatment and guess as fixed factors, and adverse effects secondary to treatment as the covariate. The adverse effect score for each subject was based on the total number of reported adverse effects attributed to treatment and ranged from 0 to 34 for the entire sample (mean ± SD = 5.3 ± 5.9, median = 4.0, mode = 0). Table 3 illustrates the degree of improvement in HDRS-17 scores on the basis of treatment assignment and treatment guess. Assigned treatment in itself had no significant association with clinical improvement ($F_2 = 0.43$, $P = .65$, $\eta_p^2 = 0.005$); however, patient guess was significantly associated with clinical improvement ($F_2 = 12.39$, $P < .001$, $\eta_p^2 = 0.118$), and there was a significant interaction between treatment and guess ($F_4 = 3.80$, $P = .005$, $\eta_p^2 = 0.076$), indicating different degrees of improvement depending upon the particular combination of treatments received and guessed. When we examined each treatment guess group separately, we observed significant treatment-related differences in clinical improvement only in the sertraline guess group ($P < .001$) (Table 3).

When we similarly examined the secondary outcome measures, CGI-S, CGI-I, GAS, and BDI, we found a pattern comparable to that for HDRS-17. Assigned treatment in itself had no significant association with clinical improvement for any of these measures ($P > .05$ for all), but patient guess was significantly associated with clinical improvement for all 4 measures ($P \leq .002$ for all). Interaction between treatment and guess was also consistent, with significant results for CGI-S ($P = .004$), CGI-I ($P = .020$), and BDI ($P = .037$), but GAS ($P = .12$) missed significance (data not shown).

Both groups that guessed active treatments had significantly greater response rates (based on ≥ 50% improvement in HDRS-17 score) than the placebo guess group (*Hypericum*: $\chi^2 = 23.85$, $P < .001$, Fisher $P < .001$, OR = 6.71 [95% CI, 2.98–14.92]; sertraline: $\chi^2 = 11.94$, $P = .001$, Fisher $P = .001$, OR = 4.10 [95% CI, 1.79–9.09]). There was also a modest but nonsignificant advantage for subjects who guessed *Hypericum* versus sertraline ($\chi^2 = 2.22$, $P = .136$, Fisher $P = .144$, OR = 1.63 [95% CI, 0.86–3.11]). Response rate patterns based on treatment assignment and guess are illustrated in Table 4. Within each guess group, significant differences in response rates between the assigned treatments were observed only for subjects who guessed sertraline ($P = .006$)—subjects who received sertraline had an OR = 4.00 (95% CI, 1.17–13.69; $P = .039$) as compared to those who received *Hypericum*, and subjects who received *Hypericum* had an OR = 0.13 (95% CI, 0.04–0.50; $P = .003$) as compared to those who received placebo.

Remission rates were 31% (n/n = 22/71) for subjects who guessed sertraline, 37% (n/n = 33/90) for those who guessed *Hypericum*, and 22% (n/n = 10/46) for those who guessed placebo. No remission-rate comparisons between treatment guess groups yielded significant differences, although there was a trend favoring subjects who guessed *Hypericum* versus placebo ($\chi^2 = 3.14$, $P = .077$, Fisher $P = .083$).

Adverse effects data were available for 195 subjects in our ITT sample. Subjects who guessed they were receiving sertraline had a mean ± SD adverse effect score of 5.0 ± 5.4; those who guessed *Hypericum* had a score of 6.6 ± 6.5; and those who guessed placebo had a score of 3.0 ± 4.6. Univariate analysis of variance for the entire sample found no significant association between adverse effect scores (independent variable) and guess (dependent variable) ($F_{23} = 1.02$,

Table 4. Response Rates[a] by Guess Group and Treatment Assignment in the Intent-to-Treat Sample (N = 207)

Treatment Guessed	Response by Guess Group, % (n/n)	Response by Treatment Received Within Guess Group: Sertraline (n = 75), % (n/n)	*Hypericum* (n = 66), % (n/n)	Placebo (n = 66), % (n/n)	Significance[b]: χ^2	*df*	*P*
Sertraline (n = 71)	56 (40/71)	62 (16/26)	29 (6/21)	75 (18/24)	10.27	2	.006
Hypericum (n = 90)	68 (61/90)	63 (26/41)	80 (20/25)	63 (15/24)	2.37	2	.305
Placebo (n = 46)	24 (11/46)	25 (2/8)	30 (6/20)	17 (3/18)	0.93	2	.628

[a]Response based on ≥ 50% decrease in 17-Item Hamilton Depression Rating Scale score.
[b]Significance values are from χ^2 analysis comparing response rates between the 3 treatment groups within each guess group. Following is the χ^2 for response rates between each guess group in the entire intent-to-treat sample: $\chi^2_2 = 23.81$, $P < .001$.

Table 5. Adverse Effects Load[a] by Guess Group and Treatment Assignment in the Intent-to-Treat Sample (N = 195)[b]

Treatment Guessed	Adverse Effects Load by Guess Group, Mean (SD)	Adverse Effects Load by Treatment Received Within Guess Group: Sertraline (n = 72), Mean (SD)	*Hypericum* (n = 64), Mean (SD)	Placebo (n = 59), Mean (SD)	Significance and Effect Size[c,d]: *F*	*df*	*P*	η_p^2
Sertraline (n = 67)	5.0 (5.4)	5.6 (5.6)	4.8 (5.5)	4.5 (5.3)	0.28	2	.76	0.009
Hypericum (n = 86)	6.6 (6.5)	7.6 (6.5)	6.3 (7.6)	5.2 (4.7)	1.00	2	.37	0.024
Placebo (n = 42)	3.0 (4.6)	6.9 (8.6)	2.1 (2.8)	2.4 (3.2)	3.28	2	.048	0.144

[a]Adverse effect scores were calculated on the basis of the number of adverse effects reported by each patient that were attributed to treatment, per data in the parent study database.
[b]Includes subjects with available adverse effects data and guess data.
[c]Significance values are from univariate analysis of variance comparing adverse effects scores between the 3 treatment groups within each guess group.
[d]One-way between-groups analysis of variance for the sample as a whole, with adverse effects as independent variable and guess as dependent variable: $F_{23} = 1.02$, $P = .44$, $\eta_p^2 = 0.121$. For treatment guess of sertraline vs placebo: $t_{107} = 1.98$, $P = .050$; for sertraline vs *Hypericum*: $t_{151} = -1.64$, $P = .102$; for *Hypericum* vs placebo: $t_{109} = 3.63$, $P < .001$.

$P = .44$, $\eta_p^2 = 0.121$). Within each guess group, the difference in adverse effect score between treatment groups reached significance only in the placebo guess group ($P = .048$) (Table 5). Comparisons of adverse effect scores between pairs of guess groups reached significance for *Hypericum* versus placebo ($P < .001$) but not for sertraline versus *Hypericum* ($P = .102$), whereas sertraline versus placebo barely achieved significance ($P = .050$). In each guess group, subjects who received sertraline had the highest adverse effect score, followed by the *Hypericum* group, with the placebo group having the lowest adverse effect score (Table 5).

Mean (SD) internal and external LOC scores were compared between guess groups. For sertraline guessers, mean internal LOC score was 48.0 ± 17.1, and mean external LOC score was 26.0 ± 11.4. For *Hypericum* guessers, mean internal LOC score was 47.0 ± 17.7, and mean external LOC score was 22.3 ± 10.7. For placebo guessers, mean internal LOC score was 47.2 ± 18.5, and mean external LOC score was 25.0 ± 11.2. We found a significant difference only in external LOC score between the sertraline guessers and *Hypericum* guessers ($t_{146} = 2.06$, $P = .041$). All other comparisons between guess groups were nonsignificant ($P > .05$).

When we compared mean (SD) LOC scores between responders and nonresponders in each guess group, we found a significant difference only in internal LOC score among those who guessed *Hypericum* (responders: mean = 44.3 ± 17.2; nonresponders: mean = 53.0 ± 17.6; $t_{82} = 2.14$, $P = .036$). For sertraline guessers (responders: mean = 44.6 ± 18.5; nonresponders: mean = 52.2 ± 14.5; $t_{59} = 1.8$, $P = .082$), we found a trend toward significance. All other comparisons of LOC scores between guess groups were nonsignificant ($P > .05$).

Age, gender, ethnicity, depression severity, and duration of index depressive episode were not significant predictors—on the basis of logistic regression and χ^2 analysis (not shown)—of treatment guessed by patients.

DISCUSSION

In our reanalysis of a large randomized controlled trial investigating the comparative efficacy of St John's wort for treatment of MDD, the magnitude of clinical improvement was strongly associated with subjects' guesses of which treatment they had received. Subjects who guessed an active treatment (sertraline or *Hypericum*) improved significantly more than those who guessed placebo, and those who guessed *Hypericum* had a modest advantage over those guessing sertraline. Considering that the original study[21] failed to demonstrate superiority for any of the actual 3 treatment arms in terms of the 2 primary outcome measures, our findings are particularly striking. There was a generally consistent pattern of improvement on the major outcome scales (HDRS-17, CGI-S, CGI-I, GAS, and BDI), with improvement favoring guessers of both active treatments over placebo and with no significant differences between *Hypericum* and sertraline guessers. This pattern was more robust and uniform than that observed in the main outcome report,[21] in which the only significant difference in outcome measures based

on treatment was observed for CGI-I, on which sertraline outperformed placebo, or in the related report by Bann and colleagues[28] on patient beliefs, in which the BDI and GAS behaved differently than the HDRS-17 and CGI.

The observed guess-related improvement may depend upon the particular combination of treatment guessed and treatment received. Subjects who both guessed and received *Hypericum* demonstrated the greatest magnitude of improvement (with a mean ± SD change in HDRS-17 of 14.72 ± 4.44) (Table 3), whereas subjects who both guessed and received placebo did the worst (mean ± SD change in HDRS-17 of 5.72 ± 7.78) (Table 3). The sertraline-treated group showed less of an HDRS-17 spread across its guess groups. This pattern within the treatment guess groups suggests a possible interaction between treatment guess and assignment in the cases of *Hypericum* and placebo.

Our findings are consistent with previous research suggesting that patients' beliefs may impact clinical outcome. To our knowledge, this is the first time that such a strong association between belief regarding ongoing medication assignment and outcome in depression has so explicitly been demonstrated, in part because previous studies have not directly queried subjects about assigned treatment belief. Bann and colleagues,[28] who also utilized the *Hypericum* Depression Trial Study Group's data set, sought to explore the association between patients' beliefs regarding the causes of their illness and clinical outcome. Using an instrument called the Explanatory Model for Depression Questionnaire, the authors found that patients who believed the causes of their depression were outside their control (external LOC) were less likely to improve.[28] We were curious as to whether the presence of a strong external LOC might be associated with clinical response in the guess groups, particularly in those who guessed placebo, but our findings did not suggest any clear association. Future studies should explore potential associations between Explanatory Model for Depression Questionnaire scores, treatment guess, and outcome.

Because beliefs regarding treatment may be influenced by the presence of side effects, we examined the relationship between subjects' guesses and the presence of adverse effects specifically attributed to treatment. In the sample as a whole, the group that guessed *Hypericum* surprisingly exhibited the highest mean ± SD total adverse effect score (6.6 ± 6.5), while those who guessed placebo had the lowest score (3.0 ± 4.6), regardless of actual treatment received (Table 5). When adverse effect data were examined by treatment group, the pattern was more in line with what we expected, with subjects who received sertraline having the highest adverse effect scores and those who received placebo having the lowest (Table 5). Only among placebo guessers did the adverse effect loads differ significantly between treatment groups ($P = .048$), suggesting that the presence or absence of side effects may have impacted patient guesses to some degree, although the findings for the sample as a whole did not suggest a significant effect overall (Tables 3 and 5).

The current reanalysis sheds new light on the parent study, which generated significant attention because of its suggestion that 2 popular and highly prescribed antidepressants might be no more effective than placebo.[29] The study's lead author was quoted as saying, "This is a classic illustration of the placebo effect confounding antidepressant trials."[29(p26)] Indeed, psychiatric research has increasingly questioned the efficacy of antidepressant medications compared to placebo, as exemplified by the recent report by Turner et al[30] that suggests that high proportions of unpublished negative trials of selective serotonin reuptake inhibitors have led to underestimates of the magnitude of the placebo response rate in the literature.

Our study suggests the possibility that patient expectations and beliefs may have a stronger relationship with clinical outcome than previously thought. Understanding the importance of patients' beliefs may be especially relevant in the area of complementary and alternative medicine, since individuals who opt for these therapies often do so on the basis of their own strongly held values, convictions, and philosophical orientations toward health and life[31] and are also more likely to believe that psychological factors can affect health.[32] As Kaptchuk has observed, "Rather than specific biological consequences, which epidemiologists designate as 'fastidious efficacy,' alternative medicine may administer an especially large dose of what anthropologists call 'performative efficacy,'" which relies on "the power of belief, imagination, symbols, meaning, expectation, persuasion, and self-relationship."[1(pp817–818)]

Given that the original study was conceived and advertised as a clinical trial of St John's wort for MDD, the subjects who ultimately enrolled may have been especially enthusiastic about alternative therapies in general and St John's wort in particular. This conclusion may be supported by the relatively higher numbers of subjects who expressed a belief that they were receiving *Hypericum* (n = 90), as opposed to sertraline (n = 71) or placebo (n = 46). Such a "selection bias" may have contributed to a closer association of patients' beliefs with clinical outcome than the actual treatment. These results could have important implications for research in depression, complementary and alternative medicine, and, perhaps, placebo-controlled trials in general, as it is difficult to imagine a reasonable way of controlling for all the beliefs of study subjects. On the other hand, these findings reflect real-world practice in which patients bring various expectations and beliefs to their treatment, which, as we have shown, are likely to be associated with their clinical outcomes.

Consistent with prior research, we failed to find demographic differences between subjects who guessed that they were receiving *Hypericum* versus those who guessed they were receiving sertraline or placebo.[1,33] However, this finding may reflect the relatively small sample size, especially with regard to ethnic breakdown of the study population. If future studies identify a correlation between gender, age, or ethnicity and predicted guesses, this finding could impact

  © COPYRIGHT 2011 PHYSICIANS POSTGRADUATE PRESS, INC.

the way in which studies are conducted, and particularly the way in which subjects for these studies are recruited and selected. Along these same lines, we found no association between depression severity or duration of the index episode and treatment guess.

This reanalysis has several limitations. Treatment guesses were elicited only after subjects had completed the acute treatment phase, and reported beliefs may have therefore been influenced by symptomatic improvement or other unknown factors. Patients who felt better at 8 weeks might be more likely to guess that they were receiving an active treatment that they wished for, suggesting that outcome predicts treatment guess rather than the converse. Prior studies with designs and outcomes similar to ours have also been limited by the difficulty of proving causality, especially in a retrospective analysis.[19] Our study did not specifically measure *expectancy*, which would more accurately describe subjects' anticipatory thoughts regarding treatment prior to beginning the study, but rather belief about treatment assignment at the end of the trial. While interpretations of clinical trials generally assume that the effects of medication and expectancy are additive, prospective studies of patient expectancy and outcome are needed, as it is unknown how medication, expectancy, and manipulation of expectancy interact with one another.[34] Future investigations would benefit from asking subjects about the reasoning behind their treatment guess, and/or by eliciting treatment guess at various stages of the study, since belief is likely to be an evolving process that may be influenced by various factors, such as side effects and/or rapidity of response, and is therefore likely to change over the course of treatment. Finally, this study did not explore the role of physician belief in clinical outcome. Data on physician belief were collected in the parent investigation and are currently being analyzed for a follow-up report.

In conclusion, our findings suggest that patient beliefs regarding their treatment may have a stronger association with clinical outcome compared to the actual medication received, although this may depend upon the particular combination of treatment guess and assignment. These findings need to be replicated in other studies, including those that do not use *Hypericum* or other natural remedies, as our sample may be less representative of individuals with depression in general. Further research into factors contributing to the so-called "performative efficacy" of medication administration is warranted. Results of future studies may have implications for utilizing the potentially therapeutic effects of patient beliefs in clinical practice.

Drug names: bupropion (Wellbutrin, Aplenzin, and others), fluoxetine (Prozac and others), sertraline (Zoloft and others).

Author affiliations: Depression Clinical and Research Program, Department of Psychiatry, Massachusetts General Hospital (MGH), Harvard Medical School, Boston.

Potential conflicts of interest: **Dr Papakostas** has served as a consultant for AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Evotec, Inflabloc, Jazz, Otsuka, Pamlab, Pfizer, Pierre Fabre, Shire, and Wyeth; has received honoraria from AstraZeneca, Bristol-Myers Squibb, Eli Lilly, Evotec, GlaxoSmithKline, Inflabloc, Jazz, Lundbeck, Otsuka, Pamlab, Pfizer, Pierre Fabre, Shire, Titan, and Wyeth; has received research support from Bristol-Myers Squibb, Forest, the National Institute of Mental Health (NIMH), Pamlab, Pfizer, and Ridge Diagnostics (formerly known as Precision Human Biolaboratories); and has served on the speakers bureaus of Bristol-Myers Squibb and Pfizer. **Dr Fava** has received research support from Abbott, Alkermes, Aspect Medical Systems, AstraZeneca, BioResearch, BrainCells, Bristol-Myers Squibb, CeNeRx BioPharma, Cephalon, Clinical Trial Solutions, Clintara, Covidien, Eli Lilly, EnVivo, Euthymics Bioscience, Forest, Ganeden, GlaxoSmithKline, Icon Clinical Research, i3 Innovus/Ingenix, Johnson & Johnson, Lichtwer, Lorex, NARSAD, National Center for Complementary and Alternative Medicine (NCCAM), National Institute on Drug Abuse, NIMH, Novartis, Organon, Pamlab, Pfizer, Pharmavite, PhotoThera, Roche, RCT Logic, Sanofi-Aventis, Shire, Solvay, Synthelabo, and Wyeth-Ayerst; has served as an advisor and consultant to Abbott, Affectis, Alkermes, Amarin, Aspect Medical Systems, AstraZeneca, Auspex, Bayer, Best Practice Project Management, BioMarin, Biovail, BrainCells, Bristol-Myers Squibb, CeNeRx BioPharma, Cephalon, Clinical Trials Solutions, CNS Response, Compellis, Cypress, Dainippon Sumitomo, DOV, Edgemont, Eisai, Eli Lilly, ePharmaSolutions, EPIX, Euthymics Bioscience, Fabre-Kramer, Forest, Genomind, GlaxoSmithKline, Grunenthal, i3 Innovus/Ingenix, Janssen, Jazz, Johnson & Johnson, Knoll, Labopharm, Lorex, Lundbeck, MedAvante, Merck, Methylation Sciences, Naurex, Neuronetics, NextWave, Novartis, Nutrition 21, Orexigen, Organon, Otsuka, Pamlab, Pfizer, PharmaStar, Pharmavite, PharmoRx Therapeutics, Precision Human Biolaboratory, Prexa, PureTech Ventures, PsychoGenics, Psylin Neurosciences, Rexahn, Ridge Diagnostics, Roche, RCT Logic, Sanofi-Aventis, Sepracor, Servier, Schering-Plough, Solvay, Somaxon, Somerset, Sunovion, Supernus, Synthelabo, Takeda, Tal Medical, Tetragenex, TransForm, Transcept, Vanda, and Wyeth-Ayerst; has received speaking and publishing honoraria from Adamed, Advanced Meeting Partners, American Psychiatric Association, American Society of Clinical Psychopharmacology, AstraZeneca, Belvoir Media Group, Boehringer-Ingelheim, Bristol-Myers Squibb, Cephalon, CME Institute/Physicians Postgraduate Press, Eli Lilly, Forest, GlaxoSmithKline, Imedex, Novartis, Organon, Pfizer, PharmaStar, MGH Psychiatry Academy/PRIMEDIA Healthcare, MGH Psychiatry Academy/Reed Elsevier, UBC, and Wyeth-Ayerst; holds equity in Compellis; holds a patent for Sequential Parallel Comparison Design (SPCD); holds a patent for research and licensing of SPCD with RCT Logic, Lippincott Williams & Wilkins, and World Scientific Publishing; has a patent application for a combination of azapirones and bupropion in major depressive disorder; and has received copyright royalties for the MGH Cognitive and Physical Functioning Questionnaire, Sexual Functioning Inventory, Antidepressant Treatment Response Questionnaire, Discontinuation-Emergent Signs and Symptoms scale, and SAFER Criteria Interview diagnostic instruments. **Dr Mischoulon** has received research support from Nordic Naturals, Fisher Wallace, and Ganeden; has served as a consultant to Bristol-Myers Squibb; has received writing honoraria from Pamlab; has received royalties from Back Bay Scientific for PMS Escape and from Lippincott Williams & Wilkins for the textbook *Natural Medications for Psychiatric Disorders: Considering the Alternatives*, for which he is co-editor; and has received honoraria from Reed Medical Education (a company working as a logistics collaborator for MGH Psychiatry Academy) (the education programs conducted by MGH Psychiatry Academy were supported through independent medical education grants from pharmaceutical companies cosupporting programs along with participant tuition; commercial entities currently supporting the MGH Psychiatry Academy are listed on the Academy's website at www.mghcme.org). **Drs Chen** and **Baer** and **Mss Youn** and **Clain** report no competing interests in relation to the subject of this article.

Funding/support: This reanalysis was supported by the MGH Depression Clinical and Research Program's clinical practice revenues. The original clinical study "Treatment of Major Depression With St John's Wort (*Hypericum*)" (clinicaltrials.gov Identifier: NCT00005013) was conducted and supported by NIMH (grant N01 MH070007) and NCCAM in collaboration with the study investigators.

Disclaimer: This article reflects the views of the authors and may not reflect the opinions or views of all the study investigators, NIMH, or NCCAM.

Previous presentations: This work was presented at the Mysell Lecture and Research Day at Harvard Medical School, March 24, 2010, Boston, Massachusetts; at MGH Clinical Research Day, May 20, 2010, Boston, Massachusetts; at the 163rd Annual Meeting of the American Psychiatric

Association, May 22–26, 2010, New Orleans, Louisiana; and at the 50th Annual Meeting of the New Clinical Drug Evaluation Unit (NCDEU), June 14–17, 2010, Boca Raton, Florida.

Acknowledgments: The authors thank the members of the *Hypericum* Depression Trial Study Group for allowing us access to the original database from their clinical trial. Information about the database and how to obtain access to it is available at http://www.nimh.nih.gov/trials/datasets/nimh-procedures-for-requesting-data-sets.shtml.

REFERENCES

1. Kaptchuk TJ. The placebo effect in alternative medicine: can the performance of a healing ritual have clinical significance? *Ann Intern Med*. 2002;136(11):817–825.
2. Sullivan MD. Placebo controls and epistemic control in orthodox medicine. *J Med Philos*. 1993;18(2):213–231.
3. Rains JC, Penzien DB. Behavioral research and the double-blind placebo-controlled methodology: challenges in applying the biomedical standard to behavioral headache research. *Headache*. 2005;45(5):479–486.
4. Kienle GS, Kiene H. The powerful placebo effect: fact or fiction? *J Clin Epidemiol*. 1997;50(12):1311–1318.
5. Hróbjartsson A, Gøtzsche PC. Is the placebo powerless? an analysis of clinical trials comparing placebo with no treatment. *N Engl J Med*. 2001;344(21):1594–1602.
6. Luparello TJ, Leist N, Lourie CH, et al. The interaction of psychologic stimuli and pharmacologic agents on airway reactivity in asthmatic subjects. *Psychosom Med*. 1970;32(5):509–513.
7. Butler C, Steptoe A. Placebo responses: an experimental study of psychophysiological processes in asthmatic volunteers. *Br J Clin Psychol*. 1986;25(pt 3):173–183.
8. Sternbach RA. The effects of instructional sets on autonomic responsivity. *Psychophysiology*. 1964;1(1):67–72.
9. Hashish I, Hai HK, Harvey W, et al. Reduction of postoperative pain and swelling by ultrasound treatment: a placebo effect. *Pain*. 1988;33(3):303–311.
10. Agras WS, Horne M, Taylor CB. Expectation and the blood-pressure-lowering effects of relaxation. *Psychosom Med*. 1982;44(4):389–395.
11. Phillips DP, Ruth TE, Wagner LM. Psychology and survival. *Lancet*. 1993;342(8880):1142–1145.
12. Moerman DE, Jonas WB. Deconstructing the placebo effect and finding the meaning response. *Ann Intern Med*. 2002;136(6):471–476.
13. Sotsky SM, Glass DR, Shea MT, et al. Patient predictors of response to psychotherapy and pharmacotherapy: findings in the NIMH Treatment of Depression Collaborative Research Program. *Am J Psychiatry*. 1991;148(8):997–1008.
14. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? a meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol*. 2009;19(1):34–40.
15. Rutherford BR, Rose SA, Sneed JR, et al. Study design affects participant expectations: a survey. *J Clin Psychopharmacol*. 2009;29(2):179–181.
16. Rutherford BR, Roose SP, Sneed J. Mind over medicine: the influence of expectations on antidepressant response. *J Am Psychoanal Assoc*. 2009;57(2):456–460.
17. Sneed JR, Rutherford BR, Rindskopf D, et al. Design makes a difference: a meta-analysis of antidepressant response rates in placebo-controlled versus comparator trials in late-life depression. *Am J Geriatr Psychiatry*. 2008;16(1):65–73.
18. Murphy PE, Fitchett G. Belief in a concerned god predicts response to treatment for adults with clinical depression. *J Clin Psychol*. 2009;65(9):1000–1008.
19. Schnoll RA, Epstein L, Audrain J, et al. Can the blind see? participant guess about treatment arm assignment may influence outcome in a clinical trial of bupropion for smoking cessation. *J Subst Abuse Treat*. 2008;34(2):234–241.
20. Hughes CW, Emslie G, Kowatch R, et al. Clinician, parent, and child prediction of medication or placebo in double-blind depression study. *Neuropsychopharmacology*. 2000;23(5):591–594.
21. *Hypericum* Depression Trial Study Group. Effect of *Hypericum perforatum* (St John's Wort) in major depressive disorder: a randomized controlled trial. *JAMA*. 2002;287(14):1807–1814.
22. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23(1):56–62.
23. First MB, Spitzer RL, Gibbon M, et al. *Structured Clinical Interview for* DSM-IV *Axis I Disorders, Patient Edition, Version 2.0*. New York, NY: Biometrics Research Department, New York State Psychiatric Institute; 1995.
24. Endicott J, Spitzer RL, Fleiss JL, et al. The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry*. 1976;33(6):766–771.
25. Guy W. *ECDEU Assessment Manual for Psychopharmacology, Revised*. *US Dept Health, Education, and Welfare publication (ADM) 76-338*. Rockville, MD: National Institute of Mental Health; 1976:218–222.
26. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4:561–571.
27. Vitiello B, Shader RI, Parker CB, et al. Hyperforin plasma level as a marker of treatment adherence in the National Institutes of Health *Hypericum* Depression Trial. *J Clin Psychopharmacol*. 2005;25(3):243–249.
28. Bann CM, Parker CB, Bradwejn J, et al. Assessing patient beliefs in a clinical trial of *Hypericum perforatum* in major depression. *Depress Anxiety*. 2004;20(3):114–122.
29. Rosack J. Controversy erupts over study of St John's Wort efficacy. *Psychiatr News*. 2002;37(10):26–48.
30. Turner EH, Matthews AM, Linardatos E, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252–260.
31. Astin JA. Why patients use alternative medicine: results of a national study. *JAMA*. 1998;279(19):1548–1553.
32. Vincent C, Furnham A, Willsmore M. The perceived efficacy of complementary and orthodox medicine in complementary and general practice patients. *Health Educ Res*. 1995;10(4):395–405.
33. Wolf S, Doering CR, Clark ML, et al. Chance distribution and placebo reactor. *J Lab Clin Med*. 1957;49(6):837–841.
34. Kirsch I. Are drug and placebo effects in depression additive? *Biol Psychiatry*. 2000;47(8):733–735.

Editor's Note: We encourage authors to submit papers for consideration as a part of our Early Career Psychiatrists section. Please contact Marlene P. Freeman, MD, at mfreeman@psychiatrist.com.

 © COPYRIGHT 2011 PHYSICIANS POSTGRADUATE PRESS, INC.