A Double-blind, Randomized Trial of St John's Wort, Fluoxetine, and Placebo in Major Depressive Disorder

Maurizio Fava, MD, Jonathan Alpert, MD, PhD,* Andrew A. Nierenberg, MD,* David Mischoulon, MD, PhD,* Michael W. Otto, PhD,* John Zajecka, MD,† Harald Murck, MD,‡ and Jerrold F. Rosenbaum, MD**

Objective: This study looks to compare the antidepressant efficacy and safety of a standardized extract of St John's wort with both placebo and fluoxetine.

Method: After a 1-week single-blind washout, patients with major depressive disorder diagnosed by Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* were randomized to 12 weeks of double-blind treatment with LI-160 St John's wort extract (900 mg/d), fluoxetine (20 mg/d), or placebo. The 17-item Hamilton Rating Scale for Depression (HAMD-17) was the primary efficacy measure, and analysis of covariance was used to compare differences in end point HAMD-17 scores across the 3 treatment groups, treating the baseline HAMD-17 as the covariate.

Results: One hundred thirty-five patients (57% women; mean age, 37.3 ± 11.0; mean HAMD-17, 19.7 ± 3.2) were randomized to double-blind treatment and were included in the intent-to-treat analyses. Analysis of covariance analyses showed lower mean HAMD-17 scores at end point in the St John's wort group (n = 45; mean ± SD, 10.2 ± 6.6) compared with the fluoxetine group (n = 47; 13.3 ± 7.3; $P < 0.03$) and a trend toward a similar finding relative to the placebo group (n = 43; 12.6 ± 6.4; $P = 0.096$). There was also a trend toward higher rates of remission (HAMD-17 <8) in the St John's wort group (38%) compared with the fluoxetine group (30%) and the placebo group (21%). Overall, St John's wort appeared to be safe and well tolerated.

Conclusion: St John's wort was significantly more effective than fluoxetine and showed a trend toward superiority over placebo. A (25%) smaller than planned sample size is likely to account for the lack of statistical significance for the advantage (indicating a moderate effect size, $d = 0.45$) of St John's wort over placebo.

(*J Clin Psychopharmacol* 2005;25:441–447)

In Western Europe, extracts from St John's wort (*Hypericum perforatum*) have been in therapeutic use for the treatment of depressed mood for hundreds of years. *Hypericum* extracts are rather complex mixtures, whose exact compositions depend on the extraction method applied. However, both hyperforin and the naphthodianthrones hypericin and its 2-hydroxymethyl derivative, pseudohypericin, often referred to as total hypericin, are considered to be the most specific ingredients.[1] However, the precise identity of the efficacious constituents of *Hypericum* extracts is not known, nor are the details of the pharmacological mechanism of antidepressant action.

Among the lay public and physicians in Europe, St John's wort (*H. perforatum*) is perceived to be effective for mild-to-moderate depression, with a benign profile of adverse drug events, including lack of sedation.[1] In support of this view, a meta-analysis of 23 European randomized trials of St John's wort extract in 1757 outpatients with mild to moderately severe depressive disorders[2] showed that, in placebo-controlled trials, St John's wort was effective in 55% of the subjects (n = 408), whereas placebo was effective in 22% of the subjects (n = 422). The same meta-analysis showed that, in the 6 active comparator trials involving a single preparation of St John's wort, a 64% response rate was observed (n = 158), compared with the 58% response rate for the antidepressant comparators (n = 159). On the other hand, Shelton et al[3] commented that most or perhaps all of the trials used in this meta-analysis had serious methodological flaws, such as a relatively short duration of the trial and failure to use standardized diagnostic practices or symptom rating instruments, thereby undermining confidence in these results.

The findings of 2 large, multicenter, placebo-controlled US studies of St John's wort have been recently published. The first study compared the efficacy and safety of a well-characterized *H. perforatum* (St John's wort) extract (LI-160) with placebo in outpatients with major depressive disorder (MDD) recruited in 11 academic medical centers in the United States.[3] The study enrolled 200 adult outpatients (mean age, 42.4 years; 67% women; 86% white) with MDD and a baseline 17-item Hamilton Rating Scale for Depression (HAMD-17)[4] score of at least 20. After a 1-week single-blind run-in of placebo, patients were randomly assigned to receive either St John's wort extract (n = 98; 900 mg/d for 4 weeks, increased to 1200 mg/d in the absence of an adequate response thereafter) or placebo (n = 102) for 8 weeks. The random coefficient analyses for the primary efficacy measure (HAMD-17) showed significant effects for time but not for treatment or time-by-treatment interaction. The proportion of participants achieving an a priori definition of response also did not differ significantly between groups (34% for St John's

*Depression Clinical and Research Program, Massachusetts General Hospital, Boston, MA; †Department of Psychiatry, Rush-Presbyterian–St Luke's Medical Center, Chicago, IL and ‡Lichtwer Pharma AG, Wallenroder Straße 8-10, D-13435 Berlin, Germany.

Received August 4, 2003; accepted after revision May 5, 2005.

Address correspondence and reprint requests to Maurizio Fava, MD, Depression Clinical and Research Program, Massachusetts General Hospital, 15 Parkman St, ACC 812, Boston, MA 02114. E-mail: mfava@partners.org.

ISSN: 0271-0749/05/2505-0441

DOI: 10.1097/01.jcp.0000178416.60426.29

wort and 22% for placebo), but the number reaching remission (HAMD-17 score of ≤7 and a Clinical Global Impression [CGI] score of 1 or 2) was significantly higher with St John's wort than with placebo ($P = 0.02$). However, these rates of remission were very low in the full intention-to-treat analysis (14/98 [14%] vs. 5/102 [5%], respectively). St John's wort was safe and well tolerated, and headache was the only adverse event (AE) that occurred with greater frequency with St John's wort than with placebo (39/95 [41%] vs. 25/100 [25%], respectively).[3] Although the authors of the study concluded that, in their study, St John's wort was not effective for treatment of major depression,[3] this conclusion is premature given that it is not uncommon for studies to fail to differentiate placebo and antidepressant treatment among agents that eventually achieve the 2 pivotal trials necessary for Food and Drug Administration approval.[5] For example, an analysis of 5 antidepressant agents found that, of 39 trials filed with the Food and Drug Administration, only 14% of these trials found active drug superior to placebo for all primary and secondary outcome measures of depression, and only 44% of studies differentiated drug from placebo on the first depression item on the HAMD.[6] Accordingly, it is fully within norm expectation for approved antidepressants for some studies to fail to differentiate active from placebo treatment.

Similar findings were obtained in the recently published multicenter study[7] which tested in MDD the efficacy and safety of the same well-characterized *H. perforatum* extract (LI-160) used in the Shelton study. The study was conducted in 12 academic and community psychiatric research clinics in the United States. Adult outpatients (n = 340) recruited between December 1998 and June 2000 with MDD and a baseline total score on the HAMD-17 of at least 20 were randomly assigned to receive St John's wort, placebo, or sertraline (as an active comparator) for 8 weeks. Based on clinical response, the daily dose of St John's wort could range from 900 to 1500 mg and that of sertraline from 50 to 100 mg. On the 2 primary outcome measures, neither sertraline nor St John's wort was significantly different from placebo. The random regression parameter estimate for mean (SE) change in HAMD-17 total score from baseline to week 8 (with a greater decline indicating more improvement) was −9.20 (0.67) (95% confidence interval, −10.51 to −7.89) for placebo vs. −8.68 (0.68) (95% confidence interval, −10.01 to −7.35) for St John's wort ($P = 0.59$) and −10.53 (0.72) (95% confidence interval, −11.94 to −9.12) for sertraline ($P = 0.18$). Full response (HAMD-17 score of ≤8 and a CGI Improvement [CGI-I] scale score of 1 or 2) occurred in 32% of the placebo-treated patients versus 24% of the St John's wort–treated patients ($P = 0.21$) and 25% of sertraline-treated patients ($P = 0.26$). Adverse effect profiles for St John's wort and sertraline differed relative to placebo. Although this study failed to support the efficacy of St John's wort in moderately severe major depression, the equally poor performance of sertraline in this trial underscores the variability in efficacy findings that are not uncommon among approved agents. The accompanying editorial by Kupfer and Frank[8] pointed out that the only 2 controlled studies of St John's wort in the United States had failed to reject the null hypothesis but had not yet provided a definitive perspective on the potential utility of St John's wort. In addition, despite the fact that the European studies had suggested efficacy of St John's wort in mild-to-moderate depression, both studies included only patients with moderate to severe MDD, as a HAMD-17 score of 20 or higher was required for entry into these 2 studies. Accordingly, there is need for further study of the antidepressant efficacy and safety of a standardized extract of St John's wort (*H. perforatum*, LI-160) relative to both placebo and fluoxetine in a population of outpatients with mild to moderately severe MDD.

MATERIALS AND METHODS

Design

This was a 12-week, randomized, active- and placebo-controlled, parallel-group, double-blind study conducted in patients with a diagnosis of MDD. The study was conducted in 2 sites (Boston and Chicago). Upon enrollment, all eligible patients were required to have discontinued any previous psychoactive medication for a specified period of time to qualify for entry into the single-blind placebo washout period of 7 days (Table 1). Patients still meeting the diagnostic inclusion criterion after the washout period entered the 12-week, acute, double-blind therapy phase, receiving 1 of the following 3 randomized double-blind treatments: (1) LI-160 St John's wort extract (300 mg thrice a day; daily dose, 900 mg), (2) fluoxetine 20 mg every day, or (3) placebo. The study was planned as a single-center trial with an anticipated total of approximately 180 enrolled patients at the Depression Clinical and Research Program of the Massachusetts General Hospital in Boston from September 1998 to September 2000. Because of difficulties in completing enrollment by the target date, a second site

TABLE 1. Treatment Schedule—Double-dummy Technique

			LI-160 Tablet		Fluoxetine Capsule	
			300 mg	Placebo	20 mg	Placebo
Washout		Placebo, TID	—	3	—	1
Acute treatment	Treatment 1	LI-160, 300 mg TID	3	—	—	1
	Treatment 2	Fluoxetine, 20 mg QD	—	3	1	—
	Treatment 3	Placebo	—	3	—	1

Figures denote number of tablets/capsules per day. The schedule for the acute therapy applies also for the responders during follow-up. TID indicates thrice a day; QD, every day.

(Department of Psychiatry, Rush-Presbyterian–St Luke's Medical Center, Chicago, Ill) was added to the Boston site. Following an interim analysis of the study performed in September 2001, the sponsor (Lichtwer Pharma AG, Berlin, Germany) opted to close the study and to proceed with the final analyses carried out on a sample size smaller (n = 135) than the one originally planned.

Eligible Patients

Patients were primarily recruited from general advertising and clinician referrals. Meeting all criteria listed below, patients of either sex and any ethnic origin with a diagnosis of MDD were included in the study. The patients were required to meet the following inclusion criteria:

a. Age, 18 to 65 years.
b. Current experience of a major depressive episode according to *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* of at least 2 weeks' duration.
c. A HAMD-17 total score $\geq$16 at both screen and baseline.
d. Negative pregnancy test within 5 days before study start in women of childbearing potential (nonchildbearing potential was defined as postmenopause for at least 1 year or surgical sterilization or hysterectomy at least 3 months before study start).
e. Use of adequate contraception in women of childbearing potential.
f. Readiness and ability on the part of the patient to comply with the physician's instructions and to fill out the self-report measures in connection with their examination at the study visits.
g. Written informed consent.

Exclusion criteria were the following:

a. Pregnancy, lactation, or nonuse of medically accepted means of contraception in women of childbearing potential.
b. Current serious suicidal or homicidal risk (according to investigator's judgment).
c. Serious or unstable medical illness including cardiovascular, hepatic, renal, respiratory, endocrine, neurological, or hematologic disease.
d. History of seizure disorder.
e. One or more of the following *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* diagnoses: organic mental disorders; substance use disorders, including alcohol, active within the last 6 months; schizophrenia; delusional disorder; psychotic disorders not elsewhere classified; bipolar disorder; or antisocial personality disorder.
f. History of multiple adverse drug reactions or allergy to the study drugs.
g. Mood-congruent or mood-incongruent psychotic features.
h. Any of the following treatments at baseline or within the specified time frame before baseline: other psychotropic drugs, 14 days; other investigational psychotropic drug, 40 days; fluoxetine, 40 days; or any other investigational drug, 1 month.
i. Unacceptability to discontinue or likelihood to need medication that is prohibited as concomitant treatment during the study.
j. Clinical or laboratory evidence of hypothyroidism.
k. Failure to respond during the course of current major depressive episode to at least 2 adequate antidepressant trials, defined as 8 weeks or more of treatment with either imipramine 150 mg or greater (or its tricyclic equivalent), phenelzine 60 mg or greater (or its monoamine oxidase inhibitor equivalent), or fluoxetine 20 mg or greater (or its selective serotonin reuptake inhibitor equivalent).
l. Any other condition which, in the investigator's judgment, may pose a significant risk to the patient's health or may decrease the chances of obtaining reliable data to achieve the objectives of the study.
m. Mental condition rendering the patient unable to understand the nature, scope, and possible risks of the study.
n. History or suspicion of unreliability, poor cooperation, or noncompliance with medical treatment.

Randomization Criteria

After the placebo washout period, all patients were randomized unless they met either 1 of the following criteria at the randomization visit: HAMD-17 less than 16 or a reduction in HAMD-17 by 25% or greater as compared with the screening visit. During the period of randomized treatment, double-blind conditions were maintained using the following double-dummy schedule (see Table 1).

For each treatment group, the following daily schedule was applied: morning, 1 tablet plus 1 capsule; midday, 1 tablet; and evening, 1 tablet. Patients were instructed to take study medications shortly before their meals. The total daily dose of active treatment was 900 mg of *Hypericum* extract (St John's wort) or 20 mg of fluoxetine.

Medications

Hypericum extract (St John's wort) (LI-160) and matching placebo were provided by Lichtwer Pharma, whereas fluoxetine and matching placebo were provided by the Massachusetts General Hospital research pharmacy, and fluoxetine capsules manufactured by Eli Lilly and Co were used. The *Hypericum* extract was standardized to between 0.12% and 0.28% hypericin, and the entire study supply came from 1 batch. The study was conducted under an investigational new drug application filed by Lichtwer Pharma.

The frequency of visits was as follows: screen (7–14 days before baseline), baseline (day 0), visit 1 (day 7 + 2 days), visit 2 (day 14 ± 2 days), visit 3 (day 28 ± 2 days), visit 4 (day 42 ± 2 days), visit 5 (day 56 ± 2 days), and visit 6 (day 84 ± 2 days).

Instruments and Efficacy Measures

On enrollment, all patients were administered the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* Axis I disorders.[9] All patients were administered at each visit the following instruments:

1. The 28-item HAMD. The 28-item HAMD[4] allows the assessment of the HAMD-17 scale, which is the primary efficacy measure of the study. This instrument was completed by the investigators based on their assessment of the patient's depressive symptoms.
2. The CGI Severity of Illness (CGI-S) and improvement (CGI-I) scales. The patients' CGI[10] was based on the investigators' assessment of the patients depressive symptoms.
3. The Beck Depression Inventory (BDI).[11] This self-rated instrument was completed by the patients based on their assessment of the severity of depression.

Safety Measures

At each visit, the following vital signs were monitored as safety parameters: blood pressure and heart rate. The following laboratory assessments (complete blood count, chemistries, and urinalysis) were performed at screen and week 12 or end point. All laboratory values as well as all electrocardiographic recordings were assessed by the investigator as to whether they had to be judged as AEs. For the recording of AEs, patients were required to report spontaneously any AEs as well as the time of onset and intensity of these events.

End Points

The primary efficacy measure for this study was the HAMD-17. This instrument was administered at each visit by the clinical study investigators. The primary efficacy end point for this study was the final HAMD-17 total score after 12 weeks of randomized treatment on end point. Secondary efficacy measures included the CGI-S, the CGI-I, and the BDI.

Evaluation of Safety

Safety analysis was performed for all patients who took at least 1 dose of study medication (all-subjects-treated group). Measures of safety included reported AEs, routine laboratory assessments (performed at the first pretreatment visit and at week 8), physical examinations (performed at the first pretreatment visit and at week 8), and the patient's vital signs, including blood pressure, heart rate, body temperature, respiration rate, and body weight (assessed at both pretreatment visits and at all study weeks that included a clinic visit).

STATISTICAL METHODS

Statistical significance was set at $P \leq 0.05$. Intent-to-treat (ITT) analyses were conducted. The primary ITT included all subjects who completed their baseline visit, were deemed eligible to continue the study, and were therefore randomized to double-blind treatment.

Baseline Clinical and Demographic Variables

Pairwise differences among treatment groups for clinical and demographic variables of the ITT sample were compared with analyses of variance (ANOVA) and Fisher exact tests.

Primary Efficacy End Point

The primary objective of the study was to evaluate the efficacy of LI-160 compared with placebo and fluoxetine for decreasing depressive symptoms in patients with MDD. One-way analysis of covariance (ANCOVA) was used to assess differences in HAMD-17 depression severity at end point (week 12). The covariate was the baseline HAMD-17, and the final HAMD-17 served as the dependent variable. These analyses were performed on the ITT populations; the last-observation-carried-forward approach was used for missing data and withdrawals.

Secondary Efficacy End Points

The secondary efficacy end points were (1) the proportion of remitters after 12 weeks of treatment (visit 6) or end point, with remission = HAMD-17 score less than 8, was assessed with pairwise Fisher exact tests; (2) the change from baseline after 12 weeks of treatment (visit 6) or end point in CGI-S was assessed with pairwise ANCOVA, using the CGI-S at baseline as the covariate; (3) the CGI-I score after 12 weeks of treatment (visit 6) or end point was assessed with pairwise ANOVA; and (4) the BDI change from baseline in the total score after 12 weeks of treatment (visit 6) or end point was assessed with pairwise ANCOVA, using the BDI at baseline as the covariate. These analyses were performed on the ITT population.

Safety

The proportion of patients reporting AEs in the ITT sample was compared across the 3 treatment groups using pairwise Fisher exact tests.

RESULTS

The ITT population included 135 outpatients with MDD (57% women; 19% minorities; mean age, 37.3 ± 11.0 years; mean HAMD-17, 19.7 ± 3.2; mean CGI-S score, 4.2 ± 0.6) who were randomized to double-blind treatment. Of those randomized, 101 (75%) were enrolled in the Boston site, and 34 (25%) in the Chicago site. There was no significant difference in pairwise comparisons of age among patients randomized to St John's wort (mean, 37.4 ± 11.7), fluoxetine (mean, 36.7 ± 9.6), or placebo (mean, 37.8 ± 12.0). Similarly, there were no significant differences in sex ratio among patients randomized to St John's wort (53% women), fluoxetine (53% women), or placebo (65% women). As shown in Tables 2 and 3, there were no significant differences at baseline in the mean HAMD-17, CGI-S, and BDI scores.

The rate of completion of the 12-week, double-blind, placebo-controlled phase was as follows: 60% (27/45) for St John's wort, 51% (24/47) for fluoxetine, and 49% (21/43) for placebo. As shown in Table 2, St John's wort treatment was associated with a significantly ($P < 0.05$) greater decrease in HAMD-17 scores compared with fluoxetine at all postbaseline visits, except visit 5 (week 8). There was also a trend ($P < 0.1$) toward a significantly greater reduction in HAMD-17 score compared with placebo at visit 1 (week 1) and visit 6/end point. There were also nonsignificantly

TABLE 2. Mean HAMD-17 Scores in the ITT Sample of Outpatients With MDD (n = 135)

	Baseline		Visit 1		Visit 2		Visit 3		Visit 4		Visit 5		Visit 6 or End Point	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
St John's wort	19.6	3.5	14.7*	6.1	12.3*	5.9	11.5*	6.9	10.2*	7.2	10.3	5.9	10.2*	6.6
	n = 45		n = 43		n = 42		n = 35		n = 34		n = 32		n = 45	
Placebo	19.9	2.9	17.1†	5.6	13.9	6.0	13.9	7.6	11.5	5.6	11.4	5.5	12.6‡	6.4
	n = 43		n = 39		n = 40		n = 35		n = 31		n = 28		n = 43	
Fluoxetine	19.6	3.1	16.7	4.4	14.7	4.1	14.3	4.7	12.4	5.5	11.2	5.8	13.3	7.3
	n = 47		n = 41		n = 36		n = 35		n = 31		n = 28		n = 47	

*$P < 0.05$, St John's wort versus fluoxetine (pairwise ANCOVA).
†$P = 0.062$, St John's wort versus placebo (pairwise ANCOVA).
‡$P = 0.096$, St John's wort versus placebo (pairwise ANCOVA).

higher rates of remission (HAMD <8) in the St John's wort group (38%) compared with the fluoxetine group (30%) and the placebo group (21%).

Table 3 reports the results of the analyses for the secondary end points. In 4 patients, the BDI at baseline was missing, so the BDI score at screen was carried forward; similarly, the BDI at end point was missing in 5 patients, so the BDI from the previous visit was carried forward. There was a trend toward a significantly ($P < 0.1$) greater reduction in CGI-S at visit 6/end point in St John's wort–treated patients compared with fluoxetine-treated patients, and the CGI-I scores were significantly ($P < 0.05$) lower at visit 6/end point in St John's wort–treated patients compared with fluoxetine-treated patients. There was also a trend ($P < 0.1$) toward a significantly lower CGI-I score among St John's wort–treated patients compared with placebo-treated patients at visit 6/end point.

There was only 1 serious AE during the course of the study: a patient randomized on St John's wort overdosed on heroin, was treated in the hospital, released, and then discontinued from the study. Overall, St John's wort appeared to be safe and well tolerated. There were no AE-related treatment discontinuations in the St John's wort–treated and placebo-treated patients, whereas 4% (2/47) of the fluoxetine-treated patients dropped out because of side effects. As shown in Table 4, there were no significant differences across treatment groups in rates of AEs, with the exception of skin rash, which occurred more frequently in patients on placebo than on fluoxetine or St John's wort. The most common AEs on St John's wort were headache (42%), dry mouth (22%), nausea (20%), gastrointestinal upset (20%), and sleepiness (18%).

DISCUSSION

In our study of outpatients with mild to moderate MDD, a well-standardized *H. perforatum* (St John's wort) extract (LI-160) was significantly more effective than fluoxetine and showed a trend toward statistically significant superiority over placebo. The main limitation of our study is that the sample size of our study (n = 135) is smaller than originally planned. As mentioned earlier, the sponsor of the study conducted an interim analysis with the intent of closing the study earlier partly because of the delay in achieving the target enrollment and partly because of the negative/inconclusive results of the 2 prior US studies. Following the interim analysis, the sponsor closed the study and opted to proceed with the final analyses carried out on a sample size smaller (n = 135) than originally planned (n = 180). Our findings are consistent with those of a meta-analysis of 23 European randomized trials of St John's wort extract in 1757 outpatients with mild to moderately severe depressive disorders.[2] The meta-analysis had shown that, in placebo-controlled trials, St John's wort was effective in 55% of the subjects (n = 408), whereas placebo was effective in 22% of the subjects (n = 422), and that, in the 6 active-comparator trials involving a single preparation of St John's wort, a 64%

TABLE 3. Mean CGI-S, CGI-I, and BDI Scores in the ITT Sample of Outpatients With MDD (n = 135)

	St John's Wort		Placebo		Fluoxetine	
	Mean	SD	Mean	SD	Mean	SD
Baseline CGI-S	4.2	0.6	4.2	0.6	4.2	0.7
	n = 45		n = 43		n = 47	
End point CGI-S	2.7*	2.3	3.0	2.2	3.1	2.3
	n = 45		n = 43		n = 47	
End point CGI-I	2.4†	2.2	2.8‡	1.2	3.0	2.2
	n = 45		n = 43		n = 47	
Baseline BDI	22.9	9.8	24.0	8.7	21.7	9.4
	n = 45		n = 43		n = 47	
End point BDI	16.1	10.7	17.9	10.8	16.6	10.3
	n = 45		n = 43		n = 47	

In 4 patients, the BDI at baseline was missing; thus, the BDI score at screen was carried forward; similarly, the BDI at end point was missing in 5 patients, so the BDI from the previous visit was carried forward.
*$P = 0.078$, St John's wort versus fluoxetine (pairwise ANCOVA).
†$P < 0.05$, St John's wort versus fluoxetine (pairwise ANOVA).
‡$P = 0.070$, St John's wort versus placebo (pairwise ANOVA).

TABLE 4. Most Frequent (>10% in at Least 1 of the Treatment Arms) Adverse Events in the ITT Sample of Outpatients With MDD (n = 135)

	Joint Pain		Cold Symptoms		Dry Mouth		Diarrhea		Muscle Pain/ Aches		Sleepiness		Flu	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
St John's wort (n = 45)	4	9	6	13	10	22	3	7	5	11	8	18	5	11
Placebo (n = 43)	7	16	4	9	4	9	4	9	3	7	3	7	1	2
Fluoxetine (n = 47)	2	4	7	15	6	13	7	15	4	9	6	13	1	2

	Headache		GIT Upset		Insomnia		Nausea		Rash		URTI	
	N	%	N	%	N	%	N	%	N	%	N	%
St John's wort (n = 45)	19	42	9	20	7	16	9	20	0	0	5	11
Placebo (n = 43)	12	28	5	12	6	14	7	16	5*	12	6	14
Fluoxetine (n = 47)	12	26	10	21	5	11	4	9	0	0	5	11

GIT indicates gastrointestinal tract; URTI, upper respiratory tract infection.
*$P < 0.05$, Placebo versus fluoxetine and versus St John's wort (pairwise Fisher exact test).

response rate was observed with St John's wort (n = 158), compared with the 58% response rate of the antidepressant comparators (n = 159). Our results are also consistent with those of a recently published multicenter European study[12] which showed that, among 375 outpatients with mild to moderate MDD (HAMD-17 score at baseline between 18 and 25), treatment with a hydroalcoholic *H. perforatum* extract was accompanied by a significantly ($P = 0.04$) greater reduction in HAMD-17 score than placebo. In this study,[12] the remission rate (defined as a HAMD-17 score <7) was also significantly ($P = 0.03$) higher for *Hypericum* (24.7%) than for placebo (15.9%).

On the other hand, our findings are not consistent with those of the 2 placebo-controlled trials of the same *H. perforatum* (St John's wort) extract (LI-160) used in our study. The first study[3] enrolled 200 adult outpatients with moderate to severe MDD. Although there was no significant difference in outcome at the end of the 8-week treatment on the primary efficacy measure between St John's wort (900–1200 mg/d) or placebo, the number reaching remission of MDD was significantly higher with St John's wort than with placebo ($P = 0.02$), with very low overall rates in the full intention-to-treat analysis (14/98 [14%] vs. 5/102 [5%], respectively).[3] The recently published multicenter study sponsored by the National Center for Complementary and Alternative Medicine and by the National Institute of Mental Health[7] provided results that were somewhat inconclusive, as both St John's wort (900–1500 mg/d) and sertraline (50–100 mg/d) were not significantly different from placebo on the 2 primary outcome measures. As noted earlier, negative studies for antidepressant medications are relatively common among agents meeting approval for Food and Drug Administration approval, and all studies on an agent need to be considered when evaluating the true effect size of the agent. The total evidence from the accumulated trials in Europe and the United States continues to suggest that St John's wort agents may offer antidepressant efficacy for some individuals with mild to moderate depression. As far as safety and tolerability are concerned, our findings are consistent with those of Shelton et al; St John's wort appeared to be safe and well tolerated, and headache was the most common AE in our trial (42%) and was the only AE that occurred with greater frequency with St John's wort than placebo in the Shelton et al study ([41%] vs. [25%], respectively).[3]

How do we explain the apparent lack of efficacy of fluoxetine in this trial? To the same extent that we used a fixed-dose approach to St John's wort (900 mg/d), we also used a fixed-dose approach to fluoxetine (20 mg/d), and data from our group suggest that a significant proportion of patients nonresponding to 20 mg/d may go on to respond when the dose is increased to 40 or 60 mg/d.[13,14] Similarly, in a recent double-blind study in MDD, where clinicians were allowed to escalate the dose greater than 20 mg/d in nonresponders to 4 weeks of fluoxetine 20 mg/d, the final mean daily dose for fluoxetine was 42 mg, with 31.3% of the patients remaining at a dose of 20 mg, while 28.4% and 40.3% receiving an increase to 40 and 60 mg, respectively.[15] These studies suggest that fluoxetine 20 mg/d, although typically considered an effective dose in the treatment of MDD, may not be an adequate dose for a significant proportion of patients having MDD. Furthermore, the smaller (25%) than planned sample size is likely to account for the lack of statistical significance for the advantage (indicating a moderate effect size, $d = 0.45$) of St John's wort over placebo.

In summary, in our study of outpatients with mild to moderate MDD, a well-standardized preparation of *H. perforatum* (St John's wort) extract (LI-160) was significantly more effective than fluoxetine and showed a trend toward superiority over placebo. Further studies are needed

to fully evaluate the antidepressant efficacy of St John's wort. NIMH, NCCAM, and the NIH Office of Disease Prevention have jointly funded a study of St John's wort compared with citalopram and placebo for the treatment of minor depression. This study will generate information about the efficacy of St John's wort for a milder form of depression and help to fill the gaps in our knowledge about the clinical use of this most important herbal product.

ACKNOWLEDGMENTS

The study was supported by a grant of Lichtwer Pharma AG (Berlin, Germany) to the Massachusetts General Hospital. The authors thank Dr John Worthington of the Massachusetts General Hospital Depression Clinical and Research Program and Drs Marc Pfeffer, Diane London, Shelley Kramer, Alexander Altschuller, and Laurie Witts of Partners Research and Education Program (PREP) for their help and support in conducting the study. The authors also thank Lee Gresham, Megan Hughes, and Pam Roffi of the Massachusetts General Hospital Depression Clinical and Research Program and Courtney Vitale of Partners Research and Education Program (PREP) for their technical assistance in conducting the study.

REFERENCES

1. Nierenberg AA, Mischoulon D, DeCecco L. St. John's wort: a critique of antidepressant efficacy and possible mechanism of action. In: Mischoulon D, Rosenbaum JF, eds. *Natural Medications for Psychiatric Disorders*. Philadelphia, PA: Lippincott Williams and Wilkins; 2002:3–12.
2. Linde K, Ramirez G, Mulrow CD, et al. St John's wort for depression—an overview and meta-analysis of randomised clinical trials. *BMJ*. 1996;313(7052):253–258.
3. Shelton RC, Keller MB, Gelenberg A, et al. Effectiveness of St John's wort in major depression: a randomized controlled trial. *JAMA*. 2001; 285(15):1978–1986.
4. Hamilton M. A rating scale for depressions. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62.
5. Otto MW, Nierenberg AA. Assay sensitivity, failed clinical trials, and the conduct of science. *Psychother Psychosom*. September–October 2002;71(5):241–243.
6. Hooper M, Amsterdam JD. Do clinical trials reflect drug potential? A review of FDA evaluation of new antidepressants. Poster presented at: NCDEU 39th Annual Meeting; June 11–14, 1998; Boca Raton, Fla.
7. Hypericum Depression Study Group. Effect of *Hypericum perforatum* (St. John's wort) in major depressive disorder: a randomized controlled trial. *JAMA*. 2002;287(14):1807–1814.
8. Kupfer DJ, Frank E. Placebo in clinical trials for depression: complexity and necessity. *JAMA*. 2002;287(14):1853–1854.
9. First MB, Spitzer RL, Gibbon M , et al. Structured Clinical Interview for *DSM-IV*–Axis I Disorders—Patient Edition (SCID-I/P, Version 2.0). New York: Biometrics Research Department, New York State Psychiatric Institute; 1995.
10. Guy W, ed. ECDEU Assessment Manual for Psychopharmacology. Rev Ed DHEW Publication No. (ADM) 7–338. National Institute of Mental Health: Rockville, MD; 1976.
11. Beck AT, Ward CE, Mendelson M. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4:561–571.
12. Lecubrier Y, Clerc G, Didi R, et al. Efficacy of St. John's wort extract WS 5570 in major depression: a double-blind, placebo-controlled trial. *Am J Psychiatry*. 2002;159:1361–1366.
13. Fava M, Rosenbaum JF, McGrath PJ, et al. A double-blind, controlled study of lithium and tricyclic augmentation of fluoxetine in treatment resistant depression. *Am J Psychiatry*. 1994;151:1372–1374.
14. Fava M, Alpert J, Nierenberg AA, et al. Double-blind study of high-dose fluoxetine vs. lithium or desipramine augmentation of fluoxetine in partial and non-responders to fluoxetine. *J Clin Psychopharmacol*. August 2002;22(4):379–387.
15. Fava M, Hoog SL, Judge RA, et al. Acute efficacy of fluoxetine versus sertraline and paroxetine in major depressive disorder including effects of baseline insomnia. *J Clin Psychopharmacol*. April 2002;22 (2):137–147.