

# Are SSRI's Superior to Placebo for Mild to Moderate Depression?

## *A Critical Evaluation of Some Research Articles*

Sam Kaspar, BSc (Hon), MSc, MD

Department of Family Medicine, McMaster University, Hamilton, Ontario

In the pharmacologic treatment of depression, the introduction of selective serotonin reuptake inhibitors (SSRI's) has been advantageous in that they have less adverse effects than older classes of antidepressants. It has been commonly concluded in studies and review articles alike that SSRI's have similar efficacy as tricyclic antidepressants (TCA's) such as imipramine, and that both classes of drugs are superior to placebo. Because of large placebo effects and various methodological issues regarding clinical trials, the present paper will critically assess three representative studies of SSRI efficacy versus placebo. Areas of focus include evaluation of study design and bias, execution of the studies, and the interpretation that was offered in each paper. It will be concluded that SSRI's are useful in severe depression. However, strongly designed studies of efficacy in mild and moderate depression are lacking, and the studies that do exist show zero or very slight advantages of pharmacologic treatment over placebo. Some serious adverse reactions such as anorgasmia, definitely offset these modest benefits in many patients with major depression of a mild to moderate severity.

### INTRODUCTION

The selective serotonin reuptake inhibitors (SSRI's) have been very useful in the pharmacologic treatment of major depressive disorder. Because there are potential adverse reactions to any medication, it is important to have evidence of efficacy of the treatment before deciding to prescribe it to a specific patient. In the following evaluations of published papers on SSRI efficacy, emphasis is placed on factors that may make the data invalid or less applicable in certain clinical situations. For example, the downfall of many studies is the pooling of all patients into an undescribed collection of depressed patients which neither reflects the same group we may face in the office, nor delineates features which may separate patients into groups of likely or unlikely responders to antidepressant medication. Where available, stratification of results has demonstrated that specific characteristics, such as severity of depression, will

determine whether these medications actually help the patient beyond having a placebo effect.

### METHODOLOGIES AND APPROACH

Medline was used to search for papers addressing SSRI efficacy in depression, particularly as compared to placebo. The search strategy included specifying Medical Subject Heading (MeSH) for depression, SSRI, fluoxetine, fluvoxamine, paroxetine, sertraline, as well as specifying publication type of randomized controlled trial, English language and restriction to abstracts that used the word "efficacy." The references were assessed mainly with respect to published criteria for the design of clinical trials (1-3). The clinical trials analyzed here all had a placebo arm. The first two papers exemplify many other reports that would have been redundant to analyze here, and were selected because they nicely demonstrate key points that are consistent in the broader literature (summarized in **Box 2**). The third paper demonstrates a weak study design.

The papers discussed below used the DSM-III or DSM-III-R criteria to define uni-

Address correspondence to:

Dr. S. Kaspar  
McMaster University, Faculty of Health Sciences, Office of Postgraduate Education,  
Room 1B7B, 1200 Main St. W, Hamilton Ontario, Canada L8N 3Z5

polar depressive episodes as the type of major affective disorder, and the depression was quantified by the Hamilton Depression score (Ham-D) in all cases (4). The first 17 items of the 21-item Ham-D are commonly used, as the final four items relate to diurnal variation, derealization, paranoia, and obsessions, and were suggested by Hamilton (1960) to be of lesser prevalence and importance in assessing depression (4). The Ham-D scoresheet is available in standard textbooks of psychiatry. Some papers used additional scales as well.

## EVALUATION OF CLAGHORN ET AL. (1992, [5]) ON PAROXETINE

This was a placebo-controlled (parallel-group), double-blind study of paroxetine for the treatment of depression. The age of patients ranged from 18 to 65 years (mean 40.6 for drug and 42.8 for placebo). The paroxetine group had more people under age 40 than over 40 (94 versus 74 subjects), and the placebo group had more females (92) than males (77), perhaps suggesting imperfect randomization.

Outpatients were recruited from four centres. Patients had to score greater than 18 on the first 17 items of the 21-item Hamilton Depression Scale (Ham-D), which has a maximal score of 50 for a single interviewer (4). The five authors of the paper included one from each of the four centres, and Dr. Dunbar from Smith Klein and Beecham Pharmaceuticals. Although a minimal score requirement of 18 (which corresponds to major depression which is mild in severity) was specified, the average Ham-D score was 27, and despite the range of scores, results were not divided or discussed in the context of severity of depression. Having a range of severities in a study is only useful if the differential outcomes of mild versus severe cases are analyzed separately. The authors also presented summary statistics for duration of the presenting episode of depression, and whether the episode was a recurrent one, but for analysis pooled results were used. The number of recurrent-episode patients (235 of 337 in total), and the duration or severity of episodes, were not systematically organized into categories of patients. Perhaps the statistical power (i.e.,  $n$  values) would have been too low for stratification because of the large number of permutations of patient histories, even though the results of a subgroup analysis would have been more useful.

Although only four patients were lost to follow-up, the completion rate (at least 28 days) was only 109/168 (65%) for paroxetine and 95/169 (56%) for placebo. Regardless, assessments of patients were made at 1, 2, 3, 4, and 6 weeks, and the authors clearly defined "complete recovery" as a total Ham-D score less than 10, or a Clinical Global Impressions (CGI) score of 1 or 2. The authors chose to present an intention-to-treat analysis, but they also gave the outcomes of the "adequate treatment group" for comparison. On careful assessment of their data handling, two things are notable: (i) a minimal time in the study was not specified before the last observation was carried forward (LOCF) for the intention to

treat analysis, and (ii) only 28 days (4 weeks) were required before assigning observations to the "adequate treatment" group. Using LOCF in fact would presumably tend to reduce the baseline-to-endpoint difference in Ham-D scores, as patients may not yet have been treated long enough to see maximal results. In the treatment group, there was a slightly larger mean reduction in Ham-D scores for the patients who stayed on paroxetine (-12.9) as compared to the mean reduction in the intention to treat analysis (-11.5). This is suspicious though, because the placebo analysis groupings showed a similar difference (-8.6 for adequate treatment and -7.1 for intention to treat). This is consistent with a purely time-dependent component of recovery.

Because anxiety often accompanies depression, the inclusion criteria also required that the Raskin Depression Scale score was greater than the Covi Anxiety scale rating, although no specific minimal or maximal scores were required. Further exclusion criteria were clearly stated (other primary psychiatric diagnosis, unstable medical condition, significant baseline lab findings, alcohol or drug abuse in past six months, pregnancy). The treatment plan included a 4-14 day placebo washout period (many of the patients were on other drugs) followed by six weeks of either drug or placebo. Six weeks is a typical duration for antidepressant trials, but as noted above, the study used LOCF on the majority of patients and defined the adequate treatment group as those who had complied for a period of only four weeks. The study also failed to use a fixed dosage, allowing the 20 mg initial dose to be altered weekly at the physician's discretion, from 10-50 mg.

Regarding the claimed efficacy of the treatment group in the intention to treat analysis, there were 38 full recoveries (of 163) in the paroxetine group and 24 (of 162) in the placebo group (42 versus 27 if using low CGI rather than low Ham-D as the criterion for full recovery). This reached statistical significance ( $p \leq 0.05$  using Fisher's exact test) but the clinical significance seems meagre with regards to the excess recovery with drug over placebo, which is only 14-15 patients (about 9% of the paroxetine patients). Taking the improved patients only, those on drug compared to those on placebo were 38:24, or about 1.58 fold, but again, the raw number of recoveries was small for both treatment regimes. In their discussion, Claghorn *et al.* (1992) claim that, when the 40% responder rate is considered in the study's "treatment resistant population," that "twice as many subjects receiving paroxetine compared with placebo made a full recovery" (5).

Co-intervention bias was not addressed in the study, and (as will be seen below) it is very common for instance to prescribe benzodiazepines to patients suffering from anxiety (Ham-D items # 9, 10, 11, worth 0-10 points combined) or sleep disorders associated with depression (Ham-D items # 4, 5, 6, worth 0-6 points combined). This obviously confounds the usage of Ham-D scores to assess changes in depression due to paroxetine, because benzodiazepines alone can reduce the Ham-D score by up to a total of 0-16 points lost on these six items.

## EVALUATION OF OTTEVANGER (1994, [6]) ON FLUVOXAMINE

This paper describes a four-week, double-blind, placebo-controlled, prospectively-randomized, parallel-group study of fluvoxamine (an SSRI), imipramine (a TCA), and placebo in subjects aged 19-70 who had a mean age of 42.6 years and were either outpatients (2 centres) or inpatients (3 centres). The similarities or differences between in- and outpatients were not commented on in the paper. The paper is a post-hoc analysis of a 338-patient study from 5 North American centres. This paper addresses many of the issues that minimized the internal validity and applicability of the study assessed above (5). A notable weakness of the Ottevanger (1994) paper (6) is that the treatment period was only four weeks. Nonetheless, most antidepressant studies are only 3-7 weeks in duration (7).

Ottevanger (6) considered three issues that were emphasized by the Consensus Committee on Clinical Trials on Antidepressants (8): (i) LOCF should be used if patients were evaluated at least once after baseline, (ii) the number of placebo responders differs in mild and severe depression, and (iii) concomitant benzodiazepine use alters the Ham-D score because of its effects on sleep and anxiety. The benzodiazepine issue was addressed by omitting from the analysis the additional 173 patients recruited at European centres, which the author wrote accounted for most of the 40% benzodiazepine co-administration rate in the study; however, the actual rate of benzodiazepine usage in the analyzed North American centres was not given. It would have been more sensible to also omit the remaining patients who received benzodiazepines. There was a 3-7 day placebo washout period, followed by randomization to a treatment group, treatment for 28 days, and subsequent intention to treat analysis; LOCF was done anytime after the placebo phase was completed.

Inclusion criteria included a minimal Ham-D score of 15 on the 17-item version of the scale. The author analyzed three subgroups of severity and presented the data according to patients with initially mild, moderate, or severe depression, who had respective Ham-D scores of 15-20, 21-25, and 26-38; the numbers of patients in each group were 100, 105, and 103, respectively. The exclusion criteria (6) were more strict, explicit and useful than those of Claghorn *et al.* (1992)(5), particularly pertaining to evaluation of previous treatment regimes (electroconvulsive therapy in past four weeks, monoamine oxidase inhibitor in past two weeks, lithium in past week, TCA in past 3 days prior to start of study). Patients were excluded from randomization if they were suicidal, which would remove item # 3 (worth 0-4 points) from the Ham-D scale. Other exclusion criteria were pregnancy, psychosis, drug or alcohol abuse, and major cardiac, renal, or hepatic disease.

Subgroup outcomes of the treatment phase revealed several interesting concepts regarding reasons for discontinuing therapy, and response rate and magnitude. In mild depression, nearly 50% of patients discontinued fluvoxamine and only about 25% of subjects discontinued placebo, possi-

bly indicating that mildly depressed patients were less likely to tolerate adverse drug effects for minimal gains in their depression. Conversely, in severe depression only about 25% of patients discontinued fluvoxamine but almost 50% discontinued placebo probably due to lack of benefit (6). Response rates in severe depression were over 63% for fluvoxamine, as compared to 23% for placebo, and the mean reduction in Ham-D score was 14.1 as opposed to 7 in placebo. "Responder" was defined as a patient with a decrease in 50% of the Ham-D score at endpoint, or a "much improved CGI score." The percent reduction of Ham-D is useful for signifying an antidepressant effect, but not for marking a full recovery; hence "responder" means there was a large improvement, and unfortunately not that a fixed endpoint was reached (e.g., score less than 10).

**Table 1:**  
Data from reference (6); (f=fluvoxamine, i=imipramine, p=placebo)

Depression	Initial Score	Actual Mean Initial (Final) Ham-D Score
Mild	15-20	f=18.0 (13.0) i=17.5(12.6) p=18.0(12.3)
Moderate	21-25	f=23.1 (13.2) i=23.2(13.1) p=23.1 (15.9)
Severe	26-38	f=29.3 (15.2) i=28.7(18.4) p=29.1 (22.1)

For mild and moderate depression, all mean final scores were clustered close together (see **Table 1**). Hence, the improvement seems to have regressed to a common endpoint. This must be interpreted with caution because of drug dosages. After day three, the 150 mg dosage was adjusted on an individual basis (up to 300 mg daily) and more severely depressed patients usually received more drug, potentially improving the average response in an unpredictable manner. There was no presentation of pill counts or analysis of dose-response relations in this study, and the percentage of patients who responded or even followed-up was not presented. Bearing in mind these caveats about drug dosages, the fluvoxamine effect on the Ham-D score was markedly superior to placebo in severe depression. The final Ham-D scores of the placebo group were on average even better than either antidepressant medication, in cases of mild depression.

Ottevanger (6) notes that the difference in effect between severe and mild depression was highly significant ( $p = 0.0001$  by ANOVA), and that cases of mild depression (Ham-D less than 20) should possibly be excluded from placebo-controlled studies in the future (or that severity of depression should at least be stratified).

Early improvement at week 1 was a sensitive indicator for final response to the SSRI (Sensitivity = 82.5%). "Improvement in week 1" was defined as a CGI score of 1, 2, or

3, and non-improvers scored 4 or 5. However, the author did not calculate positive predictive value (PPV) from his table; I calculate a PPV of 74.5%, which is reasonable for something as simple as grading improvement in the first week, and a likelihood ratio of 2.36, which is modest but good (see **Box 1**). The excess of recoveries above placebo would have been the more appropriate analysis. Too few parameters are given in the paper to make this calculation a simple matter, and it was not presented in the article explicitly.

**Box 1: Calculations**

*(based on standard formulas (3) and recovery by day 28 as the final positive outcome; see text regarding evaluation of reference [6])*

**Prevalence of recovery (Prev)** = responders at endpoint/total  
 = 57/103 = 0.553

**PPV**=(Sensitivity\*Prev)/[(Sensitivity\*Prev)+(1-Specificity)(1-Prev)]  
 =(.825)(.553)/[(.825\*.553)+(1-.65)(1-.553)] = .745

**Likelihood Ratio** = Sensitivity/(1-Specificity) = (.825)/(1-.65)  
 = (.825/.35) = 2.36

**Summary:**

**Sensitivity:** 82.5% of those recovered at 28 days, showed improvement at week 1.

**PPV:** There is a 74.5% probability that recovery criteria will be met, if there was improvement in week 1.

**Likelihood Ratio:** Week 1 improvement, for future responders, versus future non-responders is 2.36-fold.

This small study (45 patients) attempted to characterize fluvoxamine therapy with a double-blind, placebo-controlled, 28-day study in hospitalized patients at the University of Pisa, Italy. The authors found improvements on total Ham-D scores with both fluvoxamine and placebo, and found that fluvoxamine improved symptoms of retardation, cognitive disturbances, and sleep disturbances. They also used a patient self-assessment scale, CGI (more improved with fluvoxamine than placebo), and the Brief Psychiatric Rating Scale. LOCF was used if patients complied for at least 14 days. Patients were included who were females over the age of 18 (mean 53.7 in fluvoxamine group and 53.1 in placebo group), who scored at least 16 on the first 17 items of the Ham-D after the 3-7 day placebo washout period. Dosing was increased to 150 mg over the first 3 days, kept at 150 mg per day to day 6, and could be changed to a maximum of 300 mg per day “according to the response,” for days 7-28. Again, variable dosages is a confounding factor in assessing this study. Exclusion criteria were pregnancy, serious untreated and treated disease, alcoholism and drug addiction, depression secondary to another diagnosis, lithium in the past week, monoamine oxidase inhibitors, electroshock therapy in the past month, or incapacity to collaborate. Although the last

criterion omits patients who score 4/4 on item #8 (retardation) in Ham-D, Hamilton had also stated that using the rating scale was not useful in the presence of complete stupor or for mute patients (score >3), regardless of the ability to collect information (4).

Analysis of this paper will demonstrate three principal issues: (i) the division of a validated depression scale (Ham-D) into subsets of depressive symptoms, (ii) the nearly ubiquitous co-administration of benzodiazepines in some studies, and (iii) patient selection bias affecting the outcome of a clinical trial.

Firstly, the results were presented as changes in specific items on the Ham-D. The total CGI was only modestly improved by either drug or placebo, and the difference between the two curves was only present because they were already nearly a full point apart in week 1. The authors noted that fluvoxamine was statistically significantly superior to placebo ( $p < 0.05$  by Analysis of Covariance, ANCOVA) only on improving the item of “guilt feelings” but not on somatic concerns, tension, depressive mood, or motor retardation; in fact, blunted affect was improved 30% by placebo and only 20% by fluvoxamine, although this did not reach statistical significance. The total Brief Psychiatric Rating (BPR) Score, and the extracted “depressive symptoms” score that the authors fabricated, both showed percentage improvements of fluvoxamine that were greater than that of placebo, but neither comparison reached statistical significance. Using some items from the Ham-D, the authors found statistically significant improvements in cognitive disturbances and sleep disturbances ( $p < 0.05$  by ANCOVA), but not in other symptoms or the total score. Besides the scattered results and the unimpressive minimal advantages of the drug over placebo, the small sample size affords a very small power or reliability for the results. An example of the scattered results is that the anxiety item hardly differed in magnitude (and did not reach statistical significance) between placebo and drug, for both the BPR and the Ham-D scales, but showed a huge effect (with  $p < 0.05$  by ANCOVA) on the self-report symptoms inventory (which incidentally did not find a statistically significant reduction in the patient’s mood of depression). When total scores have been previously validated, it does not inspire confidence when the surveys are split up by the authors into their component questions, which alone do not indicate a constellation of symptoms that can rate depression with validity and reproducibility (4). A more acceptable approach would be to document the number of symptoms that are improved.

Flurazepam (a benzodiazepine) was prescribed in 95.7% of the fluvoxamine patients and in 100% of the placebo group. The problems with this have been described above, but because of the equal numbers of co-medications in both the drug and placebo groups, it is not possible to dismiss the whole study — but mandatory to dismiss all aspects pertaining to sleep, anxiety, and agitation — based on the confounding variable of co-intervention. In their discussion the authors suggested that benzodiazepine administration was necessary to prevent withdrawal symptoms during the study, but they never reported the pre-test prevalence of usage. The prevalence of concomitant benzodiazepine usage seems to



# Healthy lives start here.

At BD we have one mission, helping our customers live healthier lives. That's been the philosophy behind every BD product for the past 100 years. Everything we make – from insulin syringes to glucose tablets – has BD's quality assurance so you know it's superior.



Better diabetes care starts here.

**BECTON  
DICKINSON**

Becton Dickinson Canada Inc.  
2464 South Sheridan Way  
Mississauga, Ontario L5J 2M8  
[www.bd.com/diabetes](http://www.bd.com/diabetes)





vary by country, and that brings me to my final point: the site of the study as pertains to biases of patient selection (centripetal bias for tertiary care centre, popularity bias for follow-up and recruitment, referral filter bias, and accessibility of patients from some regions to the facility). The authors noted in their discussion that changes in Italian Law on Psychiatric Reform, in 1978, shifted patient management largely from the hospital to “assistance in the field,” and the ensuing reduction in numbers of hospital beds caused the inpatient population to consist mainly of the most treatment resistant, relapsing cases, or those with an inability to collaborate. I conclude that this third study is conducted and presented in such a way that evaluating from it the efficacy or effectiveness of fluvoxamine or placebo, is an utterly useless exercise.

### Box 2: Key Points

(i) Claghorn *et al.* (1992) showed a recovery rate of 38/163 for paroxetine versus 24/162 for placebo in patients having a mean Ham-D score of 27 (minimum 18). The excess of recovery rate is only 9% of the total patient pool (14/163) who responded to paroxetine above and beyond the placebo rate. Although the overall yield is still low, the number of responders is increased by 58% (14/24) with the SSRI.

(ii) Ottevanger (1994) showed that mild to moderately depressed subgroups of patients (mean Ham-D of 18 or 23 respectively) had basically zero improvement above the magnitude of the recovery with placebo. Severely depressed patients (mean Ham-D of 29) benefited from antidepressant therapy, on average.

(iii) Ottevanger (1994) noted that improvement during the first week of SSRI therapy predicted a final recovery on the drug, with a sensitivity of 82.5% and specificity of 65%, the latter being not as high presumably because of late responders on drug or placebo. With a calculated PPV of 74% and a likelihood ratio of 2.36, the early response criterion bears clinical merit.

(iv) In mild depression up to 50% of patients taking the SSRI drop out due to adverse drug effects; in severe depression up to 50% of patients on placebo drop out due to lack of effect.

(v) The paper by Conti *et al.* (1988) demonstrates several pitfalls, primarily related to referral practices and outcome measures. The latter point also raises the issue of whether a final endpoint measurement is indeed the most clinically useful parameter for judging clinical effectiveness of a course of treatment.

## DISCUSSION

These typical studies indicate that the value of SSRI's, and probably antidepressants in general, have been overestimated, mainly in the context of mild and moderate depression. A Ham-D score in this range still entails that several significant symptoms be present to some degree, but perhaps surprisingly may not rate as a severe case of major depressive disorder whose outcome score at 4-6 weeks would have been improved by drugs. A large placebo effect, especially in well-controlled studies, is a major factor, as well as concomitant

effects of the attention and counselling involved when any kind of psychiatric treatment is received. In terms of study quality, major flaws included co-administration of benzodiazepines, lack of fixed drug dosages, failure to assemble inception cohorts who were carefully assessed for history, severity of illness, and previous treatment outcomes, and especially the over-interpretation of efficacy in most cases. It would be useful to re-evaluate a larger number of papers in order to provide clear information about the actual advantage above placebo that can be expected with SSRI's, and suggest guidelines regarding what severity of depression should or should not be treated with SSRI's. In terms of balancing benefits with the risk of adverse drug effects, efficacy is of paramount importance, especially because SSRI's are presently being used for long-term maintenance therapy as well. Gastrointestinal (10) complaints and sexual dysfunction (10-14) are common (1.9 to 92.0% in various studies), and may need further treatment (i.e., reference 15).

Finally, although these “outcome” studies have been assessed here, I found no study that addressed whether patients have a subjectively easier convalescent course if they were on SSRI's versus placebo. This is an important issue when assessing patient groups who reach the same outcome measurement by the end of the study. We need to know whether the pathway to a given endpoint is altered favourably by pharmacologic therapy, because other drug-related effects such as disease-labelling, sick-role, placebo effect, and others, have psychosocial repercussions. Studies of these so-called “soft symptom” changes are lacking, and physicians judge these individually by experience and follow-up.

The energizing effects of SSRI's may cause subjective improvement which may be important to many patients. This would be a key issue to address in future studies, especially because the actual endpoint differences between drug and placebo were not impressive except for severe depression. However, the same energizing effect of SSRI's that may subjectively improve the convalescent phase, also puts suicidal patients in a dangerous position because the energy rises before the mood does, thus creating a high-risk window of suicidality for some patients (SSRI's do not increase the suicidality per se though, see references 16,17). Indeed, any adverse reaction is intensified in the eyes of both the patient and prescribing physician, when the efficacy of the drug might have been questionable in the first place for that specific patient. Large “placebo” effects lend credence to a treatment plan — especially for mild and moderate depression — which should include a variety of therapies such as psychotherapy, support, and possibly medication. This again demonstrates that follow-up is as important as the prescription, and that sometimes pill-tampering with the natural history of a common complaint such as depression may not be harmless.

## ACKNOWLEDGMENTS

I thank the people who read my paper and gave me positive comments and feedback, and also the group at the Nova Scotia Hospital who educated me with their stimulat-

ing approach to Psychiatry during my rotation there. I especially thank Dr. Alistair Munro at the Nova Scotia Hospital, and Dr. Herb Orlik and Dr. Dave Zitner who had read my paper previously.

## REFERENCES

1. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994;271(1):59-63.
2. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993;270(21):2598-601.
3. Fletcher RH, Fletcher SW, Wagner EH. *Clinical epidemiology: the essentials*. 2nd ed. Baltimore: Williams & Wilkins, 1988.
4. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatr* 1960;23:56-62.
5. Claghorn JL, Kiev A, Rickels K, Smith WT, Dunbar GC. Paroxetine versus placebo: A double-blind comparison in depressed patients. *J Clin Psychiatry* 1992;53:434-438.
6. Ottevanger EA. The efficacy of fluvoxamine in patients with severe depression. *Prog Neuro Psychopharmacol & Biol Psychiatr* 1994;8:731-740.
7. Kasper S, Fuger J, Möller H-J. Comparative efficacy of antidepressants. *Drugs* 1992;43(Suppl 2):11-23.
8. Angst, *et al.* Consensus conference on methodology of clinical trials of antidepressants, Zurich, March 1988: Report of the consensus committee. *Pharmacopsychiatry* 1989;22:3-7.
9. Conti L, Dell'Osso L, Re F, Musetti L, Battista Cassano G. Fluvoxamine maleate: Double-blind clinical trial vs placebo in hospitalized depressed patients. *Current Therapeutic Research* 1988;43:468-480.
10. Preskorn SH. Comparison of the tolerability of bupropion, fluoxetine, imipramine, nefazodone, paroxetine, sertraline, and venlafaxine. *J Clin Psychiatry* 1995;56(Suppl 6):12-21.
11. Herman JB, Brotman AW, Pollack MH, Falk WE, Biederman J, Rosenbaum JF. Fluoxetine-induced sexual dysfunction. *J Clin Psychiatry* 1990;51(1):25-7.
12. Pearlstein TB, Stone AB. Long-term fluoxetine treatment of late luteal phase dysphoric disorder. *J Clin Psychiatry* 1994;55(8):332-5.
13. Balon R, Yeragani VK, Pohl R, Ramesh C. Sexual dysfunction during antidepressant treatment. *J Clin Psychiatry* 1993;54(6):209-12.
14. Zajecka J, Fawcett J, Schaff M, Jeffriess H, Guy C. The role of serotonin in sexual dysfunction: fluoxetine-associated orgasm dysfunction. *J Clin Psychiatry* 1991;52(2):66-8.
15. Walker PW, Cole JO, Gardner EA, Hughes AR, Johnston JA, Batey SR, Lineberry CG. Improvement in fluoxetine-associated sexual dysfunction in patients switched to bupropion. *J Clin Psychiatry* 1993;54(12):459-65.
16. Crundwell JK. Fluoxetine and suicidal ideation--a review of the literature. *Int J Neurosci* 1993;8(1-2):73-84.
17. Power AC, Cowen PJ. Fluoxetine and suicidal behaviour: some clinical and theoretical aspects of a controversy. *Br J Psychiatry* 1992;161:735-41.

## AUTHOR BIOGRAPHY

Dr. Kaspar graduated from Dalhousie Medical School in 1998 and is currently a family medicine resident at McMaster University. Previously, he earned his bachelor's degree in medical biophysics at the University of Western Ontario, graduating with the gold medal. He also completed a Master's degree in physiology at Dalhousie University, and published various research articles with Dr. D. Pelzer. During medical school he held two summer research scholarships at the University of Toronto. Sam served as the Co-Editor-in-Chief (Reviews and Content) for the *Dalhousie Medical Journal* in 1996/97.