DOES ELIMINATION OF PLACEBO RESPONDERS IN DOUBLE-BLIND PLACEBO-CONTROLLED RANDOMIZED CLINICAL TRIALS OF SSRI ANTIDEPRESSANTS IN DEPRESSION INFLUENCE THE SIZE OF THE TREATMENT EFFECT? A META-ANALYTIC EVALUATION

by SANDRA LEE

A thesis
Submitted to the Faculty of Graduate Studies
In partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Community Health Sciences
Faculty of Medicine
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

Does Elimination of Placebo Responders in Double-Blind Placebo-Controlled Randomized Clinical Trials of SSRI Antidepressants in Depression Influence the Size of the Treatment Effect? A Meta-Analytic Evaluation

BY

Sandra Lee

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree

of

Master of Science

TABLE OF CONTENTS

ABSTRACT

Randomized clinical trials are accepted as the most effective method to assess the safety and effectiveness of a new drug or clinical intervention. Often, the results of a randomized clinical trial will allow a new drug to be introduced into current clinical practice. Clinical investigators and the pharmaceutical industry naturally want to optimize treatment effect. Thus, many randomized clinical trials, especially those involving psychotropic drugs, are preceded by a "placebo run-in phase", in which all subjects are given the placebo and any subject who responds to the placebo are withdrawn from the study prior to randomization.

The objective of this research is to compare the effect size of randomized controlled placebo clinical trials (in the treatment of depression) that include a placebo run-in phase with those that do not include a placebo run-in phase, using a meta-analytic approach. It is hypothesized that the size of the treatment effect will be larger in studies that eliminate placebo responders from the study after a placebo run-in phase. A literature search was carried out to find all available published randomized clinical trials involving the use of a selective serotonin reuptake inhibitor antidepressant and placebo. Data were extracted from the trials and statistical analysis was completed using the international Cochrane Collaboration Review Manager software.

The results indicate that there is no statistically significant difference in effect size between the clinical trials that have a placebo run-in phase followed by withdrawal of placebo responders and those trials that do not have such a phase. Recommendations for future research are discussed.

ACKNOWLEDGEMENTS

Thank you to my husband, Jason, my children, Veronica and Elliot, and my extended family for their patience and understanding throughout the extended time required to complete this project.

Thank you to my thesis committee: Dr. John Walker, my advisor, Dr. Tom Hassard and Dr. Shayne Taback, for their guidance, support and expertise.

A special thank you to Shannon, who took time from being a new mother to give me technical assistance with data management, and for her ongoing encouragement.

Thanks to God for giving me strength to complete this long journey.

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

Randomized clinical trials are accepted as the most valid method to assess the effectiveness of a new drug or clinical intervention. Often, the results of a randomized clinical trial will allow a new drug to be introduced into current clinical practice. Certainly, the same results will influence the prescribing practice of many clinicians. Clinical investigators and the pharmaceutical industry naturally want to optimize treatment effect. Many randomized clinical trials, especially those involving psychotropic drugs, are preceded by a "placebo run-in phase". The placebo run-in phase is a single-blind placebo period, that usually lasts 7-14 days. It occurs before randomization and all study-eligible subjects are given the placebo treatment (and taken off any antidepressant drugs) during this interval. Any responders to the placebo (i.e. symptoms improve) in this preliminary phase are withdrawn from the pool of subjects prior to random assignment and eliminated from the study. Criteria for placebo response differ for each clinical trial. A common criterion for placebo response in the treatment of depression is an improvement of $\geq$50% in the Hamilton Rating Scale for Depression (HAM-D) (Hamilton, 1960). Some studies also have a more stringent definition, that requires a final HAM-D score of $\leq$10 in addition to the 50% improvement (Brown, 1988). Placebo responders, in this situation, can be described as those subjects whose symptom profile met the study criteria at baseline assessment but no longer do so after the 7-14 day placebo run-in phase (Rabkin, McGrath, Stewart et al, 1986). Nonresponders (those subjects who did not improve in the placebo run-in phase) are then randomized into the study to receive the active drug or placebo. The effect of withdrawing these placebo responders from the study prior to randomization, and the resultant effect on the effect size of the study drug or drugs is the focus of this research.

1.2 PLACEBO RESPONSE

Placebos have undoubtedly been used for centuries for the treatment of disease, but the first paper published with reference to the placebo phenomenon was by Pepper (1945). In 1954, Gaddum stated that placebos have two real functions, one of which is to distinguish pharmacological effects from the effects of suggestion, and the other is to obtain an unbiased assessment of the result of the experiment (Gaddum, 1954). A placebo is an intervention designed to simulate medical therapy, but not believed by the investigator to be a specific therapy for the target condition (Brody, 1985).

Reasons given for the use of placebos in medical practice include: as a psychological instrument in the therapy of certain ailments arising out of psychiatric illness, as a resource in dealing with the neurotic patient, to determine the true effect of drugs apart from suggestion in experimental studies, and as a device for eliminating bias for the subject and the researcher (Beecher, 1955; Brody, 1985). A placebo response refers to any change in patient behavior or condition following the administration of a placebo (Brody, 1985). Spiro, in his book entitled *Doctors, Patients, and Placebos*, states that almost all patients will prove to be placebo reactors, if one accepts as a placebo effect responding to the physician or researcher who is helpful in a positive manner (Spiro, 1986). Placebo effects influence patient outcomes after any treatment, including surgery, which the clinician and patient believe is effective (Turner, Deyo, Loeser et al, 1994).

1.3 THE PLACEBO RESPONDER

There are two main theories about placebo responders. Beecher (1953) speculated that there is a group of patients who consistently respond to placebo. Alternatively, Gliedman (1958) have suggested that many patients are predisposed to respond to placebo, but only do so under certain conditions. Jellinek found a bimodal distribution of patients who responded to placebo for relief of headache pain in repeated trials, suggesting that consistent placebo responder and nonresponder groups exist (Jellinek,

1946). Some view placebo response as a bell-shaped curve, with a small percentage of patients never responding to placebo, a small percentage always responding to placebo, and the majority of patients responding to placebo under specific conditions of disease or treatment (Fairchild, Rush, Vasavada et al, 1986).

The original studies on placebo treatment and "placebo reactors" were done to assess pain relief (Jellinek, 1946; Beecher, 1953; Lasagna, Mosteller, von Felsinger et al, 1954). Pain is primarily a subjective phenomena and not objectively determined, so one should not carry the implication of their findings too far (Spiro, 1986). However, even diseases, that are "objectively", measured, such as hypertension and asthma, are multifactorial and have a "functional" component (Spiro, 1986). Assessment of mood and depression have some objective basis (in the different measurement scales available), but it is partly a subjective measurement, both on the part of the subject and the researcher.

Approximately one-third of any group of people will respond to placebo, regardless of what they are being tested for (Spiro, 1986). This widely accepted statistic is based on the classic article by Beecher in which a review of fifteen studies of patients suffering a variety of conditions (postoperative pain, cough, angina pectoris, headache, drug-induced mood changes, seasickness, anxiety and tension, and the common cold) revealed that, on average, 35% of patients were "satisfactorily relieved" by the placebo, with the placebo response rate ranging from 15% to 58% (Beecher, 1955).

CHAPTER 2: PLACEBO RESPONSE IN DEPRESSION

2.1 INTRODUCTION

In early antidepressant clinical trials (that mainly assessed the treatment of depression with the tricyclic antidepressants) involving patients with widely ranging symptoms and severity of depression, placebo response rates ranged from 0% to 70%, with an average across studies of 30-40% (Klerman and Cole, 1965; Rogers and Clay, 1975). These early studies indicated that diagnosis (neurotic vs. endogenous depression) and severity as somewhat predictive of placebo response. In these studies, patients with endogenous depression were consistently found to have lower placebo response rates (around 30%) and greater drug versus placebo differences than patients with reactive and neurotic depression (as high as 70% and indistinguishable from the drug response rate) (Kiloh, Ball and Garside, 1962; Raskin and Crook, 1976; Rogers and Clay, 1975).

The early studies did not include direct examination of depression severity using scales of symptom frequency and intensity that are now standard (Brown, 1988). In the early studies, it has been stated that the more severe the depression, the lower the placebo response rate (around 30%) and the less the severe the depression, the higher the response rate (around 70%). The relationship between severity of depression and placebo response has been confirmed in more recent studies. In patients with more severe depression, defined as those with HAM-D scores >20, placebo response rates were in the 30%-40% range (Brown, Dorseif and Wernicke, 1988; Fairchild, Rush, Vasavada et al, 1986). Those patients with less severe depression, defined as those with HAM-D scores <14, had placebo response rates of greater than 50% (Quitkin, Rabkin, Markowitz et al, 1987; Stewart, Quitkin, Liebowitz et al, 1983). Today, a depression of moderate or greater severity with endogenous features is considered the type of depression most likely to require and respond to antidepressants. Most recent clinical trials involving the use of antidepressant drugs routinely limit enrollment to those who meet DSM-III (American Psychiatric Association, 1980) or DSM-IV criteria (American

Psychiatric Association, 1994) for major depression (Brown, 1988). Placebo response rates are lower, but are still substantial.

The substantial placebo response rate in depression presents a dilemma to both the clinician seeking the most suitable treatment for the depressed patient and the clinical investigator attempting to assess the efficacy of a new treatment modality (Brown, 1988). It is unclear to clinicians whether a depressed patient will require an antidepressant, will do well with psychotherapy alone, or has a brief self-limiting illness requiring no treatment at all. Clinicians cannot wait six months or more to see if patients will continue to have a persistent depression before offering them pharmacological treatment. Thus, it would be ideal if clinicians and clinical investigators could predict which depressed patients are more likely to respond to placebo and which are more likely to benefit from pharmacological treatment.

2.2 CHARACTERISTICS OF THE PLACEBO RESPONDER AND THE PLACEBO NONRESPONDER IN DEPRESSION

Considerable work has been done in an attempt to characterize the so-called placebo responder. Most studies indicate that there is no specific personality characteristics or predictors for the placebo responder. Shapiro, however, has characterized placebo responders as "compliant, religious, hypochondriac, anxious, less educated, frequently using cathartics, disturbed and likely to react to drugs with atypical reactions, depressed, dependent, ideational, neurotic and extroverted" (Shapiro, 1971). Bielsky and Friedel (1976), in a review of the literature, concluded that depressed patients with neurotic, hypochondriacal, or hysterical traits; those with a history of multiple prior episodes; and those with delusions responded similarly to placebo and tricyclic antidepressants. Joyce and Paykel (1989), in a more recent review, came to similar conclusions that good premorbid personality, insidious onset of depression, psychomotor retardation, absence of psychotic features, and intermediate level of severity favour antidepressant over placebo response. It appears that gender, age and

intelligence do not contribute to the prediction of a placebo responder (Fairchild, Rush, Vasavada et al, 1986; Brown, 1988).

Rabkin et al focused on the characteristics of mild to moderately depressed patients who did and did not improve during the 10 days of placebo treatment and concluded that placebo responders are essentially similar to nonresponders (Rabkin, Stewart, McGrath et al, 1987). Brown et al examined patients with major "endogenous" depression and found no clinical difference in placebo responders and nonresponders (Brown, Dorseif and Wernicke, 1988). Fairchild et al found that placebo nonresponders have a more chronic illness, a longer time since the first depressive episode, and a longer current episode than responders (Fairchild, Rush, Vasavada et al, 1986). Downing and Rickels (1973) found that patients depressed for more than six months are less likely than patients with shorter illness to improve with placebo.

2.3 BIOLOGICAL MARKERS OF PLACEBO RESPONDERS IN DEPRESSION

There are some biological markers that may predict patients who will have a low likelihood of placebo response. Neuroendocrine abnormalities have been described in depressed patients, with cortisol hypersecretion the best documented abnormality (Joyce, 1985). A subgroup of patients with major depression have a dysregulation in the pituitary-adrenocorticol function in the direction of hyperactivity or disinhibition (Brown, Shrivastava and Arato, 1987). Cortisol hypersecretion is related to nonsuppression on a dexamethasone suppression test (DST) and patients with the most severe depressions have the highest rates of pituitary-adrenocortical hyperfunction (Arana, Baldessarini and Ornesteen, 1985). The DST result, if initially abnormal (showing cortisol nonsuppression after dexamethasone administration), usually normalizes during successful antidepressant treatment, and failure of the DST to normalize, despite clinical improvement, is associated with poor outcome and early relapse (Holsboer, Liebl and Hofshuster, 1982).

There appears to be a lower placebo response rate (after 3-6 weeks of placebo) among DST non-suppressors (8%) than among DST suppressors (59%) (Brown,

Shrivastava and Arato, 1987). Brown et al also found, for the 1 week single blind placebo run-in period, more DST suppressors (8%) were placebo run-in responders (using the criterion of 20% improvement) as compared to DST nonsuppressors (4%) (Brown, Shrivastava and Arato, 1987). Similar findings were reported by Peselow et al who found that 14% DST suppressors versus 4% DST nonsuppressors were placebo responders (Peselow, Loutin, Wolkin et al, 1986), and Coryell and Turner (1985) who found that 28% DST suppressors versus 11% of DST nonsuppressors were placebo responders, both studies drew conclusions after 1 week of placebo treatment. Brown et al also found that DST nonsuppressors did not improve with psychological treatments alone, whereas >50% of DST suppressors improve with psychological treatments (Brown, Shrivastava and Arato, 1987). Findings that DST nonsuppressors show a slightly higher rate of response to antidepressants (76-82%) than DST suppressors (64-74%) (Arana, Baldessarini and Ornesteen, 1985) suggest that DST nonsuppressors are likely to require and respond to antidepressant treatment, whereas DST suppressors are more variable in their treatment requirements.

Among outpatients with major depression, treated with placebo for 2 to 6 weeks, 45% DST suppressors recovered as opposed to 0% of DST nonsuppressors (Brown, Shrivastava and Arato, 1987), a finding that was confirmed by Peselow et al (Peselow, Loutin, Wolkin et al, 1986). DST nonsuppressors are less likely to respond to placebo or to specific psychotherapies, so the presence of an abnormal DST may indicate the need for biologic treatment (Peselow, Loutin, Wolkin et al, 1986; Brown, Shrivastava and Arato, 1987). It is also interesting that DST suppressors have a higher incidence of accompanying psychiatric disorders including personality disorders, a high incidence of stressful life events, and tendency to recurrent depression (Zimmerman, Coryell and Pfhol, 1986), which are similar to other features of placebo responders.

DST results, however, are subject to numerous confounding variables, such as the multiple medications which influence dexamethasone metabolism, individual differences in pharmacokinetics of dexamethasone, and the particular DST protocol, including dose of dexamethasone and time of blood sampling (Brown, 1988); therefore, the use of the DST test and its application in clinical and research settings requires caution. Although

the DST test was used fairly frequently in the past, its use and application has fallen out of favour (personal communication, Dr. Walker, March 2001).

Another biologic marker that has shown some promise in the prediction of antidepressant response is the sleep abnormality of a shortened rapid eye movement (REM) latency. It appears that depressed patients with a shortened REM latency have a low likelihood of response to placebo but a high likelihood of response to tricyclic antidepressants (Coble, Kupfer, Spiker et al, 1979; Svendsen and Christensen, 1981). The use of the biologic markers such as the DST or REM latency as criteria for inclusion of subjects in trials to select patients that will have a low likelihood of placebo response may increase the drug effect in the study (Joyce and Paykel, 1989). However, these selection criteria may bias the sample as to make wider generalization of the findings dubious. The use of these markers would not have applicability in clinical medicine and the general population that may require antidepressants.

2.4 THE PLACEBO RUN-IN PHASE RESPONDERS VERSUS THE POST-RANDOMIZATION PHASE PLACEBO RESPONDER

It is estimated that about 5% of patients with unipolar depression show a positive response after one week of placebo washout (Loebel, Hyde and Dunner, 1986). Quitkin et al found that 15% patients responded to placebo during a 10-day placebo washout (Quitkin, Rabkin, Markowitz et al, 1987). The actual rate of elimination of placebo responders with the placebo run-in phase in studies may be as high as 20% (Sommers-Flanagan and Sommers-Flanagan, 1996). Outpatients who "respond" during the placebo run-in phase tend to have longer episodes, a more chronic illness, and a lower initial level of symptom severity and are more likely nonendogenous. In comparison, patients who respond to placebo *after* randomization tend to have shorter current episodes and higher symptom severity at randomization (Fairchild, Rush, Vasavada et al, 1986; Rabkin, McGrath, Stewart et al, 1986; Rabkin, Stewart, McGrath, 1987). These findings, albeit from different studies, hint that those who "respond" to a placebo run-in period may not

be isomorphic with those who ultimately respond to post-randomization placebo treatment.

A study by Rabkin et al comparing characteristics of 10-day washout placebo responders to 6-week (post-randomization) placebo responders suggested that there are two types of placebo responders (Rabkin, Stewart, McGrath et al, 1987). The 10-day washout placebo responders have a milder illness (based on lower severity levels on SCL-90 factors [Derogatis, 1977] and HAM-D scores), and have a more chronic illness. There are fewer cases of primary depression in this group and they have fewer illness precipitants. The 6-week placebo responders have a shorter duration of depression, they have less time spent depressed during illness episodes and illness precipitants are more often reported. Rabkin et al also reported that there were fewer 10-day placebo washout responders in the months with the least daylight (November to February), consistent with the their hypothesis that placebo response is associated with seasonal variation. In the same study, they compared five groups of patients, 10-day placebo washout responders, 6-week (post-randomization) placebo responders, 6 week placebo nonresponders, 6-week drug responders and 6-week drug nonresponders and found no significant differences between the different groups in terms of age, gender, marital status, educational level, previous psychiatric hospitalization, prior episodes of depression, family history of depression, or prior psychiatric treatment.

Similarly, Fairchild et al found in a study of outpatients with unipolar depression, that placebo responders and placebo nonresponders do not differ significantly in age, education, marital status or occupation (Fairchild, Rush, Vasavada et al, 1986). There were no significant differences between washout placebo responders and post-randomization placebo responders, so both groups were combined as one group of placebo responders, that was used to compare with the group of placebo nonresponders. The study revealed that placebo responders had a shorter length of illness, as defined as the time in months since the onset of the first episode of major depression, but both groups of placebo responders and nonresponders reported an equivalent number of episodes of major depression (Fairchild, Rush, Vasavada et al, 1986). In the study, clinicians' ratings of symptom severity (HAM-D, Covi, and Raskin) showed no

significant differences between the two groups of placebo responders and nonresponders prior to treatment, but patients' self-reports of symptomatology before treatment revealed that responders experienced less depression according to the Carroll Rating Scale for Depression (CRS) (Carroll, Feinberg, Smouse et al, 1981) and the Beck Depression Inventory (BDI) (Beck, Ward, Mendelson et al, 1961). Placebo responders also had a higher incidence of having another psychiatric disorder (either concurrently or in the past) in addition to major depression. Although placebo responders had a shorter length of current episode and a shorter total length of illness, analysis of the data were not able to explain placebo response based on spontaneous remission in the sample (Fairchild, Rush, Vasavada et al, 1986).

Fairchild et al summarized that, endogenous subtype, presence of other psychiatric disorders (especially anxiety, substance abuse or personality disorder), and the length and chronicity of the depression significantly distinguished placebo responders from nonresponders (Fairchild, Rush, Vasavada et al, 1986). Downing and Rickels (1973) also reported that placebo was less effective in more chronic than in acute depressions. Findings of less severe depressive symptomatology (on patient self-reporting) in placebo responders is consistent with a 1973 report by Downing and Rickels in which depressed outpatients reported on the Zung self-report depression scale (Zung, 1965). Fairchild et al found that endogenously depressed patients did not respond to placebo as frequently as nonendogenously depressed patients (Fairchild, Rush, Vasavada et al, 1986). This corroborated the findings of Kiloh, Ball and Garside (1962), Raskin, Schulterbrandt, Reating et al (1967), Wittenborn and Kiremitci (1975), Bielsk and Friedel (1978), and Paykel (1972). The lack of endogenous features did not ensure placebo response as 50% of the placebo nonresponders had nonendogenous depression (Fairchild, Rush, Vasavada et al, 1986).

2.5 RELAPSE OF THE PLACEBO RUN-IN PHASE RESPONDER

There has been a paucity of research examining the subsequent outcome for placebo run-in phase responders, likely because this group is of minimal interest to the pharmaceutical industry. Without information on their subsequent clinical status, it is unclear whether their initial improvement during the 7-14 day placebo run-in phase is transient or will be sustained over time. Rabkin et al in 1986 reported a retrospective analysis of the follow-up of 10 day placebo run-in phase responders (Rabkin, McGrath, Stewart et al, 1986). In a study of 60 such patients who were followed after their elimination from the trial, 10 subjects were given psychotropic medication immediately even though their depressive symptoms were ameliorated. (The medication was given for concurrent problems of panic attacks, anxiety, bulimia, and other mood symptoms.) An additional 5 patients in that group were lost to follow-up. The remainder 45 patients constituted their sample as 10 day placebo responders and were followed up for 12 weeks. Based on their clinical status at 12 weeks, these 45 placebo responders were then classified as being relapsers (depressive symptomatology recurred) or nonrelapsers (remained well). There were 25 relapsers and 20 nonrelapsers, with 22 out of the 25 relapsers requiring medication within 6 months. The study found that relapsers had an earlier age of illness onset, had a more chronic illness history, were more likely to have received prior treatment and more often reported a family history of depressive disorder.

Overall, patients who relapsed in the study described a more chronic picture of illness with earlier onset and a greater proportion of time spent depressed. There were also more male patients and more patients who had never been married in the relapsing group. Nonaffective psychiatric disorders (substance abuse, panic disorder, ego-dystonic homosexuality, generalized anxiety disorder) were present in 64% of relapsers, but in none of the nonrelapsing group. The following diagnostic differences between the relapsing and nonrelapsing groups include: there were more nonrelapsers who met criteria for recurrent, simple, and situational depression. However, the two groups did not differ in terms of presenting depressive psychopathology and self-rated report ratings. Despite the above findings, the nonrelapsers were not entirely "well" after the

12-week study period. On a further 12-week follow-up period, after the initial 12-week study period, 25% of the nonrelapsers required non-pharmacological treatment, although they did not become worse than they were at day 10 when they were identified as placebo responders. As 40% of the relapsers relapsed before the end of two weeks and required antidepressant treatment, Rabkin et al questioned the concept of "placebo response". It has been suggested that placebo responders be only identified as such after a sustained period rather than rating on a single occasion, before excluding them from randomized clinical trials (Rabkin, McGrath, Stewart et al, 1986).

Unfortunately, the literature did not appear to have any studies analyzing placebo response in depression at varying time intervals. Most of the studies analyze data either relating to placebo response in the placebo run-in phase (7-14 days) or placebo response at the end of the study (usually 4-8 weeks).

It may be that patients who respond to placebo may be responding to something other than the placebo (Brown, 1988), such as spontaneous remission of their illness, the multiple favorable psychological effects of pill taking (Strayhorn, 1987), the elements of the treatment environment (attention, concern, opportunity to verbalize distress, or expectation of improvement). The Hawthorne effect, which is the effect of an observer on any study, and the increased attention on the patient, is well known to cause improvement (Spiro, 1986). The therapeutic alliance has been shown to affect outcome in psychotherapeutic trials. Horvath and Symonds (1991) have reported that a positive therapeutic alliance has an effect size of 0.26 on outcome in 24 psychotherapeutic trials. Similar positive correlations between the therapeutic alliance and outcome have been reported for antidepressant drug therapy (Marzialli, Marmar and Krupnick, 1981).

Quitkin et al reported in 1984, and later confirmed in 1987, that the true drug effect from antidepressants is characterized by a two-week delay in onset and persistence, and there is little evidence of the onset of antidepressant effect before two weeks (Quitkin, Rabkin, Ross et al, 1984; Quitkin, Rabkin, Markowitz et al, 1987). Therefore, early response, especially if it is not sustained, may be safely attributed to a placebo effect. In a study involving the use of phenelzine, imipramine, desipramine,

mianserin and placebo, it was found that, for each drug, onset of persistent improvement was most likely to occur during weeks 3 through 5 (Quitkin, Rabkin, Ross et al, 1984).

CHAPTER 3: THE PLACEBO RUN-IN PHASE

3.1 RATIONALE FOR THE PLACEBO RUN-IN PHASE

The origin of the introductory placebo washout technique is obscure. The 1977 *Guidelines for the Clinical Evaluation of Antidepressant Drugs* recommended that a drug-free period should precede the start of an antidepressant drug trial, but did not mention combining it with the administration of a placebo or following it with the withdrawal of placebo responders (Crout and Finkel, 1977). The placebo run-in phase is recommended by some researchers to eliminate subjects that may be non-compliant, to ensure that prior treatment(s) are washed out, to eliminate rapid remitters, to ensure a more homogeneous sample of subjects (so that fewer subjects are needed to draw a conclusion about treatment effect) and to provide a period for baseline measurement and careful medical evaluation before exposure to an experimental treatment (Pablos-Mendez, Barr and Shea, 1998).

In controlled clinical trials, much effort, usually on the part of the pharmaceutical company researching the drug, is put into detecting placebo responders and eliminating them as subjects, so that the benefits of a drug will be more evident. However, for the medical researchers running the study, there is a financial disincentive to eliminate any subjects that have been recruited, as financial reimbursement is usually based on the number of subjects who complete the study (Walker, personal communication, March 2001). Conduct of clinical trials have changed over the years and these changes may have some effect on the outcome of studies (Schweizer and Rickels, 1997). These changes include the expectation of faster recruitment, involvement of multiple sites, and employment of external contract-research organizations to monitor the progress of the studies (Schweizer and Rickels, 1997).

3.2 ARGUMENTS AGAINST THE PLACEBO RUN-IN PHASE

Although the use of the placebo run-in phase in randomized clinical trials is considered accepted standard practice, there are many who feel that this practice lowers the placebo response rates and disadvantages the placebo, thereby favoring the study drug (Senn, 1997). Skovlund (1994) argued that the single-blind placebo run-in phase should be omitted from clinical trials because it may give a distorted picture of treatment effects.

In some randomized controlled trials, the use of the placebo run-in period resulted in overestimates of the benefits of treatment (Pablos-Mendez, Barr and Shea, 1998). In the article by Pablos-Mendez et al, several studies, that incorporate the run-in period, were analyzed. The Physicians' Health Study was a large clinical trial designed to test the effects of aspirin and beta-carotene in the primary prevention of ischemic heart disease and cancer among male physicians in the United States. The study used a run-in period to screen for adherence and 33% of subjects were withdrawn from the study prior to randomization. The study found that aspirin decreased the rate of myocardial infarction with a relative risk of 0.56. If all subjects who entered the run-in period had been randomized and if it is assumed that nonadherent subjects remained nonadherent and had no treatment effect, the recalculated relative risk would be 0.71, a difference of 25%.

Another study analyzed by Pablos-Mendez et al in the article was a randomized trial comparing pindolol plus fluoxetine versus placebo plus fluoxetine in the treatment of major depression. Of the 132 eligible subjects who entered the placebo run-in period, 111 were randomized into the trial and 21 were not, 19 due to placebo response and 2 due to patient withdrawal. The pindolol/fluoxetine study found a greater proportion of responders in the fluoxetine plus pindolol group, with $p=0.04$. If the 19 placebo responders had entered the trial, the recalculated results would still have trended in favour of the fluoxetine plus pindolol group, but would not have reached statistical significance, with $p=0.09$. Pablos-Mendez et al concluded that the use of the run-in period led to an overestimation of the benefits of the treatment and recommended that

evidence from studies with run-in periods be recalculated as if the run-in period had not been used.

3.3 THE PLACEBO RUN-IN PHASE IN DEPRESSION (3 specific studies)

Although the above studies analyzed by Pablos-Mendez indicated that the use of the a placebo run-in period increased the drug-placebo difference, this procedure does not appear to provide an advantage for the active drug in randomized clinical trials for the treatment of major depression. Three studies (2 of which are meta-analyses) involving the treatment of major depression and the use of the placebo run-in period led to the conclusion that the placebo run-in period does not affect the study outcome (Greenberg, Fisher and Riter, 1995; Reimherr, Ward and Byerley, 1989; Trivedi and Rush, 1994).

3.3A Reimherr et al (1989)

Reimherr et al (Reimherr, Ward and Byerley, 1989) reported a retrospective reanalysis of their earlier study (Byerley, Reimherr, Wood et al, 1988) and found that the elimination of prerandomization placebo run-in responders reduced the drug-placebo difference and increased the placebo treatment response rates in outpatients with major depression. The original study (Byerley, Reimherr, Wood et al, 1988) was a 6-week double-blind randomized controlled placebo trial with three cells comparing treatment with fluoxetine, imipramine and placebo in depression. Subjects were recruited from local mental health centres, from private practitioners and from a limited amount of advertising. They were required to meet DSM-III criteria for major depression with a duration of at least one month and to have a HAM-D score of at least 20 to be eligible for the study. All subjects were initially given an inert placebo for 7 days. After the placebo run-in phase, those subjects whose HAM-D score improved by $\geq$20% or decreased below 20 were considered "placebo responders" and withdrawn from the study. In the original study, 18 subjects were withdrawn as they improved significantly

during the placebo run-in phase. The remaining subjects entered the actual double-blind, three-cell study and were randomized to receive either fluoxetine, imipramine or placebo. Progress was monitored in the study weekly using the 21-item HAM-D and the 7-point CGI.

The Hopkins Symptom Checklist (SCL) was also administered both at baseline and after 7 days of placebo, and weekly during the study, but it was not used as a criterion for retention in the study or in the initial analysis of the data. (In fact, the SCL was not even scored until the entire study was completed.) All subjects who completed at least two weeks of treatment were included in the data analysis. In the initial study, it was shown that both fluoxetine and imipramine were significantly superior to placebo in the treatment of depression, using the HAM-D and CGI ratings. On the basis of baseline and (6-week) endpoint HAM-D scores, both fluoxetine and imipramine treated subjects improved significantly more than placebo treated subjects. Using the CGI scores, 65% subjects treated with fluoxetine, 56% of subjects treated with imipramine, and 10% subjects treated with placebo were rated as moderately or markedly improved after 6 weeks of treatment.

Reimherr et al hypothesized that since the HAM-D was administered by a physician who was aware that the patient was on placebo after the placebo run-in phase, bias could occur. Such biases could potentially have altered the assessment; such as pressures to produce a study population could have led to the underestimation of improvement during the placebo run-in phase. As a result, a number of patients could have been entered into the study that were actually "hidden placebo responders". In the re-analysis of the earlier study, a subject was considered to be a "hidden placebo responder" if he was included in the study and had showed a >20% improvement on the depression scale of the SCL during the placebo run-in phase. Reimherr et al postulated that since the SCL was completed by the subject (who was not aware that he was on placebo tablets for the initial week of the study), it provided an alternative and perhaps less biased measure of improvement on placebo. Results on the SCL depression scale did indeed show that a significant number of subjects with very positive responses to placebo had been entered into the study. Reimherr et al determined that the subjects who

were identified as placebo responders and withdrawn from the original study had the same amount of improvement (based on improvement on the depression scale of the SCL) as the subjects who were retained in the original study but were later identified as "hidden placebo responders".

In the follow-up article, Reimherr et al then re-analyzed the results of the original study with the "hidden placebo responders" included and excluded from the analysis. Subjects were divided into three groups based on their initial response to placebo: (1) positive initial placebo responders (mean initial improvement of the SCL depression scale=21%); (2) neutral initial placebo responders (mean initial improvement on the SCL depression scale=2%); (3) negative initial placebo responders (mean initial deterioration on the SCL depression scale=-10%). The drug study results for each of these groups were then compared.

The results of the study were calculated with the total population, including and excluding the "hidden placebo responders" on two measures of improvement, the CGI (% of patients rated at end point as "much improved" or "very much improved") and the HAM-D (baseline to end point improvement score). The exclusion of "hidden placebo responders" did not improve the original results. In fact, it diminished the size of the differences in improvement between the active and placebo groups from 30% to 25% on the CGI and, surprisingly, it increased the placebo response rate observed at the end of the study from 13% to 16% based on the CGI. Similarly, the differences between the treatment and placebo groups measured by HAM-D decreased when "hidden placebo responders" were excluded from the analysis. Improvement in the placebo group decreased by 0.4 HAM-D points while improvement in the active treatment group decreased by 0.8 HAM-D points, resulting in a decrease in the t value from 2.5 ($p<0.02$) to 2.06 ($p<0.05$). Reimherr et al also concluded that the HAM-D and the SCL depression scale were measuring very closely related symptoms and the changes were similar in magnitude, based on correlation coefficients.

Although Reimherr et al initially thought that exclusion of the "hidden placebo responders" would improve the differences observed in the original analysis, the results in fact deteriorated. Reimherr et al concluded that the placebo run-in phase may have

unpredictable, possibly confounding effects as elimination of "hidden placebo responders" from the study analysis diminished differences between the active and placebo treatment groups in the study. Reimherr et al also questioned the view that placebo responsiveness is a stable characteristic of an individual patient and that it can be accurately measured by use of a placebo run-in phase. The data suggested that placebo responsiveness is not a stable characteristic but is a more complex artifact of clinical trials. Reimherr et al suggested that the placebo run-in phase may artificially suppress the level of placebo-induced improvement during the study. They also argued that the use of the placebo run-in phase is in violation of the guidelines for the protection of human subjects in that complete disclosure about the nature of the study is not revealed, thus subjects are not providing "informed consent". Although they recognize some of the practical reasons for the use of the placebo run-in phase (such as assessment of compliance and wash-out of previous medication), Reimherr et al concluded in his paper with the statement that "once the decision to enter a patient is made and the patient has started on medication, our data indicate that tampering with the study population, as occurs with a placebo washout procedure, may be detrimental to the study".

3.3B Trivedi and Rush (1994)

Trivedi and Rush (Trivedi and Rush, 1994) reported a meta-analysis of 101 randomized controlled trials that compared the efficacy of antidepressant medications in major depression, and found that a placebo run-in phase did not lower the placebo response rate, did not increase the drug-placebo difference and did not affect the drug response rate post-randomization. A search was done of Medline and Psychological Abstracts from 1975-1990. The inclusion criteria included - English language, diagnosis of major depressive or bipolar disorder, study duration of at least three weeks, use of a quantitative outcome measure, comparison between a known antidepressant drug and a placebo (another medication, or both), blinded study. They identified a total of 141 eligible, randomized placebo-controlled trials and meta-analyzed 101 of them. The analysis focused on dichotomous outcome measures (categorical scoring), based

primarily on a 50% reduction in the HAM-D as a responder. If the HAM-D was not reported, then a CGI response of 1 or 2 (markedly improved or very much improved) was counted.

The success of a treatment was reported in three different ways. First, they used the intent-to-treat analysis, which utilized all patients who improved (regardless if they remained in the study) as the numerator, and the number randomized to treatment as the denominator, which addressed the question of how many patients randomized to the treatment improved. Second, an "adequate treatment" analysis included only patients who received a predetermined minimum amount of treatment (typically 2-4 weeks for medication) as the denominator and counts those that responded as the numerator, that addressed the question of how many improve from receiving at least the minimal amount of treatment thought to be effective. Thirdly, a "completer" analysis included only those who completed the full protocol, with the numerator and denominator including only those subjects.

The meta-analysis was conducted using the Confidence Profile Method (Eddy, Hasselblad and Schacter, 1990), that uses a hierarchical Bayesian random-effects model. Findings were reported separately for inpatients and outpatients as it was felt that attrition rates, and placebo and drug response rates, may be different for the two groups. Adult and geriatric studies were combined as there was no evidence of differential responses in the two groups and the number of geriatric subjects were too few for meaningful independent analyses. The antidepressants medications were divided into groups: tricyclic medications (amitriptyline, desipramine, doxepin, imipramine, nortriptyline and protriptyline), heterocyclic medications (maprotiline, amoxapine, trazadone and buproprion), selective serotonin reuptake inhibitors (fluoxetine, paroxetine and sertraline) and MAOIs (isocarboxacid, phenelzine and tranylcypromine).

Of the 101 studies included in the meta-analysis, 50 trials had a placebo run-in phase and 51 trials did not have a placebo run-in phase. The final conclusions from the meta-analysis were that a placebo run-in phase did not reduce (or increase) the post-randomization placebo response rate, drug response rate, or drug-placebo differences. Trivedi and Rush concluded that if there are sufficient symptoms at randomization to

indicate drug treatment, the patient should be entered into the study. They proposed that the raw HAM-D score at the time of randomization should be the primary variable to dictate inclusion or exclusion, not the lack of response to the placebo run-in phase. In summary, they question the value of the single-blind placebo run-in phase and their findings lend no support to the need for such a phase in clinical trials.

3.3C Greenberg et al (1995)

The most recent meta-analysis involving placebo run-in periods and antidepressant drugs was carried out by Greenberg et al in 1995 (Greenberg, Fisher and Riter, 1995), and their findings were consistent with the two previously mentioned studies. They reported a meta-analysis of 28 double-blind, placebo-controlled trials of the effectiveness of antidepressant medications, based on a literature search of Medline and Psychological Abstracts from 1983-1992. The analysis was restricted to trials that compared placebo to only one drug, and did not include subjects with bipolar or psychotic disorders. The criteria for inclusion included studies that reported results in a fashion which enabled computation of the percentage of reduction in scores on the HAM-D from baseline to endpoint (dichotomous outcome measure). Of the 28 studies that met the criteria as outlined, 20 studies (involving a total of 848 subjects) included a placebo run-in phase and 8 studies (involving 241 subjects) did not include such a phase. The numbers of placebo response as well as response to medication were compared in the two types of studies. The percentage of dropouts were also compared. Comparisons were made with the use of t tests.

The analysis revealed that there was no significant difference between the placebo run-in and the non-placebo run-in studies in the percentage reduction in ratings on depression for subjects in the placebo groups (24.8% and 23.7% respectively). There was also no difference in the effectiveness of the antidepressant drugs in the placebo run-in and non-placebo run-in groups (48.9% and 43.3% respectively). Their analysis showed equivalent percentages of dropouts in the two groups for both the subjects assigned to the placebo groups (43.2% for placebo run-in group and 37.2% for the non-

placebo run-in group) and those assigned to the active drug groups (36.2% for the placebo run-in group and 39.9% for the non-placebo run-in group). Greenberg et al argued that the single-blind placebo run-in phase should be omitted from clinical trials as it is costly in terms of time and effort and does not serve the purpose for which it was designed (lowering the level of response to placebo and magnifying the superiority of the response to the active drug in the study). The conclusion was that the use of the placebo run-in phase may give a false sense of security about the solidity and generalizability of antidepressant drug trial results.

3.4 COMPARISON OF THESIS TO THE 3 SPECIFIC STUDIES

This thesis has similarities to the above mentioned meta-analyses in that it examined the effect of the placebo run-in period in clinical trials of the treatment of depression, but it was different in several ways. Unlike Trivedi and Rush (who drew from publications from 1975 to 1990), and Greenberg et al (who drew from publications from 1983-1992), this study included more current randomized controlled trials (up to the year 2000), but limited trials to those involving the use of selective serotonin reuptake inhibitor (SSRI) antidepressants.

SSRIs were developed and studied when there was increasing evidence indicating the role of central serotonin metabolism deficiency in the pathogenesis of depression (van Praag, 1980). SSRIs are a newer group of antidepressants that have similar effectiveness to the traditional tricyclic antidepressants, but have superior tolerability (Benfield and Ward, 1986; Benfield, Heel and Lewis, 1986; Dechant and Clissold, 1991; Murdoch and McTavish, 1992). When compared with the tricyclic antidepressants, the SSRIs have minimal adverse psychomotor effects, anticholinergic side effects, problems of weight gain or cardiotoxicity, resulting in a reduced risk of toxicity in overdose (Henry, 1992). Due to the decreased incidence of side effects with the use of SSRIs, discontinuation rates are lower with SSRIs than with tricyclic antidepressants (Montgomery, Henry, McDonald et al, 1994). As a result of their effectiveness and minimal side effect profiles, SSRIs have met wide acceptance as first-line therapy for

depression (Landen, Bjorling, Agren et al, 1998). There are currently five SSRI antidepressants available – fluoxetine, fluvoxamine, paroxetine, sertraline and citalopram (Lemberger, Fuller, Zerbe, 1985).

In contrast to the two published meta-analyses on placebo run-in periods, this study used continuous outcome measures. The published studies dichotomized the continuous measures into participants who experienced an arbitrary percentage reduction in symptoms (such as a greater than 50% reduction in the total HAM-D score). As such an approach is arbitrary and of uncertain clinical relevance, statistical power will be sacrificed. This thesis will use the primary measurement as stated in each study for analysis.

Unlike Greenberg et al, this study did not exclude trials, that compared placebo to more than one antidepressant drug, and it did not limit the analysis to the HAM-D as the measurement outcome. The inclusion of more than one active drug group in the study should decrease the clinicians' ability to identify one of the groups as representing their vested interest, which should decrease the potential for bias in the study (Greenberg, Bornstein, Greenberg et al, 1992).

CHAPTER 4: META-ANALYSIS

4.1 INTRODUCTION

In 1976, Glass coined the term "meta-analysis" to refer to "the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings" (Glass, 1976). Today, meta-analysis is viewed as a standard research process and the number of published meta-analyses has increased substantially in the past decade (Chalmers and Haynes, 1995). The use of meta-analysis is so widespread that there has been several books written on the subject of research synthesis, *The Handbook of Research Synthesis* (Cooper and Hedges, 1994) being one very good resource. In the medical field, the international Cochrane library is a well-known and widely used resource. The international Cochrane collaboration conducts research synthesis and meta-analyses, that results in high-quality systematic reviews of health-care interventions (Bero and Rennie, 1995), that are helpful for clinical decision-making. Meta-analysis, or more appropriately, research synthesis is currently defined as a review in which bias has been reduced by the systematic identification, appraisal, synthesis, and, if relevant, statistical aggregation of all relevant studies on a specific topic according to a predetermined and explicit method (Cook, Sackett and Spitzer, 1995).

There are several statistical summary units in meta-analysis (Revman 4.1 User Guide). For dichotomous data, there are the odds ratio, relative risk, and risk difference. For continuous data, there are the weighted mean difference and the standardized mean difference. One statistical unit of interest in the comparison of treatment effectiveness is the effect size (ES), expressed as the difference between treated and control group means, divided by the pooled standard deviation:

$$ES = (M_1 - M_2)/SD \qquad \text{(Glass, 1977).}$$

By using a standard unit, findings are transformed into a common metric (standard deviation units), indicating the magnitude of effect or change. An effect size of +1.25 would demonstrate a larger magnitude of change or effect than an effect size of +0.75.

Of course, primary research on a topic must exist before its synthesis can be conducted. There are several stages of research synthesis: problem formulation, data collection (literature search), data evaluation, and analysis and interpretation (Cooper and Hedges, 1994). These stages will be further developed in the Methods section of this study.

4.2 OBJECTIVES

The objective of this research is to compare the effect size of randomized controlled double-blinded placebo trials (in the treatment of depression with SSRIs) that include a placebo run-in period (that exclude placebo responders) with those that do not include a placebo run-in period. It is hypothesized that the size of the treatment effect will be larger in studies, that eliminate placebo responders from the study after a placebo run-in phase.

4.3 METHODS

4.3A Search Strategy

The search strategy aimed to identify published clinical trials of randomized controlled double-blinded clinical trials of the treatment of major depression with at least one SSRI investigation drug and a placebo. Medline and Psychlit databases were searched, employing the following key words: "depression" and "placebo" and "English in LA (language)" and ("fluoxetine" or "fluvoxamine" or "sertraline" or "paroxetine" or "citalopram"). The search was limited to studies that involve SSRI antidepressants to produce a more homogeneous sample of studies. The attempt was to identify all

appropriate trials with SSRI antidepressants since their introduction in the early 1980's to the year 2000. Great care was taken to ensure that duplication of trials did not occur.

An electronic search of the databases identified 734 potential entries. Articles that clearly did not fit the inclusion criteria as judged by titles and abstracts were rejected. There were 111 promising articles that appeared to fit the inclusion criteria on reviewing the abstract. The actual articles of these 111 trials were retrieved and examined closely for consideration of inclusion into the study and 34 were included as all the inclusion criteria were met.

Trials were also identified from hand searching of the indexes of the following major psychiatric journals for any randomized controlled placebo trials of depression and treatment with an SSRI antidepressant:

Acta Neurologica Scandanavia
Acta Psychiatrica Scandanavia
American Journal of Psychiatry
Annals of Clinical Psychiatry
Archives of General Psychiatry
Current Therapeutic Research
Biological Psychiatry
British Journal of Psychiatry
Journal of Affective Disorders
Journal of Clinical Psychiatry
Journal of Psychopharmacology

The above journals were hand searched from 1995 to 2000 with the exception of Current Therapeutic Research, which was hand searched from 1985 to 2000, as the journal is not included in the Medline or Psychlit electronic databases. Hand searching identified 28 additional trials not found previously on the electronic database search of Medline and Psychlit and all the trials were closely examined, but none of them met the inclusion criteria.

The references of all relevant meta-analyses of the treatment of depression were carefully examined. There were 143 trials identified in the references of meta-analyses, but only 8 trials met the inclusion criteria and were included in the study after close examination. In total, 42 trials met the inclusion criteria and were selected to be included in the meta-analysis. (See Appendix I for included trials and Appendix II for excluded trials.)

The international Cochrane databases were searched to determine whether a similar study (meta-analysis analyzing placebo responder elimination and effect size in the treatment of major depression) has been done. There has been no review of this type published to date. In 2000, the Cochrane library reported a review comparing SSRIs and other antidepressants in the treatment of depressive disorders (Geddes, Freemantle, Mason et al, 2000). Although this review did not approach the topic of placebo run-in periods, it provided an excellent resource for this research.

4.3B Selection Criteria

All studies were examined for inclusion into the study using a pre-developed assessment form (see Appendix III), that included the following criteria:

Inclusion criteria:

- Randomized, controlled, double-blind clinical trials
- Adult subjects, ages 18-65 years
- DSM-III or DSM-IV diagnosis of primary acute major depression
- Placebo arm
- SSRI investigational drug
- Duration of study at least 6 weeks
- Parallel design
- Continuous outcome measurement (mean change in depression scale)

Exclusion criteria:

- Concomitant primary psychiatric diagnosis other than major depression
- Concomitant medical illness
- Augmentation therapy
- Continuation therapy
- Crossover design
- Active drug wash-out phase

4.3C Quality Assessment

The quality of each selected study was assessed by two reviewers using a Jadad score (Moher, Jadad, Nichol et al, 1995). Studies were rated and given a quality rating from 1 (poorest rating) to 5 (highest rating) (Appendix IV). A Jadad score of at least 3 indicates that the study was of good quality. It was expected that all the trials would have at least a Jadad score of 3, as they were all randomized, double-blind clinical trials. Concealment of allocation was also assessed and graded:

A = adequate

B = uncertain

C = not adequate

4.3D Data Extraction

A form was developed to facilitate data extraction (Appendix V). Information extracted included study location, participant details, type and duration of intervention, primary outcome measurement, mean change in outcome (and standard deviation of the change if available) and the presence or absence of a placebo run-in period that withdraws placebo responders. Two reviewers independent of each other performed data extraction. A third reviewer compared the two independent reviews and found 86% agreement. In cases of disagreement the study was reviewed by a third reviewer, who also extracted the data. The three reviewers met to come to consensus about the data in the study. The extracted data were entered into Review Manager 4.1 for statistical analysis.

4.4 DESCRIPTION OF STUDIES

A total of 42 trials were included in the study. Appendix VI provides a summary of all the trials and the data extracted from each trial. All of the trials included 1 SSRI antidepressant with the exception of one trial that compared 2 SSRI antidepressants with placebo (Fava 1998). The data for the Fava trial were analyzed separately for each SSRI;

therefore a total of 43 sets of data were analyzed. There were 13 trials with fluoxetine, 8 trials with fluvoxamine, 14 trials with paroxetine, 6 trials with sertraline and 1 trial with citalopram. There were 31 trials with a duration of 6 weeks and 11 trials with duration of more than 6 weeks, ranging from 8-12 weeks. All of the trials, with the exception of two (Feighner, 1989 and Lapierre, 1987) were done on the outpatient population. The study size of the trials ranged from 8 to 650 participants. Thirty of the trials were published in the years 1990 to 2000, with the remaining 12 trials published prior to 1990. Twenty-seven trials had Jadad scores of 3 with 15 trials with Jadad scores of greater than 3.

All the primary outcome measurements were continuous, with 15 trials using the HAM-D 21 scale, 15 trials using an unspecified HAM-D scale, 8 trials using the HAM-D 17 scale, 1 trial using the HAM-D 13 scale, 1 trial using the HAM-D 31 scale and 2 trials using the Montgomery-Asberg depression scale (MADRS) (Montgomery and Asberg, 1979). All the trials presented the mean change in the primary measurement outcome, either in tabular or graphical format. Eleven trials reported the standard deviations of the mean change in outcome measurement directly, and 2 trials reported the standard errors of the mean change in outcome measurement, and their respective standard deviations were calculated indirectly, using the formula $SD = \sqrt{n}$ X SE (Hassard, 1991). Twenty-nine trials did not report standard deviations; therefore, estimates were obtained from other trials using the same outcome measures as outlined below.

As different measurement scales differ in length and content, they should also differ in standard deviations (personal communication, Dr.Hassard, June 2001). For each measurement scale (HAM-D 21, HAM-D 17, HAM-D unspecified and MADRS), the available standard deviations (either directly from the trials or calculated from the

standard errors from the trials) were pooled to obtain a pooled standard deviation for trials using the same measurement scale without a reported standard deviation, using the formula (Hassard, 1991):

$$\text{Pooled SD} = \sqrt{\frac{(n_1\text{-}1)SD_1^2 + (n_2\text{-}1)SD_2^2 + \ldots\ldots\ldots + (n_k\text{-}1)SD_k^2}{n_1 + n_2 + \ldots\ldots\ldots + n_k - (N)}}$$

	HAM-D 17	HAM-D 21	HAM-D	MADRS
Pooled standard deviations:	6.94	8.49	8.75	8.34

As the original Hamilton depression scale used 21 items (Hamilton, 1960), it is likely that the trials that did not specify a number on their HAM-D scale used the HAM-D 21 scale. This is supported by the similar values of the pooled standard deviations of the HAM-D 21 and the unspecified HAM-D scales. It is also important to note that the pooled standard deviation of the HAM-D 17 measurement scale is approximately 17/21 of the HAM-D 21 pooled standard deviation. The trial using the HAM-D 13 measurement scale (Walczak, 1996) did not report a standard deviation; therefore the pooled HAM-D 17 standard deviation was multiplied by 13/17 (=5.31) and was used for analysis. The trial using the HAM-D 31 measurement (Croft, 1999) did not report a standard deviation; therefore the pooled HAM-D 21 standard deviation was multiplied by 31/21 (=12.53) and was used for analysis.

Twenty-nine trials reported a placebo run-period, that withdrew placebo responders from the pool of subjects prior to randomization, and 13 trials did not report

withdrawal of placebo responders. (For the purpose of this research, it will be assumed that if no placebo run-in period is mentioned in the study, that the study did not have a placebo run-in phase and thus no subjects were withdrawn as placebo responders.) Of the 29 trials that reported withdrawal of placebo responders, only 3 specifically mentioned the number of placebo responders withdrawn. Byerley, Reimherr et al (1988) withdrew 5 participants, but did not mention the total number of subjects who entered the placebo run-in phase. Heilgenstein, Tollefson, Faries et al (1993) and Lydiard, Stahl, Hertzman et al (1997) reported withdrawing 9/164 and 81/473 subjects respectively after the placebo run-in period.

4.5 DATA SYNTHESIS

Using Review Manager 4.1 software (Revman, Cochrane Collaboration, 2000), an analysis of the standardized mean difference of the primary study outcome measure was performed. Metaview version 4.0, which is the statistical program within Review Manager 4.1, implements Hedges adjusted g for continuous outcomes, which is similar to Cohen's d, but includes an adjustment for small sample bias (Revman 4.1 User Guide, 2000). Continuous data from different measures were transformed into standard effect sizes by dividing mean values by standard deviations. In the graphical presentation of the analyses, negative standardized mean differences (falling to the left of the midline) favour the SSRI antidepressants and positive standardized mean differences (falling to the right of the midline) favour the placebo.

Forty-two trials (43 sets of results as the Fava 1988 trial included two SSRIs) contributed data to the analysis of the relative efficacy of SSRIs with placebo. Analysis

of efficacy was based on 3047 subjects treated with a SSRI antidepressant and 3740 subjects treated with a placebo. For most of the analyses, the standardized mean difference (SMD) was used to compare the effect size of the SSRIs with placebo as different outcome measurements were used in the different trials (HAM-D 21, HAM-D 17, HAM-D unspecified and MADRS). When trials using the same outcome measure are used to compare drug effect, the weighted mean difference (WMD) may be used, as the units will be equivalent.

Analyses were done using the fixed effects model and the random effects model, but the random effects model statistics will be presented as most of the analyses indicated significant heterogeneity. The main analysis with all the data will be presented first, followed by sensitivity analysis.

4.6 OVERALL SUMMARY

The SMD for SSRIs versus placebo using a fixed effects model was -0.40 (95%CI -0.45 to -0.35; Q=183.36, df=42, p<0.00001; z=15.77, p<0.00001), indicating that SSRIs are more efficacious than placebo (Figure 1).

Figure 1: Overall Data - Standard Mean Difference - Fixed

Comparison: 01 All Inclusive Data Comparisons
Outcome: 01 OVERALL DATA FIXED EFFECTS SIZE

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Fixed)	Weight %	SMD (95%CI Fixed)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		7.2	-0.48[-0.67,-0.29]
Byerley 1988	29	7.80(6.50)	32	14.40(7.70)		0.9	-0.94[-1.47,-0.41]
Cleghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		1.4	0.41[-0.01,0.83]
Cleghorn & Kiev 1992	162	8.80(8.49)	163	12.90(8.49)		5.1	-0.51[-0.73,-0.28]
Cleghorn 1992	26	5.50(8.49)	32	11.50(8.49)		0.9	-0.70[-1.23,-0.16]
Cleghorn 1992ex	27	5.49(8.31)	32	11.44(8.32)		0.9	-0.71[-1.23,-0.16]
Cohn & Crowder 1990	36	8.00(8.75)	35	9.00(8.75)		1.1	-0.34[-0.81,0.13]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		1.5	-1.19[-1.60,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		1.1	-0.53[-1.01,-0.05]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		1.1	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		3.9	-0.24[-0.49,0.02]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		2.8	-0.39[-0.70,-0.08]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		6.6	-0.50[-0.70,-0.31]
Dunlop 1991	96	5.80(5.90)	97	5.40(5.80)		2.3	0.07[-0.28,0.40]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		0.6	-0.68[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.80)		4.2	-0.26[-0.52,-0.03]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		1.4	-0.58[-1.01,-0.16]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.1	-1.61[-3.55,0.34]
Fabre 1992	36	3.08(1.35)	38	9.13(1.32)		0.3	-4.52[-5.37,-3.63]
Feva 1998	19	11.80(8.90)	55	11.10(9.40)		0.9	0.05[-0.47,0.57]
Feva 1998b	19	11.80(8.90)	54	10.80(9.50)		0.9	0.08[-0.44,0.61]
Feighner & Boyer 89	48	8.00(8.49)	52	8.80(8.49)		1.6	-0.33[-0.72,0.07]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		7.6	-0.46[-0.66,-0.30]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		6.7	-0.22[-0.42,-0.03]
Feighner 1989	12	8.00(8.75)	21	14.00(8.75)		0.5	-0.67[-1.40,0.06]
Heiligenstein 1993	26	7.16(10.25)	23	14.61(11.36)		0.8	-0.56[-1.25,-0.11]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		1.0	-0.70[-1.20,-0.20]
Lapierre 1987	20	3.00(5.94)	22	19.00(5.94)		0.4	-2.26[-3.05,-1.47]
Lydiard & Laird 89	17	10.20(8.75)	17	11.70(8.75)		0.6	-0.17[-0.84,0.51]
Lydiard & Stahl 1987	115	6.80(6.97)	119	11.10(6.87)		3.8	-0.33[-0.59,-0.07]
March 1990	12	9.50(5.54)	13	13.00(8.94)		0.4	-0.49[-1.29,0.31]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		1.3	-0.70[-1.14,-0.25]
Reimherr 1986	26	7.60(8.80)	26	11.50(7.60)		0.8	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.86(8.24)		4.5	-0.43[-0.67,-0.20]
Roth 1990	29	6.40(8.40)	21	11.10(9.00)		0.8	-0.29[-0.85,0.26]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		3.2	-0.20[-0.47,0.08]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		1.1	-0.57[-1.05,-0.09]
Silverstone 1998	118	11.10(8.49)	119	15.20(8.49)		3.8	-0.48[-0.74,-0.22]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		1.1	-0.34[-0.83,0.15]
Stark 1985	169	8.20(9.00)	165	11.00(10.10)		5.7	-0.28[-0.50,-0.06]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		4.3	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.80)	92	11.20(8.20)		2.7	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		2.1	-0.17[-0.51,0.18]
Total(95%CI)	3047		3740			100.0	-0.40[-0.45,-0.35]

Test for heterogeneity chi-square=183.36 df=42 p<0.00001
Test for overall effect z=15.77 p<0.00001

-10 -5 0 5 10
SSRI Placebo

Results were robust to the type of analysis used, with a SMD for SSRIs versus placebo, using a random effects model, of -0.47 (95%CI -0.58 to -0.36; Q=183.36, df=42, p<0.00001; z=8.11, p<0.00001), still favouring drug effect (Figure 2).

Figure 2: Overall Data - Standard Mean Difference - Random

Comparison: 01 All Inclusive Data Comparisons
Outcome: 01 OVERALL DATA RANDOM EFFECTS SIZE

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		3.2	-0.48[-0.67,-0.29]
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		2.0	-0.94[-1.47,-0.41]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		2.4	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		3.1	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		2.0	-0.70[-1.23,-0.16]
Claghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		2.0	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		2.2	-0.34[-0.81,0.13]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		2.4	-1.19[-1.60,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		2.2	-0.53[-1.01,-0.05]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		2.2	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		3.0	-0.24[-0.49,0.02]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		2.8	-0.39[-0.70,-0.09]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		3.2	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		2.7	0.07[-0.26,0.40]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		1.7	-0.66[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.60)		3.0	-0.28[-0.52,-0.03]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		2.4	-0.58[-1.01,-0.16]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.3	-1.61[-3.55,0.34]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		1.1	-4.50[-5.37,-3.63]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		2.0	0.05[-0.47,0.57]
Fava 1998b	19	11.60(8.90)	54	10.80(9.50)		2.0	0.08[-0.44,0.61]
Feighner & Boyer 89	48	6.00(8.49)	52	8.80(8.49)		2.5	-0.33[-0.72,0.07]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		3.2	-0.48[-0.66,-0.30]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		3.2	-0.22[-0.42,-0.03]
Feighner 1989	12	8.00(8.75)	21	14.00(8.75)		1.4	-0.67[-1.40,0.06]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		1.9	-0.68[-1.25,-0.11]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		2.1	-0.70[-1.20,-0.20]
Lapierre 1987	20	3.00(6.94)	22	19.00(6.94)		1.3	-2.26[-3.05,-1.47]
Lydiard & Laird 89	17	10.20(8.75)	17	11.70(8.75)		1.6	-0.17[-0.84,0.51]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		3.0	-0.33[-0.59,-0.07]
March 1990	12	9.50(6.94)	13	13.00(6.94)		1.3	-0.49[-1.29,0.31]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		2.3	-0.70[-1.14,-0.25]
Reimherr 1988	26	7.60(8.80)	26	11.50(7.60)		1.9	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		3.0	-0.43[-0.67,-0.20]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		1.9	-0.29[-0.85,0.28]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		2.9	-0.20[-0.47,0.08]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		2.1	-0.57[-1.05,-0.09]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		3.0	-0.48[-0.74,-0.22]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		2.1	-0.34[-0.83,0.15]
Stark 1985	169	8.20(9.00)	165	11.00(10.10)		3.1	-0.28[-0.50,-0.06]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		3.0	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.60)	92	11.20(8.20)		2.8	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		2.7	-0.17[-0.51,0.16]
Total(95%CI)	3047		3740			100.0	-0.47[-0.58,-0.36]

Test for heterogeneity chi-square=183.36 df=42 p<0.00001
Test for overall effect z=8.11 p<0.00001

-10 -5 0 5 10
SSRI Placebo

4.6A Studies with and without placebo run-in period

To determine a potential difference in the effect size of SSRIs in trials with a placebo run-in period, that withdraws placebo responders versus trials, that do not, the data were divided into two groups - PRI (placebo run-in that withdraws placebo responders) (Figure 3) and NPRI (no withdrawal of placebo responders) (Figure 4).

Figure 3: Overall Data PRI - Standard Mean Difference - Random

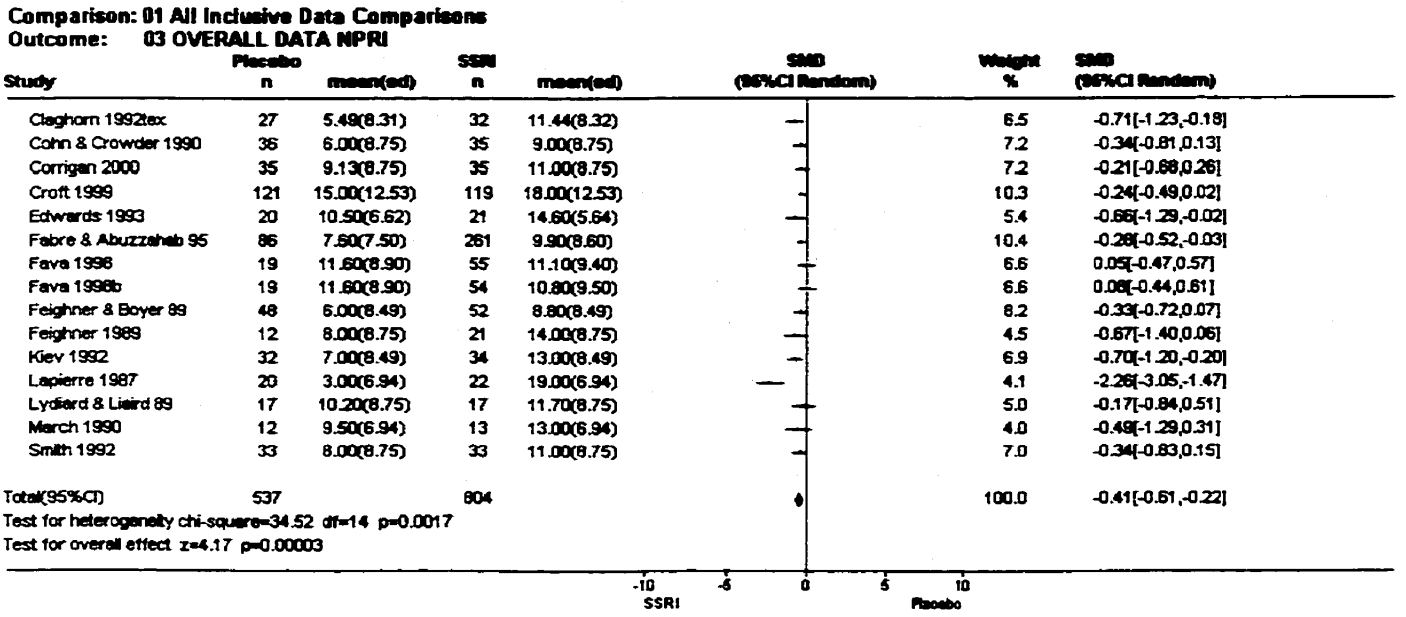
Comparison: 01 All Inclusive Data Comparisons
Outcome: 02 OVERALL DATA PRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		4.5	-0.48[-0.67,-0.29]
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		2.9	-0.94[-1.47,-0.41]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		3.4	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		4.4	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		2.9	-0.70[-1.23,-0.16]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		3.5	-1.19[-1.50,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		3.1	-0.53[-1.01,-0.05]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		4.0	-0.39[-0.70,-0.09]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		4.5	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		3.9	0.07[-0.26,0.40]
Fabre & Birkhimer 95	44	8.50(8.49)	46	13.50(8.49)		3.4	-0.58[-1.01,-0.16]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.5	-1.61[-3.55,0.34]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		1.7	-4.50[-5.37,-3.63]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		4.6	-0.48[-0.66,-0.30]
Feighner & Overo 93	129	9.30(8.49)	521	11.20(8.49)		4.5	-0.22[-0.42,-0.03]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		2.7	-0.68[-1.25,-0.11]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		4.2	-0.33[-0.59,-0.07]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		3.3	-0.70[-1.14,-0.25]
Reimherr 1988	26	7.60(8.80)	26	11.50(7.60)		2.8	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		4.3	-0.43[-0.67,-0.20]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		2.7	-0.29[-0.85,0.28]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		4.1	-0.20[-0.47,0.06]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		3.1	-0.57[-1.05,-0.09]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		4.2	-0.48[-0.74,-0.22]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		4.4	-0.29[-0.50,-0.08]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		4.3	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.60)	92	11.20(8.20)		4.0	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		3.8	-0.17[-0.51,0.18]
Total(95%CI)	2510		2936			100.0	-0.50[-0.64,-0.36]

Test for heterogeneity chi-square=147.82 df=27 p<0.00001
Test for overall effect z=6.93 p<0.00001

-10 -5 0 5 10
SSRI Placebo

Figure 4: Overall Data NPRI - Standard Mean Difference - Random



The SMD for SSRIs versus placebo in the PRI group, using a random effects model, was -0.50 (95%CI -0.64 to -0.36, Q=147.82, df=27, p<0.00001; z=6.93, p<0.00001) (Figure 3). The SMD for SSRIs versus placebo in the NPRI group, using a random effects model, was -0.41 (95%CI -0.61 to -0.22; Q=34.52, df=14, p=0.0017; z=4.17, p=0.00003) (Figure 4). In both the PRI and NPRI groups, SSRIs are shown to be more effective than placebo.

To compare the difference in the two standardized mean differences (effect size), the following formula was used:

$$Z = \frac{ES_1 - ES_2}{SE_1 + SE_2}$$ ES=effect size, SE=standard error

where $SE = \frac{\text{upper limit of CL - mean}}{1.96}$

and the two groups are significantly different if the calculated Z is greater than 1.96 (Hassard, 1991).

For the comparison of the effect sizes of SSRIs and placebo for the all-inclusive data, between the PRI and NPRI groups, using the random effects model:

$$Z = \frac{0.50 - 0.41}{0.071 + 0.097} = 0.536 \text{ (N.S.)}$$

As 0.536 is not >1.96, there was no significant difference in the effect sizes between the PRI and NPRI groups in the all-inclusive data.

4.7 Additional Comparisons

Heterogeneity (within and between study variation) was assessed using the Chi-squared test of heterogeneity (Oxman, 1995). If heterogeneity exists, then an explanation, such as potential clinical differences in the studies, should be sought to explain the heterogeneity (Thompson, 1994). There was substantial heterogeneity in the data as indicated by the large (very significant) Q in all of the above analyses. The primary analyses used random effects model, that takes into account within-study sampling error and between-studies variation in the assessment of uncertainty and provides wider confidence limits to the effect size and hence a more conservative result (Cooper and Hedges, 1994).

The following sensitivity analyses were undertaken as there was statistical heterogeneity:

1. Comparison of PRI and non-PRI studies using the HAMD-D 21 using weighted mean differences
2. Study duration 6 weeks versus study duration > 6 weeks
3. Recent Studies
4. Selected study size (eliminating the very large and very small studies)
5. Jadad score = 3 versus Jadad score > 3

4.7A Comparison of PRI and NPRI Studies using the HAMD-D 21 using weighted mean differences

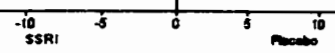
The trials that used HAM-D 21 and HAM-D unspecified were grouped (as it appears that those trials not specifying the HAM-D scale are likely using HAM-D 21) to compare effect sizes. There were 31 trials that used the HAM-D 21 measurement scales. The weighted mean difference (WMD) for SSRIs versus placebo, using the random effects model, was -3.65 (95%CI -4.58 to -2.71; Q=126.24, df=30, p<0.00001, z=7.65, p<0.00001), which indicates that there was a 3.65 difference in HAM-D units favouring SSRIs for efficacy for all the trials that used the HAM-D 21 scale for outcome measurement (Figure 5).

Figure 5: HAM-D 21 - Weighted Mean Difference - Random

Comparison: 08 HAMD & HAMD-21 Comparisons
Outcome: 01 HAM-D 21 COMPARISONS WEIGHTED MEAN DIFFERENCE

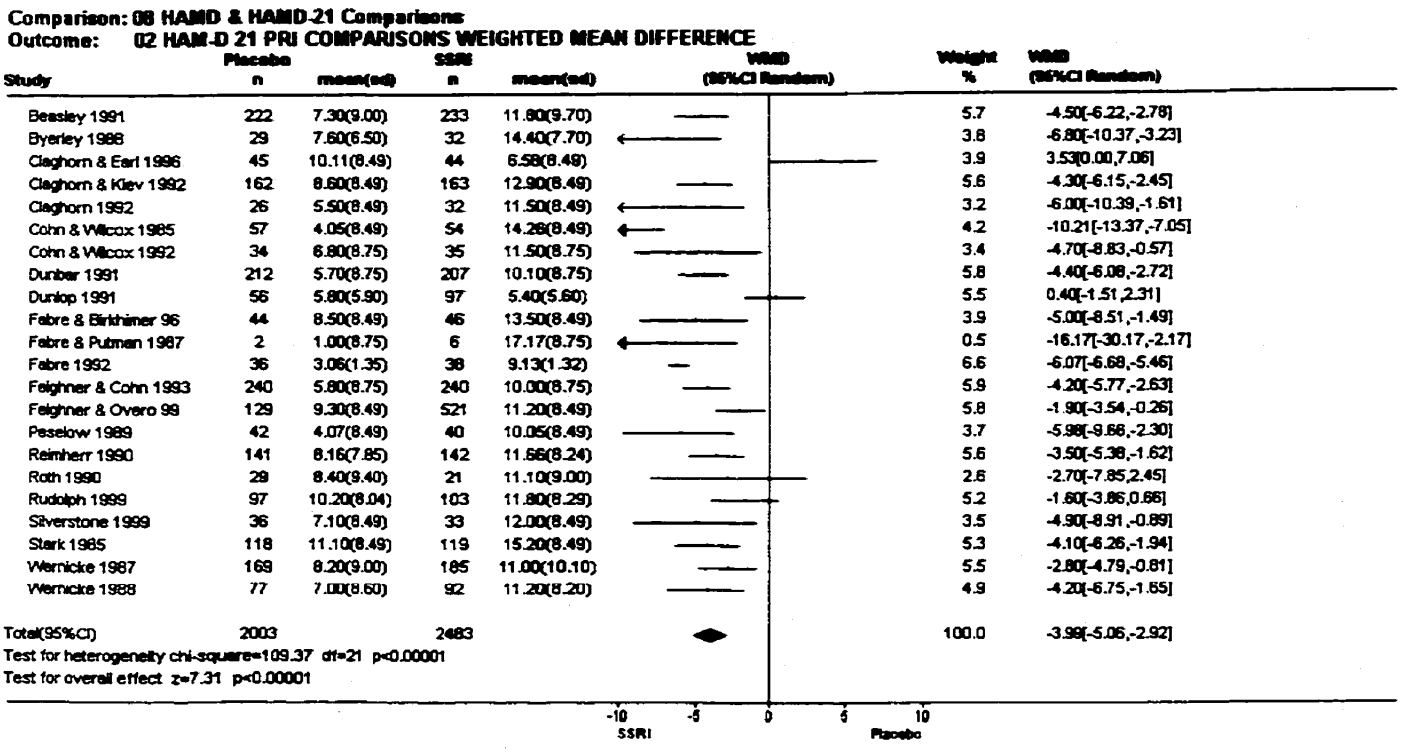
Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		4.5	-4.50[-6.22,-2.78]
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		3.0	-6.80[-10.37,-3.23]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		3.0	3.53[0.00,7.06]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		4.4	-4.30[-6.15,-2.45]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		2.4	-6.00[-10.39,-1.61]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		2.6	-3.00[-7.07,1.07]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		3.3	-10.21[-13.37,-7.05]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		2.6	-4.70[-8.83,-0.57]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		4.5	-4.40[-6.08,-2.72]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		4.3	0.40[-1.51,2.31]
Fabre & Birkhimer 95	44	8.50(8.49)	46	13.50(8.49)		3.0	-5.00[-8.51,-1.49]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.4	-16.17[-30.17,-2.17]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		5.2	-6.07[-6.68,-5.46]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		2.3	0.50[-4.21,5.21]
Fava 1998b	19	11.80(8.90)	54	10.80(9.50)		2.2	0.80[-3.94,5.54]
Feighner & Boyer 89	48	6.00(8.49)	52	8.80(8.49)		3.2	-2.80[-6.13,0.53]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		4.6	-4.20[-5.77,-2.63]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		4.5	-1.90[-3.54,-0.26]
Feighner 1989	12	8.00(8.75)	21	14.00(8.75)		1.6	-6.00[-12.21,0.21]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		2.6	-6.00[-10.10,-1.90]
Lydiard & Laird 89	17	10.20(8.75)	17	11.70(8.75)		1.7	-1.50[-7.38,4.38]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		2.9	-5.98[-9.66,-2.30]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		4.3	-3.50[-5.38,-1.62]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		2.0	-2.70[-7.85,2.45]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		4.0	-1.60[-3.86,0.66]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		2.7	-4.90[-8.91,-0.89]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		4.1	-4.10[-6.26,-1.94]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		2.5	-3.00[-7.22,1.22]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		4.3	-2.80[-4.79,-0.81]
Wernicke 1987	77	7.00(8.80)	92	11.20(8.20)		3.8	-4.20[-6.75,-1.65]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		3.4	-1.50[-4.50,1.50]
Total(95%CI)	2267		2887			100.0	-3.65[-4.58,-2.71]

Test for heterogeneity chi-square=126.24 df=30 p<0.00001
Test for overall effect z=7.85 p<0.00001



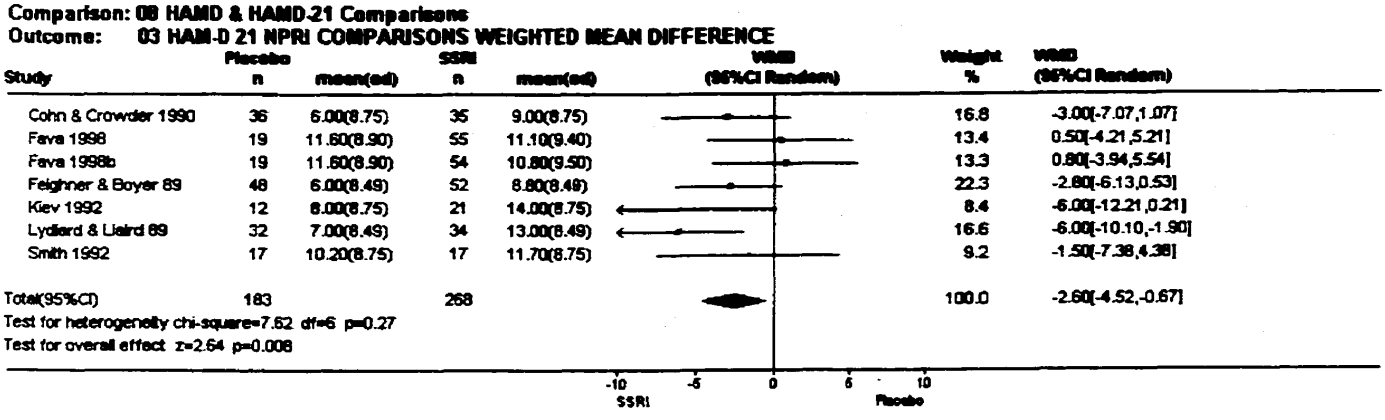
The WMD for SSRIs versus placebo in the PRI group was –3.99 (CI-5.06 to –2.92; Q=109.37, df=21, p<0.00001; z=7.31, p<0.00001) (Figure 6).

Figure 6: HAM-D 21 PRI - Weighted Mean Difference - Random



The WMD for SSRIs versus placebo in the NPRI group was –2.60 (CI-4.52 to –0.67; Q=7.62, df=6, p=0.27; Z=2.64, p=0.008)(Figure 7). It is noted that the PRI/HAM-D 21 group is heterogeneous, but the NPRI/HAM-D 21 group is homogenous, as indicated by the small, non-significant Q value for the NPRI group.

Figure 7: HAM-D 21 NPRI - Weighted Mean Difference - Random



To compare the PRI and NPRI groups within the HAM-D 21 group:

$$Z = \frac{3.99 - 2.60}{0.546 + 0.985} = 0.908 \text{ (NS)}$$

There was no significant difference in the effect sizes between the PRI and the NPRI studies using the HAM-D 21.

4.7B Study duration 6 weeks versus study duration > 6 weeks

Data were separated into trials of duration of 6 weeks and trials of greater than 6 weeks. There were 31 trials of duration of 6 weeks and there were 12 trials of duration of greater than 6 weeks. The SMD for SSRIs versus placebo, using a random effects model, in the trials of 6 weeks' duration was -0.54 (95%CI -0.69 to -0.39; Q=169.69, df=30, p<0.00001; z=7.07, p<0.00001) (Figure 8).

Figure 8: 6 Weeks - Standard Mean Difference - Random

Comparison: 02 6 Weeks or Less
Outcome: 01 6 WEEKS

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		4.2	-0.48[-0.67,-0.29]
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		2.9	-0.94[-1.47,-0.41]
Cleghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		3.3	0.41[-0.01,0.83]
Cleghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		4.1	-0.51[-0.73,-0.28]
Cleghorn 1992	26	5.50(8.49)	32	11.50(8.49)		2.9	-0.70[-1.23,-0.16]
Cleghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		2.9	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		3.1	-0.34[-0.81,0.13]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		3.4	-1.19[-1.60,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		3.1	-0.53[-1.01,-0.05]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		3.8	-0.39[-0.70,-0.09]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		4.2	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		3.7	0.07[-0.26,0.40]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		2.5	-0.66[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.60)		4.0	-0.28[-0.52,-0.03]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.5	-1.61[-3.55,0.34]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		1.8	-4.50[-5.37,-3.63]
Feighner & Boyer 89	48	6.00(8.49)	52	8.80(8.49)		3.4	-0.33[-0.72,0.07]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		4.2	-0.48[-0.66,-0.30]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		4.2	-0.22[-0.42,-0.03]
Feighner 1999	12	8.00(8.75)	21	14.00(8.75)		2.2	-0.67[-1.40,0.06]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		3.0	-0.70[-1.20,-0.20]
Lapierre 1987	20	3.00(6.94)	22	19.00(6.94)		2.0	-2.26[-3.05,-1.47]
Lydiard & Laird 89	17	10.20(8.75)	17	11.70(8.75)		2.4	-0.17[-0.84,0.51]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		3.2	-0.70[-1.14,-0.25]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		2.8	-0.29[-0.85,0.26]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		3.1	-0.57[-1.05,-0.09]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		3.1	-0.34[-0.83,0.15]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		4.2	-0.29[-0.50,-0.08]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		4.1	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.60)	92	11.20(8.20)		3.8	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		3.7	-0.17[-0.51,0.18]
Total(95%CI)	2272		2886			100.0	-0.54[-0.69,-0.39]

Test for heterogeneity chi-square=169.69 df=30 p<0.00001
Test for overall effect z=7.07 p<0.00001

-10 -5 0 5 10
SSRI Placebo

The SMD for SSRIs versus placebo, using a random effects model, for studies of duration more than 6 weeks was -0.34 (95%CI -0.44 to -0.23; Q=11.41, df=11, p=0.41; z=6.48, p<0.00001) (Figure 9).

Figure 9: More Than 6 Weeks - Standard Mean Difference - Random

Comparison: 03 More than 6 Weeks
Outcome: 01 MORE THAN 6 WEEKS

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		4.6	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		15.0	-0.24[-0.49,0.02]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		5.7	-0.58[-1.01,-0.16]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		3.8	0.05[-0.47,0.57]
Fava 1998b	19	11.60(8.90)	54	10.80(9.50)		3.7	0.08[-0.44,0.61]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		3.2	-0.68[-1.25,-0.11]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		14.6	-0.33[-0.59,-0.07]
March 1990	12	9.50(6.94)	13	13.00(6.94)		1.6	-0.49[-1.29,0.31]
Reimherr 1988	26	7.60(8.80)	28	11.50(7.60)		3.4	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		17.3	-0.43[-0.67,-0.20]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		12.7	-0.20[-0.47,0.08]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		14.6	-0.48[-0.74,-0.22]
Total(95%CI)	775		854			100.0	-0.34[-0.44,-0.23]

Test for heterogeneity chi-square=11.41 df=11 p=0.41
Test for overall effect z=6.48 p<0.00001

-10 -5 0 5 10
SSRI Placebo

Comparing these two groups (6 weeks and more than 6 weeks):

$$Z = \frac{0.40 - 0.34}{0.031 + 0.051} = 1.098 \text{ (NS)}$$

There was no significant difference in effect size for SSRIs between the 2 groups separated by duration of study.

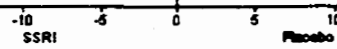
On separation of the data for PRI versus NPRI groups in studies of duration 6 weeks, using the random effects model, the SMD for SSRIs versus placebo in the PRI group, was -0.53 (95%CI -0.72 to -0.35; Q=143.24, df=20, p<0.00001; z=5.70, p<0.00001) (Figure 10).

Figure 10: 6 Weeks PRI - Standard Mean Difference - Random

Comparison: 02 6 Weeks or Less
Outcome: 02 6 WEEKS PRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		5.9	-0.46[-0.67,-0.29]
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		4.1	-0.94[-1.47,-0.41]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		4.7	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		5.7	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		4.1	-0.70[-1.23,-0.16]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		4.8	-1.19[-1.60,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		4.4	-0.53[-1.01,-0.05]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		5.4	-0.39[-0.70,-0.09]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		5.8	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		5.2	0.07[-0.26,0.40]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.8	-1.61[-3.55,0.34]
Fabre 1992	36	3.08(1.35)	38	9.13(1.32)		2.6	-4.50[-5.37,-3.63]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		5.9	-0.46[-0.66,-0.30]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		5.9	-0.22[-0.42,-0.03]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		4.6	-0.70[-1.14,-0.25]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		3.9	-0.29[-0.85,0.28]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		4.4	-0.57[-1.05,-0.09]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		5.8	-0.29[-0.50,-0.08]
Welczak 1996	200	6.00(5.31)	99	7.00(5.31)		5.6	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.60)	92	11.20(8.20)		5.3	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		5.1	-0.17[-0.51,0.18]
Total(95%CI)	1941		2358			100.0	-0.53[-0.72,-0.35]

Test for heterogeneity chi-square=143.24 df=20 p<0.00001
Test for overall effect z=5.70 p<0.00001



The SMD for SSRIs versus placebo in the NPRI group, using the random effects model, was -0.57 (95%CI -0.84 to -0.30; Q=26.20, df=9, p=0.0019; z=4.17, p=0.00003) (Figure 11).

Figure 11: 6 Weeks NPRI - Standard Mean Difference - Random

Comparison: 02 6 Weeks or Less
Outcome: 03 6 WEEKS NPRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Claghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		10.1	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		11.0	-0.34[-0.81,0.13]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		8.7	-0.68[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.80(7.50)	261	9.90(8.60)		14.5	-0.26[-0.52,-0.03]
Feighner & Boyer 89	49	6.00(8.49)	52	8.80(8.49)		12.2	-0.33[-0.72,0.07]
Feighner 1989	12	8.00(8.75)	21	14.00(8.75)		7.4	-0.67[-1.40,0.06]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		10.5	-0.70[-1.20,-0.20]
Lapierre 1987	20	3.00(6.94)	22	19.00(6.94)		6.8	-2.35[-3.05,-1.47]
Lydiard & Laird 89	17	10.20(8.75)	17	11.70(8.75)		8.1	-0.17[-0.84,0.51]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		10.7	-0.34[-0.83,0.15]
Total(95%CI)	331		528			100.0	-0.57[-0.84,-0.30]

Test for heterogeneity chi-square=26.20 df=9 p=0.0019
Test for overall effect z=4.17 p=0.00003

-10 -5 0 5 10
SSRI Placebo

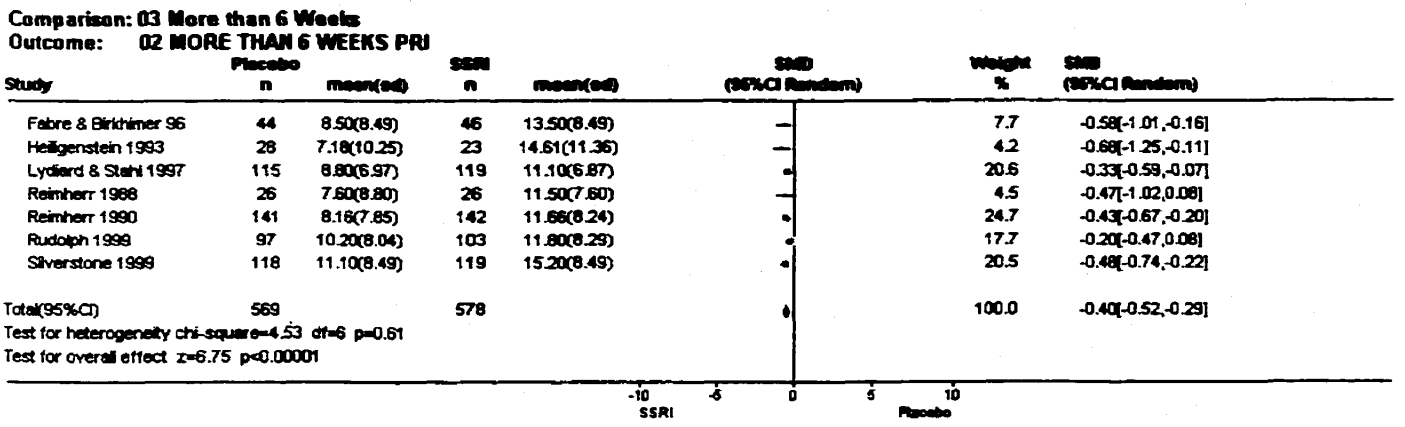
On comparing the two groups, the PRI and NPRI groups (in studies of duration 6 weeks or less):

$$Z = \frac{0.57 - 0.53}{0.092 + 0.138} = 0.174 \text{ (NS)}$$

There was no significant difference between the effect sizes for PRI and NPRI groups in studies of duration 6 weeks.

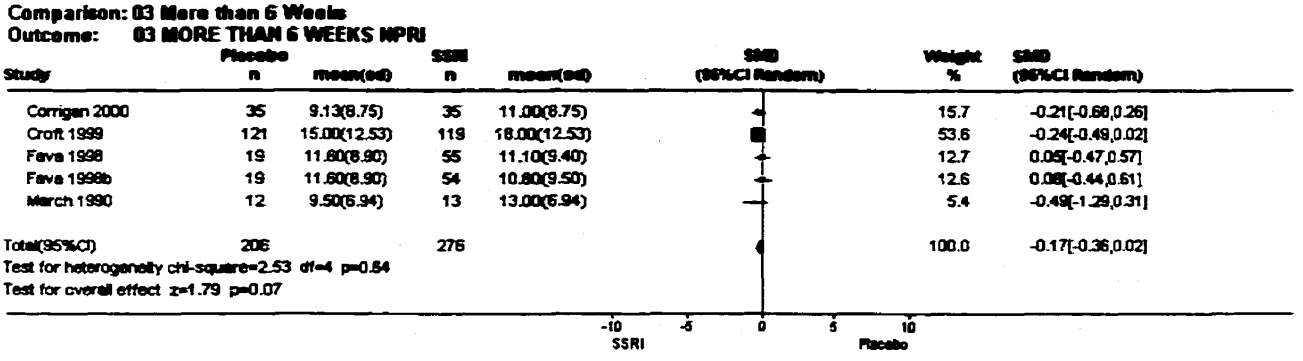
It is noted that the subgroup of trials with duration of more than 6 weeks is homogeneous. Using the random effects model, the SMD for PRI/more than 6 weeks group is -0.40 (95%CI -0.52 to -0.29; Q=4.53, df=6, p=0.61; z=6.75, p<0.00001) (Figure 12).

Figure 12: More Than 6 Weeks PRI - Standard Mean Difference - Random



The SMD for NPRI/more than 6 weeks group is -0.17 (95%CI -0.36 to -0.02; Q=2.53, df=4, p=0.64; z=1.79, p=0.07) (Figure 13).

Figure 13: More Than 6 Weeks NPRI - Standard Mean Difference - Random



On comparing PRI and NPRI groups in studies duration of more than 6 weeks:

$$Z = \frac{0.40 - 0.17}{0.056 + 0.077} = 1.73 \text{ (NS)}$$

There was no significant difference in effect sizes for PRI versus NPRI groups for studies of duration more than 6 weeks.

4.7C Recent Studies

There were 31 trials published from 1990 to 2000. Using the random effects model, the SMD for SSRIs versus placebo was –0.41 (95%CI –0.54 to –0.29; Q=135.25, df=30, p<0.00001; z=6.41, p<0.00001) (Figure 14).

Figure 14: Recent Studies - Standard Mean Difference - Random

Comparison: 10 Studies between 1990 & 2000
Outcome: 02 RECENT STUDIES

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		4.3	-0.46[-0.67,-0.29]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		3.1	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		4.1	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		2.6	-0.70[-1.23,-0.16]
Claghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		2.6	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		2.9	-0.34[-0.81,0.13]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		2.8	-0.53[-1.01,-0.05]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		2.9	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		4.0	-0.24[-0.49,0.02]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		3.7	-0.39[-0.70,-0.09]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		4.2	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		3.6	0.07[-0.26,0.40]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		2.2	-0.66[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.60)		4.0	-0.26[-0.52,-0.03]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		3.1	-0.58[-1.01,-0.16]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		1.5	-4.50[-5.37,-3.63]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		2.6	0.05[-0.47,0.57]
Fava 1998b	19	11.60(8.90)	54	10.80(9.50)		2.6	0.08[-0.44,0.61]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		4.3	-0.46[-0.66,-0.30]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		4.3	-0.22[-0.42,-0.03]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		2.4	-0.68[-1.25,-0.11]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		2.7	-0.70[-1.20,-0.20]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		3.9	-0.33[-0.59,-0.07]
March 1990	12	9.50(6.94)	13	13.00(6.94)		1.6	-0.49[-1.29,0.31]
Reimherr 1990	141	8.18(7.85)	142	11.68(8.24)		4.1	-0.43[-0.67,-0.20]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		2.4	-0.29[-0.85,0.28]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		3.8	-0.20[-0.47,0.08]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		2.8	-0.57[-1.05,-0.09]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		3.9	-0.48[-0.74,-0.22]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		2.8	-0.34[-0.83,0.15]
Walczak 1996	200	6.00(5.31)	98	7.00(5.31)		4.0	-0.19[-0.43,0.05]
Total(95%CI)	2500		3090			100.0	-0.41[-0.54,-0.29]

Test for heterogeneity chi-square=135.25 df=30 p<0.00001
Test for overall effect z=6.41 p<0.00001

-10 -5 0 5 10
SSRI Placebo

The SMD for the PRI group for trials from 1990 to 2000 was –0.46 (95%CI –0.63 to –0.30; Q=122.95, df=19, p<0.00001; z=5.44, p<0.00001) (Figure 15).

Figure 15: Recent Studies PRI - Standard Mean Difference - Random

Comparison: 10 Studies between 1990 & 2000
Outcome: 01 RECENT STUDIES PRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		6.0	-0.48[-0.67,-0.29]
Claghorn & Earl 1995	45	10.11(8.49)	44	6.58(8.49)		4.6	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		5.8	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		3.9	-0.70[-1.23,-0.16]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		4.2	-0.53[-1.01,-0.05]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		5.4	-0.39[-0.70,-0.09]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		6.0	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		5.2	0.07[-0.26,0.40]
Fabre & Birkhimer 95	44	8.50(8.49)	46	13.50(8.49)		4.6	-0.58[-1.01,-0.16]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		2.3	-4.50[-5.37,-3.63]
Feighner & Cohn 1993	240	5.80(8.75)	240	10.00(8.75)		6.0	-0.48[-0.66,-0.30]
Feighner & Overo 99	129	9.30(8.49)	521	11.20(8.49)		6.0	-0.22[-0.42,-0.03]
Heiligenstein 1993	26	7.18(10.25)	23	14.61(11.36)		3.7	-0.68[-1.25,-0.11]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		5.6	-0.33[-0.59,-0.07]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		5.8	-0.43[-0.67,-0.20]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		3.7	-0.29[-0.85,0.28]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		5.5	-0.20[-0.47,0.08]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		4.2	-0.57[-1.05,-0.09]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		5.6	-0.48[-0.74,-0.22]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		5.7	-0.19[-0.43,0.05]
Total(95%CI)	2060		2398			100.0	-0.46[-0.63,-0.30]

Test for heterogeneity chi-square=122.95 df=19 p<0.00001
Test for overall effect z=5.44 p<0.00001

-10 -5 0 5 10
SSRI Placebo

The SMD for the NPRI group for trials from 1990 to 2000 was –0.30 (95%CI –0.43 to –0.17; Q=10.45, df=10, p=0.4; z=4.61, p<0.00001) (Figure 16).

Figure 16: Recent Studies NPRI - Standard Mean Difference - Random

Comparison: 10 Studies between 1990 & 2000
Outcome: 03 RECENT STUDIES NPRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Claghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		5.8	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		7.3	-0.34[-0.81,0.13]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		7.3	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		23.0	-0.24[-0.49,0.02]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		4.1	-0.66[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.60)		24.6	-0.28[-0.52,-0.03]
Fava 1998	19	11.50(8.90)	55	11.10(9.40)		6.0	0.05[-0.47,0.57]
Fava 1998b	19	11.60(8.90)	54	10.80(9.50)		5.9	0.08[-0.44,0.61]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		6.5	-0.70[-1.20,-0.20]
March 1990	12	9.50(6.94)	13	13.00(6.94)		2.6	-0.49[-1.29,0.31]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		6.8	-0.34[-0.83,0.15]
Total(95%CI)	440		692			100.0	-0.30[-0.43,-0.17]

Test for heterogeneity chi-square=10.45 df=10 p=0.4
Test for overall effect z=4.61 p<0.00001

-10 -5 0 5 10
SSRI Placebo

On comparing the PRI and NPRI groups for trials from 1990 to 2000:

$$Z = \frac{0.46 - 0.30}{0.168 + 0.066} = 0.68 \text{ (N.S.)}$$

There was no significant difference in effect sizes between the PRI and NPRI groups in the subgroup of more recent studies.

4.7D Selected study size

The four largest trials (n=419-650) and the four smallest trials (n=8-35) were excluded from the analysis (Beasley 1991, Dunbar 1991, Fabre & Putman 1987, Fava 1998b, Feighner & Cohn 1993, Feighner & Overo 1999, Lydiard & Liard 1989 and March 1990). The SMD for SSRIs versus placebo, using the random effects model, was –0.51 (95%CI –0.65 to –0.36; Q=171.87, df=33, p<0.00001; z=6.87, p<0.00001) (Figure 17). SSRIs are significantly more effective than placebo.

Figure 17: Selected Study Size - Standard Mean Difference - Random

Comparison: 09 Select Study Size Comparisons
Outcome: 01 SELECTED STUDY SIZE

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		2.6	-0.94[-1.47,-0.41]
Cleghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		3.0	0.41[-0.01,0.83]
Cleghorn & Kiev 1992	162	8.80(8.49)	163	12.90(8.49)		3.6	-0.51[-0.73,-0.28]
Cleghorn 1992	26	5.50(8.49)	32	11.50(8.49)		2.6	-0.70[-1.23,-0.16]
Cleghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		2.6	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		2.8	-0.34[-0.81,0.13]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.26(8.49)		3.0	-1.19[-1.60,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		2.7	-0.53[-1.01,-0.05]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		2.8	-0.21[-0.68,0.25]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		3.5	-0.24[-0.49,0.02]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		3.4	-0.39[-0.70,-0.09]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		3.3	0.07[-0.26,0.40]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		2.3	-0.66[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.60)		3.6	-0.28[-0.52,-0.03]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		3.0	-0.58[-1.01,-0.16]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		1.6	-4.50[-5.37,-3.63]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		2.6	0.05[-0.47,0.57]
Feighner & Boyer 89	48	6.00(8.49)	52	8.80(8.49)		3.1	-0.33[-0.72,0.07]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		2.5	-0.68[-1.25,-0.11]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		2.7	-0.70[-1.20,-0.20]
Lapierre 1987	20	3.00(6.94)	22	19.00(6.94)		1.8	-2.26[-3.05,-1.47]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		3.5	-0.33[-0.59,-0.07]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		2.9	-0.70[-1.14,-0.25]
Reimherr 1988	26	7.60(8.80)	26	11.50(7.60)		2.5	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		3.6	-0.43[-0.67,-0.20]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		2.5	-0.29[-0.85,0.26]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		3.5	-0.20[-0.47,0.08]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		2.7	-0.57[-1.05,-0.09]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		3.5	-0.48[-0.74,-0.22]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		2.7	-0.34[-0.83,0.15]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		3.7	-0.29[-0.50,-0.08]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		3.6	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.60)	92	11.20(8.20)		3.4	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		3.2	-0.17[-0.51,0.18]
Total(95%CI)	2182		2428			100.0	-0.51[-0.65,-0.36]

Test for heterogeneity chi-square=171.87 df=33 p<0.00001
Test for overall effect z=6.87 p<0.00001

-10 -5 0 5 10
SSRI Placebo

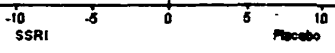
The SMD for the PRI group for selected study size was -0.53 (95%CI -0.71 to -0.34; Q=140.86, df=22, p<0.00001; z=5.63, p<0.00001) (Figure 18).

Figure 18: Selected Study Size PRI - Standard Mean Difference - Random

Comparison: 09 Select Study Size Comparisons
Outcome: 02 SELECTED STUDY SIZE PRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Byerley 1988	29	7.60(6.50)	32	14.40(7.70)		3.8	-0.94[-1.47,-0.41]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		4.3	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		5.1	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		3.8	-0.70[-1.23,-0.16]
Cohn & Wilcox 1985	57	4.05(8.49)	54	14.28(8.49)		4.3	-1.19[-1.60,-0.79]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		4.0	-0.53[-1.01,-0.05]
Doogan 1994	90	12.10(8.34)	83	15.40(8.34)		4.8	-0.39[-0.70,-0.09]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		4.7	0.07[-0.26,0.40]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		4.3	-0.58[-1.01,-0.16]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		2.4	-4.50[-5.37,-3.63]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		3.6	-0.68[-1.25,-0.11]
Lydiard & Stahl 1997	115	8.80(6.97)	119	11.10(6.87)		5.0	-0.33[-0.59,-0.07]
Peselow 1989	42	4.07(8.49)	40	10.05(8.49)		4.2	-0.70[-1.14,-0.25]
Reimherr 1988	26	7.60(8.80)	26	11.50(7.60)		3.7	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		5.1	-0.43[-0.67,-0.20]
Roth 1990	29	8.40(9.40)	21	11.10(9.00)		3.6	-0.29[-0.85,0.28]
Rudolph 1999	97	10.20(8.04)	103	11.80(8.29)		4.9	-0.20[-0.47,0.08]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		4.0	-0.57[-1.05,-0.09]
Silverstone 1999	118	11.10(8.49)	119	15.20(8.49)		5.0	-0.48[-0.74,-0.22]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		5.1	-0.29[-0.50,-0.08]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		5.0	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.80)	92	11.20(8.20)		4.8	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		4.6	-0.17[-0.51,0.18]
Total(95%CI)	1705		1729			100.0	-0.53[-0.71,-0.34]

Test for heterogeneity chi-square=140.86 df=22 p<0.00001
Test for overall effect z=5.63 p<0.00001



The SMD for the NPRI group for selected study size was -0.46 (95%CI-0.69 to -0.23; Q=30.66, df=10, p=0.0007; z=3.92, p=0.00009) (Figure 19).

Figure 19: Selected Study Size NPRI - Standard Mean Difference - Random

Comparison: 09 Select Study Size Comparisons
Outcome: 03 SELECTED STUDY SIZE NPRI

Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (95%CI Random)	Weight %	SMD (95%CI Random)
Claghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		8.3	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		9.1	-0.34[-0.81,0.13]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		9.1	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		12.4	-0.24[-0.49,0.02]
Edwards 1993	20	10.50(6.62)	21	14.60(5.64)		7.0	-0.86[-1.29,-0.02]
Fabre & Abuzzahab 95	86	7.60(7.50)	261	9.90(8.60)		12.6	-0.28[-0.52,-0.03]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		8.4	0.05[-0.47,0.57]
Feighner & Boyer 89	48	6.00(8.49)	52	8.80(8.49)		10.2	-0.33[-0.72,0.07]
Kiev 1992	32	7.00(8.49)	34	13.00(8.49)		8.7	-0.70[-1.20,-0.20]
Lapierre 1987	20	3.00(6.94)	22	19.00(6.94)		5.4	-2.26[-3.05,-1.47]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		8.9	-0.34[-0.83,0.15]
Total(95%CI)	477		699			100.0	-0.46[-0.69,-0.23]

Test for heterogeneity chi-square=30.66 df=10 p=0.0007
Test for overall effect z=3.92 p=0.00009

-10 -5 0 5 10
SSRI Placebo

On comparing the PRI and NPRI groups for selected study size:

$$Z = \frac{0.53 - 0.46}{0.097 + 0.117} = 0.327 \text{ (N.S.)}$$

There was no significant difference in effect sizes between the PRI and NPRI groups in the subgroup of select study size.

4.7E Jadad score = 3/>3

Trials were also separated based on their Jadad scores. As all the trials are randomized controlled placebo clinical trials, they all have a minimum Jadad score of 3. Twenty-eight trials had Jadad score of 3 and 15 trials had Jadad scores of greater than 3. Using the random effects model, the SMD for SSRIs versus placebo in trials with a Jadad score of 3 was –0.43 (95%CI –0.58 to –0.27; Q=131.39, df=27, p<0.00001; z=5.45, p<0.00001) (Figure 20).

Figure 20: JADAD Score = 3 - Standard Mean Difference - Random

Comparison: 11 JADAD SCORES 3
Outcome: 01 JADAD SCORES = 3

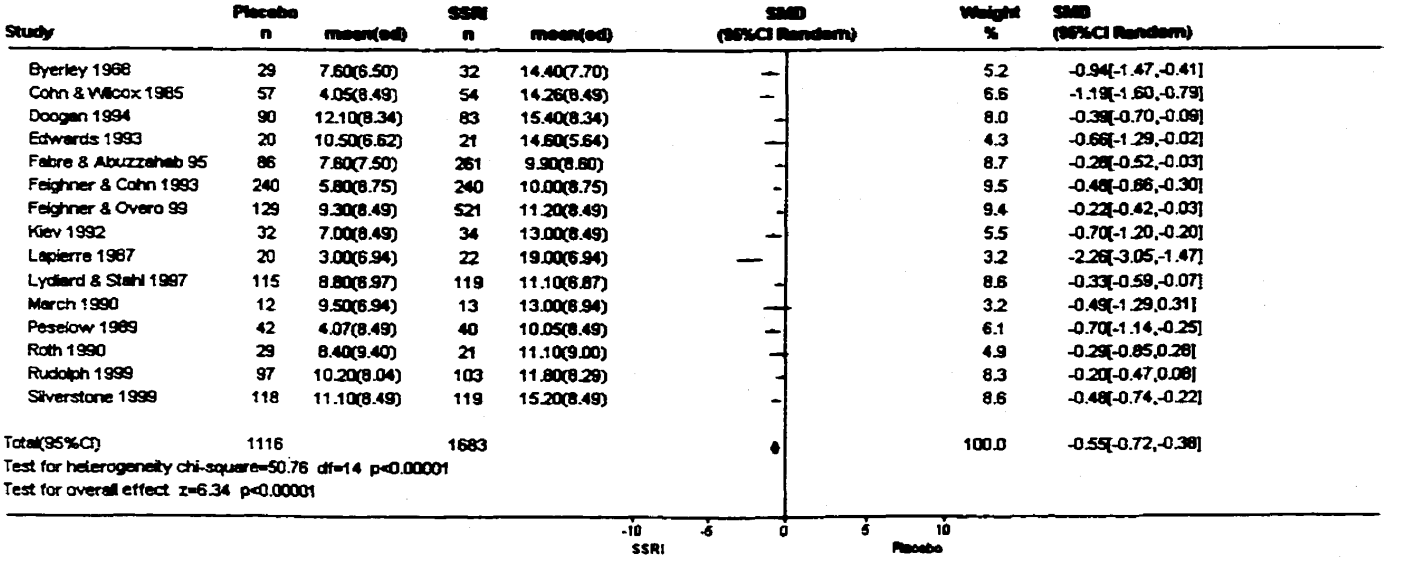
Study	Placebo n	Placebo mean(sd)	SSRI n	SSRI mean(sd)	SMD (96%CI Random)	Weight %	SMD (96%CI Random)
Beasley 1991	222	7.30(9.00)	233	11.80(9.70)		4.8	-0.48[-0.67,-0.29]
Claghorn & Earl 1996	45	10.11(8.49)	44	6.58(8.49)		3.7	0.41[-0.01,0.83]
Claghorn & Kiev 1992	162	8.60(8.49)	163	12.90(8.49)		4.7	-0.51[-0.73,-0.28]
Claghorn 1992	26	5.50(8.49)	32	11.50(8.49)		3.2	-0.70[-1.23,-0.16]
Claghorn 1992tex	27	5.49(8.31)	32	11.44(8.32)		3.2	-0.71[-1.23,-0.18]
Cohn & Crowder 1990	36	6.00(8.75)	35	9.00(8.75)		3.5	-0.34[-0.81,0.13]
Cohn & Wilcox 1992	34	6.80(8.75)	35	11.50(8.75)		3.4	-0.53[-1.01,-0.05]
Corrigan 2000	35	9.13(8.75)	35	11.00(8.75)		3.5	-0.21[-0.68,0.26]
Croft 1999	121	15.00(12.53)	119	18.00(12.53)		4.5	-0.24[-0.49,0.02]
Dunbar 1991	212	5.70(8.75)	207	10.10(8.75)		4.8	-0.50[-0.70,-0.31]
Dunlop 1991	56	5.80(5.90)	97	5.40(5.60)		4.2	0.07[-0.26,0.40]
Fabre & Birkhimer 96	44	8.50(8.49)	46	13.50(8.49)		3.7	-0.58[-1.01,-0.16]
Fabre & Putman 1987	2	1.00(8.75)	6	17.17(8.75)		0.6	-1.61[-3.55,0.34]
Fabre 1992	36	3.06(1.35)	38	9.13(1.32)		1.9	-4.50[-5.37,-3.63]
Fava 1998	19	11.60(8.90)	55	11.10(9.40)		3.2	0.05[-0.47,0.57]
Fava 1998b	19	11.60(8.90)	54	10.80(9.50)		3.2	0.08[-0.44,0.61]
Feighner & Boyer 89	48	6.00(8.49)	52	8.80(8.49)		3.8	-0.33[-0.72,0.07]
Feighner 1999	12	8.00(8.75)	21	14.00(8.75)		2.4	-0.67[-1.40,0.06]
Heiligenstein 1993	28	7.18(10.25)	23	14.61(11.36)		3.0	-0.68[-1.25,-0.11]
Lydiard & Laird 89	17	10.20(8.75)	17	11.70(8.75)		2.6	-0.17[-0.84,0.51]
Reimherr 1988	26	7.60(8.80)	26	11.50(7.60)		3.1	-0.47[-1.02,0.08]
Reimherr 1990	141	8.16(7.85)	142	11.66(8.24)		4.6	-0.43[-0.67,-0.20]
Shrivastava 1992	36	7.10(8.49)	33	12.00(8.49)		3.4	-0.57[-1.05,-0.09]
Smith 1992	33	8.00(8.75)	33	11.00(8.75)		3.4	-0.34[-0.83,0.15]
Stark 1985	169	8.20(9.00)	185	11.00(10.10)		4.7	-0.29[-0.50,-0.08]
Walczak 1996	200	6.00(5.31)	99	7.00(5.31)		4.6	-0.19[-0.43,0.05]
Wernicke 1987	77	7.00(8.60)	92	11.20(8.20)		4.3	-0.50[-0.81,-0.19]
Wernicke 1988	48	5.70(8.60)	103	7.20(9.10)		4.1	-0.17[-0.51,0.18]
Total(95%CI)	1931		2057			100.0	-0.43[-0.58,-0.27]

Test for heterogeneity chi-square=131.39 df=27 p<0.00001
Test for overall effect z=5.45 p<0.00001

-10 -5 0 5 10
SSRI Placebo

The SMD for SSRIs versus placebo in trials with a Jadad score of greater than 3 was –0.55 (95%CI –0.72 to –0.38; Q=50.76, df 14, p<0.00001; z=6.34, p<0.00001) (Figure 21).

Figure 21: JADAD Score > 3 - Standard Mean Difference - Random



On comparing the trials with a Jadad score of 3 with those having a Jadad score of greater than 3:

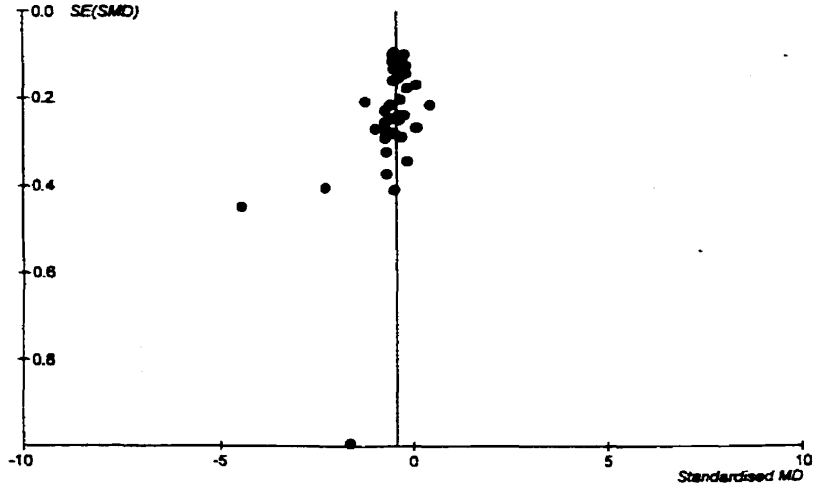
$$Z = \frac{0.55 - 0.43}{0.087 + 0.082} = 0.710 \text{ (N.S.)}$$

There was no significant difference between the effect sizes for the groups with a Jadad score of 3 and those with a Jadad score of greater than 3.

4.8 Funnel Plot

On examination of the funnel plot for all the trials of the standardized mean difference versus the standard error of the standardized mean difference (Figure 22), the presence of the symmetrical, inverted funnel-shaped plot suggested that there is no evidence of publication bias in the meta-analysis trials, except for absence of small trials that show placebo effect.

Figure 22: Funnel plot of SMD vs standard error of SMD



There is a clustering of large studies with a relatively small standard error and a wide scattering of small studies with a larger standard error. There are no trials in the bottom right hand corner of the plot, indicating absence of small trials, that did not show any experimental drug effect, and, as a result, were likely not published. As many trials are funded by the pharmaceutical industry, any trials indicating lack of drug effect will likely not be published. It may also be possible that some trials were published in a language other than English; therefore, was not included in the meta-analysis.

CHAPTER 5 DISCUSSION

The meta-analysis indicates that there is no impact of the placebo run-in phase and its subsequent withdrawal of placebo responders in clinical trials of SSRI antidepressants for depression. The hypothesis of a larger drug effect size with trials that eliminate placebo responders prior to randomization is not supported by the data. There is no significant difference in the effect sizes between the trials with a placebo run-in phase (and subsequent elimination of placebo responders) and those without such a procedure.

Trivedi and Rush proposed that the reason a placebo run-in phase failed to affect placebo response rates or drug-placebo differences post-randomization was because there is often an equal delay in time in studies that do not have a placebo run-in phase, as patients will have several visits before they are entered into the study. This delay, with or without a placebo, may induce an improvement in symptoms, either due to the passage of time, or interaction with the investigator. They also suggested that measurement may present a problem in some patients, in that the reliability of the (17-item) HAM-D score is in the order of ± 2, so that a clinically meaningless "drop" can lead to subject exclusion (Trivedi and Rush,1994).

From a scientific standpoint, there is no reason to utilize the placebo run-in phase to eliminate placebo responders as it is costly in terms of time and effort. Some authors (Senn, 1997; Skovlund, 1994) have raised ethical concerns about the use of procedures involving dropping early responders. Research participants are not informed of this procedure; therefore, there is incomplete disclosure of the nature of the study. Informed consent implies that prospective subjects are advised of all procedures of the study. It is

likely that information about the placebo run-in period is often either withheld or presented in a somewhat misleading manner to the subject.

There were 42 trials included in the meta-analysis. It is easier to detect genuine differences if there is a larger body of evidence available for analysis (Hassard, 1991). If the number of trials being analyzed were to increase (as may be available in the future), statistical power will also increase to the point of detecting significant smaller differences between the trials with and without placebo run-in periods.

In the real world, many clinical trials will be funded by the pharmaceutical industry in which larger effect sizes for the investigational drug will be desirable. The trials with a placebo run-in phase and elimination of placebo responders prior to randomization trended to result in a larger (albeit non-significant) drug effect size. The meta-analysis indicated that selective serotonin reuptake inhibitors are more effective than placebo in the treatment of depression, but it appears that this effect is more substantial in the trials that eliminated placebo responders. This larger effect size will be more impressive in the post-research marketing process.

It is likely that the practice of using placebo run-in periods and placebo responder elimination will continue. If an investigator uses this practice in the research, it is recommended that all details relating to this practice be explicitly stated in the publication of the trial. In the review, only three trials that eliminated placebo responders stated the number of placebo responders withdrawn. Of those three trials, only two of them explicitly stated the number of subjects that entered into the placebo run-in phase. It is pertinent that researchers state, in the publication, the number of subjects that entered the placebo run-in phase, if one exists, and the number and reasons subjects were eliminated

prior to randomization. It is also recommended that study means and standard deviations be included in the publication.

There are no specific requirements for reporting of research and studies in medical journals. Some researchers are very meticulous in stating the number of placebo responders withdrawn from their study and some do not even mention a run-in period. The Consolidated Standards of Reporting Trials (CONSORT) statement provides helpful recommendations, but is not accepted by all journals (Moher, Schulz, Altman,2001). Both the original CONSORT statement, published in 1995 (Begg, Cho, Eastwood et al,1996) and the revised version, published in 2001 (Moher, Schulz, Altman,2001) comprise a checklist and flow diagram to ensure the clear reporting of key elements of clinical trials.

Meta-analytic research will be improved if standards were upgraded for the reporting of primary research as recommended by the most recent CONSORT statement. It would be helpful if editors of medical journals would request that the placebo run-in phase and its resultant withdrawal of placebo responders be mentioned specifically in submitted articles. Lastly, the terms *placebo run-in phase* and *placebo responder* (or similar terms) should be made an indexable term and included in the checklist for reporting clinical trials in the biomedical literature (Begg, 1996).

APPENDIX I INCLUDED TRIALS

Beasley Jr CM, Sayler ME, Bosomworth JC, Wernicke JF. High-dose fluoxetine: efficacy and activating-sedating effects in agitated and retarded depression. *J Clin Psychopharmacol* 1991;11:166-174.

Byerley WF, Reimherr FW, Wood DR, Grosser BI. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *J Clin Psychopharmacol* 1988;8:112-115.

Claghorn J. A double-blind comparison of paroxetine and placebo in the treatment of depressed outpatients. *Int Clin Psychopharmacol* 1992;6 Suppl 4:25-30.

Claghorn JL. The safety and efficacy of paroxetine compared with placebo in a double-blind trial of depressed outpatients. *J Clin Psychiatry*1992;53 Suppl:33-35.

Claghorn JL, Earl CQ, Walczak DD, Stoner KA, Wong LF, Kanter D et al. Fluvoxamine maleate in the treatment of depression: a single-center, double-blind, placebo-controlled comparison with imipramine in outpatients. *J Clin Psychopharmacol* 1996;16:113-120.

Claghorn JL, Kiev A, Rickels K, Smith TW, Dunbar GC. Paroxetine versus placebo: a double-blind comparison in depressed patients. *J Clin Psychiatry* 1992;53:434-438.

Cohn JB, Crowder JE, Wilcox CS, Ryan PJ. A placebo- and imipramine-controlled study of paroxetine. *Psychopharmacol Bull* 1990;26:185-189.

Cohn JB, Wilcox C. A comparison of fluoxetine, imipramine, and placebo in patients with major depressive disorder. *J Clin Psychiatry* 1985;46:26-31.

Cohn JB, Wilcox CS. Paroxetine in major depression: a double-blind trial with imipramine and placebo. *J Clin Psychiatry* 1992;53 Suppl:52-56.

Corrigan MH, Denahan AQ, Wright CE, Ragual RJ, Evans DL. Comparison of pramipexole, fluoxetine, and placebo in patients with major depression. *Depress Anxiety* 2000;11:58-65.

Croft H, Settle Jr E, Houser T, Batey SR, Donahue RMJ, Ascher JA. A placebo-controlled comparison of the antidepressant efficacy and effects on sexual functioning of sustained-release bupropion and sertraline. *Clin Ther* 1999;21:643-658.

Doogan DP, Langdon CJ. A double-blind, placebo-controlled comparison of sertraline and dothiepin in the treatment of major depression in general practice. *Int Clin Psychopharmacol* 1994;9:95-100.

Dunbar GC, Cohn JB, Fabre LF, Feighner JP, Fieve RR, Mendels J et al. A comparison of paroxetine, imipramine and placebo in depressed out- patients. *Br J Psychiatry* 1991;159:394-398.

Dunlop SR, Dornseif BE, Wernicke JF, Potvin JH. Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. *Psychopharmacol Bull* 1990; 26:173-180.

Edwards JG, Goldie A. Placebo controlled trial of paroxetine in depressive illness. *Human Psychopharmacol* 1993;8:203-209.

Fabre L, Birkhimer LJ, Zaborny BA, Wong LF, Kapik BM. Fluvoxamine versus imipramine and placebo: a double-blind comparison in depressed patients. *Int Clin Psychopharmacol* 1996;11:119-127.

Fabre LF. A 6-week, double-blind trial of paroxetine, imipramine, and placebo in depressed outpatients. *J Clin Psychiatry* 1992;53 Suppl:40-43.

Fabre LF, Abuzzahab FS, Amin M, Claghorn JL, Mendels J, Petrie WM et al. Sertraline safety and efficacy in major depression: a double blind fixed dose comparison with placebo. *Biol Psychiatry* 1995;38:592-602.

Fabre LF, Putman HP. A fixed-dose clinical trial of fluoxetine in outpatients with major depression. *J Clin Psychiatry* 1987;48:406-408.

Fava M, Amsterdam JD, Deltito JA, Salzman C, Schwaller M, Dunner DL. A double-blind study of paroxetine, fluoxetine, and placebo in outpatients with major depression. *Ann Clin Psychiatry* 1998;10:145-150.

Feighner JP, Boyer WF, Meredith CH, Hendrickson GG. A placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. *Int Clin Psychopharmacol* 1989;4:239-244.

Feighner JP, Boyer WF, Merideth CH, Hendrickson GG. A double-blind comparison of fluoxetine, imipramine and placebo in outpatients with major depression. *Int Clin Psychopharmacol* 1989;4:127-134.

Feighner JP, Cohn JB, Fabre Jr LF, Fieve RR, Mendels J, Shrivastava RK et al. A study comparing paroxetine placebo and imipramine in depressed patients. *J Affect Disord* 1993;28:71-79.

Feighner JP, Overo K. Multicenter, placebo-controlled, fixed-dose study of citalopram in moderate-to-severe depression. *J Clin Psychiatry* 1999;60:824-830.

Heiligenstein JH, Tollefson GD, Faries DE. A double-blind trial of fluoxetine, 20 mg, and placebo in out-patients with DSM-III-R major depression and melancholia. *Int Clin Psychopharmacol* 1993;8:247-251.

Kiev A. A double-blind, placebo-controlled study of paroxetine in depressed outpatients. *J Clin Psychiatry* 1992;53 Suppl:27-29.

Lapierre YD, Browne M, Horn E, Oyewumi LK, Sarantidis D, Roberts N et al. Treatment of major affective disorder with fluvoxamine. *J Clin Psychiatry* 1987;48:65-68.

Lydiard RB, Laird LK, Morton Jr WA, Steele TE, Kellner C, Laraia MT et al. Fluvoxamine, imipramine, and placebo in the treatment of depressed outpatients: effects on depression. *Psychopharmacol Bull* 1989;25:68-70.

Lydiard RB, Stahl SM, Hertzman M, Harrison WM. A double-blind, placebo-controlled study comparing the effects of sertraline versus amitriptyline in the treatment of major depression. *J Clin Psychiatry* 1987;58:484-491.

March JS, Kobak KA, Jefferson JW, Mazza J, Greist JH. A double-blind, placebo-controlled trial of fluvoxamine versus imipramine in outpatients with major depression. *J Clin Psychiatry* 1990;51:200-202.

Peselow ED, Filippi AM, Goodnick P, Barouche F, Fieve RR. The short- and long-term efficacy of paroxetine HCl: A. Data from a 6- week double-blind parallel design trial vs. imipramine and placebo. *Psychopharmacol Bull* 1989;25:267-271.

Reimherr FW, Byerley WF, Ward MF, Lebegue BJ, Wender PH. Sertraline, a selective inhibitor of serotonin uptake, for the treatment of outpatients with major depressive disorder. *Psychopharmacol Bull* 1988;24:200-205.

Reimherr FW, Chouinard G, Cohn CK, Cole JO, Itil TM, LaPierre YD et al. Antidepressant efficacy of sertraline: a double-blind, placebo- and amitriptyline-controlled, multicenter comparison study in outpatients with major depression. *J Clin Psychiatry* 1990;51 Suppl B:18-27.

Roth D, Mattes J, Sheehan KH, Sheehan DV. A double-blind comparison of fluvoxamine, desipramine and placebo in outpatients with depression. *Prog Neuropsychopharmacol Biol Psychiatry* 1990;14:929-939.

Rudolph RL, Feiger AD. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord* 1999;56:171-181.

Shrivastava RK, Shrivastava SHP, Overweg N, Blumhardt CL. A double-blind comparison of paroxetine, imipramine, and placebo in major depression. *J Clin Psychiatry* 1992;53 Suppl:48-51.

Silverstone PH, Ravindran A. Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. *J Clin Psychiatry* 1999;60:22-28.

Smith WT, Glaudin V. A placebo-controlled trial of paroxetine in the treatment of major depression. *J Clin Psychiatry* 1992;53 Suppl:36-39.

Stark P, Hardison D. A review of multicenter controlled studies of fluoxetine vs. imipramine and placebo in outpatients with major depressive disorder. *J Clin Psychiatry* 1985;46:53-58.

Walczak DD, Apter JT, Halikas JA, Borison RL, Carman JS, Post GL et al. The oral dose-effect relationship for fluvoxamine: a fixed-dose comparison against placebo in depressed outpatients. *Ann-Clin-Psychiatry* 1996;8:139-151.

Wernicke JF, Dunlop SR, Dornseif BE, Bosomworth JC, Humbert M. Low-dose fluoxetine therapy for depression. *Psychopharmacol Bull* 1988;24:183-188.

Wernicke JF, Dunlop SR, Dornseif BE, Zerbe RL. Fixed-dose fluoxetine therapy for depression. *Psychopharmacol Bull* 1987;23:164-168.

APPENDIX II EXCLUDED TRIALS

Aguglia, Casacchia, Cassano, Faravelli, Ferrari, Giordano, Pancheri, Ravizza, Trabucchi, Bolino, &et. Double-blind study of the efficacy and safety of sertraline versus fluoxetine in major depression. *Int Clin Psychopharmacol.* 93;8:197-202.

Altamura, Montgomery, &Wernicke. The evidence for 20 mg a day of fluoxetine as the optimal dose in the treatment of depression. *British-Journal-of-Psychiatry.* 88;153:109-112.

Amin, Ananth, &Coleman. Fluvoxamine: anitidepressant effects confirmed in a placebo-controlled international study. *Clinical Neuropharmacology.* 84;7:580-581.

Amore, Bellini, Beradi, Berlinzani, Cervino, Ferrari, &Innamorati. Double-blind comparison of fluvoamine and impramine in depressed patients. *Current Therapeutic Research.* 89;46:815-820.

Anderson. Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *J Affect Disord.* 2000;58:19-36.

Anonymous. Citalopram: clinical effect profile in comparison with clomipramine. A controlled multicenter study. *Psychopharmacology.* 86;90:131-138.

Anonymous. A double-blind multi-center trial of fluoxetine and dothiepin in major depressive illness. *International Clinical Psychopharmacology.* 88;18:289-299.

Ansseau, Darimont, Lecoq, De Nayer, Evrard, Kremer, Devoitille, Dierick, Mertens, Mesotten, &et. Controlled comparison of nefazodone and amitriptyline in major depressive inpatients. *Psychopharmacology (Berl).* 94;115:254-60.

Ansseau, Gabriels, &Loyens. Controlled comparison of paroxetine and fluvoxamine in major depression. *Human Psychopharmacology.* 94;9:329-336.

Ansseau, Papart, Troisfontaines, Bartholome, Bataille, Charles, Schittecatte, Darimont, Devoitille, De Wilde, &et. Controlled comparison of milnacipran and fluoxetine in major depression. *Psychopharmacology (Berl).* 94;114:131-7.

Antidepressant Drug Trial Group. A double-blind multi-centre trial of fluoxetine and dothiepin in major depressive illness. South Wales *Int Clin Psychopharmacol.* 88;3:75-81.

Arminen, Ikonen, Pulkkinen, Leinonen, Mahlanen, Koponen, Kourula, Ryyppo, Korpela, Lehtonen, &et. A 12-week double-blind multi-centre study of paroxetine and imipramine in hospitalized depressed patients. *Acta Psychiatr Scand.* 94;89:382-9.

Asberg, Eriksson, Martensson, Traskman-Bendz, &Wagner. Therapeutic effects of serotonin uptake inhibitors in depression. *J Clin Psychiatry.* 86;47 Suppl:23-35.

Bakish, Cavazzoni, Chudzik, Ravindran, &Hrdina. Effects of selective serotonin reuptake inhibitors on platelet serotonin parameters in major depressive disorder. *Biol Psychiatry.* 97;41:184-90.

Baldwin, Hawley, Abed, Maragakis, Cox, Buckingham, Pover, &Ascher. A multicenter double-blind comparison of nefazodone and paroxetine in the treatment of outpatients with moderate-to-severe depression. *J Clin Psychiatry.* 96;57 Suppl 2:46-52.

Bascara. A double-blind study to compare the effectiveness and tolerability of paroxetine and amitriptyline in depressed patients. *Acta Psychiatr Scand Suppl.* 89;350:141-2.

Battegay, Hager, &Rauchfleisch. Double-blind comparative study of paroxetine and amitriptyline in depressed patients of a university psychiatric outpatient clinic (pilot study). *Neuropsychobiology.* 85;13:31-7.

Baumann, Nil, Souche, Montaldi, Baettig, Lambert, Uehlinger, Kasas, Amey, &Jonzier-Perey. A double-blind, placebo-controlled study of citalopram with and without lithium in the treatment of therapy-resistant depressive patients: a clinical, pharmacokinetic, and pharmacogenetic investigation. *J Clin Psychopharmacol.* 96;16:307-14.

Beasley, Dornseif, Pultz, Bosomworth, &Sayler. Fluoxetine versus trazodone: efficacy and activating-sedating effects. *J Clin Psychiatry.* 91;52:294-9.

Beasley, Holman, &Potvin. Fluoxetine compared with imipramine in the treatment of inpatient depression. A multicenter trial. *Ann Clin Psychiatry.* 93;5:199-207.

Beasley, Sayler, Bosomworth, &Wernicke. High-dose fluoxetine: efficacy and activating-sedating effects in agitated and retarded depression. *J Clin Psychopharmacol.* 91;11:166-74.

Beasley, Sayler, &Potvin. Fluoxetine versus amitriptyline in the treatment of major depression: a multicenter trial. *Int Clin Psychopharmacol.* 93;8:143-9.

Benkert, Szegedi, Wetzel, Staab, Meister, &Philipp. Dose escalation vs. continued doses of paroxetine and maprotiline: a prospective study in depressed out-patients with inadequate treatment response. *Acta Psychiatr Scand.* 97;95:288-96.

Bennie, Mullin, &Martindale. A double-blind multicenter trial comparing sertraline and fluoxetine in outpatients with major depression. *J Clin Psychiatry.* 95;56:229-37.

Berman, Anand, Cappiello, Miller, Hu, Oren, &Charney. The use of pindolol with fluoxetine in the treatment of major depression: final results from a double-blind, placebo-controlled trial. *Biol Psychiatry.* 99;45:1170-7.

Berman, Darnell, Miller, Anand, &et al. Effect of pindolol in hastening response to fluoxetine in the treatment of major depression: A double blind, placebo controlled trial. *American-Journal-of-Psychiatry.* 97;154:37-43.

Bersani, Rapisarda, Ciani, Bertolino, &Sorge. A double-blind comparative study of sertraline and amitriptyline in outpatients with major depressive episodes. *Human Psychopharmacology.* 94;9:63-68.

Bezchlibnyk-Butler, Aleksic, &Kennedy. Citalopram--a review of pharmacological and clinical effects. *Journal of Psychiatry and Neuroscience.* 2000;25:241-254.

Bialik, Ravindran, Bakish, &Lapierre. A comparison of placebo responders and nonresponders in subgroups of depressive disorder. *J Psychiatry Neurosci.* 95;20:265-70.

Bignamini&Rapisarda. A double-blind multicentre study of paroxetine and amitriptyline in depressed outpatients. Italian Paroxetine Study Group. *Int Clin Psychopharmacol.* 92;6 Suppl 4:37-41.

Boerner&Moller. The importance of new antidepressants in the treatment of anxiety/depressive disorders. *Pharmacopsychiatry.* 99;32:119-26.

Borup, Meidahl, Petersen, Vangtorp, &le Fevre Honore. An early clinical phase II evaluation of paroxetine, a new potent and selective 5HT-uptake inhibitor in patients with depressive illness. *Pharmacopsychiatria.* 82;15:183-6.

Bouchard, Delaunay, Delisle, Grasset, Mermberg, Molczadzki, Pagot, Richou, Robert, Ropert, &et. Citalopram versus maprotiline: a controlled, clinical multicentre trial in depressed patients. *Acta Psychiatr Scand.* 87;76:583-92.

Bougerol, Uchida, Gachoud, Kohler, &Mikkelsen. Efficacy and tolerability of moclobemide compared with fluvoxamine in depressive disorder (DSM III). A French/Swiss double-blind trial. *Psychopharmacology (Berl).* 92;106 Suppl:S102-8.

Bourin. [Serotonin and antiserotonergic agents]. *Rev Prat.* 90;40:1711-3.

Bowden, Schatzberg, Rosenbaum, Contreras, Samson, Dessain, &Sayler. Fluoxetine and desipramine in major depressive disorder. *J Clin Psychopharmacol.* 93;13:305-11.

Boyer&Feighner. An overview of fluoxetine, a new serotonin-specific antidepressant. *Mt Sinai J Med.* 89;56:136-40.

Bramanti, Ricci, Roncari, Bilone, Inga, Teti, De Critofaro, Ceccarelli, Di Perri, &Candela. An Italian multicenter experience with fluvoxamine, a new antidepressant drug, versus imipramine. *Current Therapeutic Research.* 88;43:718-725.

Bremner. Fluoxetine in depressed patients: a comparison with imipramine. *J Clin Psychiatry.* 84;45:414-9.

Bremner, Abrahams, Crupie, McCawley, Proctor, &Sathananthan. Multicenter double-blind comparison of nomifensine and imipramine for efficacy and safety in depressed outpatients. *J Clin Psychiatry.* 84;45:56-9.

Bressa, Brugnoli, &Pancheri. A double-blind study of fluoxetine and imipramine in major depression. *Int Clin Psychopharmacol.* 89;4 Suppl 1:69-73.

Byerley, Reimherr, Wood, &Grosser. Fluoxetine, a selective serotonin uptake inhibitor, for the treatment of outpatients with major depression. *J Clin Psychopharmacol.* 88;8:112-5.

Casciano, Arikian, Tarride, Doyle, &Casciano. A pharmacoeconomic evaluation of major depressive disorder (Italy). *Epidemiol Psichiatr Soc.* 99;8:220-31.

Chouinard. A double-blind controlled clinical trial of fluoxetine and amitriptyline in the treatment of outpatients with major depressive disorder. *J Clin Psychiatry.* 85;46:32-7.

Chouinard, Saxena, Belanger, Ravindran, Bakish, Beauclair, Morris, Vasavan Nair, Manchanda, Reesal, Remick, &O'Neill. A Canadian multicenter, double-blind study of paroxetine and fluoxetine in major depressive disorder. *J Affect Disord.* 99;54:39-48.

Christiansen, Behnke, Black, Ohrstrom, Bork-Rasmussen, &Nilsson. Paroxetine and amitriptyline in the treatment of depression in general practice. *Acta Psychiatr Scand.* 96;93:158-63.

Claghorn. A double-blind comparison of paroxetine and placebo in the treatment of depressed outpatients. *Int Clin Psychopharmacol.* 92;6 Suppl 4:25-30.

Claghorn. The safety and efficacy of paroxetine compared with placebo in a double- blind trial of depressed outpatients. *J Clin Psychiatry.* 92;53 Suppl:33-5.

Claghorn, Earl, Walczak, Stoner, Wong, Kanter, &Houser. Fluvoxamine maleate in the treatment of depression: a single-center, double-blind, placebo-controlled comparison with imipramine in outpatients. *J Clin Psychopharmacol.* 96;16:113-20.

Claghorn&Feighner. A double-blind comparison of paroxetine with imipramine in the long- term treatment of depression. *J Clin Psychopharmacol.* 93;13:23S-27S.

Claghorn, Kiev, Rickels, Smith, &Dunbar. Paroxetine versus placebo: a double-blind comparison in depressed patients. *J Clin Psychiatry.* 92;53:434-8.

Clerc, Ruimy, &Verdeau-Palles. A double-blind comparison of venlafaxine and fluoxetine in patients hospitalized for major depression and melancholia. The Venlafaxine French Inpatient Study Group. *Int Clin Psychopharmacol.* 94;9:139-43.

Cohn, Crowder, Wilcox, &Ryan. A placebo- and imipramine-controlled study of paroxetine. *Psychopharmacol Bull.* 90;26:185-9.

Cohn&Wilcox. A comparison of fluoxetine, imipramine, and placebo in patients with major depressive disorder. *J Clin Psychiatry.* 85;46:26-31.

Cohn&Wilcox. Paroxetine in major depression: a double-blind trial with imipramine and placebo. *J Clin Psychiatry.* 92;53 Suppl:52-6.

Coleman&Block. Fluvoxamine maleate, a serotonergic antidepressant; a comparison with chlorimipramine. *Prog Neuropsychopharmacol Biol Psychiatry.* 82;6:475-8.

Conti, Dell'Osso, &Re. Fluoxamine maleate: double-blind clinical trial vs placebo in hospitalized depressed patients. *Current Therapeutic Research.* 88;43:468-480.

Coppen&Bailey. Enhancement of the antidepressant action of fluoxetine by folic acid: A randomised, placebo controlled trial. *Journal-of-Affective-Disorders.* 2000;60:121-130.

Corne&Hall. A double-blind comparative study of fluoxetine and dothiepin in the treatment of depression in general practice. *Int Clin Psychopharmacol.* 89;4:245-54.

Corrigan, Denahan, Wright, Ragual, &Evans. Comparison of pramipexole, fluoxetine, and placebo in patients with major depression. *Depress Anxiety.* 2000;11:58-65.

Coryell. Augmentation strategies for inadequate antidepressant response: a review of placebo-controlled studies. *Ann Clin Psychiatry.* 2000;12:141-6.

Croft, Settle, Houser, Batey, Donahue, &Ascher. A placebo-controlled comparison of the antidepressant efficacy and effects on sexual functioning of sustained-release bupropion and sertraline. *Clin Ther.* 99;21:643-58.

Crow, Daly, Fatemi, MacGraw, &Meller. Citalopram: a rapid-onset selective serotonin reuptake inhibitor? *Can J Psychiatry.* 99;44:1051-2.

Cunningham. Depression in the medically ill: choosing an antidepressant. *J Clin Psychiatry.* 94;55 Suppl A:90-7; discussion 98-100.

Cunningham, Borison, Carman, Chouinard, Crowder, Diamond, Fischer, &Hearst. A comparison of venlafaxine, trazodone, and placebo in major depression. *J Clin Psychopharmacol.* 94;14:99-106.

Danish University Antidepressant Group. Paroxetine: a selective serotonin reuptake inhibitor showing better tolerance, but weaker antidepressant effect than clomipramine in a controlled multicenter study. *J Affect Disord.* 90;18:289-99.

de Jonghe, Ravelli, &Tuynman-Qua. A randomized, double-blind study of fluoxetine and maprotiline in the treatment of major depression. *Pharmacopsychiatry.* 91;24:62-7.

de Wilde, Mertens, Overo, &Petersen. Citalopram versus mianserin. A controlled, double-blind trial in depressed patients. *Acta Psychiatr Scand.* 85;72:89-96.

De Wilde, Spiers, Mertens, Bartholome, Schotte, &Leyman. A double-blind, comparative, multicentre study comparing paroxetine with fluoxetine in depressed patients. *Acta Psychiatr Scand.* 93;87:141-5.

Debus, Rush, Himmel, Tyler, Polatin, &Weissenburger. Fluoxetine versus trazodone in the treatment of outpatients with major depression. *J Clin Psychiatry.* 88;49:422-6.

Dechant&Clissold. Paroxetine. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic potential in depressive illness. *Drugs.* 91;41:225-53.

deJonghe, Swinkels, &Tuynman-Qua. Randomized double-blind study of fluoxamine and maprotiline in treatment of depression. *Pharmacopsychiatry.* 91;24:21-27.

DeWilde&Doogan. Fluvoxamine and chlorimipramine in endogenous depression. *Journal of Affective Disorders.* 82;4:249-259.

DeWilde, Mertens, &Wakelin. Clinical trials of fluoxamine vs chlorimipramine with single and three times daily dosing. *British Journal of Clinical Pharmacology.* 83;15:427S-431S.

Dick&Ferrero. A double-blind comparative study of the clinical efficacy of fluvoxamine and chlorimipramine. *Br J Clin Pharmacol.* 83;15 Suppl 3:419S-425S.

Dominguez, Goldstein, Jacobson, &Steinbook. A double-blind placebo-controlled study of fluvoxamine and imipramine in depression. *J Clin Psychiatry.* 85;46:84-7.

Doogan&Caillard. Sertraline in the prevention of depression. *Br J Psychiatry.* 92;160:217-22.

Doogan&Langdon. A double-blind, placebo-controlled comparison of sertraline and dothiepin in the treatment of major depression in general practice. *Int Clin Psychopharmacol.* 94;9:95-100.

Dowling, Webb, Halpin, &Sangiwa. Fluoxetine: a comparative study with dothiepin. *Irish Journal of Psychiatry.* 90:3-7.

Dunbar, Claghorn, Kiev, Rickels, &Smith. A comparison of paroxetine and placebo in depressed outpatients. *Acta Psychiatr Scand.* 93;87:302-5.

Dunbar, Cohn, Fabre, Feighner, Fieve, Mendels, &Shrivastava. A comparison of paroxetine, imipramine and placebo in depressed out- patients. *Br J Psychiatry.* 91;159:394-8.

Dunlop, Dornseif, Wernicke, &Potvin. Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. *Psychopharmacol Bull.* 26:173-80.

Dunner&Dunbar. Optimal dose regimen for paroxetine. *J Clin Psychiatry.* 92;53 Suppl:21-6.

Edwards&Anderson. Systematic review and guide to selection of selective serotonin reuptake inhibitors. *Drugs.* 99;57:507-33.

Edwards&Goldie. Placebo controlled trial of paroxetine in depressive illness. *Human-Psychopharmacology-Clinical-and-Experimental.* 93;8:203-209.

Einarson, Arikian, Casciano, &Doyle. Comparison of extended-release venlafaxine, selective serotonin reuptake inhibitors, and tricyclic antidepressants in the treatment of depression: a meta-analysis of randomized controlled trials. *Clin Ther.* 99;21:296-308.

Ekselius, von Knorring, &Eberhard. A double-blind multicenter trial comparing sertraline and citalopram in patients with major depression treated in general practice. *Int Clin Psychopharmacol.* 97;12:323-31.

Emrich, Berger, Riemann, &von Zerssen. Serotonin reuptake inhibition vs. norepinephrine reuptake inhibition: a double-blind differential-therapeutic study with fluvoxamine and oxaprotiline in endogenous and neurotic depressives. *Pharmacopsychiatry.* 87;20:60-3.

Fabre&Crismon. Efficacy of fluoxetine in outpatients with major depression. *Current Therapeutic Reserach.* 85;37:115-123.

Fabre, Birkhimer, Zaborny, Wong, &Kapik. Fluvoxamine versus imipramine and placebo: a double-blind comparison in depressed patients. *Int Clin Psychopharmacol.* 96;11:119-27.

Fabre. A 6-week, double-blind trial of paroxetine, imipramine, and placebo in depressed outpatients. *J Clin Psychiatry.* 92;53 Suppl:40-3.

Fabre, Abuzzahab, Amin, Claghorn, &et al. Sertraline safety and efficacy in major depression: A double blind fixed dose comparison with placebo. *Biological-Psychiatry.* 95;38:592-602.

Fabre&Putman. A fixed-dose clinical trial of fluoxetine in outpatients with major depression. *J Clin Psychiatry.* 87;48:406-8.

Fabre, Scharf, &Itil. Comparative efficacy and safety of nortriptyline and fluoxetine in the treatment of major depression: a clinical study. *J Clin Psychiatry.* 91;52 Suppl:62-7.

Fava, Amsterdam, Deltito, Salzman, Schwaller, &Dunner. A double-blind study of paroxetine, fluoxetine, and placebo in outpatients with major depression. *Ann Clin Psychiatry.* 98;10:145-50.

Fava, Judge, Hoog, Nilsson, &Koke. Fluoxetine versus sertraline and paroxetine in major depressive disorder: changes in weight with long-term treatment. *J Clin Psychiatry.* 2000;61:863-7.

Fava, Rosenbaum, Hoog, Tepner, Kopp, &Nilsson. Fluoxetine versus sertraline and paroxetine in major depression: tolerability and efficacy in anxious depression. *J Affect Disord.* 2000;59:119-26.

Fawcett, Zajecka, Kravitz, Jeffriess, &Scorza. Fluoxetine versus amitriptyline in adult out-patients with major depression. *Current Therapeutic Reserach.* 89;45:821-832.

Feiger, Kiev, Shrivastava, Wisselink, &Wilcox. Nefazodone versus sertraline in outpatients with major depression: focus on efficacy, tolerability, and effects on sexual function and satisfaction. *J Clin Psychiatry.* 96;57 Suppl 2:53-62.

Feighner. A comparative trial of fluoxetine and amitriptyline in patients with major depressive disorder. *J Clin Psychiatry.* 85;46:369-72.

Feighner. A double-blind comparison of paroxetine, imipramine and placebo in depressed outpatients. *Int Clin Psychopharmacol.* 92;6 Suppl 4:31-5.

Feighner&Boyer. Paroxetine in the treatment of depression: a comparison with imipramine and placebo. *Acta Psychiatr Scand Suppl.* 89;350:125-9.

Feighner&Boyer. Paroxetine in the treatment of depression: a comparison with imipramine and placebo. *J Clin Psychiatry.* 92;53 Suppl:44-7.

Feighner, Boyer, Meredith, &Hendrickson. A placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. *Int Clin Psychopharmacol.* 89;4:239-44.

Feighner, Boyer, Merideth, &Hendrickson. A double-blind comparison of fluoxetine, imipramine and placebo in outpatients with major depression. *Int Clin Psychopharmacol.* 89;4:127-34.

Feighner, Cohn, Fabre, Fieve, Mendels, Shrivastava, &Dunbar. A study comparing paroxetine placebo and imipramine in depressed patients. *J Affect Disord.* 93;28:71-9.

Feighner, Gardner, Johnston, Batey, Khayrallah, Ascher, &Lineberry. Double-blind comparison of bupropion and fluoxetine in depressed outpatients. *J Clin Psychiatry.* 91;52:329-35.

Feighner&Overo. Multicenter, placebo-controlled, fixed-dose study of citalopram in moderate-to-severe depression. *J Clin Psychiatry.* 99;60:824-30.

Ferreri. Fluoxetine versus amineptine in the treatment of outpatients with major depressive disorders. *Int Clin Psychopharmacol.* 89;4 Suppl 1:97-101.

Ferreri, Lavergne, Berlin, Payan, &Puech. Benefits from mianserin augmentation of fluoxetine in patients with major depression non-responders to fluoxetine alone. *Acta Psychiatr Scand.* 2001;103:66-72.

Fieve, Goodnick, Peselow, &Schlegel. Fluoxetine response: endpoint vs pattern analysis. *Int Clin Psychopharmacol.* 86;1:320-3.

Fieve, Goodnick, Peselow, Barouche, &Schlegel. Pattern analysis of antidepressant response to fluoxetine. *J Clin Psychiatry.* 86;47:560-2.

Folkerts, Michael, Tolle, Schonauer, Mucke, &Schulze-Monking. Electroconvulsive therapy vs. paroxetine in treatment-resistant depression -- a randomized study. *Acta Psychiatr Scand.* 97;96:334-42.

Fontaine. The efficacy and safety of sertraline versus imipramine in outpatients with major depression: a six month double blind parallel multicenter study. *European Neuropsychopharmacology.* 1:75.

Franchini, Gasperini, &Smeraldi. A 24-month follow-up study of unipolar subjects: a comparison between lithium and fluvoxamine. *J Affect Disord.* 94;32:225-31.

Fredman, Fava, Kienke, White, Nierenberg, &Rosenbaum. Partial response, nonresponse, and relapse with selective serotonin reuptake inhibitors in major depression: a survey of current "next- step" practices. *J Clin Psychiatry.* 2000;61:403-8.

Fudge, Perry, Garvey, &Kelly. A comparison of the effect of fluoxetine and trazodone on the cognitive functioning of depressed outpatients. *J Affect Disord.* 90;18:275-80.

Gagiano. A double-blind comparison of paroxetine and fluoxetine in patients with major depression. *British Journal of Clinical Research.* 93;4:145-152.

Gagiano, Muller, Fourie, &Le Roux. The therapeutic efficacy of paroxetine: (a) an open study in patients with major depression not responding to antidepressants; (b) a double-blind comparison with amitriptyline in depressed outpatients. *Acta Psychiatr Scand Suppl.* 89;350:130-1.

Gasperini, Gatti, Bellini, Anniverno, &Smeraldi. Perspectives in clinical psychopharmacology of amitriptyline and fluvoxamine. A double-blind study in depressed inpatients. *Neuropsychobiology.* 92;26:186-92.

Gattaz, Vogel, Kick, &Kohnen. Moclobemide versus fluoxetine in the treatment of inpatients with major depression. *J Clin Psychopharmacol.* 95;15:35S-40S.

Geerts, Bruynooghe, De Cuyper, Demeulemeester, &Haazen. Moclobemide versus fluoxetine for major depressive episodes. *Clin Neuropharmacol.* 94;17 Suppl 1:S50-7.

Gelenberg&Ottosson. "Paroxetine versus placebo: A double blind comparison in depressed patients": Duplicate publication. *Journal-of-Clinical-Psychiatry.* 95;56:81.

Ginestet. Fluoxetine in endogenous depression and melancholia versus clomipramine. *Int Clin Psychopharmacol.* 89;4 Suppl 1:37-40.

Gonella, Baignoli, &Ecari. Fluvoxamine and imipramine in the treatment of depressive patients: a double-blind controlled study. *Curr Med Res Opin.* 90;12:177-84.

Goodnick, Fieve, Peselow, Barouche, &Schlegel. Double-blind treatment of major depression with fluoxetine: use of pattern analysis and relation of HAM-D score to CGI change. *Psychopharmacol Bull.* 87;23:162-3.

Gravem, Amthor, Astrup, Elgen, Gjessing, Gunby, Pettersen, Kyrdalen, Vaadal, Ofsti, &et. A double-blind comparison of citalopram (Lu 10-171) and amitriptyline in depressed patients. *Acta Psychiatr Scand.* 87;75:478-86.

Grimsley&Jann. Paroxetine, sertraline, and fluvoxamine: new selective serotonin reuptake inhibitors. *Clin Pharm.* 92;11:930-57.

Guelfi, Ansseau, Corruble, Samuelian, Tonelli, Tournoux, &Pletan. A double blind comparison of the efficacy and safety of milnacipran and fluoxetine in depressed inpatients. *International-Clinical-Psychopharmacology.* 98;13:121-128.

Guelfi, Dreyfus, &Pichot. A double-blind controlled clinical trial comparing fluvoxamine with imipramine. *Br J Clin Pharmacol.* 83;15 Suppl 3:411S-417S.

Guelfi, Strub, &Loft. Efficacy of intravenous citalopram compared with oral citalopram for severe depression. Safety and efficacy data from a double-blind, double- dummy trial. *J Affect Disord.* 2000;58:201-9.

Guy, Wilson, Ban, King, Manov, &Fjetland. A double-blind clinical trial of fluvoxamine and imipramine in patients with primary depression. *Psychopharmacol Bull.* 84;20:73-8.

Haffmans, Timmerman, &Hoogduin. Efficacy and tolerability of citalopram in comparison with fluvoxamine in depressed outpatients: a double-blind, multicentre study. The LUCIFER Group. *Int Clin Psychopharmacol.* 96;11:157-64.

Hamilton, Jones, Hoda, Keane, Majid, &Zaidi. Flupenthixol and fluvoxamine in mild to moderate depression: a comparison in general practice. *Pharmatherapeutica.* 89;5:292-7.

Hanson&Bussiere. Predicting relapse: a meta-analysis of sexual offender recidivism studies. *J Consult Clin Psychol.* 98;66:348-62.

Harris, Szulecka, &Anstee. Fluoxamine versus amitriptyline in depressed hospital out-patients: a multicenter double-blind comparative trial. *British Journal of Clinical Reasearch.* 91;2:89-99.

Heiligenstein, Tollefson, &Faries. A double-blind trial of fluoxetine, 20 mg, and placebo in out-patients with DSM-III-R major depression and melancholia. *Int Clin Psychopharmacol.* 93;8:247-51.

Heiligenstein, Tollefson, &Faries. Response patterns of depressed outpatients with and without melancholia: A double blind, placebo controlled trial of fluoxetine versus placebo. *Journal-of-Affective-Disorders.* 94;30:163-173.

Hirschfeld. Efficacy of SSRIs and newer antidepressants in severe depression: Comparison with TCAs. *Journal-of-Clinical-Psychiatry.* 99;60:326-335.

Hochstrasser, Isaksen, Koponen, Lauritzen, Mahnert, Rouillon, Wade, Andersen, Pedersen, Swart, &Nil. Prophylactic effect of citalopram in unipolar, recurrent depression: placebo-controlled study of maintenance therapy. *Br J Psychiatry.* 2001;178:304-10.

Itil, Shrivastava, Mukherjee, Coleman, &Michael. A double-blind placebo-controlled study of fluvoxamine and imipramine in out-patients with primary depression. *Br J Clin Pharmacol.* 83;15 Suppl 3:433S-438S.

Judd, Moore, Norman, Burrows, Gupta, &Parker. A multicentre double blind trial of fluoxetine versus amitriptyline in the treatment of depressive illness. *Aust N Z J Psychiatry.* 93;27:49-55.

Kasper, Moller, Montgomery, &Zondag. Antidepressant efficacy in relation to item analysis and severity of depression: a placebo-controlled trial of fluvoxamine versus imipramine. *Int Clin Psychopharmacol.* 95;9 Suppl 4:3-12.

Keegan, Bowen, Blackshaw, Saleh, Dayal, Remillard, Shrikhande, Cebrian Perez, &Boulton. A comparison of fluoxetine and amitriptyline in the treatment of major depression. *Int Clin Psychopharmacol.* 91;6:117-24.

Keller, Kocsis, Thase, Gelenberg, Rush, Koran, Schatzberg, Russell, Hirschfeld, Klein, McCullough, Fawcett, Kornstein, LaVange, &Harrison. Maintenance phase efficacy of sertraline for chronic depression: a randomized controlled trial. *JAMA.* 98;280:1665-72.

Kelly, Perry, Holstad, &Garvey. Serum fluoxetine and norfluoxetine concentrations and antidepressant response. *Ther Drug Monit.* 89;11:165-70.

Kiev. A double-blind, placebo-controlled study of paroxetine in depressed outpatients. *J Clin Psychiatry.* 92;53 Suppl:27-9.

Kiev&Feiger. A double-blind comparison of fluvoxamine and paroxetine in the treatment of depressed outpatients. *J Clin Psychiatry.* 97;58:146-52.

Klok, Brouwer, van Praag, &Doogan. Fluvoxamine and clomipramine in depressed patients. A double-blind clinical study. *Acta Psychiatr Scand.* 81;64:1-11.

Kornstein, Schatzberg, Thase, Yonkers, McCullough, Keitner, Gelenberg, Davis, Harrison, &Keller. Gender differences in treatment response to sertraline versus imipramine in chronic depression. *Am J Psychiatry.* 2000;157:1445-52.

Kuha, Mehtonen, Henttonen, &Naarala. The efficacy of fluoxetine versus maprotiline in depressed patients and by dose. *Nord. Psykiatr. Tidsskr.* 91;45:109-117.

Kuhs&Rudolf. A double-blind study of the comparative antidepressant effect of paroxetine and amitriptyline. *Acta Psychiatr Scand Suppl.* 89;350:145-6.

Laakmann, Blaschke, Engel, &Schwarz. Fluoxetine vs amitriptyline in the treatment of depressed out-patients. *Br J Psychiatry Suppl.* 88:64-8.

Landen, Bjorling, Agren, &Fahlen. A randomized, double-blind, placebo-controlled trial of buspirone in combination with an SSRI in patients with treatment-refractory depression. *J Clin Psychiatry.* 98;59:664-8.

Landen, Eriksson, Agren, &Fahlen. Effect of buspirone on sexual dysfunction in depressed patients treated with selective serotonin reuptake inhibitors. *J Clin Psychopharmacol.* 99;19:268-71.

Lapierre, Browne, Horn, Oyewumi, Sarantidis, Roberts, Badoe, &Tessier. Treatment of major affective disorder with fluvoxamine. *J Clin Psychiatry.* 87;48:65-8.

Lapierre, Rastogi, &Singhal. Fluvoxamine influences serotonergic system in the brain: neurochemical evidence. *Neuropsychobiology.* 83;10:213-6.

Laursen, Mikkelsen, Rasmussen, &le Fevre Honore. Paroxetine in the treatment of depression--a randomized comparison with amitriptyline. *Acta Psychiatr Scand.* 85;71:249-55.

Lecrubier, Pletan, Solles, Tournoux, &Magne. Clinical efficacy of milnacipran: placebo-controlled trials. *Int Clin Psychopharmacol.* 96;11 Suppl 4:29-33.

Levine, Deo, &Mahadevan. A comparative trial of a new antidepressant, fluoxetine. *Br J Psychiatry.* 87;150:653-5.

Levine, Deo, &Mahadevan. A comparative trial of a new antidepressant, fluoxetine. *Int Clin Psychopharmacol.* 89;4 Suppl 1:41-5.

Levitan, Shen, Jindal, Driver, Kennedy, &Shapiro. Preliminary randomized double-blind placebo-controlled trial of tryptophan combined with fluoxetine to treat major depressive disorder: antidepressant and hypnotic effects. *J Psychiatry Neurosci.* 2000;25:337-46.

Loeb, Albano, &Gandolfo. Fluoxetine versus imipramine. *Int Clin Psychopharmacol.* 89;4 Suppl 1:75-9.

Lonnqvist, Sihvo, Syvalahti, &Kiviruusu. Moclobemide and fluoxetine in atypical depression: a double-blind trial. *J Affect Disord.* 94;32:169-77.

Lonnqvist, Sintonen, Syvalahti, Appelberg, Koskinen, Mannikko, Mehtonen, Naarala, Sihvo, Auvinen, &et. Antidepressant efficacy and quality of life in depression: a double-blind study with moclobemide and fluoxetine. *Acta Psychiatr Scand.* 94;89:363-9.

Loo, Saiz-Ruiz, Costa, Ansseau, Herrington, Vaz-Serra, Dilling, &de Risio. Efficacy and safety of tianeptine in the treatment of depressive disorders in comparison with fluoxetine. *J Affect Disord.* 99;56:109-18.

Lydiard, Laird, Morton, Steele, Kellner, Laraia, &Ballenger. Fluvoxamine, imipramine, and placebo in the treatment of depressed outpatients: effects on depression. *Psychopharmacol Bull.* 89;25:68-70.

Lydiard, Stahl, Hertzman, &Harrison. A double-blind, placebo-controlled study comparing the effects of sertraline versus amitriptyline in the treatment of major depression. *J Clin Psychiatry.* 97;58:484-91.

Manna, Martucci, &Agnoli. Double-blind controlled study on the clinical efficacy and safety of fluoxetine vs clomipramine in the treatment of major depressive disorders. *Int Clin Psychopharmacol.* 89;4 Suppl 1:81-8.

Manniche&Bech. Paroxetine and imipramine in the treatment of depressive patients in psychiatric practice. *Acta Psychiatrica Scandanavia.* 92;86:437-444.

March, Kobak, Jefferson, Mazza, &Greist. A double-blind, placebo-controlled trial of fluvoxamine versus imipramine in outpatients with major depression. *J Clin Psychiatry.* 90;51:200-2.

Marchesi, Ceccherininelli, Rossi, &Maggini. Is anxious agitated major depression responsive to fluoxetine? A double blind comparison with amitriptyline. *Pharmacopsychiatry.* 98;31:216-221.

Martensson, Nyberg, Toresson, Brodin, &Bertilsson. Fluoxetine treatment of depression. Clinical effects, drug concentrations and monoamine metabolites and N-terminally extended substance P in cerebrospinal fluid. *Acta Psychiatr Scand.* 89;79:586-96.

Martin, Hilton, Kerry, &Richards. General practitioners' perceptions of the tolerability of antidepressant drugs: a comparison of selective serotonin reuptake inhibitors and tricyclic antidepressants. *BMJ.* 97;314:646-51.

Masco&Sheetz. Double-blind comparison of fluoxetine and amitriptyline in the treatment of major depressive illness. *Advances in Therapy.* 85;2:275-284.

McGrath, Stewart, Janal, Petkova, Quitkin, &Klein. A placebo-controlled study of fluoxetine versus imipramine in the acute treatment of atypical depression. *Am J Psychiatry.* 2000;157:344-50.

Mendels. Clinical experience with serotonin reuptake inhibiting antidepressants. *J Clin Psychiatry.* 87;48 Suppl:26-30.

Mertens&Pintens. Paroxetine in the treatment of depression. A double-blind multicenter study versus mianserin. *Acta Psychiatr Scand.* 88;77:683-8.

Michelson, Fava, Amsterdam, Apter, Londborg, Tamura, &Tepner. Interruption of selective serotonin reuptake inhibitor treatment. Double-blind, placebo-controlled trial. *Br J Psychiatry.* 2000;176:363-8.

Miller, Naylor, Murtagh, &Winslow. A double-blind comparison of paroxetine and placebo in the treatment of depressed patients in a psychiatric outpatient clinic. *Acta Psychiatr Scand Suppl.* 89;350:143-4.

Moller, Berzewski, Eckmann, Gonzalves, Kissling, Knorr, Ressler, Rudolf, Steinmeyer, Magyar, &et. Double-blind multicenter study of paroxetine and amitriptyline in depressed inpatients. *Pharmacopsychiatry.* 93;26:75-8.

Montgomery. The efficacy of fluoxetine as an antidepressant in the short and long term. *Int Clin Psychopharmacol.* 89;4 Suppl 1:113-9.

Montgomery. Serotonin, sertraline and depression. *Journal of Psychopharmacology.* 95;9:179-184.

Montgomery&Djarv. The antidepressant efficacy of citalopram. *Int Clin Psychopharmacol.* 96;11 Suppl 1:29-33.

Montgomery, Dufour, Brion, Gailledreau, Laqueille, Ferrey, Moron, Parant-Lucena, Singer, Danion, &et. The prophylactic efficacy of fluoxetine in unipolar depression. *Br J Psychiatry Suppl.* 88:69-76.

Montgomery, Henry, McDonald, Dinan, Lader, Hindmarch, Clare, &Nutt. Selective serotonin reuptake inhibitors: meta-analysis of discontinuation rates. *Int Clin Psychopharmacol.* 94;9:47-53.

Montgomery, McAuley, Rani, Roy, &Montgomery. A double blind comparison of zimelidine and amitriptyline in endogenous depression. *Acta Psychiatr Scand Suppl.* 81;290:314-27.

Montgomery, Rasmussen, Lyby, Connor, &Tanghoj. Dose response relationship of citalopram 20 mg, citalopram 40 mg and placebo in the treatment of moderate and severe depression. *Int Clin Psychopharmacol.* 92;6 Suppl 5:65-70.

Montgomery, Rasmussen, &Tanghoj. A 24-week study of 20 mg citalopram, 40 mg citalopram, and placebo in the prevention of relapse of major depression. *Int Clin Psychopharmacol.* 93;8:181-8.

Montgomery&Dunbar. Paroxetine is better than placebo in relapse prevention and the prophylaxis of recurrent depression. *International-Clinical-Psychopharmacology.* 93;8:189-195.

Moon&Vince. Treatment of major depression in general practice: a double-blind comparison of paroxetine and lofepramine. *Br J Clin Pract.* 96;50:240-4.

Moon, Jago, Wood, &Doogan. A double-blind comparison of sertraline and clomipramine in the treatment of major depressive disorder and associated anxiety in general practice. *Journal of Psychopharmacology.* 94;8:171-176.

Muijen, Roy, Silverstone, Mehmet, &Christie. A comparative clinical trial of fluoxetine, mianserin and placebo in depressed outpatients. *Acta Psychiatr Scand.* 88;78:384-90.

Mullin, Pandita-Gunawardena, &Whitehead. A double-blind comparison of fluvoxamine and dothiepin in the treatment of major affective disorder. *Br J Clin Pract.* 88;42:51-5.

Mundo, Guglielmo, &Bellodi. Effect of adjuvant pindolol on the antiobsessional response to fluvoxamine: a double-blind, placebo-controlled study. *Int Clin Psychopharmacol.* 98;13:219-24.

Nathan, Perel, Pollock, &Kupfer. The role of neuropharmacologic selectivity in antidepressant action: fluvoxamine versus desipramine. *J Clin Psychiatry.* 90;51:367-72.

Nemeroff, Ninan, &Ballenger. Double-blind multicenter comparison of fluvoxamine versus sertraline in the treatment of depressed outpatients. *Depression.* 95;3:163-9.

Nielsen, Behnke, Arup, Christiansen, Geisler, Ipsen, Maach-Moller, &Ohrberg. A comparison of fluoxetine and imipramine in the treatment of outpatients with major depressive disorder. *Acta Psychiatr Scand.* 93;87:269-72.

Nielsen, Morsing, Petersen, Larsen, Moller, Manniche, &Skausig. Paroxetine and imipramine treatment of depressive patients in a controlled multicentre study with plasma amino acid measurements. *Acta Psychiatr Scand.* 91;84:233-41.

Nierenberg, Farabaugh, Alpert, Gordon, Worthington, Rosenbaum, &Fava. Timing of onset of antidepressant response with fluoxetine treatment. *Am J Psychiatry.* 2000;157:1423-8.

Noguera, Altuna, Alvarez, Ayuso, Casais, &Udina. Fluoxetine vs. clomipramine in depressed patients: a controlled multicentre trial. *J Affect Disord.* 91;22:119-24.

Norton, Sireling, Bhat, Rao, &Paykel. A double-blind comparison of fluvoxamine, imipramine and placebo in depressed patients. *J Affect Disord.* 84;7:297-308.

Ohrberg, Christiansen, Severin, Calberg, Nilakantan, Borup, Sogaard, Larsen, Loldrup, Bahr, &et. Paroxetine and imipramine in the treatment of depressive patients in psychiatric practice. *Acta Psychiatr Scand.* 92;86:437-44.

Ontiveros. A double-blind comparative study of paroxetine and fluoxetine in out-patients with depression. *British Journal of Clinical Research.* 97;8:23-32.

Ottevanger. The efficacy of fluvoxamine in patients with severe depression. *British Journal of Clinical Research.* 91;2:125-132.

Ottevanger. Fluvoxamine and clomipramine in depressed hospitaiised patients: results from a randomised, double-blind study. *Encephale.* 95;21:317-21.

Patris, Bouchard, Bougerol, Charbonnier, Chevalier, Clerc, Cyran, Van Amerongen, Lemming, &Hopfner Petersen. Citalopram versus fluoxetine: a double-blind, controlled, multicentre, phase III trial in patients with unipolar major depression treated in general practice. *Int Clin Psychopharmacol.* 96;11:129-36.

Perez&Ashford. A double-blind, randomized comparison of fluvoxamine with mianserin in depressive illness. *Curr Med Res Opin.* 90;12:234-41.

Perez, Gilaberte, Faries, Alvarez, &Artigas. Randomised, double-blind, placebo-controlled trial of pindolol in combination with fluoxetine antidepressant treatment. *Lancet.* 97;349:1594-7.

Perry. Pharmacotherapy for major depression with melancholic features: relative efficacy of tricyclic versus selective serotonin reuptake inhibitor antidepressants. *J Affect Disord.* 96;39:1-6.

Perry, Garvey, Kelly, Cook, Dunner, &Winokur. A comparative trial of fluoxetine versus trazodone in outpatients with major depression. *J Clin Psychiatry.* 89;50:290-4.

Peselow, Filippi, Goodnick, Barouche, &Fieve. The short- and long-term efficacy of paroxetine HCl: A. Data from a 6- week double-blind parallel design trial vs. imipramine and placebo. *Psychopharmacol Bull.* 89;25:267-71.

Poelinger&Haber. Fluoxetine 40 mg vs maprotiline 75 mg in the treatment of out-patients with depressive disorders. *Int Clin Psychopharmacol.* 89;4 Suppl 1:47-50.

Poirier&Boyer. Venlafaxine and paroxetine in treatment-resistant depression. Double-blind, randomised comparison. *Br J Psychiatry.* 99;175:12-6.

Rapaport, Coccaro, Sheline, Perse, Holland, Fabre, &Bradford. A comparison of fluvoxamine and fluoxetine in the treatment of major depression. *J Clin Psychopharmacol.* 96;16:373-8.

Rapaport&Judd. Minor depressive disorder and subsyndromal depressive symptoms: functional impairment and response to treatment. *J Affect Disord.* 98;48:227-32.

Ravindram, Teehan, Bakish, Yatham, O-Reilly, Fernando, Manchanda, Charbonneau, &Buttars. The impact of setraline, desipramine, and placebo on psychomotor functioning in depression. *Human Psychopharmacology.* 95;10:273-281.

Ravindran, Judge, Hunter, Bray, &Morton. A double-blind, multicenter study in primary care comparing paroxetine and clomipramine in patients with depression and associated anxiety. Paroxetine Study Group. *J Clin Psychiatry.* 97;58:112-8.

Reimherr, Amsterdam, Quitkin, Rosenbaum, Fava, Zajecka, Beasley, Michelson, Roback, &Sundell. Optimal length of continuation therapy in depression: a prospective assessment during long-term fluoxetine treatment. *Am J Psychiatry.* 98;155:1247-53.

Reimherr, Byerley, Ward, Lebegue, &Wender. Sertraline, a selective inhibitor of serotonin uptake, for the treatment of outpatients with major depressive disorder. *Psychopharmacol Bull.* 88;24:200-5.

Reimherr, Chouinard, Cohn, Cole, Itil, LaPierre, Masco, &Mendels. Antidepressant efficacy of sertraline: a double-blind, placebo- and amitriptyline-controlled, multicenter comparison study in outpatients with major depression. *J Clin Psychiatry.* 90;51 Suppl B:18-27.

Remick, Campos, Keller, Misri, Loomer, &Westin. A comparison of the safety and efficacy of zimeldine amd maprotiline in depressed outpatients. *Current Therapeutic Research.* 84;36:77-85.

Remick, Claman, Reesal, Gibson, &Carter. Comparison of fluoxetine and desipramine in depressed outpatients. *Current Therapeutic Research.* 53:457-465.

Remick, Keller, Gibson, &Carter. A comparison between fluoxetine and doxepin in depressed patients. *Current Therapeutic Research.* 46:842-848.

Remick, Reesal, Allen, Claman, Ramirez, Perry, &Kelle. Comparison of fluoxamine and amitriptyline in depressed outpatients. *Current Therapeutic Research.* 94;55:243-250.

Reynaert, Parent, Mirel, Janne, &Haazen. Moclobemide versus fluoxetine for a major depressive episode. *Psychopharmacology (Berl).* 95;118:183-7.

Rickels, Amsterdam, Clary, Fox, Schweizer, &Weise. The efficacy and safety of paroxetine compared with placebo in outpatients with major depression. *J Clin Psychiatry.* 92;53 Suppl:30-2.

Rickels, Amsterdam, Clary, Fox, Schweizer, &Weise. A placebo-controlled, double-blind, clinical trial of paroxetine in depressed outpatients. *Acta-Psychiatr-Scand-Suppl.* 89;350:117-23.

Rickels, Amsterdam, &Avallone. Fluoxetine in major depression: a controlled study. *Current Therapeutic Research.* 86;39:559-563.

Robertson, Abou-Saleh, Harrison, Nairac, Edwards, Lock, Burns, &Katona. A double-blind controlled comparison of fluoxetine and lofepramine in major depressive illness. *Journal of Psychopharmacology.* 94;8:98-103.

Robinson, Schultz, Castillo, Kopel, Kosier, Newman, Curdue, Petracca, &Starkstein. Nortriptyline versus fluoxetine in the treatment of depression and in short term recovery after stroke: A placebo controlled, double blind study. *American-Journal-of-Psychiatry.* 2000;157:351-359.

Ropert. Fluoxetine versus clomipramine in major depressive disorders. *Int Clin Psychopharmacol.* 89;4 Suppl 1:89-95.

Rosenberg, Damsbo, Fuglum, Jacobsen, &Horsgard. Citalopram and imipramine in the treatment of depressive patients in general practice. A Nordic multicentre clinical study. *Int Clin Psychopharmacol.* 94;9 Suppl 1:41-8.

Roth, Mattes, Sheehan, &Sheehan. A double-blind comparison of fluvoxamine, desipramine and placebo in outpatients with depression. *Prog Neuropsychopharmacol Biol Psychiatry.* 90;14:929-39.

Rudolph&Feiger. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord.* 99;56:171-81.

Rudorfer&Potter. Antidepressants. A comparative review of the clinical pharmacology and therapeutic use of the 'newer' versus the 'older' drugs. *Drugs.* 89;37:713-38.

Saletu, Grunberger, Anderer, Linzmayer, &Zyhlarz. Comparative pharmacodynamic studies with the novel serotonin uptake- enhancing tianeptine and -inhibiting fluvoxamine utilizing EEG mapping and psychometry. *J Neural Transm.* 96;103:191-216.

Schlake, Kuhs, Rolf, Bosse, Schuhknecht, Rudolf, &Brune. Platelet 5-HT transport in depressed patients under double-blind treatment with paroxetine versus amitriptyline. *Acta Psychiatr Scand Suppl.* 89;350:149-51.

Shaw, Thomas, Briscoe, Watkins, Crimmins, Harris, Lovett, Raj, Lloyd, Osborne, &et. A comparison of the antidepressant action of citalopram and amitriptyline. *Br J Psychiatry.* 86;149:515-7.

Shrivastava, Shrivastava, Overweg, &Blumhardt. A double-blind comparison of paroxetine, imipramine, and placebo in major depression. *J Clin Psychiatry.* 92;53 Suppl:48-51.

Silverstone&Ravindran. Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. Venlafaxine XR 360 Study Group. *J Clin Psychiatry.* 99;60:22-8.

Smeraldi. Amisulpride versus fluoxetine in patients with dysthymia or major depression in partial remission: a double-blind, comparative study. *J Affect Disord.* 98;48:47-56.

Smith&Glaudin. A placebo-controlled trial of paroxetine in the treatment of major depression. *J Clin Psychiatry.* 92;53 Suppl:36-9.

Sonawalla, Spillmann, Kolsky, Alpert, Nierenberg, Rosenbaum, &Fava. Efficacy of fluvoxamine in the treatment of major depression with comorbid anxiety disorders. *J Clin Psychiatry.* 99;60:580-3.

Song, Freemantle, Sheldon, House, Watson, Long, &Mason. Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability. *BMJ.* 93;306:683-7.

Stage, Bech, Gram, Kragh-Sorensen, Rosenberg, &Ohrberg. Are in-patient depressives more often of the melancholic subtype? Danish University Antidepressant Group. *Acta Psychiatr Scand.* 98;98:432-6.

Stark&Hardison. A review of multicenter controlled studies of fluoxetine vs. imipramine and placebo in outpatients with major depressive disorder. *J Clin Psychiatry.* 85;46:53-8.

Stott, Blagden, &Aitken. Depression and associated anxiety in primary care: A double-blind comparison of paroxetine and amotriptyline. *European Neuropsychopharmacology.* 93;3:324-325.

Stratta, Bolino, Cupillari, &Casacchia. A double-blind parallel study comparing fluoxetine with imipramine in the treatment of atypical depression. *Int Clin Psychopharmacol.* 91;6:193-6.

Stuppaeck, Geretsegger, Whitworth, Schubert, Platz, Konig, Hinterhuber, &Fleischhacker. A multicenter double-blind trial of paroxetine versus amitriptyline in depressed inpatients. *J Clin Psychopharmacol.* 94;14:241-6.

Tamminen&Lehtinen. A double-blind parallel study to compare fluoxetine with doxepin in the treatment of major depressive disorders. *Int Clin Psychopharmacol.* 89;4 Suppl 1:51-6.

Tan&Levin. Citalopram in the treatment of depression and other potential uses in psychiatry. *Pharmacotherapy.* 99;19:675-89.

Taneri&Kohler. Fluoxetine versus nomifensine in outpatients with neurotic or reactive depressive disorder. *Int Clin Psychopharmacol.* 89;4 Suppl 1:57-61.

Tapani&Lehtinen. A double-blind parallel study to compare fluoxetine with doxepin in the treatment of major depressive disorders. *International Journal of Psychopharmacology.* 89;4:51-56.

Tignol. A double-blind, randomized, fluoxetine-controlled, multicenter study of paroxetine in the treatment of depression. *J Clin Psychopharmacol.* 93;13:18S-22S.

Tignol, Stoker, &Dunbar. Paroxetine in the treatment of melancholia and severe depression. *Int Clin Psychopharmacol.* 92;7:91-4.

Timmerman, de Beurs, Tan, Leijnse-Ybema, Sanchez, Hopfner Petersen, &Cohen Stuart. A double-blind comparative clinical trial of citalopram vs maprotiline in hospitalized depressed patients. *Int Clin Psychopharmacol.* 87;2:239-53.

Tollefson, Greist, Jefferson, Heiligenstein, Sayler, Tollefson, &Koback. Is baseline agitation a relative contraindication for a selective serotonin reuptake inhibitor: a comparative trial of fluoxetine versus imipramine. *J Clin Psychopharmacol.* 94;14:385-91.

Tollefson, Rampey, Beasley, Enas, &Potvin. Absence of a relationship between adverse events and suicidality during pharmacotherapy for depression. *J Clin Psychopharmacol.* 94;14:163-9.

Tollefson&Sayler. Course of psychomotor agitation during pharmacotherapy of depression: Analysis from double-blind controlled trials with fluoxetine. *Depression and Anxiety.* 96-97;4:294-311.

Tollefson, Holman, Sayler, &Potvin. Fluoxetine, placebo, and tricyclic antidepressants in major depression with and without anxious features. *Journal-of-Clinical-Psychiatry.* 94;55:50-59.

Tome, Cloninger, Watson, &Isaac. Serotonergic autoreceptor blockade in the reduction of antidepressant latency: personality variables and response to paroxetine and pindolol. *J Affect Disord.* 97;44:101-9.

Tome&Isaac. Cost-benefit & cost-effectiveness analysis of the rapid onset of selective serotonin reuptake inhibitors by augmentation. *Int J Psychiatry Med.* 97;27:377-90.

Tome&Isaac. Cost effectiveness study of a year follow-up of selective serotonin reuptake inhibitor (SSRI) and augmentor combination compared with SSRI and placebo. *Int Clin Psychopharmacol.* 98;13:175-82.

Tome&Isaac. One year real world prospective follow-up study of a major depressive episode of patients treated with paroxetine and pindolol or paroxetine for 6 weeks. *Int Clin Psychopharmacol.* 98;13:169-74.

Tome, Isaac, Harte, &Holland. Paroxetine and pindolol: a randomized trial of serotonergic autoreceptor blockade in the reduction of antidepressant latency. *Int Clin Psychopharmacol.* 97;12:81-9.

Trindade, Menon, Topfer, &Coloma. Adverse effects associated with selective serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis. *CMAJ.* 98;159:1245-52.

Upward, Edwards, Goldie, &Waller. Comparative effects of fluoxetine and amitriptyline on cardiac function. *Br J Clin Pharmacol.* 88;26:399-402.

Van Moffaert, Bartholome, &Cosyns. A controlled comparison of sertraline and fluoxetine in acute and continuation treatment of major depression. *Human Psychopharmacology.* 95;10:393-405.

Versiani, Ontiveros, Mazzotti, Ospina, Davila, Mata, Pacheco, Plewes, Tamura, &Palacios. Fluoxetine versus amitriptyline in the treatment of major depression with associated anxiety (anxious depression): A double blind comparison. *International-Clinical-Psychopharmacology.* 99;14:321-327.

Walczak, Apter, Halikas, Borison, Carman, Post, Patrick, Cohn, Cunningham, Rittberg, Preskorn, Kang, &Wilcox. The oral dose-effect relationship for fluvoxamine: a fixed-dose comparison against placebo in depressed outpatients. *Ann-Clin-Psychiatry.* 96;8:139-51.

Waldinger, Hengeveld, Zwinderman, &Olivier. Effect of SSRI antidepressants on ejaculation: a double-blind, randomized, placebo-controlled study with fluoxetine, fluvoxamine, paroxetine, and sertraline. *J Clin Psychopharmacol.* 98;18:274-81.

Ware. Fluvoxamine: a review of the controlled trials in depression. *J Clin Psychiatry.* 97;58 Suppl 5:15-23.

Weilburg, Rosenbaum, Biederman, Sachs, Pollack, &Kelly. Fluoxetine added to non-MAOI antidepressants converts nonresponders to responders: a preliminary report. *J Clin Psychiatry.* 89;50:447-9.

Wernicke, Bosomworth, &Ashbrook. Fluoxetine at 20 mg per day: The recommended and therapeutic dose in the treatment of depression. *International-Clinical-Psychopharmacology.* 89;4:63-67.

Wernicke, Dunlop, Dornseif, Bosomworth, &Humbert. Low-dose fluoxetine therapy for depression. *Psychopharmacol Bull.* 88;24:183-8.

Wernicke, Dunlop, Dornseif, &Zerbe. Fixed-dose fluoxetine therapy for depression. *Psychopharmacol Bull.* 87;23:164-8.

Wheatley, van Moffaert, Timmerman, &Kremer. Mirtazapine: efficacy and tolerability in comparison with fluoxetine in patients with moderate to severe major depressive disorder. Mirtazapine- Fluoxetine Study Group. *J Clin Psychiatry.* 98;59:306-12.

Williams, Edwards, Newburn, Mullen, Menkes, &Segkar. A double-blind comparison of moclobemide and fluoxetine in the treatment of depressive disorders. *Int Clin Psychopharmacol.* 93;7:155-8.

Young, Coleman, &Lader. A controlled comparison of fluoxetine and amitriptyline in depressed out-patients. *Br J Psychiatry.* 87;151:337-40.

Zanardi, Franchini, Gasperini, Perez, &Smeraldi. Double-blind controlled trial of sertraline versus paroxetine in the treatment of delusional depression. *Am J Psychiatry.* 96;153:1631-3.

APPENDIX III INCLUSION CRITERIA

Exclusion/Inclusion Criteria for articles for meta-analysis of RCT for drugs for depression/placebo wash-out

Authors _____
Title _____
Journal _____

		Notes
Population 18-65 years (human)	☐	
DSM-III or DSM-IV diagnosis of major depression (acute)	☐	
No concomitant psychiatric diagnosis (anxiety permitted)	☐	
No concomitant medical illness	☐	
Randomized, controlled, double-blinded clinical trial	☐	
Placebo (inert) arm	☐	
SSRI antidepressant	☐	
Single drug (no augmentation)	☐	
Duration of study at least 6 weeks	☐	
Acute treatment (not continuation)	☐	
Parallel design (no crossover)	☐	
No active drug wash-out period	☐	
Outcome (minimum mean change/ continuous outcome)	☐	
INCLUDE (all criteria met)	☐	

APPENDIX IV JADAD SCORE CRITERIA

JADAD SCORE

Authors _____

Title _____

Journal _____

	Notes
Was the study described as RANDOMIZED?	☐
Was the study described as DOUBLE-BLIND?	☐
Was there a description of WITHDRAWALS and DROPOUTS?	☐
+/- 1 point for appropriateness of RANDOMIZATION	☐
+/- 1 point for appropriateness of DOUBLE-BLINDING	☐

TOTAL SCORE

APPENDIX V DATA EXTRACTION FORM

Study Characteristics to enter

Author: _____

Title: _____

Journal: _____

	Methods
Duration	
JADAD Score (1-5)	
Primary Measurement	

	Participants
Number of Participants	
Settings OP/IP	

	Interventions	**Dosage of Drug**
Active Drug vs Placebo		

	Concealment
A Adequate	
B Unclear	
C Inadequate	
D Not Used	

Outcomes

	Treatment Group
N - Size of Study on the Drug	
Mean - End point minus baseline change	
Standard Deviation	

	Placebo Group
N - Size of Study on the Drug	
Mean - End point minus baseline change	
Standard Deviation	

	PRI
PRI phase (yes or no)	

APPENDIX VI MASTER LIST OF ALL INCLUDED TRIALS/DATA

Beasley 1991	6 weeks JADAD 3 HAMD	455 Participants OP	Fluoxetine 60 mg	233	11.8	9.7	222	7.3	9	Yes	A
Byerley 1988	6 weeks JADAD 5 HAMD-21	61 Participants OP	Fluoxetine 80 mg	32	14.4	7.7	29	7.6	6.5	Yes	A
Claghorn & Earl 1996	6 weeks JADAD 3 HAMD-21	89 Participants OP	Fluvoxamine 150 mg	44	6.58	8.49	45	10.11	8.49	Yes	B
Claghorn & Kiev 1992	6 weeks JADAD 3 HAMD-21	325 Participants OP	Paroxetine 50 mg	163	12.9	8.49	162	8.6	8.49	Yes	B
Claghorn 1992	6 weeks JADAD 3 HAMD-21	58 Participants OP	Paroxetine 50 mg	32	11.5	8.49	26	5.5	8.49	Yes	B
Claghorn 1992tex	6 weeks JADAD 3 HAMD-17	59 Participants OP	Paroxetine 30 mg	32	11.44	8.32	27	5.49	8.31	No	B
Cohn & Crowder 1990	6 weeks JADAD 3 HAMD	71 Participants OP	Paroxetine 50 mg	35	9	8.75	36	6	8.75	No	B
Cohn & Wilcox 1985	6 weeks JADAD 4 HAMD-21	112 Participants OP	Fluoxetine 80 mg	54	14.26	8.49	57	4.05	8.49	Yes	B
Cohn & Wilcox 1992	6 weeks JADAD 3 HAMD	69 Participants OP	Paroxetine 50 mg	35	11.5	8.75	34	6.8	8.75	Yes	B
Corrigan 2000	8 weeks JADAD 3 HAMD-17	70 Participants OP	Fluoxetine 20 mg	35	11	8.75	35	9.13	8.75	No	B
Croft 1999	8 weeks JADAD 3 HAMD-31	240 Participants OP	Sertraline 200 mg	119	18	12.53	121	15	12.53	No	B
Doogan1994	6 weeks JADAD 5 MADRS	173 Participants OP	Setraline 100 mg	83	15.4	8.34	90	12.1	8.34	Yes	A

Study ID	Methods	Participants	Interventions	SSRI N	SSRI Mean	SSRI SD	Placebo N	Placebo Mean	Placebo SD	PRI	Allocation concealment
Dunbar 1991	6 weeks JADAD 3 HAMD	419 Participants OP	Paroxetine 50 mg	207	10.1	8.75	212	5.7	8.75	Yes	B
Dunlop 1990	6 weeks JADAD 3 HAMD	65 Participants OP	Fluoxetine 60 mg	97	5.4	5.6	56	5.8	5.9	Yes	B
Edwards 1993	6 weeks JADAD 4 HAMD-17	41 Participants OP	Paroxetine 30 mg	21	14.6	5.64	20	10.5	6.62	No	B
Fabre & Abuzzahab 1995	6 weeks JADAD 4 HAMD-17	288 Participants OP	Sertraline 200 mg	261	9.9	8.6	86	7.6	7.5	No	B
Fabre & Birkhimer 1996	8 weeks JADAD 3 HAMD-21	90 Participants OP	Fluvoxamine 150 mg	46	13.5	8.49	44	8.5	8.49	Yes	B
Fabre & Putman 1987	6 weeks JADAD 3 HAMD	8 Participants OP	Fluoxetine 60 mg	6	17.17	8.75	2	1	8.75	Yes	B
Fabre 1992	6 weeks JADAD 3 HAMD-21	74 Participants OP	Paroxetine 50 mg	38	9.13	1.32	36	3.06	1.35	Yes	B
Fava 1998	12 weeks JADAD 3 HAMD-21	74 Participants OP	Paroxetine 55 mg	55	11.1	9.4	19	11.6	8.9	No	B
Fava 1998b	12 weeks JADAD 3 HAMD-21	35 Participants OP	Fluoxetine 80 mg	54	10.8	9.5	19	11.6	8.9	No	B
Feighner & Boyer 1989	6 weeks JADAD 3 HAMD-21	100 Participants OP	Fluoxetine 80 mg	52	8.8	8.49	48	6	8.49	No	B
Feighner & Cohn 1993	6 weeks JADAD 4 HAMD	480 Participants OP	Paroxetine 50 mg	240	10	8.75	240	5.8	8.75	Yes	B
Feighner & Overo 1999	6 weeks JADAD 4 HAMD-21	650 Participants OP	Citralopram 60 mg	521	11.2	8.49	129	9.3	8.49	Yes	B
Feighner 1989	6 weeks JADAD 3 HAMD	33 Participants IP	Fluvoxamine 300 mg	21	14	8.75	12	8	8.75	No	B

Study ID	Methods	Participants	Interventions	SSRI N	SSRI Mean	SSRI SD	Placebo N	Placebo Mean	Placebo SD	PRI	Allocation concealment
Heiligenstein 1993	8 weeks JADAD 3 MADRS	51 Participants OP	Fluoxetine 20 mg	23	14.61	11.36	28	7.18	10.25	Yes	B
Kiev 1992	6 weeks JADAD 4 HAMD-21	66 Participants OP	Paroxetine 50 mg	34	13	8.49	32	7	8.49	No	B
Lapierre 1987	6 weeks JADAD 4 HAMD-17	42 Participants IP	Fluvoxamine 300 mg	22	19	6.94	20	3	6.94	No	B
Lydiard & Liaird 1989	6 weeks JADAD 3 HAMD	34 Participants OP	Fluvoxamine 300 mg	17	11.7	8.75	17	10.2	8.75	No	B
Lydiard & Stahl 1997	8 weeks JADAD 4 HAMD-17	261 Participants OP	Sertraline 200 mg	119	11.1	6.87	115	8.8	6.97	Yes	B
March 1990	8 weeks JADAD 5 HAMD-17	25 Participants OP	Fluvoxamine 300 mg	13	13	6.94	12	9.5	6.94	No	A
Peselow 1989	6 weeks JADAD 4 HAMD-21	82 Participants OP	Paroxetine 50 mg	40	10.05	8.49	42	4.07	8.49	Yes	B
Reimherr & Byerley 1988	8 weeks JADAD 3 HAMD-17	52 Participants OP	Sertraline 200 mg	26	11.5	7.6	26	7.6	8.8	Yes	B
Reimherr & Chouinard 1990	8 weeks JADAD 3 HAMD	283 Participants OP	Sertraline 200 mg	142	11.66	8.24	141	8.16	7.85	Yes	B
Roth 1992	6 weeks JADAD 4 HAMD	56 Participants OP	Fluvoxamine 300 mg	21	11.1	9	29	8.4	9.4	Yes	B
Rudolph 1999	8 weeks JADAD 4 HAMD-21	200 Participants OP	Fluoxetine 60 mg	103	11.8	8.29	97	10.2	8.04	Yes	A
Shrivastava 1992	6 weeks JADAD 3 HAMD-21	69 Participants OP	Paroxetine 50 mg	33	12	8.49	36	7.1	8.49	Yes	B
Silverstone 1999	12 weeks JADAD 4 HAMD-21	237 Participants OP	Fluoxetine 60 mg	119	15.2	8.49	118	11.1	8.49	Yes	B

Study ID	Methods	Participants	Interventions	SSRI N	SSRI Mean	SSRI SD	Placebo N	Placebo Mean	Placebo SD	PRI	Allocation concealment
Smith 1992	6 weeks JADAD 3 HAMD	66 Participants OP	Paroxetine 50 mg	33	11	8.75	33	8	8.75	No	B
Stark 1985	6 weeks JADAD 3 HAMD	354 Participants OP	Fluoxetine 80 mg	185	11	10.1	169	8.2	9	Yes	B
Walczak 1996	6 weeks JADAD 3 HAMD-13	299 Participants OP	Fluvoxamine 150 mg	99	7	5.31	200	6	5.31	Yes	B
Wernicke 1987	6 weeks JADAD 3 HAMD	169 Participants OP	Fluoxetine 40 mg	92	11.2	8.2	77	7	8.6	Yes	B
Wernicke 1988	6 weeks JADAD 3 HAMD	151 Participants OP	Fluoxetine 60 mg	103	7.2	9.1	48	5.7	8.6	Yes	B

REFERENCES

American Psychiatric Association. *DSM-III: Diagnostic and statistical manual of mental disorders.* 3rd ed. Washington: The American Psychiatric Association, 1980.

American Psychiatric Association. *DSM-IV: Diagnostic and statistical manual of mental disorders.* 4th ed. Washington: The American Psychiatric Association, 1994.

Arana GW, Baldessarini RJ, Ornesteen M. The dexamethasone supression test for diagnosis and prognosis in psychiatry: commentary and review. *Arch Gen Psychiatry* 1985;42:1193-1204.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561-571.

Begg C, Cho M, Eastwood S. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996;276:637-639.

Beecher HK. The powerful placebo. *JAMA* 1955;159:1602-1606.

Beecher HK, Keats AS, Mosteller F, Lasagna L. The effectiveness of oral analgesics (morphine, codeine, acetylsalicylic acid) and the problem of placebo "reactors" and "non-reactors". *J Pharmacol Exp Therap* 1953;109:393-400.

Benfield P, Hill RC, Lewis SP. Fluoxetine: a review of its pharmacodynamic and pharmacokinetic properties and therapeutic efficacy in depressive illness. *Drugs* 1986;32:481-508.

Benfield P, Ward A. Fluvoxamine: a review of its pharmacodynamic and pharmacokinetic properties and therapeutic efficacy in depressive illness. *Drugs* 1986;32:313-334.

Bero L, Rennie D. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *JAMA* 1995;274:1935-1938.

Bielsky RJ, Friedel RO. Prediction of tricyclic antidepressant response. *Arch Gen Psychiatry* 1978;33:1479-1489.

Brody H. Placebo effect: an examination of Grunbaum's definition. In: White L, Tursky B, Schwartz GE, eds. *Placebo: theory, research, and mechanisms.* New York: Guilford Press, 1985.

Brown WA. Predictors of placebo response in depression. *Psychopharmacol Bull* 1988;24:14-17.

Brown WA, Dornseif BE, Wernicke JF. Placebo response in depression: A search for predictors. *Psychiatry Res* 1988;26:259-264.

Brown WA, Shrivastava RK, Arato M. Pre-treatment in pituitary-adrenocortical status and placebo response in depression. *Psychopharmacol Bull* 1987;23:155-159.

Byerley WF, Reimherr FW, Wood DR, Grosser BI. Fluoxetine, a selective serotonin uptake inhibitor for the treatment of outpatients with major depression. *J Clin Psychopharmacol* 1988;8:112-115.

Carroll BJ, Feinberg M, Smouse PE, Rawson SG, Grechen JF. The Carroll Rating Scale for Depression I: development, reliability, and validity. *Br J Psychiatry* 1981;138:194-200.

Chalmers I, Haynes RB. Reporting, updating and correcting systematic reviews of the effects of health care. In: Chalmers I, Altman DG, editors. *Systematic reviews.* London: BMJ Publishing Group, 1995:86-95.

Coble PA, Kupfer DJ, Spiker DG, Neil JF, McPartland RJ. EEG sleep in primary depression: a longitudinal placebo study. *J Affect Disord* 1979;1:131-138.

Cook DJ, Sackett DL, Spitzer W. Methodologic guidelines for systematic reviews of randomized controlled trials in health care from the Potsdam consultation on meta-analysis. *J Clin Epidemiol 1995*;48:167-171.

Cooper H, Hedges LV, eds. *The handbook of research synthesis.* New York: Russell Sage Foundation, 1994.

Coryell W, Turner R. Outcome with desipramine therapy in subtypes of nonpsychotic neurotic depression. *J Affect Disord* 1985;9:149-154.

Crout JR, Finkel MJ. *Guidelines for the clinical evaluation of antidepressant drugs.* (DHEW Publication #77-3042) Rockville: U.S. Food and Drug Administration, 1977.

Dechant KL, Clissold SP. Paroxetine: a review of its pharmacodynamic and pharmacokinetic properties and therapeutic potential in depressive illness. *Drugs* 1991;41: 225-253.

Derogatis LR. The SCL-90-R: administration, scoring and procedures manual-I. *Clinical Psychomotor Research*, Baltimore 1977.

Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974;19:1-15.

Downing RW, Rickels K. Predictors of response to amitriptyline and placebo in three outpatient treatment settings. *J Nerv Ment Dis* 1973;156:109-129.

Eddy DM, Hasselblad V, Schacter R. A Bayesian method for synthesizing evidence: the confidence profile method. *Int J Technol Assess Health Care* 1990;6:31-35.

Fairchild CJ, Rush J, Vasavada N, Giles D, Khatami M. Which depressions respond to placebo? *Psychiatry Res* 1986;18:217-226.

Gaddum JH. Walter Ernest Dixon memorial lecture. *Clin Pharmacology* 1954;47:195-204.

Geddes JR, Freemantle N, Mason J, Eccles P, Boynton J. SSRIs versus other antidepressants for depressive disorder (Cochrane Review). In: *The Cochrane Library*,Issue 4,2000. Oxford: Update software.

Glass GV. Primary, secondary and meta-analysis. *Educ Res* 1976;5:3-8.

Glass GV. Integrating findings: the meta-analysis of research. *Rev Educ* 1977;5:351-379.

Gliedman LH, Nash EH Jr, Imber SD, Stone AR, Frank JD. Reduction of symptoms by pharmacologically inert substances and by short-term psychotherapy. *Arch Neurol Psychiatry* 1958;79:345-351.

Greenberg RP, Bornstein RF, Greenberg MD, Fisher S. A meta-analysis of antidepressant outcome under "blinder conditions". *J Consult Clin Psychology* 1992;60:664-669.

Greenberg RP, Fisher S, Riter JA. Placebo washout if not a meaningful part of antidepressant drug trials. *Percept Mot Skills* 1995;81:688-690.

Guy W. *ECDEU assessment manual for psychopharmacology.* Washington: Superintendent of Documents, U.S. Government Printing Office, U.S. Department of Health, Education and Welfare Publications, Washington, DC, No. 76-338.

Hamilton M. A rating scale for depression. *J Neurol Neurosurg and Psychiatry* 1960;23:56-62.

Hassard TH. *Understanding biostatistics.* St.Louis: Mosby Year Book, 1991.

Hassard TH, director of biostatistics unit, Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba.

Henry JA. Toxicity of antidepressants. *Int Clin Psychopharm* 1992;6(suppl):22-27.

Holsboer F, Liebl R, Hofshuster E. Repeated dexamenthasone suppression test during depressive illness: normalization of test result compared with clinical improvement. *J Affect Disord* 1982;4:93-101.

Horvath AO, Symonds BD. Relationship between working alliance and outcome in psychotherapy: a meta-analysis. *J Counsel Psychol* 1991;38:139-149.

Jellinek EM. Clinical tests on comparative effectiveness of analgesic drugs. *Biometric Bull* 1946;2:87-91.

Joyce PR. Neuroendocrine changes in depression. *Aust NZ J Psychiatry* 1985;19:120-127.

Joyce PR, Paykel ES. Predictor of drug response in depression. *Arch Gen Psychiatry* 1989;46:89-99.

Kiloh LC, Ball JRB, Garside RF. Prognostic factors in the treatment of depressive states with imipramine. *BMJ* 1962;1:1225-1227.

Klerman GL, Cole JO. Clinical pharmacology of imipramine and related antidepressant compounds. *Pharmacol Rev* 1965;17:101-104.

Landen M, Bjorling G, Agren H, Fahlen T. A randomized, double-blind, placebo-controlled trial of buspirone in combination with an SSRI in patients with treatment-refractory depression. *J Clin Psychiatry* 1998;59:664-668.

Lasagna L, Mosteler F, von Felsinger JM, Beecher HK. A study of the placebo response. *Am J Med* 1954;16:770-779.

Lemberger L, Fuller RW, Zerbe RL. Use of specific serotonin uptake inhibitors as antidepressants. *Clin Neuropharmacol* 1985;8:299-317.

Loebel AD, Hyde TS, Dunner DL. Early placebo response in anxious and depressed patients. *J Clin Psychiatry* 1986;47:230-233.

Marzialli E, Marmar C, Krupnick J. Therapeutic alliance scales: development and relationship to psychotherapy outcome. *Am J Psychiatry* 1981;138:361-364.

Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials. *Control Clin Trials* 1995;16:62-73.

Moher D, Schulz KF, Altman D (for the CONSORT group). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987-1990.

Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382-389.

Montgomery SA, Henry J, McDonald G, Dinan T, Lader M, Hindmarch I et al. Selective serotonin reuptake inhibitors: meta-analysis of discontinuation rates. *Int Clin Psychopharm* 1994;9:47-53.

Murdoch D, McTavish D. Sertraline: a review of its pharmacodynamic and pharmacokinetic properties and therapeutic potential in depression and obsessive-compulsive disorder. *Drugs* 1992;44:604-624.

Oxman A. The Cochrane Collaboration, 1995.

Pablos-Mendez A, Barr G, Shea S. Run-in periods in randomized trials: implications for the application of results in clinical practice. *JAMA* 1998;279:222-225.

Paykel ES. Depressive typologies and response to amitriptyline. *Br J Psychiatry* 1972;120:147-156.

Pepper OHP. A note of placebo. *Ann J Pharm* 1945;117:409-412.

Peselow ED, Loutin A, Wolkin A, Rohrs C, Novatt A, Siekierski J, Retrosen J. The dexamethasone suppression test and response to placebo. *J Clin Psychopharmacology* 1986;6:286-291.

Quitkin FM, Rabkin JD, Markowitz JM, Stewart JW, McGrath PJ, Harrison W. Use of pattern analysis to identify true drug response. *Arch Gen Psychiatry* 1987;44:259-264.

Quitkin FM, Rabkin J, Ross D, Stewart JW. Identification of true drug response to antidepressants. *Arch Gen Psychiatry* 1984;41:782-786.

Rabkin JG, McGrath P, Stewart JW, Harrison W, Markowitz JS, Quitkin FM. Follow-up of patients who improved during placebo washout. *J Clin Psychopharmacol* 1986;6:274-278.

Rabkin JG, Stewart JW, McGrath PJ, Markowitz JS, Harrison W, Quitkin FM. Baseline characteristics of 10-day placebo washout responders in antidepressant trials. *Psychiatry Res* 1987;21:9-22.

Raskin A, Crook TH. The endogenous neurotic distinction as a predictor of response to antidepressant drugs. *Psychol Med* 1976;6:59-70.

Raskin A, Schulterbrandt J, Reating N, Rice CE. Factors of psychopathology in interview, ward behaviors and self-report ratings of hospitalized depressives. *J Consult Psychology* 1967;*31:270-278.*

Reimherr FW, Ward MF, Byerley WF. The introductory placebo washout: a retrospective evaluation. *Psychiatry Res* 1989;30:191-199.

Revman 4.1 User Guide. The Cochrane Collaboration, June 27, 2000.

Rogers SC, Clay PM. A statistical review of controlled trials of imipramine and placebo in the treatment of depressive illness. *Br J Psychiatry* 1975;127:599-603.

Schweizer E, Rickels K. Placebo response in generalized anxiety: its effect on the outcome of clinical trials. *J Clin Psychiatry* 1997;58 (Suppl 11):30-38.

Senn S. Are placebo run ins justified? *BMJ* 1997;314:1191-1193.

Shapiro AK. *The placebo response: modern perspectives in world psychiatry.* Vol.2. Edinburgh: Oliver and Boyd, 1971.

Skovlund E. Should the single-blind placebo responder exclusion phase be omitted from clinical trials? *Nordic J Psychiatry* 1994;48:159-165.

Stewart JW, Quitkin FM, Liebowitz M, McGrath PJ, Harrison WM, Klein DF. Efficacy of desipramine in depressed outpatients: response to Research Diagnostic Criteria diagnosis and severity of illness. *Arch Gen Psychiatry* 1983;40:202-207.

Sommers-Flanagan J, Sommers-Flanagan R. Efficacy of antidepressant medication with depressed youth: what psychologists should know. *Pro Psychol Res Prac* 1996;26:155-153.

Spiro HM. *Doctors, patients, and placebos*. New Haven: Yale University Press, 1986.

Stewart JW, Quitkin F, Liebowitz M, McGrath PJ, Harrison WM, Klein DF. Efficacy of desipramine in depressed outpatients: Response according to Research Diagnostic Criteria diagnosis and severity of illness. *Arch Gen Psychiatry* 1983;40:202-207.

Strayhorn JM. Control groups for psychosocial intervention outcome studies. *Am J Psychiatry* 1987;144:275-282.

Svendsen K, Christensen PG. Duration of REM sleep latency as predictor of effect of antidepressant therapy. *Acta Psychiatr Scand* 1981;64:238-243.

Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated? *BMJ* 1994; 309: 1351-1353.

Trivivedi MH, Rush J. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmcology* 1994;11:33-43.

Turner JA, Deyo RA, Loeser JD, von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. *JAMA* 1994;271:1609-1614.

Van Praag HM. Central monoamine metabolism in depressions. I. Serotonin and related compounds. *Compr Psychiatry* 1980;21:30-43.

Walker J, clinical psychologist, St. Boniface General Hospital, Winnipeg, Manitoba.

Wittenborn JR, Kiremitci N. A comparison of antidepressant medications in neurotic and psychotic patients. *Arch Gen Psychiatry* 1975;32:1172-1176.

Zimmerman M, Coryell W, Pfhol B. The validity of the dexamethasone suppression test as marker for endogenous depression. *Arch Gen Psychiatry* 1986;43:347-355.

Zung WWK. A self-rating depression scale. *Arch Gen Psychiatry* 1965;12:63-70.