uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Dorian Deshauer

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)

GRADE / DEGREE

Department of Epidemiology and Community Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Enrichment-Discontinuation Designs In Psychiatric Drug Maintenance Therapy

TITRE DE LA THÈSE / TITLE OF THESIS

Jeremy Grimshaw

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

David Moher

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Paul Hébert

Peter Tugwell

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Enrichment-Discontinuation Designs In Psychiatric Drug Maintenance Therapy

Dorian Deshauer

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements for the
MSc Degree

Department of Epidemiology and Community Medicine

Ottawa, Ontario, Canada

January, 2007

Your file Votre référence
ISBN: 978-0-494-49186-7
Our file Notre référence
ISBN: 978-0-494-49186-7

Enrichment-Discontinuation Designs in Psychiatric Drug Maintenance Therapy

Dorian Deshauer

Supervisor: Jeremy Grimshaw
Thesis Committee: David Moher, Dean Fergusson

Abstract:

Objective: This thesis tests the validity of a randomized controlled trial variant, the enrichment-discontinuation design (also called a randomized discontinuation design) against gold-standard classic RCT's using published trials of psychiatric drug therapy. **Methods:** A series of systematic reviews were conducted to identify all maintenance trials of mood stabilizers and antidepressants. Planned comparisons between enrichment discontinuation trials and gold standard classic RCT's were conducted. Finally, a sample of research literature was reviewed to identify the extent to which the limits of enrichment discontinuation trials were identified. **Results:** There was a non-statistically significant trend favoring drugs used in open phases vs classic RCT's. The lack of extended classic RCT's of psychiatric drugs is poorly recognized. **Conclusion:** Enrichment discontinuation designs dominated the evidence base supporting psychiatric drug maintenance. Bias cannot be ruled out. Limits to the design are poorly recognized in the literature. For further details, refer to the executive summary.

Executive Summary:

Background: The Enrichment-Discontinuation (ED) design is a randomized controlled trial (RCT) variant commonly used in psychiatric drug maintenance studies. It is efficient relative to classic RCT's, since it can demonstrate efficacy while randomizing fewer patients. ED trials only randomize patients who have already recovered on an open intervention. Efficiency comes at the cost of vulnerability to bias from drug withdrawal-rebound effects as well as bias from differential attrition favoring the open intervention. In a 1991 paper, Greenhouse et al cited these vulnerabilities as limiting the stand-alone value of ED designs in psychiatric drug maintenance trials. They proposed that psychiatric ED trials should incorporate several treatment strategies to address questions of bias. While their method has not been used at the individual trial level, it is well suited to assessing a collection of trials in which alternative treatment strategies have been used. **Objective:** The primary objective of this thesis was to test for bias in psychiatric drug maintenance ED trials using published data. A secondary objective was to identify the extent to which the unique interpretive limitations of ED designs are identified in the psychiatric literature. Method: Through a systematic literature review, psychiatric drug maintenance trials of mood stabilizers for DSM-IV bipolar disorder (lithium and anticonvulsants) and antidepressants for DSM-IV major depressive disorder (selective serotonin reuptake inhibitors (SSRI's), tricyclic antidepressants (TCA's) and monoamine oxidase inhibitors (MAOI's)) were identified. A series of comparisons between trial effect sizes under contrasting treatment strategies adapted from Greenhouse et al were constructed using meta-analysis. The identified ED trials were reviewed for discussions of design limitations, as were a selection of treatment guidelines, key textbooks and systematic reviews covering psychiatric drug maintenance. **Results:** 14 trials of mood stabilizers and 43 trials of antidepressant medications were identified for meta-analysis. Of these, 10 (71 %) of the mood stabilizer trials and 41(95 %) of the antidepressant maintenance trials used ED designs. There was evidence of bias in placebo controlled lithium maintenance trials using ED designs; lithium enriched populations OR 22.0; 95% CI 7.0-68.7 (favouring lithium) vs anticonvulsant enriched populations OR 1.9; 95% CI 1.2-2.8 (favouring lithium). This was not replicated in head to head trials of lithium vs anticonvulsants using alternate enrichment strategies. Among antidepressant trials, estimates of bias were inconsistent and were limited by a paucity of alternative designs. Methodological limitations of ED designs were discussed in approximately half of the included trials. At the level of key psychiatric texts, treatment guidelines and systematic reviews, discussions of methodological limitations focused on mood stabilizer trials for bipolar disorder, with comparatively sparse discussions of ED antidepressant trials. **Conclusion:** ED designs dominate the randomized trial literature supporting psychiatric drug maintenance therapy and the paucity of alternative treatment strategies limits the assessment of bias using the published literature. Little is known about the natural course of placebo responders with bipolar disorder or major depressive disorder. Uncertainty about differential attrition favoring open interventions and withdrawal-rebound symptoms remains unresolved. This in turn limits the confidence with which recommendations for drug maintenance can be made.

Table of Contents

Acknowledgement

Joy Albuquerque (referred to in chapter 3 as JA) assisted with data abstraction for the lithium review and for the review of ED trials in general medicine in chapter 5.

Chapter 1

Introduction

1.1 Overview

This thesis focuses on the application of a randomized controlled trial (RCT) variant, the enrichment-discontinuation (ED) design, to psychiatric drug maintenance therapies. The two-phase design, described in chapter 2, consists of an extended open run-in during which time the target population receives the study intervention (phase 1), followed by a randomization phase (phase 2) in which 'responders' to the study intervention (the study population) are randomized to ongoing phase 1 intervention, placebo, or an active competitor. In contrast, maintenance trials using the classic RCT design randomize the target population to a study intervention, placebo or active competitor and follow cases recovering under each experimental condition. Under certain assumptions (chapter 2.4a), ED trials can give valid response estimates while randomizing fewer patients than would be required using the classic RCT design. In the context of psychiatric drug maintenance trials, it is unclear whether these assumptions are met (143, 176, 197). If, for example, there is significant departure from the assumption of phase 1 non-differential attrition or if phase 2 drug withdrawal-rebound symptoms are sufficient, the design may overestimate the treatment effect (chapter 2) (197).

Introduced to psychiatry in 1970 as a 'discontinuation trial', this RCT variant has been widely adopted for testing maintenance therapy with mood stabilizing drugs (chapter 3) and antidepressants (chapter 4). In clinical practice, maintenance drug therapy is recommended

following recovery from a second or third episode of DSM-IV major depressive disorder and after a single episode of mania in DSM-IV bipolar disorder (chapter 5) (365).

The term 'maintenance' is used throughout the thesis to refer to drug therapy taken for more than six months after recovery from an acute clinical episode. Maintenance therapy often extends for a lifetime, with the purpose of preventing future clinical decompensations.

I decided to investigate the prevalence and application of ED designs in psychiatric drug maintenance trials after a systematic review of lithium therapy found that the design favored phase 1 interventions, thus raising concerns of bias (89). The thesis adapts Greenhouse's method (chapter 2.4) (143) for evaluating ED trials using published data from maintenance trials across two classes of psychiatric drugs – mood stabilizers (used to treat DSM-IV bipolar disorder) and antidepressants (as applied to DSM-IV major depressive disorder).

Because the validity of ED trials depends on a special set of assumptions, we were also interested in how often these assumptions are recognized and discussed in clinical trials using the design, in systematic reviews of psychiatric drug maintenance therapy, as well as in key psychiatric textbooks and treatment guidelines (chapter 5).

1.2 Definition of terms: Classic RCT's and RCT variants

A cornerstone of medical evidence is the classic parallel two-arm Randomized Controlled Trial (from here on, referred to as the 'Classic RCT'). Classic RCT's draw directly from a target

population, and randomly assign patients to a study intervention or to a comparator. The fundamental advantage of classic RCT's is their potential to sort therapeutic signals from the noise of confounders. While data from classic RCT's can potentially generate the highest level of confidence (219), modifications have been proposed to overcome specific challenges in applying the design. In a summary paper, MacLehose et al (219) reviewed ten modifications of the classic RCT. They classified the modifications as 'hybrids' "if they were intended to provide both randomized and non-randomized estimates of effectiveness" and classified them as 'RCT variants' "if they adhered to the principle of randomization" (MacLehose p.57). I will use MacLehose's definition of 'RCT variant' throughout the thesis.

1.3 Outline of the thesis

Chapter 2 defines enrichment discontinuation trials, tracing its origins in psychiatry to a controversy regarding prophylactic lithium therapy for bipolar disorder (formerly 'manic depression'). We build on previous work by Greenhouse et al (143), proposing a method for assessing bias relative to the design's application in psychiatric drug maintenance trials.

Chapter 3 identifies the prevalence and context in which ED trials have been used within the class of mood stabilizers including lithium and the anticonvulsants divalproate, carbamazepine and lamotrigine. The method proposed in chapter 2 is then applied to the trials.

Chapter 4 describes the prevalence and critically evaluates the application of ED designs in maintenance trials of antidepressants including selective serotonin reuptake inhibitors (SSRI's) and tricyclic antidepressants (TCA's).

Chapter 5 explores a general impression that the limits inherent to ED trials are not widely discussed in the psychiatric literature. As a test of this impression, the chapter reviews a selection of treatment guidelines, textbook chapters, systematic reviews and published maintenance trials using ED designs with a view to identifying whether this RCT variant and its extensive application to psychiatric drug trials is recognized and its limitations discussed.

Finally, chapter 6 wraps up with an overall summary of the thesis findings, their limitations, and some thoughts on how gaps in the literature left by the dominance of ED trials might be approached.

1.4 Relevance of the topic

Enrichment-discontinuation designs have been widely adopted in psychiatric maintenance drug trials, yet their suitability for this application has received little systematic study. If the design proves to be biased when applied to maintenance psychiatric medications, the evidence base supporting maintenance medication may require re-interpretation, and steps may be required to clarify the resulting uncertainty.

Chapter 2

Enrichment-Discontinuation Designs: Applications, controversy and an interpretive approach

2.1 Definition of Enrichment-Discontinuation design

ED trials are two-phase designs in which an intervention of interest is used for an indefinite period before randomization to placebo or competitor, so that only remitted cases already exposed to the intervention are randomized (figure 1)(197). Maintenance of response is retrospectively defined in terms of cases relapsing when treatment is removed.

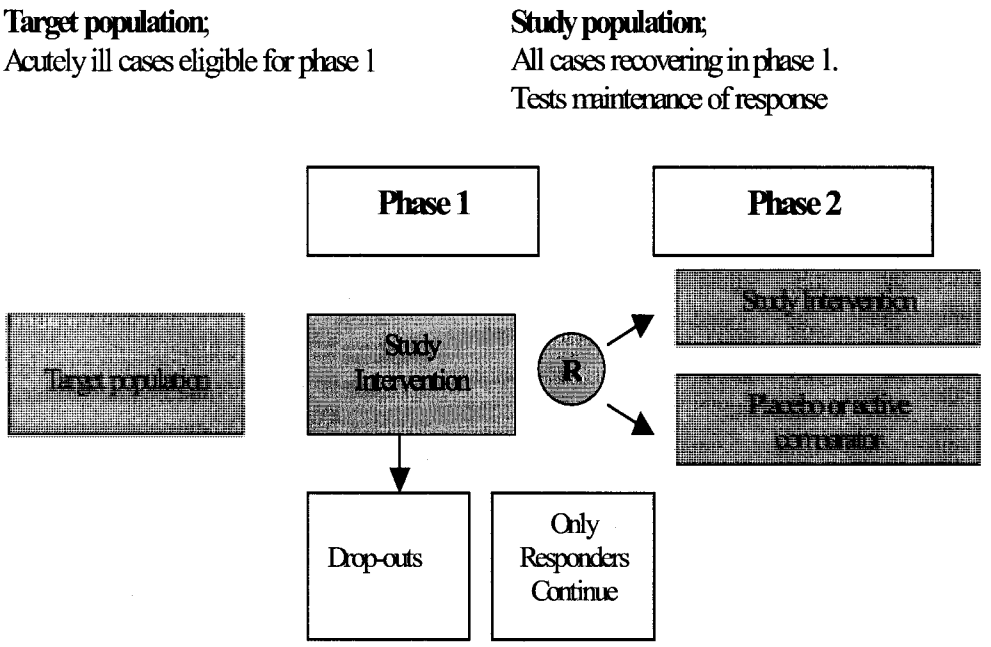


Figure 1: Enrichment-Discontinuation (ED) design.

2.2 Advantages of ED designs

By studying only recovered cases, enrichment discontinuation trials can demonstrate efficacy while randomizing up to 80% fewer cases than classic RCT's of the target population. The open phase acts as a pre-randomization filter for non-responders and non-compliers (197). Relative efficiency depends on the proportion of medication responders to placebo responders in the target population as well as the sensitivity and specificity of the open treatment filter for identifying responders (retrospective definition of 'responder' explained in more detail below, section 2.4). The design has been proposed for applications where the classic RCT is considered unfeasible for logistic reasons (high drop-out rates or low response rates). It may also be useful where acute phase placebo treatment is considered unethical. As an example of the latter, Kopec et al cited a trial of hydroxychloroquine for systemic lupus, in which an ED design was chosen over a classic RCT because placebo treatment for the target population was considered unacceptable by both physicians and patients (197).

ED designs should be distinguished from classic RCT's preceded by brief placebo run-ins. Brief placebo run-ins are generally thought to be unbiased compliance filters, as demonstrated in a systematic review of SSRI trials (206) and an evaluation of several general medical interventions (258).

In contrast with classic RCT's using brief placebo run-ins, ED trials deal with maintenance of response as defined retrospectively in terms of relapses when active treatment is withdrawn (see section 2.X below). This contrasts with maintenance trails using the classic RCT approach, where the maintenance of long-term response is studied prospectively in relation to placebo or other comparator in the target population.

2.3Special limits of ED trials

The efficiency associated with ED trials comes at a cost, since ED designs are difficult to interpret as stand-alone trials. Key assumptions must be met for applicability to the target population as well as for its validity relative to the study population. Departure from these assumptions can lead to an over-estimation of benefits and under-estimation of toxicities.

2.3a Key assumptions related to phase 1

In phase 1, the most important assumption is that the open phase filters latent responders proportionally to their distribution in the target population. Departure from this assumption (also referred to as non-differential attrition) in phase 1 can lead to an inaccurate estimate of effect size relative to a classic RCT of the target population. Retention of proportionally more treatment responders to spontaneous responders tends to inflate the phase 2 effect size. Analogously, if cases experiencing adverse effects in phase 1are not proportionately filtered, the design will tend to under-estimate toxicity. The practical relevance of differential attrition in phase 1 is a matter of debate. Kopec et al cite Friedman as interpreting phase 1 differential attrition as selection bias (122), while Quitkin et al (270) support the position that the design replicates practical experience, since … "subjects who appear to respond to initial, open treatment are more likely to receive the medication of interest in normal clinical practice". From the latter perspective, the phase 1 filter introduces a question of applicability but not bias, since the study population consists only of responders to the intervention.

Sometimes, ED trials are used to establish superiority of one drug over another, with a competitor to the open phase intervention introduced after randomization. This maneuver is however intuitively perplexing. If phase 1 selects for treatment response and medication compliance, it seems counterintuitive to draw comparisons between responders to phase 1 interventions and comparators introduced in phase 2 after randomization. This may be of particular relevance when comparators have a therapeutic time lag, as is the case for many psychiatric drugs. An example of this adaptation of the ED design is Calabrese et al (2003) (50), a maintenance trial in bipolar disorder. Patients with DSM-IV bipolar who had recovered from acute symptoms while taking lamotrigine, an anticonvulsant used in the treatment of bipolar disorder, were randomized to placebo substitution, lithium substitution or ongoing treatment with lamotrigine. The resulting effect size for lithium was much smaller than expected based on previous trials which used lithium enrichment (89).

2.3b Assumptions regarding phase 2

Switching from active medication to placebo can trigger rebound-withdrawal symptoms, inflating the apparent medication effect in phase 2. If withdrawal effects are difficult to sort out from recurrence (139, 143, 287), there is a potential for misclassification bias. This has been discussed by Greenhouse et al (143) in the context of ED trials of lithium maintenance, where it was observed that the prophylactic advantages of lithium in ED trials are concentrated in the first 8 weeks following phase 2, consistent with withdrawal misclassification. Moncrieff (239) has made a more forceful argument for rebound bias in ED lithium trials based on the same data, but she provides no specific method for avoiding misclassification of early recurrences.

In contrast to the case of mood stabilizers (including lithium and anticonvulsants) where the extent of withdrawal-rebound effects remains controversial, antidepressant withdrawal effects have received more empirical attention. Antidepressant withdrawal symptoms severe enough to increase depression-rating scores past recurrence thresholds appear to affect 20-30% of patients abruptly discontinued from SSRI's. The intensity of withdrawal symptoms appears to differ within the class of SSRI's. In addition to a worsening of mood, the symptoms include dizziness, nausea, headache, sleep disturbance, and irritability. A rating scale has been developed for measuring the symptoms (25, 287).

Challenges blinding patients to active medications in psychiatric trials have been discussed in the context of antidepressant trials using the classic RCT design (104), but the complex relationship between long-term compliance in ED trials, patient expectations for receiving the experimental intervention, and response has received little systematic study.

There is evidence suggesting that patients in placebo controlled classic SSRI randomized trials consistently detect treatment allocation to placebo, resulting in a 'wish bias' favoring the investigational drug (27). ED trials may be even more difficult to blind adequately, since all patients are exposed the study drug for an extended period during phase 1. Phase 2 randomization may then effectively un-blind patients, changing their threshold for symptom reporting and their motivation to remain in the trial.

2.4 Methodological studies of ED trials

Greenhouse et al

Greenhouse et al (1991) (143) used an original data re-analysis of a pivotal lithium ED trial to estimate the impact of mis-classifying early phase 2 recurrences as recurrences due to a lack of effective treatment rather than as recurrences due to drug withdrawal effects. Reasoning that in psychiatric trials it is difficult or impossible to causally differentiate drug withdrawal-rebound recurrences (and therefore design bias) from recurrences due to the removal of effective therapy, they excluded recurrences in the first 8 weeks following phase 2. The result was that a previously positive 'maintenance effect' disappeared. Because of uncertainty classifying early recurrences, they concluded that ED designs were inherently difficult to interpret. They proposed that future maintenance trials should incorporate a series of six treatment strategies, assigned from the start, and aimed at clarifying uncertainty around specific selection effects and withdrawal-rebound bias (Figure 2). Strategies 1 and 5 amount to classic RCT's, while strategies 3 and 6 allow hypothesis generation about withdrawal-rebound. Strategies 2 and 4 aggregate specific responder selection effects and withdrawal effects relative to phase 1 interventions.

Treatment strategy	Preliminary Phase	Maintenance Phase
1	A	A
2	A	B
3	A	P
4	B	A
5	B	B
6	B	P

Figure 2: Trial design for assessing ED trials as proposed by Greenhouse et al (143). A and B indicate types of treatment and P indicates placebo.

Kopec et al

Kopec et al (197) used a modeling exercise to compare the required sample size in a classic RCT with that in an ED trial under idealized assumptions, developing specific terminology for ED trials in the process. Under the following assumptions (2.4a), Kopec et al showed that the relative response rates observed in phase 2 of an ED trial will be identical to the relative response rates observed in a classic RCT of the same target population in which responders were followed forward. That is, under these assumptions, ED trials are shown to be efficient and valid relative to a gold standard RCT. They used phase 2 results as a 'gold standard' by which to gauge the sensitivity and specificity of phase 1 as a 'medication response filter' relative to source populations.

2.4a Assumptions under which relative rates of the study population of an ED trial will be identical to the relative rate of responders from a classic RCT of the target population. Adapted from Kopec et al (197):

Phase 1 assumptions:

- Non-differential attrition of 'medication responders' and 'placebo responders' in phase 1 compared to phase 2. Preferential selection of medication responders inflates response rate in phase 2 relative to a classic RCT.

- Placebo response is not prevented by receiving active medication. Deviation favors active medication.

Phase 2 assumptions:

- Treatment is non-curative. Deviation tends to be conservative. As cure rate or spontaneous response rate increases, the phase 2 response rate will decrease.

- Response rate in phase 2 placebo group not affected by non-compliance. Deviation is conservative if non-compliance is more common in the medication arm of phase 2.

- Perfect agreement in the response observed in the enrichment and the discontinuation phases. Deviation tends to be conservative if treatment effect fades over time.

- Rebound-withdrawal effects, if present, can be differentiated from a return of the condition being treated. Deviation tends to favor active medication and results from a mis-classification bias.

- Blinding is feasible. Deviation tends to favor active medication, resulting from an expectation or 'wish bias'.

2.4b Definition of terms in Kopec's model:

- R_0 Denotes the latent 'placebo response' rate in the source population
- R Denotes the latent 'medication response' rate in the source population
- RD Denotes the response rate difference between active and placebo treatment in the source population
- r_0 Denotes the observed placebo response rate in phase 2
- *r* Denotes the observed medication response rate in phase 2
- *rd* Denotes the response rate difference between active and placebo treatment in phase 2

2.4c Relating Kopec's model to this thesis:

When the assumptions in 2.4a are met, the relative response rates in a classic RCT of the target population will be the same as the relative maintenance of response rates from an ED trial, even though the rate difference in phase 2 will almost always be greater than the rate difference in phase 1 when the study intervention is more effective than placebo. In Kopec's terminology (2.4b), when $1>R>R_0$, the ED design will produce a higher rate difference (*rd*) than classic RCT's, since $rd = R-R_O = 1-R_O /R$ (as demonstrated by Kopec et al p. 963)[1].

Further, the relative maintenance of response rates among responders in a classic RCT of the target population will be the same as the relative maintenance of response rate from an ED trial. For example, assume that the treatment response rate R in a classic RCT is 0.5 and that the placebo response rate R_O is 0.3. Under the assumptions in 2.4a, if this same cohort were to enter an ED trial, the placebo response rate in phase 2 would be $r_0=R_0/R$ (= .3/.5 = .6) (Kopec p. 963). The rate difference in phase 2 would be 1-.6=.4 , in contrast with .5-.3 = .2 in the classic RCT,

[1] In ED trials, the initial 'response' rate *r* in phase 2 is 1.0 by definition, since only responders continue past the open phase.

increasing the power of ED trials in comparison to a classic RCT conducted in the same cohort. At the same time, the rate ratios (established under assumptions 2.4a) will remain the same; $r/ r_0 = R/R_O$ (Kopec formula A.7 (197)). As applied to this thesis, departure from $r/ r_0 = R/R_O$ could be consistent with departure from assumptions in 2.4a.

2.3c Retrospective definition of 'Medication Responder'

The term 'medication responder' is ambiguous as operationalized in ED trials, since it relies on the open phase filter to identify a latent class of responders and phase 2 as the gold standard definition of response. That is, the latent class of responders is defined retrospectively relative to a failure rate following discontinuation of the phase 1 intervention. As demonstrated by Kopec et al (198), enrichment designs are limited to an approximation of open phase sensitivity relative to medication responders surviving to the discontinuation arm, but calculations of specificity cannot be made without highly idealized assumptions (*2.4a*). To illustrate this, consider Kopec's model (figure 3), which uses the discontinuation phase as the gold standard for medication response.

Discontinuation phase responder (proportion)

Open Phase Responder (proportion)

	YES	NO	
YES	f_1	f_2	f_1+f_2
NO	f_3	f_4	f_3+f_4
	R	1 -R	1.0

R= proportion of true responders in the open phase assuming phase II is the gold standard
f_1, f_2, f_3, f_4 denote proportions.

Figure 3: Classifying open (phase 1) and discontinuation (phase 2) responders (Kopec et al (197))

The model rests on the assumption that f_4 can be inferred from f_1 and f_2[2]. In ED trials, this assumption may not hold. The specificity of the open phases for defining medication responders cannot be determined, since the second row in fig. 2, designated by f_3 and f_4 assumes latent constructs that remain, by definition, untested by ED designs[3]. 1-R defined in terms of the discontinuation arm (ie. *r*/R =1) rests on the assumption of perfect compliance and perfect agreement in the response observed in the open enrichment phase and the discontinuation phase. However the assignment of a value of 1 to *r*/R, as Kopec et al note, is unlikely to hold in an experimental context, and *r*/R is central to quantifying the sensitivity and specificity of open phases relative to the source population. If the construct of 'medication responder' requires an

[2] Specificity of $r = SP_r = f_4 / (1-R)$, Sensitivity of $r = SE_r = f_1 / R$ and $f_1 + f_2 = (SE)_r(R) + (1-SP_r)(1-R)$ Kopec et al p. 970.

[3] There are intuitive parallels to the logical term, 'begging the question'. A statement begs the question if it contains a definite premise or move that would not be accepted by any reasonable person who is initially prone to deny the conclusion (30).

estimate of sensitivity and specificity to become clinically meaningful, then ED trials can provide only partial empirical support for the designation.

2.2 Clinical applications

Applications in psychiatry

ED designs have been used for over 30 years in psychiatric maintenance drug trials. In 1993, Kopec et al (197) identified 9 ED designs for psychiatric indications: 3 in schizophrenia (99, 228, 292), one in panic disorder (226), one in bulimia nervosa (354), one for anxiety (281), one for social phobia (214) and two in depression (271, 284). The number of ED trials has increased substantially over the past decade.

Application in general medicine

MacLehose et al (219) attributed the ED design to Amery and Dony (1975) (17) for its application to angina treatments, where relapse is identifiable shortly following placebo substitution. Shepherd (310) referred to the use of the design in a 1965 report, where it was used to test prednisone maintenance for ulcerative colitis (207). Kopec et al identified the use of ED designs in 25 studies of general medical interventions including cardiology (N=13) (91, 92, 98, 103, 115, 120, 135, 205, 263, 264, 291, 338, 347), gastroenterology (N=4) (95, 237, 327, 328), rheumatology (N=2) (10, 336), neurology (N=1) (60), oncology (N=2) (67, 136), immunology (N=1) (167), endocrinology (N=1) (119), respirology (N=1) (319) and gynecology (N=1) (320).

2.3 Historical background of ED designs for psychiatric drug maintenance trials

Lithium and the Shepherd/Schou debate

ED trials of psychiatric drug maintenance therapy first appeared in 1970, when two separate lithium maintenance trials were published under the headings "Double Blind Discontinuation" (21) and "Double Blind" (231). Subsequent to 1970, the term 'discontinuation' does not appear in conjunction with ED maintenance trials of psychiatric drugs, which are referred to simply as 'RCT's' (19, 40, 42, 50, 70, 89, 113, 132, 178, 179, 268, 299, 324)(see also Chapters 3 and 4, thesis)[4]. The first ED trials in psychiatry were designed to address significant controversy about lithium prophylaxis in manic depression. As a reflection of this controversy, Healey cites a long and acrimonious debate between Michael Shepherd, a prominent British epidemiologist and psychiatrist and the well-known Danish psychiatrist, Mogens Schou (162).

Following a series of crossover trials conducted in the 1950's and 1960's (140, 222, 298), Schou concluded that lithium was effective in the treatment of acute mania. In 1967, he published the first maintenance trial of lithium as a prophylactic drug for manic depression and recurrent unipolar depression (22) using what he referred to as a 'mirror image design'(162). This design is currently known by the designation 'quasi-experiment' using 'one-group posttest-only design' with multiple substantive posttests (308). Only patients who had already been on lithium for a minimum 12 consecutive months were retrospectively evaluated, based on chart review, using the number of episodes requiring admission or 'extra supervision' as the outcome variable. Schou took the results as strong evidence that lithium was useful as a prophylactic treatment for

[4] A possible exception to this trend exists in the recent publication of a 'discontinuation trial' with the stated objective of comparing withdrawal effects between two SSRI's. (Baldwin 2006). Although this was not technically a maintenance trial, it did use the ED design and supports the role of rebound/withdrawal effects when the design is applied to antidepressants.

manic depression. Shepherd begged to differ, and in a methodological paper entitled *Prophylactic Lithium: Another Therapeutic Myth?* (31), he identified a range of problems with the mirror image design including selection bias and observer bias.

Schou responded by conducting an ED trial in which patients who had been stable on lithium therapy for up to 7 years were randomized to discontinuation or ongoing lithium. Of the patients randomized to lithium, none relapsed, while among those randomized to placebo, 90% relapsed within six months. Schou speculated about a specific subtype of manic depression characterized by 'lithium response', a subtype that might be identified through a distinct lithium response and eventually through genetic testing (162). This speculation was popularized in the 1970's, leading to attempts at developing lithium response models (7, 149) based on retrospective definitions of 'responders'[5]. The same year, Shepherd wrote a thinly veiled polemic against Shou's work, suggesting (as he had done openly in less formal settings) that investigator enthusiasm played a big role in his trial outcomes (310). Shepherd and Schou held contrasting views of interpretations of maintenance trials throughout their lives. The debate finally polarized around questions of selection bias and observer bias, and ultimately reflects two ways of interpreting the open phases of ED trials – as 'drug-specific' response filters (Schou) (154) as opposed to 'non-specific' compliance and response filters (Shepherd). Callahagn and Berrios, have interpreted the Shepherd/Schou debate as a proxy for broader issues in psychiatry; a hope of finding specific medication effects (Schou) (154, 162) and a view of psychiatric illnesses as complex, multi-determined syndromes for which medication-specific effects are likely to be modest (Shepherd)

[5] In a contemporary version of Shepherd's position, Senn argues that if the category of responder varies randomly within individuals, attempts at finding specific genetic markers of response will be ill-fated (306).

(56). The Shepherd-Schou debate and its relevance to patient self-selection in the open phases of ED trials has not yet been resolved (176).

2.6 A simple method for critically evaluating ED designs using published trial data

While the treatment matrix proposed by Greenhouse et al faces practical barriers for an individual trial, it may be possible, at the level of a body of literature taken as a whole, to apply comparisons between interventions or classes of interventions. Using this approach, each cell in the design is filled by a trial or a group of trials. Comparisons between ED trials and classic RCT's (figure 2, treatment strategies 1,3,5,6) have practical value in contextualizing phase 2 effect sizes which, if presented on their own, may suggest a considerably more optimistic outcome than can be expected in practice. Comparisons between trials using alternative 'enrichment strategies' may reflect a responder bias favoring phase 1 interventions (Figure 2, treatment strategies 2 and 4). Figure 4 gives an example of this method applied to hypothetical data in which trials using an ED design have a higher response ratio than a corresponding maintenance trial in which responders from a classic RCT are followed forward.

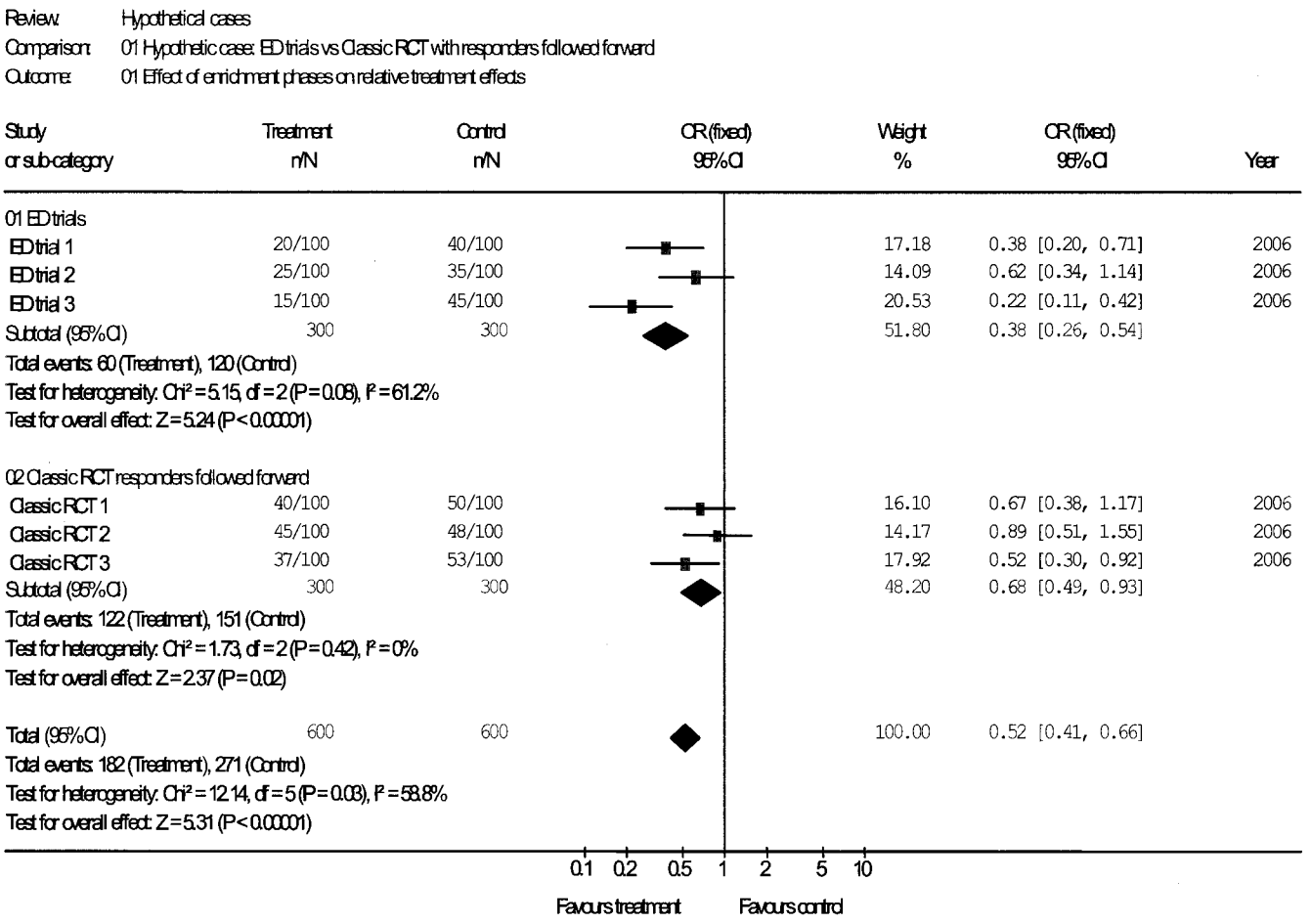


Figure 4: Hypothetical comparison between trials using ED design vs trial using classic RCT design to test maintenance drug therapy. Example suggests a trend favoring ED designs.

2.6a Assumptions for critically evaluating ED trials using published data:

1. Treatment is non-curative.
2. Medication efficacy is the same throughout the course of the syndrome (for both acute and maintenance phases).
3. Patient heterogeneity between comparator trials is evenly distributed between trials using alternative treatment strategies.

4. When pooling trials in which different interventions from the same family are pooled (ex. pooling SSRI's as a group, Tricyclic Antidepressants as a group, anticonvulsants as a group), heterogeneity between idiosyncratic drug effects is distributed evenly across comparisons.

2.6b 'Direct comparisons' (same source-population)

Trials incorporating both classic RCT and discontinuation arms allow an estimate of the aggregate effects of selection bias (phase 1 differential attrition) and withdrawal-rebound effects relative to the same source population. 'Direct comparisons' approximate Greenhouse's treatment strategies 1 and 3. As shown in figure 5, direct comparisons contain an ED trial as well a classic RCT.

.

Direct comparison (trial within a trial): better for assessing bias,
Fair representation of responders guaranteed, same target population

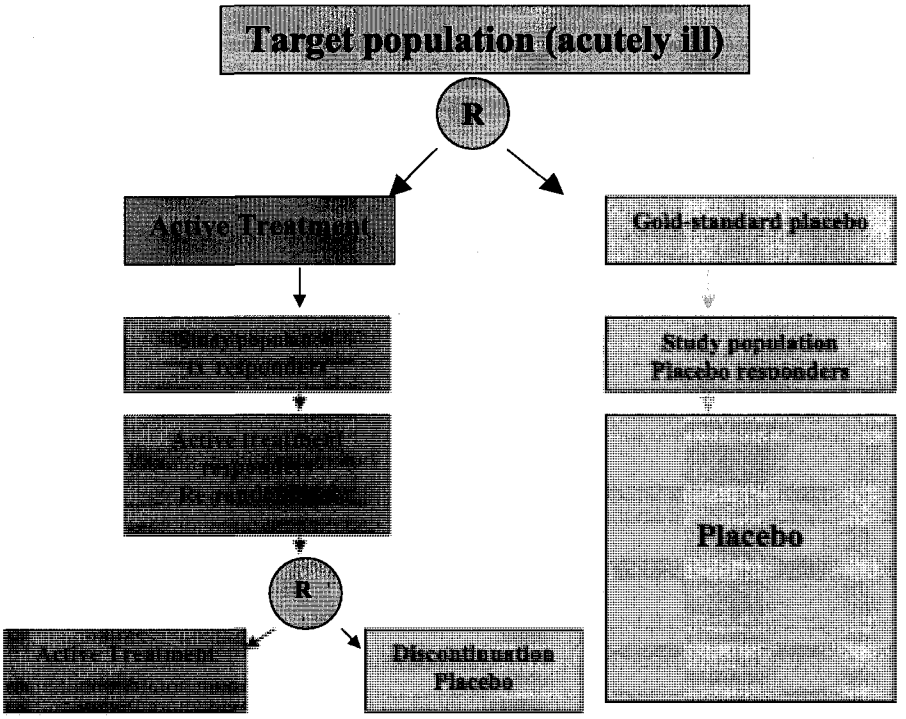


Figure 5: Trial-within-trial design as direct comparison using the same source population. Phase 2 response rates are compared to response rates defined according to the classic RCT treatment strategy.

2.6c Indirect comparisons (different source population)

'Indirect comparisons' use different source populations. The comparison aims to contextualize reported effect sizes in ED trials relative to responders continued in classic RCT's (Indirect comparison Figure 5a). Similar effect sizes between responders in classic RCT's and ED designs would suggest that aggregate 'enrichment effects' (intervention-specific responder selection and withdrawal-rebound) are small. A larger effect in phase 2 of an ED trial would be consistent with departure from an assumption in 2.4a (suggestive of bias), but may also reflect heterogeneous illness courses in source populations upon which the comparison is based.

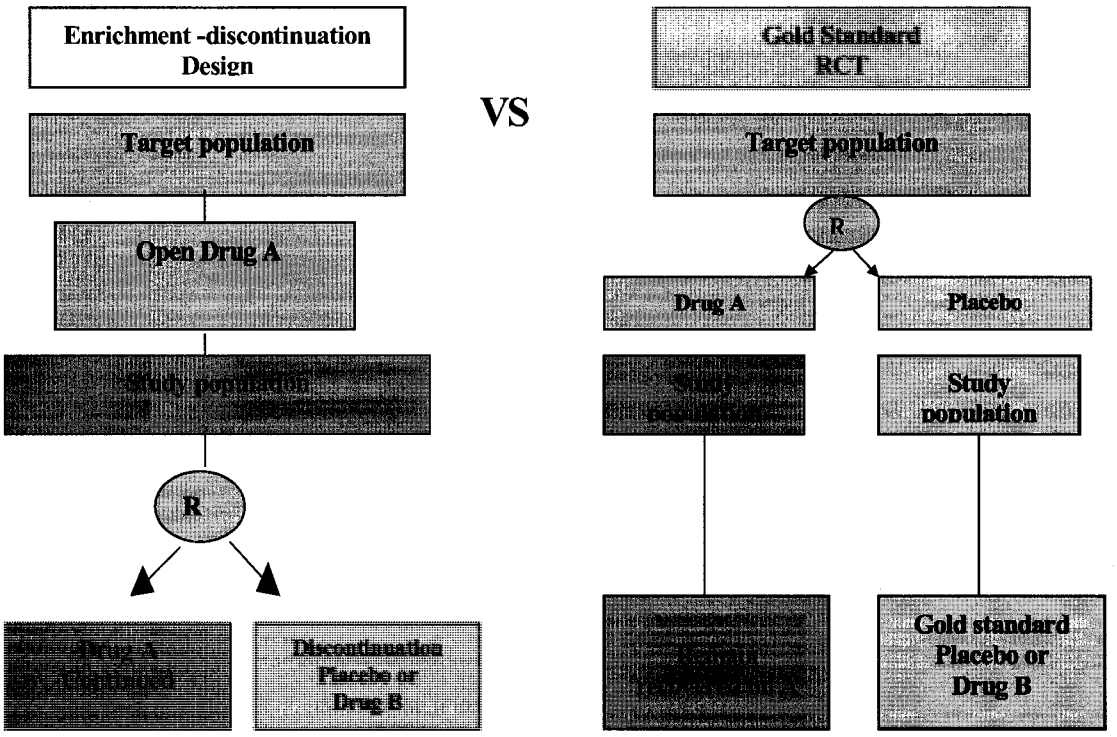


Figure 5a: Indirect comparison, ED trial vs classic RCT

The advantage of indirect comparisons is the potentially greater number of trials available for evaluation, allowing stratification according to run-in duration and number of pre-randomization drop-outs, factors that may reflect open phase differential attrition.

A draw-back to indirect comparisons is potential confounding due to heterogeneous source populations and heterogeneous drug effects.

Summary

ED designs, if unbiased in the context of psychiatric maintenance drug therapy, should yield similar rate ratios to classic RCT's, even though their rate differences will be greater. In this chapter, we have proposed a simple method for testing the assumption of equal rate ratios in ED trials and classic RCT's as well as between ED trials in which the use of contrasting enrichment strategies may lead to selection bias favoring phase 1 interventions. Our effort is modeled after an untested proposal by Greenhouse et al. Key factors that might underlie different observed rate ratios include phase 1 non-differential attrition, medication withdrawal-rebound symptoms in phase 2, and heterogeneity in the source populations from which the data was generated.

Chapter 3

Enrichment-Discontinuation designs in maintenance therapy for bipolar disorder

3.1 Introduction:

3.1a Bipolar disorder

DSM-IV bipolar disorder is a phasic condition in which periods of depression alternate with mania or hypomania. The diagnostic category is sub-divided into bipolar type I, bipolar type II, and bipolar not otherwise specified, a residual category for sub-threshold cases (15). The North American point-prevalence of bipolar type I disorder is .5-2.4% and for bipolar type II disorder it is .2-5.0% (355). Bipolar disorder represents a substantial personal, family and societal burden, reflected in part by chronic relationship strains, unemployment rates in the 60% range and an annual suicide rate of 15-20 times the North American general population rate of 13-15/100,000 (201). Implicit in managing bipolar disorder is a set of interventions believed to reduce suffering in the acute phase, with maintenance of stability over the long run. These interventions include periodic hospitalization, psychological therapies (365), financial assistance and maintenance medications.

3.1b Drug maintenance for bipolar disorder

Maintenance medication is recommended by North American practice guidelines following a single manic episode (14, 365). By convention, maintenance begins 16-24 weeks after recovery from an acute decompensation and can continue indefinitely (117).

Up until the mid 1980's, lithium carbonate was the 'default' treatment for both the acute and maintenance drug therapy of bipolar disorder.

Background of lithium and bipolar disorder:

Lithium, either in its bromide, chloride, or carbonate salt, has been used therapeutically since the mid 19th century. Initially found in spa waters and used as a non-specific tonic for an array of conditions including gout and diabetes, its systematic use in manic depression began in the late 1940's (21). Randomized trials of lithium monotherapy in manic depression conducted in the 1950's and 60's were primarily 'mirror image trials' (now called crossover trials), in which patients' illness courses were observed on and off lithium (112, 140, 222, 298, 361). Because of long-held concerns with withdrawal bias and blinding (239), these early trials are mentioned only historically (143). Lithium maintenance therapy co-evolved with the concept of manic depression. For some psychiatrists, recovery while taking lithium became *defacto* evidence of a latent 'lithium deficiency' (154). While this tendency is no longer current, its influences on psychiatrist's beliefs and actions is reflected by studies attempting to correlate trace lithium levels in drinking water with psychiatric hospital admissions and crime rates (83, 177, 256, 262, 300, 300, 350). Hopes of finding specific predictors of lithium response have met with modest success (7, 39, 146, 147, 149, 193, 311).

Lithium's maintenance efficacy relative to placebo was first tested in a series of ED trials beginning in 1970 (21). These trials involved randomizing patients who had recovered while receiving 'open' lithium to either ongoing lithium or placebo discontinuation (of

lithium). After lithium's maintenance superiority to placebo gained acceptance based in part on ED trials, it became the reference compound for future trials.

Background of anticonvulsants and bipolar disorder:

Anticonvulsants, particularly divalproate, carbamazepine and lamotrigine, were the first widely adopted alternatives to lithium maintenance therapy for bipolar disorder. Divalproate (Divalproex, Epival) is one of the most commonly used maintenance therapies in North America for bipolar disorder. Prescription sales rival those for lithium in the maintenance phase of bipolar disorder (44). The parent compound of divalproate, valproic acid was synthesized in 1882, and may be the first effective synthetic psychotropic drug (44). It was initially approved by the FDA for absence seizures in 1978 and has also been widely used in complex partial seizures, generalized tonic-clonic seizures and myoclonic seizures. Its use for bipolar disorder gradually increased as an 'off-label' indication (51, 52, 159, 172, 203, 215, 229, 269, 294, 346). A large placebo controlled trial of divalproate in acute mania was conducted in 1994 (40), followed by an ED trial using divalproate in the maintenance phase of bipolar disorder (41). Divalproate is believed to have particular therapeutic advantages in rapidly cycling bipolar disorder (defined in DSM-IV as four or more acute episodes of depression or mania in a year) (296). Other anticonvulsants have been reported to have efficacy in the maintenance of bipolar disorder. The most promising of these appears to be lamotrigine (43). Since it may be efficacious in bipolar depression, the use of lamotrigine is on the increase. Carbamazepine, another anticonvulsant, has also been used since the 1970's, but has remained a third-line treatment option. Evidence for topiramate and gabapentin in bipolar

disorder remains disappointing and it is unlikely that these anticonvulsants will gain widespread use. FDA approval of divalproate and lamotrigine in the maintenance therapy of bipolar disorder was granted based on ED trials using anticonvulsant enrichment with randomization to ongoing anticonvulsant, lithium, or placebo.

Other maintenance drugs used in bipolar disorder:

A third class of maintenance medications, the atypical antipsychotics, is also used in bipolar drug maintenance therapy[6]. Evidence is preliminary and includes ED trials using anticonvulsants or lithium as active comparators (339).

3.1c Chapter Objectives

1. Identify the prevalence of ED designs relative to classic RCT's in the evidence base supporting drug maintenance for bipolar disorder.
2. Using direct comparisons and indirect comparisons as outlined in chapter 2, test for bias relative to trials in which responders from classic RCT's are followed into continuation and maintenance phases.
3. Using indirect comparisons (groups of trials in which phases 1 and 2 interventions differ), identify possible 'responder bias' favoring phase 1 interventions across the evidence base for maintenance therapy in bipolar disorder.

[6] Atypical antipychotic medications used in bipolar disorder include olanzapine (Zyprexa), quetiapine (Seroquel) and risperidone (Risperdal). The evidence base for these compounds in bipolar disorder is in its nascency (61).

3.2 Methods

Participants: All patients with a diagnosis of DSM-IV, DSM-III, DSM-III-R, RDC, or Feighner defined bipolar disorder, type I or type II. No age restrictions. Inpatient or outpatient settings.

Interventions: Drug treatment of a minimum 6 months following recovery from the most recent episode of either mania or depression. The brief use (maximum 2 weeks) of one 'rescue medication' of either sedative hypnotic class (ex. benzodiazepine, chloral hydrate) or antipsychotic class (ex. haloperidol) in addition to the maintenance drug was permitted in the first month following recovery from the most recent episode.

Comparators: Lithium, Divalproex, Lamotrigine, Carbamazepine, Olanzapine, Quetiapine, Risperidone, Placebo

Outcomes: Binary outcomes; Number of cases recurrence free and completing study; Drop-outs due to adverse effects; Drop-outs for all reasons other than recurrence.

Study design: RCT's including ED trials and classic RCT's: placebo controlled and active comparator.

3.2b Search Strategy

Lithium: Through the OVID interface, we applied the Cochrane RCT search filter (1) to search MEDLINE, EMBASE, and PSYCHINFO between 1966 and May 2005 for all available English language randomized studies using the text word "lithium". The Cochrane Central Register of Controlled Trials was also searched using the text word "lithium". No attempt was made to locate unpublished trials through pharmaceutical manufacturers. References between 1949 and 1966 were found by careful tracing of

bibliographies from reports published between 1970 and 1980. The Jadad (173) scale and Schulz (302) component for assessment of randomization, blinding, dropouts withdrawals and concealment of allocation were completed by DD and JA (Joy Albuquerque). Standardized case report forms were used for data abstraction (DD and JA). Trial funding sources were recorded.

Divalproate: Through the OVID interface, MEDLINE and PSYCHINFO were searched for English language studies published between 1966 and May 2005 using the keywords [epival or divalproex or divalproate or valproic acid] AND [depression or bipolar disorder]. The Cochrane Central Register of Controlled Trials was also searched using the text words divalproate, divalproex, epival and valproic acid. In addition, the bibliographies of 10 review articles known to DD (24, 38, 44, 53, 81, 82, 188, 189, 221, 248) were checked to identify RCT's missed by the electronic strategy. Trial funding sources were recorded. One reviewer, DD, completed the data abstraction. No attempt was made to locate unpublished trials through pharmaceutical manufacturers. Authors of the STEP-BD (US) project (G. Sachs) and the BALANCE trial (UK) (J. Geddes) were contacted to ensure that no relevant studies were missed (DD personal communications).

Carbamazapine, lamotrigine, olanzapine, quetiapine and risperidone:
RCT's of a minimum 6 months applying these compounds were identified through published systematic reviews. Through the OVID interface, MEDLINE and the Cochrane Collaboration's Database of Systematic Reviews were searched using keywords 'bipolar disorder' and each of the compounds (carbamazapine, lamotrigine, olanzapine, quetiapine

and risperidone) between 2000 and May 2005. For the MEDLINE search, the 'review' and 'English language' filters were used. Only reviews in which the terms 'systematic review' or 'maintenance' appeared in the abstract were explored further. Bibliographies were searched for other reviews that may have been missed using this strategy. Trials of a minimum 6 months were screened and included if a binary outcome was provided. Trial funding sources were recorded. No attempt was made to locate unpublished trials. One reviewer (DD) completed this portion of the search.

Data Synthesis:

Using Revman version 4.2, a meta-analysis software package, binary outcomes were combined to give an overall effect estimate, examining the proportions free of depression, hypomania or mania, and free of any relapse at study endpoint. We assumed that the probability of relapse was approximately constant between 6-12 and12-24 months after the last episode (137), and therefore felt it would be valid to combine these outcomes. We treated the anticonvulsants divalproate, carbamazepine and lamotrigine as a class for indirect comparisons, since they are treated similarly in clinical settings (212, 365). Early dropouts were recorded where possible. The data from the primary studies was combined using a fixed effects model if I^2 < than 50%. Comparisons were constructed using the method outlined in chapter 2.

3.3 Results

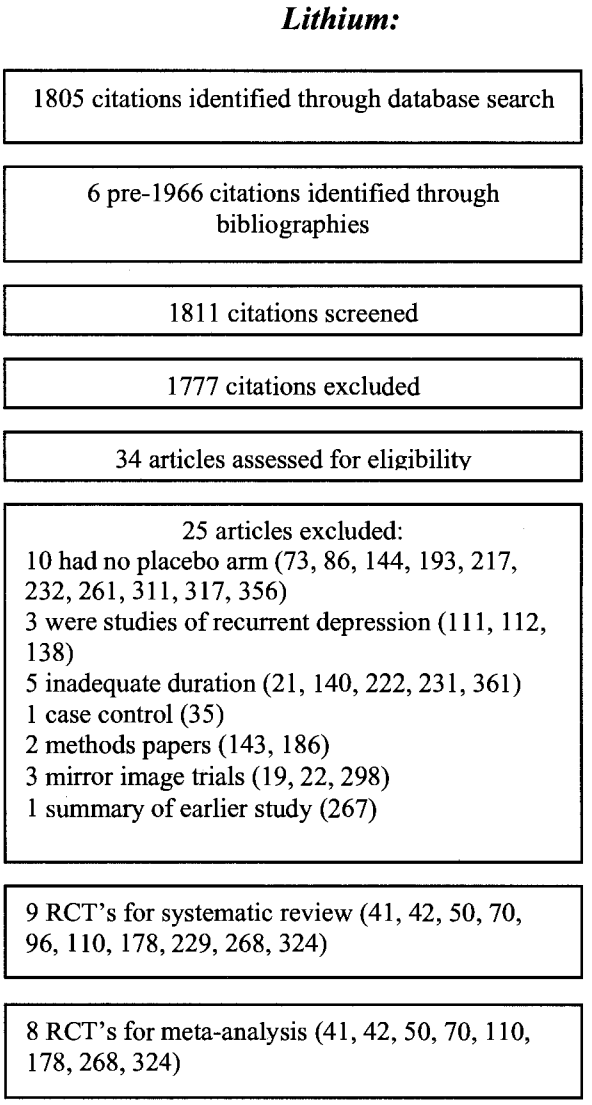


Figure 1: Flow diagram for selection of placebo controlled lithium RCT's

The lithium database search identified 1805 citations (figure 1). Bibliographies yielded 6 pre-1966 trials. A total of 34 papers were identified for closer examination. Of these, 10 were superiority trials without placebo arms (73, 86, 144, 193, 217, 232, 261, 311, 317, 356), three were studies of recurrent depression (111, 112, 138), five were of inadequate

duration (21, 140, 222, 231, 361), one was a case-control study (35) two were methodological papers (143, 186), three used a mirror image design (19, 22, 298) and one was a summary of earlier studies (267). This left nine studies meeting the inclusion criteria for this review (41, 42, 50, 70, 96, 110, 178, 268, 324). Dunner (96) did not provide the number surviving at endpoint and could not be included in the meta-analysis.

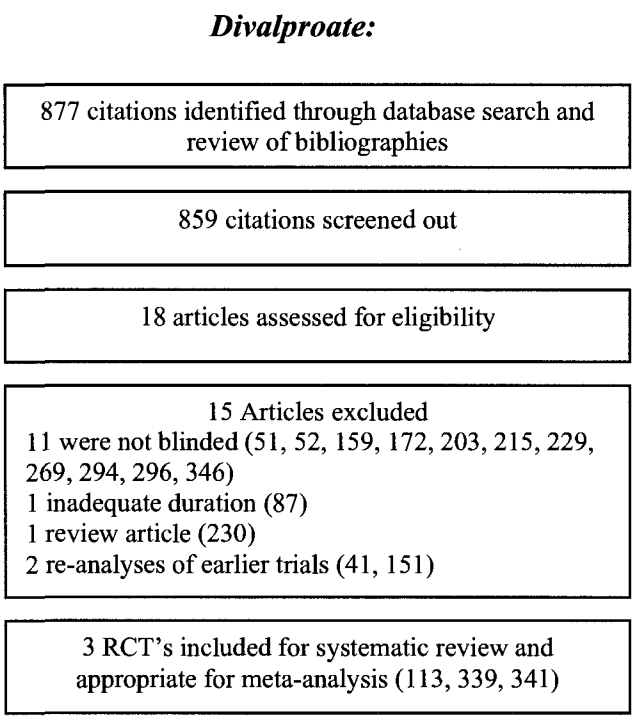


Figure 2: Flow diagram for selection of maintenance RCT's of divalproate.

877 citations for divalproate were identified after removing duplicates (figure 2). Of the 18 maintenance studies selected for closer examination, 15 were excluded; 11 were not blinded (51, 52, 159, 172, 203, 215, 229, 269, 294, 296, 346), one was of inadequate duration (87), one was a review article (230) and two were re-analyses of previous

studies (41, 151). This left three eligible trials of divalproate maintenance in bipolar disorder (113, 339, 341).

Carbamazapine, lamotrigine, olanzapine, quetiapine and risperidone:

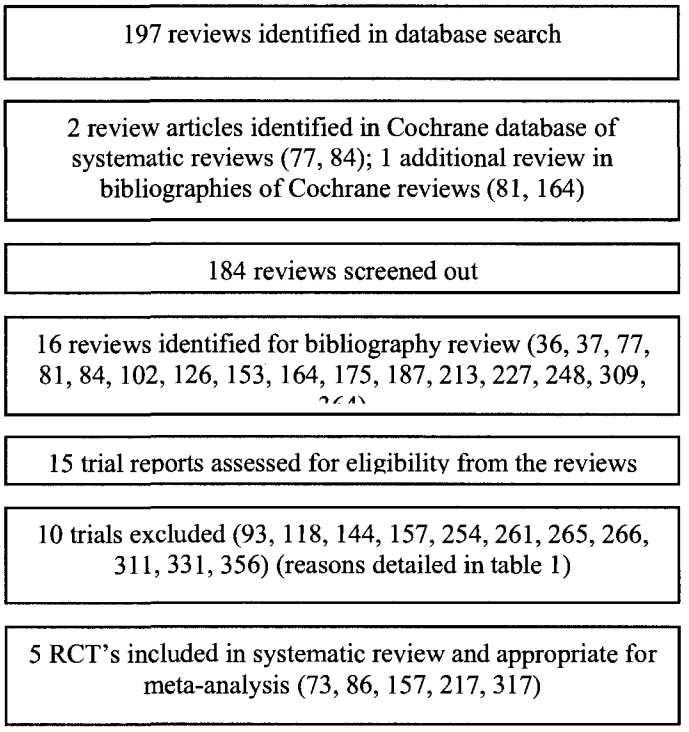


Figure 3: Flow diagram for selection of maintenance RCT's of carbamazapine, lamotrigine, olanzapine, quetiapine and risperidone.

The MEDLINE search yielded 197 reviews, of which 13 used the terms 'maintenance' and 'systematic review' in the abstract (36, 37, 102, 126, 153, 164, 175, 187, 213, 227, 248, 309, 364) (figure 3). The Cochrane Database of Systematic Reviews yielded two reviews (carbamazepine and anticonvulsants) (77, 84) and three protocols (lamotrigine, olanzapine and risperidone) (5, 65, 220, 278). There were no reviews of quetiapine. Bibliographies of Cochrane reviews yielded an additional systematic review (81, 164). From these, 15 trials were identified,

of which 5 were included (73, 86, 157, 217, 317), and 10 were excluded for reasons outlined in table 1 (93, 118, 144, 157, 254, 261, 265, 266, 311, 331, 356).

Table 1: Characteristics of excluded trials

Study	Population	Blinding?	Comparison
Placedi 1986 1year (261)	In & outpatients schizoaffective, major affective; 41/43 female, CPZ allowed; enriched	yes	CBZ vs Li
Okuma 1981 1 year (254)	In & outpatients ICD 9 manic or bipolar psychosis	no	CBZ vs placebo
Watkins 1987 (356)	Recurrent affective disorder, some were in long-term remission, some acutely manic	yes	Li vs CBZ
Stoll 1985 (331)	Mania, schizoaffective	not discussed	Li vs CBZ
Post 1983 (266)	Case series N=7 lithium resistant rapid cycling BP	no	CBZ used
Greil 1997 (144)	MAP study BP patients, multicenter, 175 followed Li vs CBZ	no	CBZ vs Li
Simhandahl 1993 (311)	Open trial of CBZ slow-release on bipolar prophylaxis	no	CBZ vs Li
Post 1990 (265)	Refractory affective disorders hospitalized 15 BP, 5 atypical BP, 2 unipolar	no	CBZ vs previous treatment
Frankenburg 1988 (118)	Bipolar Disorder Retrospective study	n/a	CBZ open
Di Costanzo 1991 (93)	Rapid cycling BP disorder; Retrospective study	n/a	CBZ vs Li+CBZ

Study	Conclusion	Enrichment	Reasons for exclusion
Placedi 1986 1year	Differential 'response patterns"	yes	Vague outcomes
Okuma 1981 1 year	CBZ>placebo	unclear	Outcomes unclear
Watkins 1987	Li>CBZ	yes/no	Vague inclusion, unclear outcomes, antidepressants and antipsychotics allowed
Stoll 1985	CBZ>Li	no	Vague, mixed population, unclear outcomes, allows neuroleptics, no discussion of randomization blinding or allocation
Post 1983	CBZ works for Depression and Mania	unclear	Outcomes unclear, open design
Greil 1997	'Selective responsiveness'	no	Large study with open assessments. Results support prevailing hypotheses CONSIDER for discussion -- largest Euro study
Simhandahl 1993	Li=CBZ	Unclear	Open label
Post 1990	CBZ effective in some cases	unclear	Open label, case series
Frankenburg 1988	CBZ may be effective	unclear	Retrospective review
Di Costanzo 1991	Combination therapy gives faster results	n/a	Retrospective chart review

3.3a Trials available for Indirect Comparisons

Lithium vs Placebo:

Studies are summarized in table 2 according to design, sample size, source population, diagnostic breakdown, outcome measures, rescue medications, use of pre-randomization open phases and quality rating. Where the patient recruitment base was not reported, no comment is recorded. Overall, there were nine trials that included 1432 patients randomized to lithium or placebo, 671 men 761 women, mean age at entry 44.2 years. Diagnostic categories used included Bipolar I (N=1010) (41, 42, 50), Bipolar II (N=89) (96, 179), Bipolar I and II (N=81) (110) , Bipolar (N=135) (268), Manic Depressive (N= 52) (324), and Recurrent Affective Disorders (N=65) (70). Calabrese (2003) (50) was a multi-center, international collaboration. All other included studies except Coppen (UK, 1971) (29) were conducted within the United States. Phase 1 dropouts, when recorded, ranged from 35% - 52% (table 2). Six of the trials were published on or before 1982 whereas three were published between 2000 and 2003. All of the trials published since 1982 were funded by industry whereas the earlier ones were investigator driven. There was an 18-year gap between reports of placebo controlled lithium maintenance trials (1982-2000).

***Table 2**: Study characteristics, included trials involving lithium and placebo*

Source	Population (N before dropouts)	Phase 1 dropouts n (%); Taper (y/n)	N randomized	Primary Study Aim	Outcome summary
Calabrese et al, 2003 (22)	BP I, recently depressed (966)	501 (52%); y	463	Test Lamotrigine	Survival time to episode
Bowden et al, 2003 (19)	BP I, recently hypomanic or manic (349)	174 (50%); y	175	Test Lamotrigine	Survival time to episode
Bowden et al, 2000 (18)	BP I, recently manic (571)	199 (35%); y	372	Test Divalproex	Survival time to episode
Kane et al, 1982 (67)	BP II, general outpatients (not stated)	49 (unknown); n	49	Test Lithium	# episodes during study
Dunner et al, 1976 (39)	RDC BP II, State Hospital (not stated)	40 (unknown); n/a	40	Test Lithium	mean # episodes/year, # relapse free
Fieve et al, 1976 (41)	RDC BP I and II (not stated)	81 (unknown); n/a	81	Test Lithium	mean # episodes/year, # relapse free
Prien et al, 1973 (100)	RDC BP, recent depression, VA hospitals (138)	70 (50%); n	135	Test Lithium	number of episodes, # relapse free
Stallone et al, 1973 (111)	Manic Depressive (not stated)	52 (unknown); n	52	Test Lithium	mean # episodes/year, # relapse free
Coppen et al, 1971 (29)	Recurrent Affective Disorders (not stated)	65 (unknown); n	65	Test Lithium	% time ill, # relapse free

Source	Event threshold	Open Phase	Rescue Meds	Jadad	Schultz	Funding
Calabrese et al, 2003	Clinical relapse	16 wks on lamotrigine	Chloral hydrate 2g/d, lorazepam 1mg/d or equivalent	4	Adequate	Commercial
Bowden et al, 2003	Clinical relapse	16 weeks on lamotrigine	Chloral hydrate 2g/d, lorazepam 1mg/d or equivalent	4	Adequate	Commercial
Bowden et al, 2000	Mania or depression rating scores	3 months open stabilization	Lorazepam to 6mg/d*14d, month 1 then up to 7days****	4	Adequate	Commercial
Kane et al, 1982	RDC MDD*1wk, md*4wks, any mania, hm*1wk	No ('hypo-manic' group only)	None	2	Adequate	Unclear
Dunner et al, 1976	Clinical relapse	No	None during maintenance, all during relapse	2	Adequate	NIH
Fieve et al, 1976	Clinical relapse	No	None during maintenance, all during relapse	3	Unclear	NIH
Prien et al, 1973	Clinical need for additional treatment	4 months, considered 'early phase'	None during maintenance, all during relapse	3	Adequate	NIH
Stallone et al, 1973	Clinical relapse	some' received lithium	Amitriptyline or imipramine for depression, CPZ for mania	2	Adequate	MRC
Coppen et al, 1971	7 point scale, global assessment	16 week open phase	None during maintenance, all during relapse	2	Adequate	MRC

* MDD: Major Depressive Disorder, md: Minor Depressive Disorder, hm: Hypomania
** Scale for Randomization and Blinding. Maximum 5
***Scale for Concealment of Allocation
****10mg haloperidol for one week in first month

All trials allowed some form of co-intervention (for example, 80% of relapses in Prien 1973 were hospitalized). It was not possible to assess the impact of these co-interventions because the episode duration was not specified, and details and duration of co-interventions were generally not given (table 1). Bowden 2003 (42) allowed short-term intermittent chloral hydrate or benzodiazepines, as did Calabrese 2003 (50). Bowden 2000 (41) allowed lorazepam to 6mg/day for the first 14 days, as well as haloperidol up to 10mg per day during the first month (table 2).

In the nine included reports, three outcomes were provided. The first, time to relapse of threshold symptoms, was used in Bowden 2000 and 2003 (41, 42), Calabrese 2003 (50) and Kane 1982 (179). The second and third outcomes evaluated the mean number of relapses per patient year or the mean percentage time spent as an inpatient, as described in Dunner 1976 (96), Fieve 1976 (110), Prien 1973 (268), Stallone 1973 (324) and Coppen 1971 (70).

It was found that outcome measures were inconsistent across the trials, and therefore, some suitable measure of between-study comparison had to be identified. We identified "event-free" as our endpoint and this outcome could be combined across all studies. "Event free" refers to the absence of relapse deemed to require interventions other than lithium.

Lithium vs. monotherapy with divalproate, carbamazepine, lamotrigine, olanzapine, quetiapine or risperidone:

Included trials with their published outcomes are summarized in table 3. Trial endpoints were triggered by threshold symptoms on either a depression or mania scale, or the need for added interventions, as determined by investigators. While 7 of 12 trials reported median survival time, none provided confidence intervals for these times. Only the number of participants recurrence free at endpoint was available for pooled analysis.

Maintenance studies included 5 trials comparing lithium and CBZ (167 patients) (73, 86, 157, 217, 317), 2 trials comparing divalproate and lithium (431 patients) (41, 113) and 2 trials comparing lamotrigine with lithium (638 patients) (42, 50). There were no monotherapy maintenance trials comparing lithium with quetiapine or risperidone. One trial, discussed below under the heading 'lithium vs. combination' (341), involved olanzapine but the comparator was olanzapine combined with either lithium or divalproate. All trials used the DSM-III, DSM-IIIR or DSM-IV classification systems, and were published between 1988 and 2005. With the exception of Coxhead 1992 (73), the carbamazepine trials were conducted on hospital-based patients. Two of the five trials used lithium enrichment prior to randomizing to carbamazepine (73, 86). All trials allowed rescue medications from the antipsychotic class.

Table 3: Study characteristics, included superiority trials (primary objective to test non-lithium mood stabilizer)

Study	Entry threshold	Population	Primary aim	Comparisons	# Recurrence free at endpoint
Findling 2005, 18months (113)	Childrens depression scale<40, YMRS<12.5 , Child GAS>51	youths 5-17 BP1or 2 DSM4	DVPx superiority	DVPx vs Li	yes
Tohen 2004, 18months (339)	YMRS>16	Partially LI or DVPx resistant BPI manic or mixed (outpatients!)DSM4	combo vs monotherapy	monotherapy vs combo	yes
Tohen 2003, 47 wks (341)	YMRS>20	251 BP1 acute mania or mixed DSM4	test olanzapine	olanzapine vs DVPx	yes
Calabrese 2003 18month (50)	CGI-S score of 3 (mildly ill) or less X4wks	BP1 recently depressed (within past 2 months) DSM4	test lamotrigine	lamotrigine vs Li	yes
Hartong 2003, 2yrs (157)	acute or prophylactic	94 DSM-3R BP1+2	avoid enrichment	CBZ vs Li	yes
Bowden 2003, 18 months (42)	CGI-S score of 3 (mildly ill) or less X4wks	BP1 recently manic or hypomanic (DSM4)	test lamotrigine	lamotrigine vs Li	yes
Bowden 2000, 1 year (41)	Prophylactic YMRS<11, DSS<13, GAS>60	BPI most recent manic within past 3 months DSM4	DVPx vs Li	DVPx vs Li	yes
Calabrese 2000, 6 month (55)	lamotrigine>100mg+HAM-D<14 and MRS<12	BP1 or 2 rapid cycling (DSMIV)	test lamotrigine	lamotrigine vs placebo	yes
Denicoff 1997, 3 year (86)	4wks stable	DSM3R BP1or2 60%rapid cycle	CBZ vs Li vs Both (crossover)	CBZ,Li,Both	yes
Coxhead 1992, (73)	Stable long-term on lithium	Li Responsive DSM III BP	Test CBZ in BP prophylaxis	CBZ vs Li	yes
Small 1991, 2 years (acute and prophylaxis) (317)	DSM3R BPI GAS<60, SDMS-D&M manic scale>7	Hospitalized mania (referred) most had tried Li	Test CBZ acute and prophylactic	Li, CBZ	yes
Luznat 1988, 1year (217)	Bech-Rafaelson score >10, DSM3 BP, SA	admitted manic patients	Test CBZ acute and prophylactic	Li, CBZ	responders'

Table 3 (continued)

Study	Median time to relapse	Median Survival Time	Survival reporting	Active rescue meds	Hospitalization as trigger	Med use triggers end
Findling 2005, 18months	yes	yes	yes	ongoing ritalin allowed in a third	yes, clinical intervention	yes
Tohen 2004, 18months	yes	yes	yes, several variables	loraz 2mg/d X5consec or 60 days tot.	no	yes
Tohen 2003, 47wks	yes	yes	partial	lorazepam 2mg/day or benztropine	no	yes
Calabrese 2003 18month	yes	yes	yes	benzos only	yes, clinical intervention	yes
Hartong 2003, 2yrs		not reported in paper	yes, diagram, partial	max benzo oxazepam to 100mgX14d	no	yes
Bowden 2003, 18 months	yes	yes	yes	benzos only	yes, clinical intervention	yes
Bowden 2000, 1 year	yes	yes	yes, to mania and dep	lorazepam 6mg/day X14 days, Haldol 10 2wks month1	could trigger termination	yes
Calabrese 2000, 6 month	yes	yes	yes	lorazepam 2mg/day	yes, clinical intervention	yes
Denicoff 1997, 3 year	Mean only	Mean only	yes	Yes,++ adjuvant meds	yes	Failure of adjuvant
Coxhead 1992,	no	(graph only)	Yes (graph only)	Only temazepam	yes	yes
Small 1991, 2years (acute and prophylaxis)	no	no	no	chloral hydrate or amobarb	already hospitalized	yes
Luznat 1988, 1year	no	no	no	benzos, neuroleptics, antidepressants	no	records amount of meds

Table 3 (continued)

Study	**Depression score triggers end**	**Mania scale triggers end**	*Clinical relapse as endpoint*	**DSM-IV syndrome severity**	**Enrichment phase (phase 1)**	**YMRS or MRS**	**BRMAS (mania scale)**	**BRMES (depression)**	**Retrospective life chart**
Findling 2005, 18months	no	no	yes	no	up to 20 wks. Combo lithium and divalproex	yes	no	no	no
Tohen 2004, 18months	yes, HAM-D remission<8	yes. YMRS, remission<12	yes	no	Yes, all got Olanz+MS acutely, early monotherapy at investigators discretion	yes	no	no	no
Tohen 2003, 47wks	yes, HAM-D remission<8	yes. YMRS, remission<12	yes	Yes, for depression and mania	No, Acutely randomized	yes	no	no	no
Calabrese 2003, 18month	no	no	yes	no	yes, up to 16 wks lamotrigine	yes	no	no	no
Hartong 2003, 2yrs	no	no	yes	no	No acute or prophylactic	no	yes	yes	no
Bowden 2003, 18 months	no	no	yes	no	yes, up to 16 wks lamotrigine	yes	no	no	no
Bowden 2000, 1 year	yes, DSS >25	yes, YMRS>16	yes	no	3 month open with DVPX or Li at investigators discretion (half got each)	yes	no	no	no
Calabrese 2000, 6 month	no	no	yes	no	yes, 4 weeks (only 1/3 completed)	yes	no	no	yes
Denicoff 1997, 3 year	One of many outcomes	no	yes	no	Yes, mean 104 days	yes	no	no	yes
Coxhead 1992, 12 months	no	no	yes	no	yes	no	no	no	no
Small 1991, 2years (acute and prophylaxis)	no	no	yes, also no response at 8wks	no	No (2 week run-in off meds)	yes	no	no	no
Luznat 1988, 1year	no (but noted in follow-up)	no, but used in evaluating f/u	'poor' vs 'satisfactory' responders	no	no	no	no	no	no

Table 3 (continued)

Study	**GAS**	**DSS (Depressive syndrome scale)**	**Bech Raefelson mania ating**	**21-item Ham-D**	**CPRS (comprehensive scale)**	**CGI**	**KSADS**	**SADS**	**PANSS**	**Funding**
Findling 2005, 18months	(CGAS)	no	no	Children's scale	no	no	yes	no	no	Commercial
Tohen 2004, 18months	no	no	no	yes	no	yes	no	no	no	Commercial
Tohen 2003, 47wks	no	no	no	yes	no	yes	no	no	yes	Commercial
Calabrese 2003, 18month	yes	no	no	yes	no	yes	no	no	no	Commercial
Hartong 2003, 2yrs	no	no	no	no	yes	no	no	no	no	Unclear
Bowden 2003, 18 months	yes	no	no	yes	no	yes	no	no	no	Commercial
Bowden 2000, 1 year	yes	yes	no	no	no	no	no	yes	no	Commercial
Calabrese 2000, 6 month	yes	no	no	yes	no	yes	no	yes	no	Commercial
Denicoff 1997, 3 year	no	no	no	yes	no	yes	no	no	no	Unclear
Coxhead 1992, 12 months	no	no	yes	yes	no	no	no	no	no	Unclear
Small 1991, 2years (acute and prophylaxis)	yes	no	no	yes	no	yes	no	yes	no	Unclear
Luznat 1988, 1year	no	no	yes	17 item	no	no	no	no	no	Unclear

Table 3 (continued)

Study	N Entered	N Randomized Taper y/n	Adverse event reporting	Baseline characteristics	*Comments*
Findling 2005, 18months	161	60; y	yes	yes	10% completed. Lithium and divalproex roughly equal
Tohen 2004, 18months	160*	99; y	AIMS, BARS, adverse events well reported	yes, given in previous paper on management of acute mania	
Tohen 2003, 47wks	251	251; y	yes. Inc. movement disorder. Very detailed	yes. 76% rapid cycling. 62%psychotic, 20% treatment resistant	comparative efficacy for prophylaxis better studied in trial starting with stable remission
Calabrese 2003, 18month	966	463; y	yes	several illness course and demographic details provided	Full remission at entry not necessary. Definition of remission very different from 70's Li studies
Hartong 2003, 2yrs	94	94; n	4-point scale, D/C sec to S/E details given		
Bowden 2003, 18 months	349	175; y	yes, details given	several illness course and demographic details provided	Goodwin has done a pooled analysis of the two lamotrigine studies
Calabrese 2000, 6 month	324	182; y	yes	documented	full remission at entry not necessary. Definition of remission very different from 70's Li studies
Bowden 2000, 1 year	571	372; y	yes details given	Only for intent to treat sample. Enrichment group not characterized	full remission at entry not necessary. Definition of remission very different from 70's Li studies
Denicoff 1997, 3 year (planned)	Not stated	52; n/a	minimal	Very ill, descriptors provided,	Very ill population. Confusing design, incomplete reporting Cannot combine in meta-analysis
Coxhead 1992, 12 months	**	31; n	some	Highly selected population of lithium responders	Female to Male ratio = 2:1 May be confounded by lithium rebound. Should not combine
Small 1991, 2years (acute and prophylaxis)	74	52; n/a	yes	yes	very ill population, chronically mentally ill.
Luznat 1988, 1year	54	54; n	limited	yes	very ill, SMI population. Use of unrecorded rescue meds potentially confounding. Limited study

- *Extension of acute study. Only excellent responders to combination therapy could enter this next phase.
- ** 145 assessed, completely stable DSM III BP on lithium. Most patients declined participation

Lithium vs Combination

There were an inadequate number of reports involving any combination of lithium, divalproex and/or olanzapine to test the impact of enrichment phases on trial outcomes. In total, three randomized trials were located. These included two ED trials and one classic RCT. The first ED trial was conducted on adolescents diagnosed with bipolar disorder, comparing Lithium vs Divalproex after a combined divalproex/lithium run-in (N=60) (113). In this trial, only 15% completed, and the outcome did not favor either intervention (OR 1.00; 95% CI 0.19-5.40). The second ED trial included adults with bipolar disorder, 35% of whom had a history of substance abuse in the 6 months preceding the trial (340).. The comparisons were olanzapine plus (divalproex or lithium) vs monotherapy olanzapine, monotherapy divalproex or monotherapy lithium or ongoing treatment with combination of olanzapine and (lithium or divalproex) (N=99 patients). In this trial, 21% of cases completed the planned 18 month trial. Cases allocated to the phase 1 intervention in phase 2 had a higher chance of remaining in the trial without relapse than those assigned to monotherapy (OR 3.93; 95% CI 1.31-11.80). One classic RCT randomized manic patients to olanzapine or divalproex and followed responders for 1 year. 15% of the inception cohort completed the trial, which did not favor either intervention (OR 1.01; 95% CI 0.32-3.22).

3.3b *Indirect comparisons*

Among the identified mood stabilizer maintenance reports, no 'direct comparison' trials were found. Comparisons below are therefore all of the indirect type.

Lithium vs Placebo Indirect comparisons:

There were no 'direct comparison' placebo controlled lithium trials. For indirect comparisons, Prien, Stallone and Coppen (70, 268, 324) were considered suitable for combining, as they all used lithium enrichment on a similar, RDC-defined bipolar population. The trials by Kane (179) and Fieve (110) used a classic RCT design and did not allow use of lithium prior to the study. Bowden, Calabrese et al (41, 42, 50) used divalproex or lamotrigine enrichment, respectively and both trials used time to relapse as their primary outcome.

To test the impact of using different interventions in phase 1 as opposed to phase 2, placebo controlled trials using lithium in phase 1 were combined (figure 4, labeled as 'lithium enrichment') as well as those using a competitor in phase 1 with randomization to lithium ('lamotrigine enrichment' in figure 4) and trials using classic RCT designs (designated in figure 4 as 'no enrichment'). Also included in figure 4 is an indication of potential diagnostic heterogeneity between the trials, specified as 'RDC or Feighner criteria' or DSM-IV criteria (more on this in the discussion).

As illustrated by the indirect comparisons in figure 4, the use of phase 1 lithium 'enrichment' (OR = 22.0; 95% CI 7.0-68.7) is associated with a significantly larger effect of lithium relative to placebo in phase 2, compared with trials in which 'lamotrigine enrichment' is used in phase 1 with a subsequent randomization to lithium or placebo in phase 2 (OR = 1.9; 95% CI 1.2-2.8). The classic

RCT group shows an intermediate effect between the lithium enriched and anticonvulsant enriched trials (OR = 3.18; 95% CI 0.65-15.46).

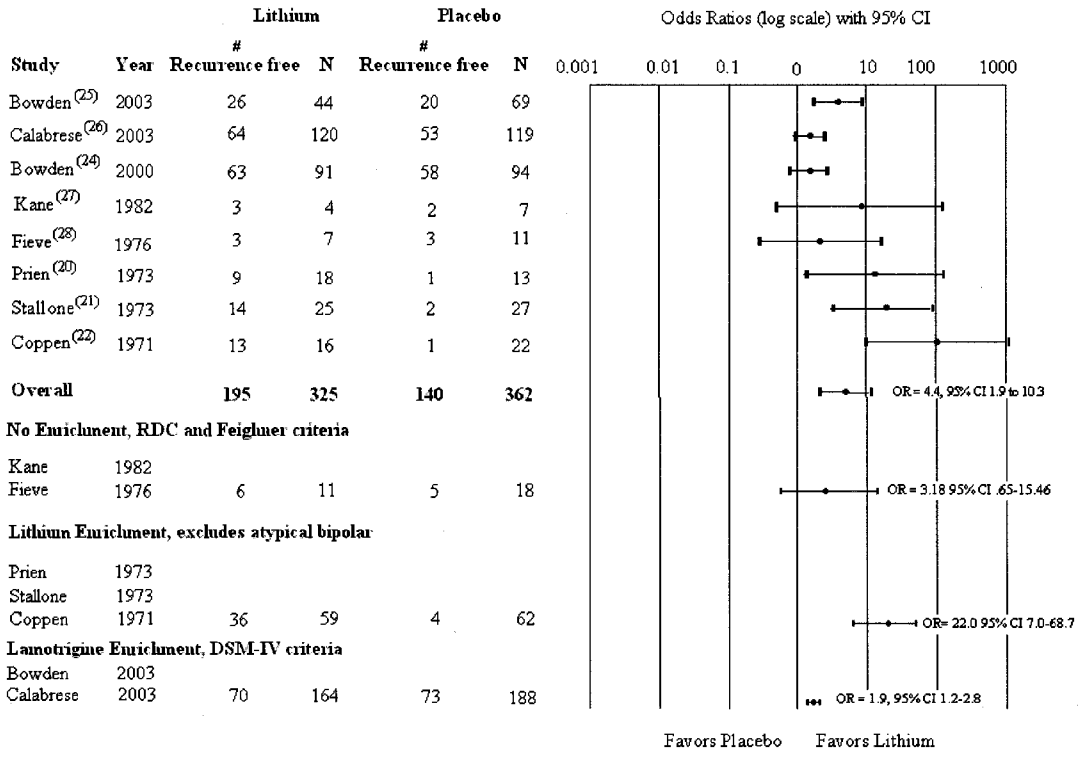


Figure 4: Indirect comparisons: lithium monotherapy vs placebo trials. Figure shows that lithium enriched trials strongly favor lithium while anticonvulsant (lamotrigine) enriched trials randomized to lithium or placebo show a lower effect. Classic RCT's comparing lithium with placebo are intermediate to lithium enriched and anticonvulsant enriched trials. (Please note reference numbers refer to bibliography of Deshauer et al 2005 (89) attached at end of thesis chapter 6).

Publication Bias; Sensitivity analysis

As illustrated in figures 5 and 6 respectively, this cohort effect can also be explained by publication bias. Figure 5 shows that small trials with large effect sizes are unbalanced by small trials showing small effects. Figure 6 shows inconsistent trial quality, with lower quality trials yielding larger effect sizes. The quality assessments were generally higher in more recent trials than in older trials. We

rated patient blinding as inadequate for all trials, since no attempt was made to address the problem of patients' identifying their allocation to placebo or active medication. All patients reaching the double blind phase had received active medications known to have easily identifiable side effects. In the context of antidepressant trials, an analogous situation may contribute to a 'wish bias' against placebo treatment (27).

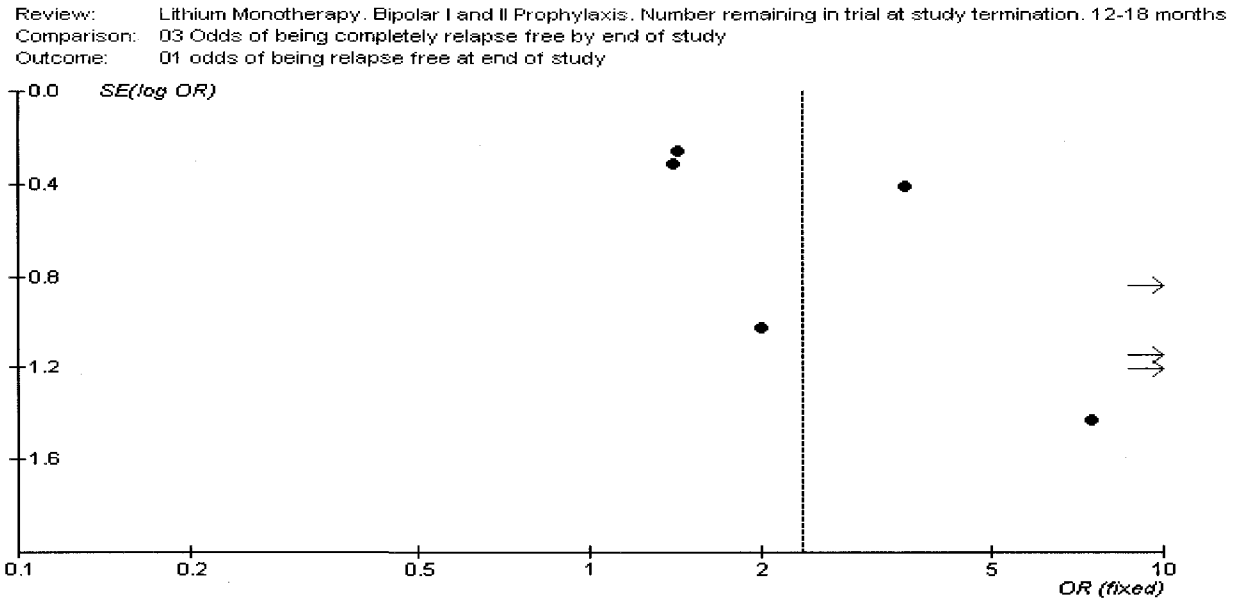


Figure 5: Asymmetric funnel plot consistent with publication bias.

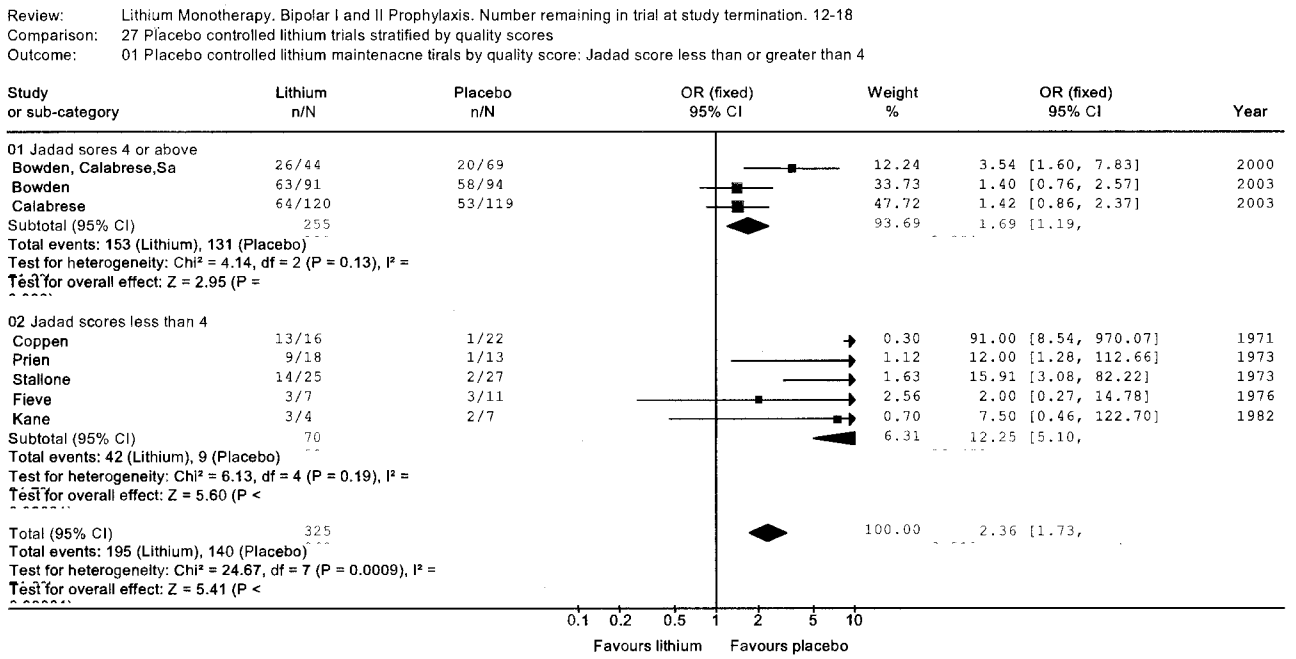


Figure 6: Alternate 'quality' explanation for cohort effect among placebo controlled lithium reports, grouped by Jadad score (<4, 4 or more) showing greater effect size with lower quality score.

Lithium vs anticonvulsants Indirect comparisons:

A total of 8 trials were available for indirect comparisons involving lithium and an anticonvulsant. These included three classic RCT's ('non-enriched' in figure 7) (157, 217, 317) (Hartong 2003, Luznat 1988 and Small 1991), in which lithium or carbamazepine were assigned during an acute decompensation at the time of admission to hospital and all responders to initial treatment allocation were followed forward. The overall comparison between lithium and carbamazepine failed to demonstrate superiority of either intervention (OR =0.80; 95% CI: 0.36-1.79). ED trials using lithium enrichment and randomizing to lithium or anticonvulsant did not favor of lithium (OR 1.27;

95% CI 0.52-3.10). Similarly, ED trials using anticonvulsant enrichment with randomization to anticonvulsant or lithium did not favor anticonvulsant maintenance in phase 2.

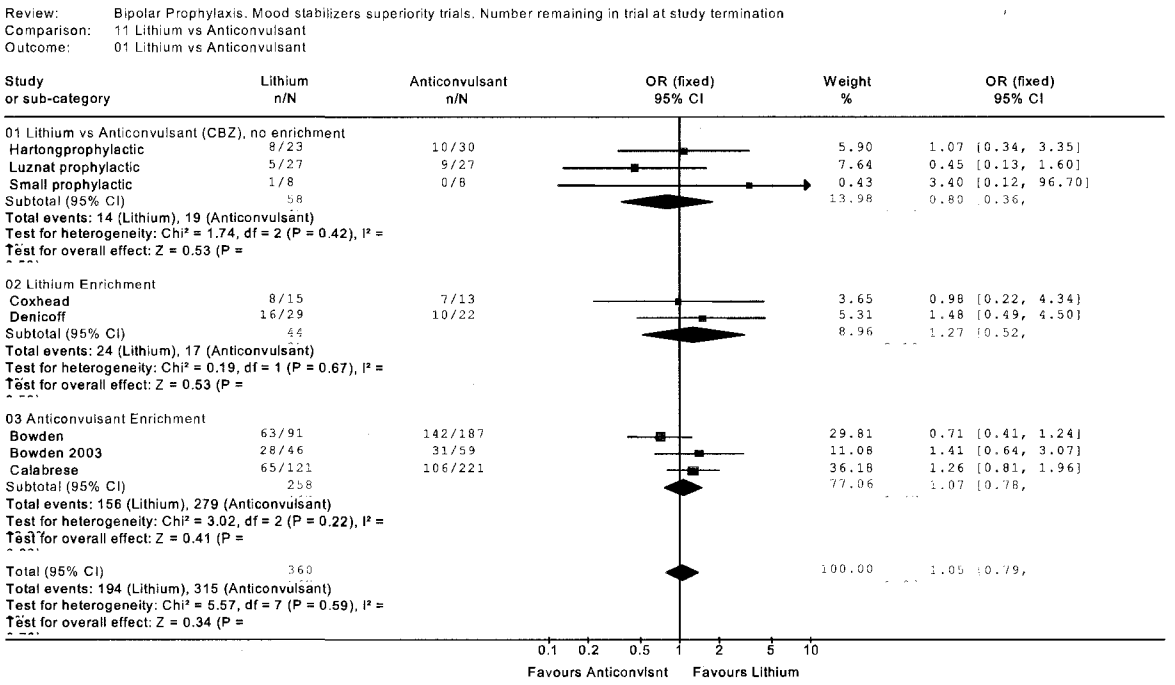


Figure 7: Indirect comparisons: lithium vs anticonvulsant; classic RCT design, lithium enrichment and anticonvulsant enrichment. No indication of bias favoring phase-1 interventions

3.4 Discussion:

In this chapter, we attempted to use published trials of drugs commonly applied in the maintenance therapy of bipolar disorder to construct a comparative framework similar to that proposed by Greenhouse et al (143), to test for bias in ED designs. Because of the relatively small number of maintenance trials available for the task, we were limited to indirect comparisons of placebo controlled lithium trials and lithium-anticonvulsant comparisons. ED designs were used in the majority of mood stabilizer maintenance trials conducted over the past 3 decades and in all maintenance trials for bipolar disorder conducted since 1995, limiting the classic RCT comparisons to five small reports.

We hypothesized that placebo controlled lithium trials in which phase 1 lithium 'responders' ('lithium enriched' trials) were randomized to ongoing lithium or placebo in phase 2 would favor lithium relative to ED trials in which phase 1 consisted of anticonvulsant 'responders' ('anticonvulsant enriched') who are then randomized to lithium or placebo in phase 2. Among placebo controlled lithium maintenance trials, the use of lithium enrichment in phase 1 with randomization to placebo or ongoing lithium in phase 2 is associated with a significantly larger effect relative to placebo than is observed in trials using anticonvulsant enrichment followed by randomization to lithium or placebo. This is not replicated among head to head trials comparing lithium and anticonvulsant under alternative enrichment strategies, suggesting that alternative explanations for the apparent 'enrichment effect' in placebo controlled lithium trials should be considered.

Alternative explanations for an apparent 'enrichment effect' among placebo controlled lithium maintenance trials include publication bias among early reports (all of the pre-1982 reports were of small trials, unbalanced by reports failing to demonstrate superiority to placebo), secular changes in healthcare and in trial recruitment thresholds (41, 146) (there is an 18 year gap in reports of placebo controlled lithium trials between 1982 and 2000), and inconsistent trial quality including the absence of tapering in trials before 1982[7]. For example, we have replicated the 'enrichment' effect in placebo controlled lithium trials, stratifying trials by quality scores, yielding a cohort effect favoring lithium in small, lower quality reports; OR = 11.71 (95% CI; 3.56-38.52) vs OR = 1.79 (95% CI; 1.06-3.01).

A further explanation is diagnostic drift, small changes in the way bipolar patients are classified over the three decades covered by the review. Earlier trials used the RDC or Feighner criteria, which may exclude cases that are currently included as 'bipolar' under the DSM-IV criteria. The clinical relevance of small changes in diagnostic criteria to treatment response in bipolar disorder is controversial (148).

The adequacy of blinding is an important component of all RCT's, and inadequate blinding has been associated with bias favoring investigational drugs (302). ED trials raise special problems with blinding, since all patients are exposed to the open phase

[7] In the course of preparing this review, we came across reports of lithium withdrawal-rebound symptoms suggesting that the early, enthusiastic reports of lithium's efficacy may have been due in part to acute medication withdrawal (63, 143, 239). All three larger trials used a tapering period.

intervention for an extended period of time before phase 2 randomization. Potential limitations with blinding inherent to ED trials may mean that both patients and investigators, through observing prominent and often distinctive side effects (ex. lithium), are effectively unblinded in phase 2. This in turn has the potential to interact with patient-related factors (wish bias (27)) as well as investigator-related factors (for example, conflict of interest). The latter is of potential relevance to ED trials of mood stabilizers. Since 1995 all trials included in our report were industry sponsored, and 70% of the 2307 patients included in published reports since 1973 have been enrolled in industry-sponsored trials.

High dropout rates are universal among the maintenance trials reviewed here, and must be kept in mind in all discussions of the generalizability of long-term psychiatric drug trials. Phase 1 drop-outs ranged from 35-50% and post randomization drop-rates ranged from 70-85% (of the remaining inception cohorts). These rates exceed the 15% maximum recommended by the Agency for Health Care Policy Research for studies over 3 months (9). High drop-out rates render trials difficult to interpret and raise questions about differential attrition (responder or 'tolerator' selection). One approach to detecting differential attrition in ED trials is to look at phase 2 drop-out rates for reasons other than relapse. If the ratio of drop-outs for reasons other than relapse between ED trials tends to favor the enrichment drug, this may suggest a 'tolerator selection' effect. We used this approach with the available data (figure 8, Appendix) and found a non-significant trend toward lower drop-out rates in phase 2 of ED trials favoring enrichment drugs (ie. favoring lithium in lithium enriched trials randomized to ongoing lithium or placebo and

favoring placebo in trials in which anticonvulsant was used in phase 1 followed by randomization to lithium or placebo) (43, 50).

A number of general methodological issues limited both our data gathering process and the inferences we could draw. As an observational study, this study is prone to selection bias and observer bias, particularly since one reviewer did the majority of trial identification and data abstraction. It is also possible, particularly in the reviews of carbamazepine, lamotrigine, olanzapine, quetiapine and risperidone that our strategy for identifying systematic reviews missed publications, since the term 'systematic review' has not yet been indexed in the National Library of Medicine (247). The potential disadvantages of the strategy are however counterbalanced by the high quality of Cochrane reviews, which were included in the identified reports. Finally, our search is now over a year old and may have missed recent contributions to the literature.

At the level of the data, alternative treatment strategies were sparse, limiting available comparisons to three classic RCT's comparing anticonvulsant and lithium, two lithium enriched ED trials in which lithium enriched cohorts were randomized to ongoing lithium or anticonvulsant substitution, and three ED trials in which anticonvulsant enriched cohorts were randomized to ongoing anticonvulsant, lithium or placebo. If selection effects were large, we may have seen outcomes favoring the enrichment compound. The fact that this was not observed suggests either that specific selection effects, if present, are confounded by factors including overlapping efficacy between active comparators, delayed time to effect for comparators introduced after phase 1, and differential harms.

Whether using a different outcome, such as time to relapse would have changed the results is unknown. Using the number of cases relapse-free at trial completion, may be too coarse to pick up small phase 1 selection effects between drugs, since it provides no information about the majority of cases who relapse or drop out before the planned trial endpoint (208).

3.5 Summary

We used data from published mood stabilizer trials in a series of comparisons testing for bias in ED designs applied to the population of bipolar patients treated with drug maintenance therapy. ED designs dominated the evidence base, and there was a paucity of classic RCT's for drawing comparisons.

Among placebo controlled lithium maintenance trials, it appeared that the use of ED designs favors the enrichment drug relative to placebo, but this finding was not replicated in a set of active comparator trials of lithium and anticonvulsants. A number of plausible alternative explanations for the observed effect were proposed. We questioned whether our chosen outcome was sensitive enough to detect 'enrichment effects' stemming from the use of ED designs. In addition, we questioned the confounding effects of inconsistent trial quality on the published data.

Appendix

Review: Lithium Monotherapy. Bipolar I and II Prophylaxis. Number remaining in trial at study termination. 12-18 months
Comparison: 25 Drops for reasons other than relapse
Outcome: 01 Drops for reasons other than relapse. Lithium vs placebo

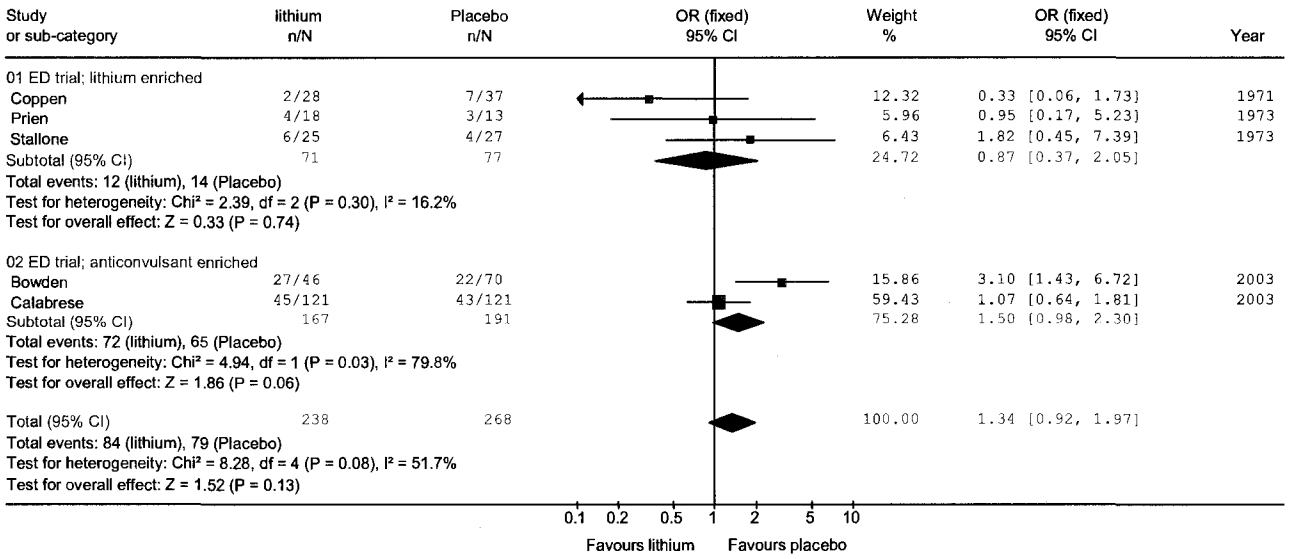


Fig. 8: Phase 2 drop-outs for reasons other than relapse in placebo controlled lithium trials under alternative enrichment strategies. Trials with anticonvulsant enrichment in phase 1 show higher lithium drop-outs relative to placebo in phase 2 than trials using lithium enrichment in phase 1. This would be consistent with a phase 1 'tolerator selection'.

Chapter 4

Enrichment-Discontinuation designs in antidepressant maintenance trials

4.1a Introduction:

Major depression:

Major depression, defined by RDC, ICD or DSM criteria, is a complex syndrome with an incidence of approximately 3/1000 per year (Epidemiological Catchment Area Study; Baltimore estimate based on 15 years) (97), a lifetime prevalence of 16% (191) and a 30 day prevalence of 4.9% (National Comorbidity Survey) (32). The gap between incidence and prevalence reflects chronicity and recurrence (166). Among patients receiving standard treatment in secondary or tertiary care psychiatric settings, recurrence rates are 41% at 12 months, 59% at 24 months and 74% within 5 years (321).

The concept of a specific drug treatment for the syndrome of depression evolved gradually over the first half of the 20th century and is well documented in the historical literature (161). Less well documented is the evolution of drug treatment from a time-limited intervention reserved for the most severe cases (56, 366) to the current view of lifetime maintenance for the 'prevention' of future episodes. Maintenance is defined as the period extending beyond 16 weeks after complete symptom resolution. The 12-16 weeks following recovery is designated 'continuation therapy'(117).

The RCT evidence supporting maintenance antidepressants relies heavily on trials using ED designs. These trials report 12-24 month recurrence rates of 20% among cases

receiving ongoing medications and 40% among patients randomized to placebo discontinuation in phase 2 (129).

Historical overview of drug maintenance therapy for depression:

MAOI'S:

Of the three classes of antidepressant medication covered in this thesis, the Monoamine Oxidase Inhibitors (MAOI's) are the oldest group of drugs, having evolved from tuberculostatic agents. Their application in psychiatry increased in the 1950's following an article reporting improved moods among tuberculosis patients treated with iproniazid (161). MAOI's have been largely supplanted due to serious adverse effects, forming a third or fourth line treatment for cases not responding to less harmful interventions.

TCA's:

The second largest group of antidepressants, the tricyclic antidepressants (TCA's) first became available in 1957 with the launch of imipramine. This was followed by the launch of its metabolite desipramine, as well as amitriptyline and its metabolite nortriptyline. Tricyclic antidepressants dominated the drug treatment of depression for almost 3 decades until the launch of fluoxetine (Prozac) in 1987. TCA's have more adverse effects than the group of SSRI's, and their use is declining.

SSRI's:

SSRI use has increased dramatically over the past decade in North America. From 1991 to 1997, the annual number of antidepressant prescriptions written by primary care providers in the United States increased from nearly 25 million to over 50 million, while

psychiatrist initiated prescriptions increased from 15 to 33 million. (Coyne p. 799) (74, 163). The prevalence of maintenance antidepressant use now exceeds the point prevalence of depression in some populations (74, 182, 301).

4.1b Chapter Objectives:

4. Identify the prevalence of ED designs relative to classic RCT's in the evidence base supporting drug maintenance for major depressive disorder.
5. Using direct comparisons and indirect comparisons as outlined in chapter 2, test for bias relative to continuation or maintenance trials using the classic RCT design, in which cases recovering while on placebo are compared with cases recovering on medications.
6. Using indirect comparisons, identify possible responder bias favoring phase 1 run-in interventions where the phase 1 and phase 2 interventions differ.

4.2 Methods

4.2a *Participants:* All patients with a diagnosis of DSM-IV, DSM-III, III-R, RDC major depression or major depressive disorder, No age restrictions. Inpatient or outpatient settings.

Interventions: Drug treatment of a minimum 20 weeks following recovery from the most recent episode of depression. No co-interventions permitted.

Comparators: Monotherapy SSRI (fluoxetine, sertraline, fluvoxamine, paroxetine, citalopram, escitalopram, venlafaxine), buproprion, MAOI including moclobemide,

tranylcypramine, TCA (imipramine, desipramine, nortriptyline, amitriptyline, clomipramine, doxapine), placebo.

Outcomes: Binary outcomes; Cumulative relapses reported as last observation carried forward (LOCF), Drop-outs due to adverse effects, Drop-outs for all reasons other than recurrence.

Study design: RCT's including RCT variants: placebo controlled and active comparator.

4.2b Search strategy and data abstraction

SSRI continuation and maintenance trials of at least 20 weeks duration were identified from an exhaustive systematic review of all available SSRI trial reports (including unpublished trials) by Fergusson et al 2005 (107). Continuation and maintenance trials of tricyclic antidepressants and Monoamine Oxidase Inhibitors of at least 20 wks were identified from a high quality systematic review by Geddes et al (129). Trial bibliographies were checked for further references.

In addition to trials from Fergusson et al (107) and Geddes et al (129), MEDLINE was searched through the OVID interface using the key words Depression AND [fluoxetine or citalopram or escitalopram or sertraline or paroxetine or fluvoxamine or venlafaxine or duloxetine or bupropion or moclobamide or hypericum or desipramine or clomipramine] between January 2004 and April 2006.

Standardized case report forms were used for data abstraction and the Jadad scale for randomization, blinding and description of withdrawals (36) as well as the Schulz

component for allocation concealment (37) were used. A request for unpublished data providing categorical outcomes for one classic RCT with a continuation phase for responders (323) was partially fulfilled by the manufacturer. Further requests for unpublished survival data from two 'direct comparison' trials (101, 244) were declined. The search, data abstraction and quality rating was done by one reviewer (DD).

Data synthesis

Using Revman version 4.2, a meta-analysis software package, binary outcomes were combined to give an overall effect estimate. We assumed that the probability of relapse was approximately constant between 6-12 and 12-24 months after the last episode (137), and therefore felt it would be valid to combine these outcomes. We treated the SSRI's as a class for indirect comparisons, since they have been shown to be of similar efficacy across a wide array of clinical settings (20). We also combined the serotonin norepinephrine inhibitor (SNRI) venlafaxine with the SSRI's, even though there are claims that venlafaxine is marginally more efficacious than the broad class of SSRI's in the acute management of depression (318). TCA and MAOI reports were combined because they were conducted in the same decade and are believed to have similar efficacy (20). Early dropouts were recorded where possible. The data from the primary studies was combined using a fixed effects model if I^2 < than 50%. Comparisons were constructed using the method outlined in chapter 2.

4.3 Results

4.3a Identified Trials

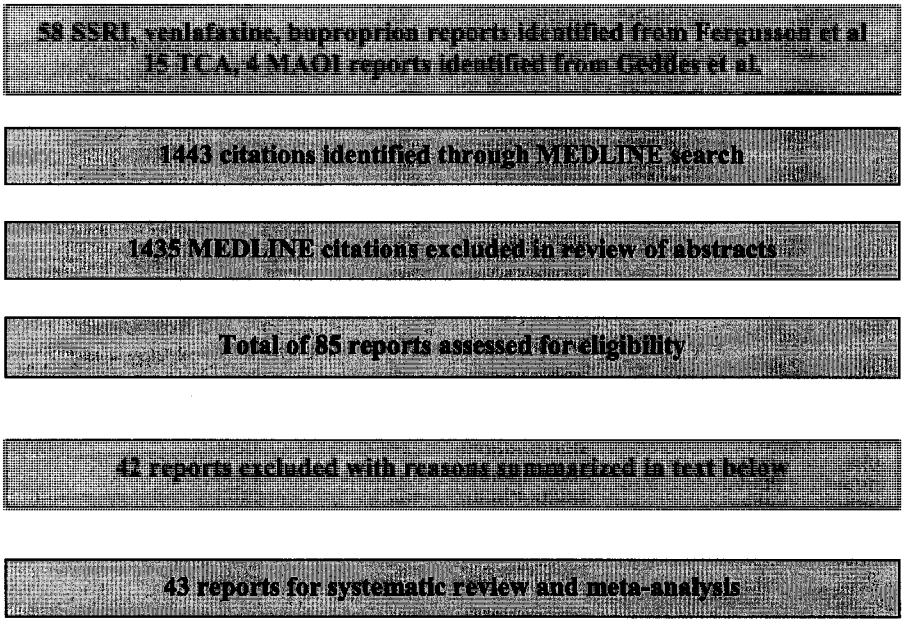


Figure 1: Flow diagram for selection of antidepressant continuation and maintenance RCTs. References to Geddes, Fergusson and all included and excluded trials along with reasons for exclusion are provided in the text below.

From Fergusson et al (107), reports from 58 placebo-controlled trials of SSRI's, venlafaxine and bupropion greater than 20 weeks were identified by DF and reviewed by DD. From Geddes et al (129), reports of 15 TCA trials and 4 MAOI trials were identified and reviewed (28, 69, 71, 80, 117, 131, 138, 155, 178, 184, 194, 196, 200, 236, 255, 267, 279, 283, 326). Of the reports from Fergusson et al, thirty five were excluded based on psychiatric diagnoses other than major depression (33, 34, 46, 62, 68, 76, 78, 79, 127, 141, 145, 158, 185, 199, 204, 224, 225, 234, 252, 273-275, 280, 285, 286, 290, 293, 305, 325, 330, 342, 351, 352, 357, 360). This left 23 trials (29, 85, 94, 101, 106, 130, 134,

165, 169, 190, 192, 195, 209, 241-244, 276, 277, 282, 289, 295, 315, 323, 335, 348, 358, 363) from Fergusson et al. Of the reports from Geddes et al, four were excluded; one MAOI trial due to inadequate duration (80) ; one MAOI trial due to the study diagnosis of dysthymia (155); one amitriptyline trial on the basis of inadequate reporting (326), and one imipramine trial (200), which was a 5 year follow-up of another trial already reported in this review (117) leaving 15 reports included (28, 69, 71, 80, 117, 131, 138, 155, 178, 184, 194, 196, 200, 236, 255, 267, 279, 283, 326).

From the MEDLINE strategy, 1443 citations were identified and abstracts were electronically reviewed. From these abstracts, 8 reports were selected for closer review (25, 101, 209, 243, 272, 276, 307, 315), of which 3 were excluded; 2 due of a lack of placebo arm (25, 307), and one due to inclusion of 'minor depression' (272). This left a total of 43 included trials. Enrichment-discontinuation designs accounted for 41 of 43 maintenance trials, two of which incorporated a discontinuation arm to an acutely randomized trial (101, 244). Tables 1 and 2 summarizes all included reports.

Table 1: Summary of all included TCA and MAOI trials

TCA/MAOI Trial (Duration post randomization)	TCA Maintenance Trials, Demographics, Clinical Features
Reynolds 1999 (1-3 yrs) (279)	Recurrent depression, non-psychotic, non-dysthyimc;Universtiy geropsych clinic; 1yr cumulative recurrences at 1,2,3 yrs; Age over 60, (mean 67), 25%M, 75%F;
Kocis 1996 (2 yrs) (196)	Chronic depression DSM IIIR dysthymia, double depression, chronic depression outpatients; 55%F, 50% axis II, 35% have anxiety disorders; Mean HDRS 20, mean age 37, 15% married, Half had never tried antidepressants. 28 wk acute and continuation open phases. Med taper was used after randomization.
Old age group 1993 (2 yrs) (255)	Major Depression, first episode not excluded. 15 centers, Age >60, 73% women; outcomes available across wide number of months
Robinson (Prien) 1991 (2yrs) (284)	RDC depression referred to Marshall university Outpatients for, at least 1 prior episode. HAM-D >18, GAS<60, recent episode>20wks ago exclude history of mania, substance in prior year, mean age 43, M=9 F=38
Frank 1990 (3yrs) (117)	RDC Recurrent Major depressive disorder 14% BPII, HRSD >17, Raskin >7, ; 3 yr intended f/u. only 17% starting maintenance phase failed to complete cumulative recurrences available at 1,2,3yrs
Georgotas 1989 (1yr) (131)	Major depression, HAMD>16, University hospital outpatients. Age >55;mean age 65; 28F, 23M No exclusion for 1st episode
Cook 1986 (6 months) (69)	Major Depression. Elderly VA outpatients, mean age 63. Treated with TCA>1yr without recurrence. N=15 All Males; 6 month f/u
Glen 1984 (3yr) (138)	Recurrent 'primary' unipolar depression 9 UK centers, 71 psychiatrists, Intended 3yr; median age 51 M=29 F=107 % relapse free available for every 6 months to 36 months
Prien, 1984 (2yrs) (267)	RDC Major depressive disorder, recurrent affective disorders including bipolar disorder. 24% inpatients; Age 21-60, Raskin score>7, GAS<60 at entry, NIH study 5 US centers,. M=113 F= 230 2 yr maintenance study outcomes at 1 and 2 yrs
Kane, 1982 (2yrs) (179)	RDC recurrent unipolar. Outpatients Mean age 47.5, M=10, F=17; intended 2yrs.
Bialos, 1982 (6 months) (28)	Major depression in long-term remission on amitriptyline (ave 3.7 yrs). M=14, F=3. Placebo substituted for amitriptyline in half. 6 months f/u Mean age=57
Coppen 1978 (1yr) (71)	Major depression; 23 inpatients and 9 outpatients treated with amitriptyline for depression. 11 first episode, 21 recurrent; mean age 53.. M=16, F=16
Klerman 1974 (9months) (194)	DSM III neurotic depression; Outpatients, Two large US centers; Median age 'late 30's' Raskin score min7; 278 females with. No exclusion for first episode Most had 1 episode, 5 bipolar.
Mindham 1973 (6 months) (236)	Moderate to severe MDD fully remitted on imip or ami. 3-10 wks. In or outpt. 60% had no prior depressions 8 UK centers, Mean age 47. M=33 F=59
Kay 1970 (6months) (184)	Post-ECT maintenance for unipolar depression 7UK hospitals, mostly in-patients age 20-75; 50% between 40-59; M=48 F=84

Table 1: continued

TCA/MAOI trial	Comparisons	Same intervention in phase 1 and phase 2?	When randomized?
Reynolds (1999)	Nortrip, IPT, med clinic + placebo, IPT+nortryp Maintenance 3 years planned	Yes. Nortryp + IPT, 16wks.	After unspecified acute period and 16 wks continuation with Nortryp + IPT
Kocsis (1996)	Desipramine vs Placebo	yes, open 28 wks.	after 28 wks
Old age group (1993)	Dothiapin, Placebo	8wks continuation on open compound then to Dothiapin or placebo	Any compound, ECT allowed acutely (duration unspecified). Randomized after 8wks stable on meds.
Robinson (Prien) (1991)	Phenelzene 60,45, placebo	Yes, Phenelzene at entry and 16 wk continuation	After 16 wks continuation
Frank (1990)	5 arms, IPT, IPT+imip, imp, IPT + placebo, placebo, mean age 40 22%M 78%F	After recovery of min. 20 wks following ongoing treatment with imip + IPT	after min 20wks acute and continuation (10wks euthymic)
Georgotas (1989)	Nortip, Phenelzene, Placebo	Yes (TCA or MAO given acutely). 23 wks open	Complex design. 7wks acute open, 16 wks continuation open then randomized.
Cook (1986)	TCA (dox, ami, imi, des) vs Placebo	All were stable long-term then randomized to placebo	After long-term wellness (unspecified, over a year)
Prien, (1984)	Imip, Li, Placebo	Yes, Imip + Li for 2 months then randomized	after acute preference phase (treating psychiatrist chose). Duration unspecified
Glen (1984)	Amitriptyline, Li, Placebo	Open phase at psychiatrist preference. 73% received tricyclic ad's	1-40wks, median 8.5wks after acute episode
Kane, (1982)	Imip, Li, Placebo	yes, 6wks open imip. 6 months open uncontrolled treatment (unspecified)	after min. 6 months euthymic
Bialos, (1982)	ongoing amitriptyline maintenance vs placebo mean age 57	Yes. average 3.7 yrs on amitrip maintenance prior to discontinuation	in long-term stable maintenance
Coppen (1978)	Amitriptyline, placebo	yes, only responders to amitrip were included	after recovery (HAM-D <6 for at least 6 wks)
Klerman (1974)	Amitriptyline, placebo 4-6wks	Yes; 6wk open amitriptyline	after 6wk acute phase
Mindham (1973)	amitryp, imip, placebo	Yes 3-10 month run-in imip or amitryp	after successful 3-10 month run-in response
Kay (1970)	Amitriptyline, Diazepam	No (immediately followed ECT)	After ECT

Table 1: continued

TCA/MAOI trial	N acute	N continuation	N maintenance (if distinction made)	Pre-discontinuation Taper
Reynolds 1999	180	140	124	6wk
Kocis 1996	129	66	50	Yes, 4 wks
Old age group 1993	219	unknown	69	Not stated
Robinson (Prien) 1991	88	73	47	Yes, 3wks
Frank 1990	280(5arms)	157	128	Yes, 3wks
Georgotas 1989	Unknown	unknown	52	Not stated
Cook 1986	Unknown	n/a	15	Yes 4 or 8 wks
Glen 1984	Unknown	136		Yes, 2 wk
Prien, 1984	343	150complted 8 wks	99 finished 8wks	Not stated
Kane, 1982	Unknown	unknown	27	Not stated
Bialos, 1982	Unknown	unknown	17	Yes, 3 wks
Coppen 1978	Unknown	32		Not stated
Klerman 1974	278	150		Not stated
Mindham 1973	not stated	92		Yes, duration unclear
Kay 1970	132			N/A

TCA/MAOI trial	% completing open phase	% completing Randomized phase	How cumulative relapses reported	Cumulative relapse	# completing
Reynolds 1999	68.9	Unclear	LOCF	Yes	no
Kocis 1996	81	63	LOCF	yes	no
Old age group 1993	31.5	Unclear	LOCF	yes	no
Robinson (Prien) 1991	83	18	LOCF,CA	yes	no
Frank 1990	Unknown	82	LOCF	yes	yes
Georgotas 1989	Unknown	Unclear	Unclear	yes	yes
Cook 1986	Unknown	83	CA	yes	no
Glen 1984	Unknown	28	LOCF	yes	no
Prien, 1984	28.9	37	LOCF,CA	yes	yes
Kane, 1982	Unknown	59	LOCF	yes	yes
Bialos, 1982	Unknown	89	CA	yes	yes
Coppen 1978	Unknown	Unclear	Unclear	yes	yes
Klerman 1974	54	72	Unclear	yes	yes
Mindham 1973	Unknown	Unclear	LOCF	yes	no
Kay 1970	Unknown	Unclear	LOCF	yes	yes

LOCF=Last observation carried forward

Table 1: continued

TCA/MAOI trial	Survival analysis? Mean, median time, SD?	Adverse effect reporting	Schultz	Allocation concealment
Reynolds 1999	graph only	yes, includes mania	4	Adequate
Kocis 1996	graph	yes	4	Adequate
Old age group 1993	graph only. Cox regression	no	3	Unclear
Robinson (Prien) 1991	graph only	yes	3	Unclear
Frank 1990	yes, median time +/- SE	some; specifies no mania	3	Adequate
Georgotas 1989	graph only	some	2	Unclear
Cook 1986	no	no	2	Unclear
Prien, 1984	graph only	yes, detailed reporting	4	Adequate
Glen 1984	graph only	yes, 2 cases of mania	4	Adequate
Bialos, 1982	graph only. Half recurred in 3-15 wks	no	2	Unclear
Kane, 1982	no	no	2	Adequate
Coppen 1978	no	yes, detailed reporting	2	Adequate
Klerman 1974	no	yes, some	2	Unclear
Mindham 1973	graph only	detailed reporting	2	Unclear
Kay 1970	no	yes, 2 mania, 3 suicide	2	Unclear

Table 1: continued

TCA/MAOI trial	Outcomes	Number completing/#randomized	Funding
Reynolds 1999	Clinical endpoint. Multiple scales done	36/107 completed 3 yr study	NIH
Kocis 1996	clinical, scale		NIH
Old age group 1993	MADRS>10	58/79 randomized completed 2yrs	Industry
Robinson (Prien) 1991	Clinical endpoint. Multiple scales done	8 of 47randomized	NIMH
Frank 1990	clinical recurrence + HRSD >15	22/128 completed 3 yrs	NIMH
Georgotas 1989	HAM + clinical recurrence	19/51 completed 1 year	NIMH
Cook 1986	Clinical endpoint. MADRS, HAMD,	6/15 completed 6 months	Unclear
Prien, 1984	clinical depression + GAS<60	50/150	NIMH
Glen 1984	clinical relapse requiring more than benzos	18/67 in amitrip or placebo finished	MRC
Kane, 1982	clinical, use of meds	All have follow-up (N=12) at 1 year	Foundation
Bialos, 1982	clinical determination and self-report	N=17 full f/u at 6 months	Unknown
Coppen 1978	recurrence requiring hospitalization	N=32. Full f/u for 12 months	MRC
Klerman 1974	clinical relapse, HAMD, Hopkins scale	106/150 completed 9 months	NIMH
Mindham 1973	meds, clinical judgment	unclear	MRC
Kay 1970	clinical relapse	53/132 completed	Foundation

Table 2: Summary of SSRI trials

SSRI trial (duration post-randomization)	SSRI Placebo-controlled Maintenance Trials: Overview
Montgomery (2004) (1 yr) (243)	DSM-III R Major Depressive Disorder responding to 6 months open venlafaxine; >1 episode in last 5 yrs. 75 had 3 or fewer episodes in 5 yrs (including index episode); Mean age 43; M=62 F=138 Multicenter Europe and USA. Chloral hydrate or benzos permitted as rescue meds. Inter-rater meetings arranged.
Rapaport (2004) (36 wks) (276)	DSM IV Major Depressive Disorder; designated as recurrent depression but first episode not explicitly excluded. Maintenance trial follows 8wks RCT on escitalopram. Responders (incl.placebo responders) got 8 wks open escitalopram; 36 wk maintenance. Age 18-81; mean age 42; Maintenance randomized M=107 F=167; 53 US centers, (9 month maintenance double blind phase)
Emslie (2004) (32 wk) (101)	Child and adolescent depression, multicenter US (15 sites), mean age 13, 1 wk placebo run-in, Authors note potential rebound in maintenance phase placebo. 9 wk open, 10 wk dose titration, 32 wk maintenance. Both analyses available. Acutely randomized
Simon (2004) (6 months) (315)	DSM IV Major Depressive Disorder 21 item HAM-D over 19, multicenter 4-10 day placebo run-in, then open 8wks venlafaxine XR flex dose 75-225. Responders after 8 wks HAM-D <10, CGI <4, 2wk taper. Mean age (maintenance) 42, max age 79, F=188 M=104; number of prior episodes not listed; 6 month trial
Lepine (2004) (18 months) (209)	Recurrent Major Depressive Disorder, in sustained remission treated with medication other than sertraline. Approximately half had 3-5 lifetime episodes of depression. Half had more than 5 prior episodes. Before entering washout period, 75% were taking antidepressants, mostly fluoxetine. 2 month open placebo run-in, then 18 months sertraline or placebo. Mean age randomized 46, M=85 F=203. French multicenter
Gelenberg (2003) (13 months) (130)	Recurrent MDD with incomplete remission or dysthymia. Entry HAMD-24 of 20. Mean duration of MDD 8 years, dysthymia mean 23 years. N total (randomized) = 160; mean age 44, F= 108; multicenter secondary and tertiary care patients; USA; Randomized acutely to open nefazodone plus psychotherapy or nefazodone without psychotherapy or psychotherapy alone. Responders to either nefazodone or nefazodone plus psychotherapy randomized to placebo discontinuation and 52 week follow-up.
Wilson (2003) (2 yr) (363)	Major Depression; age over 65, mean 78; outpatients. UK multicenter; If HRSD <50%, then 20 wks continuation phase. Randomized M=33, F=80; open phase 8 wks sertraline. ** This was the first episode for 72% of participants.
Weihs (2002) 11 months (358)	Recurrent Major Depressive Disorder; HAMD-21 item score min. 18. Duration of current episode 2-6 months in 53%; number of prior episodes = 1-2 in 45%. USA multicenter. F=66%. 8 wk acute open treatment followed by placebo discontinuation, 11 month follow-up.
Klysner (2002) (1 yr) (195)	Unipolar Major Depressive Disorder MADRS >22. 93% had no prior episodes in the past 5 years outpatients >age 65; mean age 75. Single center Denmark. 20 or 30 mg citalopram. Open citalopram 8wks then 16wks open continuation then 48wks randomized M=36 F=154
Hypericium study group (2002) (18 wks) (169)	DSM IV Major Depression. Not excluding 1st episode. 340 Outpatients started acute RCT. Minimum HAM-D score 20 (moderate depression). Responders to any arm of the acute trial (N=129 of 245 completing the first 8 wks) continued for a further 18 wks.
Gilaberte (2001) (1 yr) (134)	Recurrent Major Depressive Disorder, DSM IIIR; min 1 prior in past 5 yrs, Mean age 44, 78% F, prev ep 2.3, SD1.2, prev suicide attempts, 7.1%, 40% fam Hx MDD, HAMD baseline 24, CGI baseline 5, 10 sites, Spain; 48 wk maintenance following 32 wks open.
Hochstrasser (2001) (12-19 months) (165)	Recurrent unipolar depression; in and out patients; Mean 2 episodes in past 5 yrs (range 1-10) 54 centers UK and continental Europe; 6-9wks open citalopram, Randomized M=76 F=188; mean age 43.
Dekker (2000) (5 months) (85)	Major Depressive Disorder. Unusually high drop-outs; Proportion first episode unclear; Outpatients, single clinic, Amsterdam; Mean age 37 sd 10, 62%F, 5 month maintenance RCT 16wks open
Schmidt (2000) (25 wks) (295)	DSMIV, Major Depressive Disorder, CGI>4; first episode excluded if prior non-response to prozac or 2 or more AD's, high quality company sponsored trial; mean age 40, 65%F, 27%, 13 wks open then 25 wk to test prozac once per week vs 20mg per day
Stahl (2000) (6 months) (323)	DSM IV Major Depressive Disorder X 2months, min 22 on HAM D 17, No co-morbidity, no suicide attempts in past 12 months, HAMD suicide score<3, no prior failure on SSRI. 50% first episode; Acutely randomized maintenance trial; 8 US centers, 1wk placebo run-in then 24 wks; age 18-60, mean age 39, 60%F

Table 2: continued

SSRI trial (duration post-randomization)	SSRI Placebo-controlled Maintenance Trials: Overview
Rouillon (2000) (1 yr) (289)	Recurrent Major Depressive Disorder, minimum 1 episode in past 2 yrs. 75% inpatients, 25% outpatients at trial entry; HDRS <13 by end of 6 wks open; F:M ratio 2:1 mean age 45;20% had been hospitalized for a previous episodes . 104 centers 40% had 2 or fewer prior depressive episodes
Feiger (1999) (9 months) (106)	Major Depressive Disorder (Non-psychotic)) 40% had no prior episodes, 40% had 1-2 prior episodes. Responders to 16 wk single blind nefazodone. 5 sites outpatients, . Mean age 41. randomized M=37 F=94
Versiani (1999) (1 yr) (348)	Recurrent DSM IIIR Major Depressive Disorder , HAM-D>18; median 3 prior episodes, age 18-65, mean age 43; Randomized M=76, F=210. International multicenter trial 1 year f/u post randomization response defined as reduction in HAM-D or MADRS by 50%; 6wk open
Keller (1998) (18 months) (190)	Chronic Depression; Half had double depression, chronic major depression in half; mean 2 prior MDD; 45% had personality disorder, 60% had prior psychotherapy, 20% had prior antidepressants; 28wks open sertraline; 12 US sites. Maintenance phase M=55 F=106; mean age 40. Full remission (HAM-D<7 or 50% reduction in score) flexible dosing sertraline 50-200mg/d.
Terra (1998) (1 yr) (335)	Recurrent Major Depressive Disorder, min 2 prior episodes in past 5 yrs (Mean 3.5 sd 1.4); MADRS 10 or less, CGI 1 or 2 ; mean age randomized 45; M=54, F=150; multicenter French trial (63 sites)., 6wks open acute then 4.5 months continuation open then randomized 1 year follow-up
Reimherr (1998) (1 yr) (277)	Major Depressive Disorder or BPII; depressed 1 month; 12-14 wk open prozac; 5 US tertiary care outpatient clinics; entered 50wk f/u. F=141, M=57; Mean age 40. Objective to determine optimal duration of maintenance (point of diminishing returns) No exclusion of first episode depression, mean number with first episode not given.
Robert (1995) (6 months) (282)	Major depression; Number of prior episodes not stated responders to 8wk open citalopram. median age 48; Randomized F=162 M=64 multicenter French trial
Kishimoto (1994) (Planned 18 months) (192)	Recurrent depression min 3 episodes in past 2 yrs. 4 Japanese hospitals, 22 patients stable on mainserin (10) or TCA(12) randomized to placebo discontinuation; Planned 18 months but most recurred in first 4 months. F=19 M=3. Mean age 56 mean 2.5 episodes in past 2 years.
Montgomery (1993) (1 yr) (242)	DSM IIIR Major Depressive Disorder, 2+ current episode in past 4 yrs. Paroxetine F=132 M=40 5 psych. Outpatient departments, UK. Many had 4+ episodes in past 4 yrs.
Montgomery (1993) (6 months) (244)	DSM IIIR Major depressive disorder, recurrent Maintenance followed 6 wk acute rct (508 inpts, outpts, daypts from 18 centers). 207 (F=135, M=72) responders continued (60 responded to placebo). Then 147 were randomized to citalopram 20 or 40 mg or placebo. Number of prior episodes not stated
Doogan (1992) (11 months) (94)	Recurrent Major Depressive Disorder; Mean age 51 (range 19-78); Randomized F=204 M=91 Benzos and other drugs allowed in open phase; European multicenter, outpatients. 8wk open sertraline then 44 wks maintenance. 3/4 recurrent depression
Montgomery (1988) (1yr) (241)	DSM III Major Depressive Disorder HAMD>18. Outpatient responders to 6 months open fluoxetine. Min 1 prior MDD in 5 yrs Age and sex not specified mean 2.5 episodes in past 5 yrs (SD 1.6). French multicenter.
Bjork (1983) (Intended 18 months) (29)	Recurrent major depression. 20 on prior lithium with min. 3 depressions in the 4 years before lithium started OR treated for at least 3 depressions with antidepressants in the past 4 yrs. Many appear to have been on very long-term maintenance over 10yrs with lithium. . Intended 18 months study, most dropped earlier. Minimum 4 months recovered prior to study. mean age 51, range 33-69; M=7 F=31 (in randomized sample). Outcomes avail at 6,9,12,18 months single center study

Table 2: continued

SSRI trial	Comparisons	Enriched with study compound?	When randomized?
Montgomery 2004	venlafaxine vs placebo	yes, 8 wks acute and early continuation	after 8 wks acute and continuation
Rapaport 2004	escitalopram vs placebo	yes, 8wks RCT then all responders to open escitalopram X8wks	after 8 wks acute and 8 wks open continuation
Emslie 2004	prozac 20,60,placeb	no, acutely randomized to placebo	sub-group of med responders randomized at 19 wks
Simon 2004	venlafaxine XR vs placebo	yes; 8 wks acute open then randomized	after acute and early continuation
Lepine 2004	sertral 50 or 100 vs placebo after 2 month placebo run-in	No	after 2 month placebo run-in
Gellenberg 2003	Nefazodone vs placebo	Yes, 28wks. Half had CBT plus nefazodone	Responders to 28wks therapy Nefaz+/- CBT
Wilson 2003	Sertraline vs placebo	yes, open 8wks and continuation 16 wks	after continuation
Weihs 2002	Buproprion SR vs placebo	Yes, 8 wks	8 wks
Klysner 2002	Citalopram placebo	yes, only responders to citalopram 8wks acute and 16wks continuation	after 16 wks open continuation
Hypericium study group 2002	Hypericium, Sertraline, Placebo	No. Trial is continuation of acute RCT	Acute randomization
Hochstrasser 2001	Citalopram, placebo	yes responders to citalopram acute and continuation randomized	after acute and 16 wk continuation (randomized in maintenance)
Gilaberte 2001	Prozac, placebo	yes, responders to 32 wks open	after response to 32 wks
Dekker 2000	Prozac, placebo	yes, only responders to open prozac for 16 wks were randomized to 5 mnths	after 16 wks acute open
Schmidt 2000	Prozac 20 daily, placebo, prozac weekly	yes, responders to 13 wk open	responders to 13 wks
Stahl 2000	sert, cital, placeb	no	acute randomization
Rouillon 2000	Milnacipran	yes open milnacipran for 6 months	after 6 months open
Feiger 1999	nefazodone vs placebo	yes, only nefazodone acute responders randomized	after initial single blind 16 wk trial.
Versiani 1999	Reboxetine, Placebo	yes, only open responders to 6 wk acute reboxetine	after 6 wks acute open reboxetine

Table 2: continued

SSRI trial	Comparisons	Enriched with study compound?	When randomized?
Terra 1998	Fluvoxamine, Placebo	**yes, only responders to open 6 months fluvoxamine**	after 6 months open
Keller 1998	maintenance sertraline vs placebo	yes, Double blind acute randomization for 12 wks	after 4 month continuation
Reimherr 1998	fluoxetine/placebo at 14, 38, and 50 wks.	yes, 12-14 wks open prozac	acter acute response
Montgomery 1993	paxil vs placebo 1yr	Yes	after acute response
Montgomery 1993	citalopram vs placebo (note: 60 acute placebo responders excluded from analysis)	yes, only citalopram acute responders were randomized (expect very good outcome)	after acute response, only those on citalopram were randomized
Robert 1995	citalopram vs placebo. Only responders on open citalopram included	yes, only open citalopram responders randomized	after 8 wks acute treatment
Kishimoto 1994	Mainserin vs placebo	yes, mean 5 months well on mainserin or TCA	half had prior mainserin
Doogan 1992	Sertraline vs placebo	1wk placeb then 8wks sert then 44 wks maintenance	afer 1wk placeb, 8wk sert (+/- benzo and 'other')
Montgomery 1988	fluoxetine/placebo	yes, 6 months acute and continuation before randomizing	after continuation
Bjork 1983	Zimeldine vs placebo	no. half had lithium, half had 1-5wk antidepressant washout prior to prophylaxis	in prophylactic stage

Table 2: continued

SSRI trial	N acute	N continuation	N maintenance	% Completing open phase
Montgomery (2004)	495		235	47.4
Rapaport (2004)		502	274	54.6
Emslie (2004)	219	158	75	N/A
Simon (2004)	490	292		59.6
Lepine (2004)	placeb=371		288	77.6
Gellenberg (2003)	681	unclear	160	160
Wilson (2003)	254		113	44.4
Wiehs (2002)	816	423		63%
Hypericium study group (2002)	340	129		N/A
Klysner (2002)	230	172	121	52.6
Gilaberte (2001)	253	206	140	55.3
Hochstrasser (2001)	427	327	264	61.8
Schmidt (2000)	932		501	53.8
Stahl (2000)	323			N/A
Dekker (2000)	147	48	30	20.4
Rouillon (2000)	500		214	42.8
Feiger (1999)	467		at 16 wks, 131	28.5
Versiani (1999)	358	286		79.9
Terra (1998)	436		204	46.8
Keller (1998)	309	209	161	52.1
Reimherr (1998)	839 entered acute open 12-14 wks	395 randomized to 4groups after stabilization,	not defined	47.1
Robert (1995)	391	226		57.8
Kishimoto (1994)	N/A	N/A	26	
Montgomery (1993)	tot=172	135		78.4
Montgomery (1993)	508	207 (only 147 on cital randomized, 60 were on placebo)		40.7
Doogan (1992)	480	295		61.4
Montgomery (1988)	465	254	220	47.3
Bjork (1983)	unknown	unknown	40 (2 dropped early)	47.3

Table 2: continued

SSRI trial	%completing RCT(both phases)	Cumulative relapse	Analysis Used in report LOCF?	# completing	Funding
Montgomery (2004)	36.4	yes	SA, LOCF	yes	Industry
Rapaport (2004)	44.9	yes	SA, Unclear	yes	Industry
Emslie (2004)	N/A	yes	CA, LOCF	yes	Industry
Simon (2004)	26	yes	CA, LOCF	yes	Industry
Lepine (2004)	57.3	yes	Unclear, SA	yes	Industry
Gellenberg (2003)	11.6%	yes	CA,LOCF,CRA	yes	Industry
Wilson (2003)	27	yes	SA, unclear	yes	Industry
Weihs (2002)	12.4%	yes	LOCF,SA	yes	Industry
Klysner (2002)	33.1	yes	SA, Unclear	yes	Industry
Hypericium study group (2002)	24	yes	CA, LOCF	yes	Foundation
Hochstrasser (2001)	9.5	yes	SA, Unclear	no	Industry
Gilaberte (2001)	55.7	yes	LOCF, CA	yes	Industry
Schmidt (2000)	38	yes	LOCF	yes	Industry
Stahl (2000)	N/A	yes	LOCF	yes	Industry
Dekker (2000)		yes	LOCF	no	Industry
Rouillon (2000)		yes	LOCF, SA	no	Industry
Versiani (1999)	51.4	yes	LOCF, SA	yes	Industry
Feiger (1999)	55	yes	LOCF, SA	yes	Industry
Terra (1998)	76	yes	LOCF, SA	yes	Industry
Montgomery (1988)	47.7	yes	LOCF	no	Industry
Reimherr (1998)		yes	CA	no	Industry
Keller (1998)	36.7	yes	SA, LOCF	yes	Industry
Robert (1995)		yes	LOCF, Unclear	no	Industry
Kishimoto (1994)	Unclear	yes	EA	no	Unknown
Montgomery (1993)	64.4	yes	LOCF	no	Industry
Montgomery (1993)	57.1	yes	Unclear	no	Industry
Doogan (1992)	50	yes	LOCF	yes	Industry
Bjork (1983)	Unclear	yes	LOCF, SA	yes	Unknown

LOCF= Last observation carried forward
CA = Completer analysis
SA = Survival analysis
CRA = Competing risk analysis
'Unclear' refers to a lack of clarity on how cumulative relapses were reported, as observed cases or as 'last observation carried forward'.

SSRI studies	Survival analysis? Mean, meidian time, SD?	Adverse effect reporting?	Jadad score (Max 5))	Concealment (as per Schultz)
Montgomery (2004)	Graphs, lots of excellent comparisons	Yes	3	Unclear
Rapaport (2004)	Graphs	Yes detailed	3	Unclear
Emslie (2004)	Graphs	Detailed reporting	3	Unclear
Simon (2004)	Graphs, tables	Yes, detailed	3	Unclear
Lepine (2004)	Graphs only	Yes	4	Adequate
Gellenberg (2003)	Survival analysis, graphs	Yes	4	Adequate
Wilson (2003)	Summaries at 4,8,12,48 and 100 wks	No	4	Adequate
Weihs (2002)	Yes, graphs and summary	Yes	3	Adequate
Klysner (2002)	Graph only	Yes, detailed	3	Unclear
Hypericium study group (2002)	Graph only	Yes, detailed	3	Adequate
Hochstrasser (2001)	Graph only	Yes, detailed includes manic	4	Unclear
Gilaberte (2001)	Graphs	Detailed reporting	3	Unclear
Dekker (2000)	Graph only	Minimal	2	Adequate
Schmidt (2000)	Graphs	Detailed reporting	3	Unclear
Stahl (2000)	Graphs	Detailed reporting	3	Unclear
Rouillon (2000)	Graph only	Yes	3	Unclear
Feiger (1999)	Graph only	Yes	3	Unclear
Versiani (1999)	Graph only	Yes, detailed	3	Unclear
Terra (1998)	Graph only	Yes, no mania reported	3	Unclear
Keller (1998)	Graphs only	Yes, detailed	3	Unclear
Reimherr (1998)	secondary paper McGrath has mean, SD at various endpoints. Recurrences at 14, 36, 50wks	No	3	Unclear
Robert (1995)	Graph only	Yes	3	Unclear
Kishimoto (1994)	Graph only	Minimal to none	2	Unclear
Montgomery (1993)	Graph only. Recurrences at 16 and 52 wks	Yes	3	Unclear
Montgomery (1993)	Graph only	Yes	4	Adequate
Doogan (1992)	Graphs only	Yes	3	Unclear
Eric (1991)				
Montgomery (1988)	Graph only, recurrences given quarterly	Some. 2/456 hypoman.acute	3	Unclear
Bjork (1983)	Graphs, tables	Summary. 1 mania reported	2	Unclear

Table 2: continued

SSRI studies	Recurrence Threshold	Taper/Discussion or rebound
Montgomery (2004)	CGI 4 or more	2 wk taper. Separate analysis over 4 wks post D/C
Rapaport (2004)	MADRS >22	No taper. Does not discuss early rebound but graph suggests.
Emslie (2004)	scales, cumulative recurrences	Early relapses post d/c described. Significant potential for rebound
Simon (2004)	DSM IV MDD + CGI 4 or more OR two consec CGI 4 or more OR final CGI-S 4 or more and quit study	28 days post discontinuation analyzed separately. 2 wk taper
Lepine (2004)	clinical recurrence	Rebound/taper not an issue
Gellenberg (2003)	Committee consensus of MDD plus HAMD-24 minimum 16 at two consecutive visits	No
Wilson (2003)	HDRS over 13 + DSM III R depression	½ relapses in 1st 26 wks. No taper indicated
Wiehs (2002)	Clinical determination of need for additional treatment	No
Klysner (2002)	MADRS>21	Cumulative relapses given at wk 4,8,12,24, and 48. No increased recurrences or adverse events at time of randomization. No taper
Hypericium study group (2002)	17 item HAM-D 20+ AND CGI=4+ on 2 or more consecutive visits. Psychosis or suicidal ideation exclusionary criteria	No abrupt discontinuation
Hochstrasser (2001)	MADRS>21	Dose-related relapse symptoms. Citalopram at 20,40 and 60 lead to progressively steeper loss post discontinuation. No taper
Gilaberte (2001)	clinical relapse, scale,	No taper. No discussion of potential rebound
Dekker (2000)	clinical relapse, threshold on HAMD	No taper
Schmidt (2000)	clinical scales(several), cumulative recurrence	No taper. No discussion of rebound
Stahl (2000)	clinical and scales, cumulative reccurance	Taper/Rebound not relevant
Rouillon (2000)	clinical recurrence + HDRS over 18	Minimal rebound apparent. No taper
Versiani (1999)	Increase of HAM-D or MADRS by 50%	No taper. No discussion of rebound
Feiger (1999)	HAMD >18 on two consecutive visits OR early discontinuation due to lack of efficacy.	No taper, but apparently no or little rebound effect in first months. Potential rebound discussed
Terra (1998)	Clinical diagnosis of depression (five symptoms of DSM III-R criteria, suicidal behaviour)	Luvox tapered to 100 mg/d 4 wks prior to placebo discontinuation
Reimherr (1998)	Clinical criteria for Major Depression on 2 weeks at any assessment OR HAM-D scores over 14 on 3 consecutive weeks	No taper. Raises question of definition of relapse vs recurrence with levels of recovery
Keller (1998)	full MDD, CGI over 4, HAMD increase of 4,	Sertraline taper at 50mg/wk. Some reference to potential rebound
Robert (1995)	MADRS >24, +clinical judgment	No apparent rebound, but survival curve drops sharply post randomization. No taper
Kishimoto (1994)	HDRS over 10	7/13 relapsed in 1st 2 months of withdrawal. ½ were stabilized on mainserin before trial. No taper

SSRI studies	Recurrence Threshold	Taper/Discussion or rebound
Montgomery (1993)	clinical recurrence, need for more meds	High losses immediately post withdrawal. No taper
Montgomery (1993)	MADRS over 22	Continuation of acute RCT. Some exploration of rebound. No taper
Doogan (1992)	CGI 4 or more	No discussion of rebound. No taper
Montgomery (1988)	clinical, HAMD, CGI ratings every 3 months	States withdrawal effects within 3 months are not significant
Bjork (1983)	2 psychiatrists; RDC MDD; >9 on CPRS-D	No taper. No discussion of rebound

Reported binary outcomes:

Outcome data was available at 6-9, 12, 18, 24 and 36 months (Table 3). TCA and MAOI trials had limited reporting of the number completing and number dropping out for reasons other than relapse or recurrence.

	TCA/MAOI trials N=15	SRI/SNRI Trials N=29
Cumulative # recurring/relapsing	15	28
Number completing	13	28
Number stopping, reasons other than relapse/recurrence with detailed breakdown	2	25
6 –9 month cumulative recurrence	7	11
12 month cumulative recurrence	6	17
18 month cumulative recurrence	1	5
24 month cumulative recurrence	8	1
36 month cumulative recurrence	3	0

Table 3: Summary of available outcomes.

4.3b Trials available for direct and indirect comparisons

TCA's and MAOI's:

Among TCA and MAOI trials, there were no trials suitable for combining in either direct or indirect comparisons (Table 4). With the exception of 2 amitriptyline trials for which inadequate reporting was available (138, 184), all TCA and MAOI trials used the study drug in phase 1 and randomized to either ongoing study drug or placebo.

	Imipramine	**Desipramine**	Amitriptyline	Nortriptyline	Phenelzine	Dothiapin
# of trials (N)	5 (228)	1 (50)	6 (298)	2 (104)	2 (98)	1
Open phase <16 wks	3	0	4	2	2	1
Open phase >16 wks	2	1	2	0	0	0
Same intervention, phase I and II	5	1	4	2	2	1
Different intervention in phase I vs phase II	0	0	2*	0	0	0

Table 4: Breakdown of TCA and MAOI trials by use of open phases and duration of open phases.
* Inadequate reporting of phase 1 interventions to include in indirect comparisons.

SSRI's:

SSRI's, (including SNRI's and buproprion trials) (table 5) provided 4 potential trials for an indirect comparisons. Three of these had adequate reporting to carry out the comparison. There were two direct comparison trials. The single trial using different interventions in phases 1 and 2 was effectively a crossover trial, as it enrolled patients

treated with fluoxetine maintenance (half life over 10 days) to 2 months of open placebo in phase 1 and sertraline, a similar SSRI in phase 2.

	Bupropion SR	Fluoxetine	Paroxetine	Fluvoxamine	Venla-faxine	Citalopram or escitalopram	Milnacipran	Sertraline	Nefaza-done	Zimeldine or Mainserin
# of trials (N)	1 (423)	5 (1100)	1 (135)	1 (203)	2 (517)	4 (817)	1 (214)	6 (1271)	2 (296)	2 (64)
Classic RCT	0	1	0	0	0	1	0	2	0	0
Direct comparison design	0	1	0	0	0	1	0	0	0	0
Open phase >16wks	0	3	1	0	1	2	1	1	2	2
Open phase <16wks	1	1	0	1	1	1	0	3	0	0
Different Phase I vs phase II medication	0	0	0	0	0	0	0	1*	0	0

Table 5: Breakdown of SSRI, SNRI and buproprion maintenance trials by use of open phases and duration of open phases. Two trials included more than one antidepressant arm. *1 trial (Lepine et al(209)) resembles a crossover trial (described separately below).

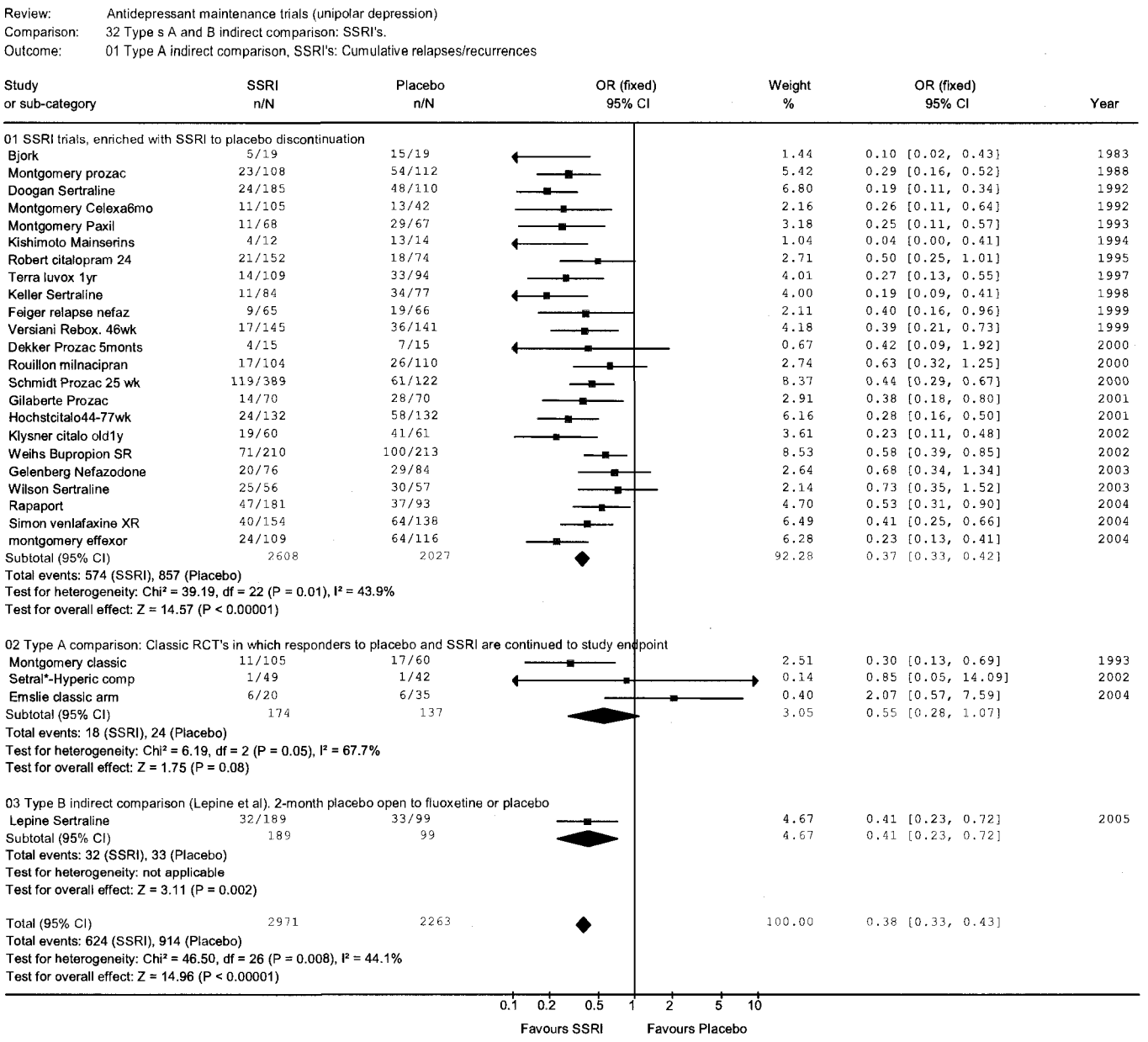
4.3c Indirect Comparisons:

A limited number of indirect comparisons were possible, summarized in figure 2. Three reports were available to make up an indirect comparison[8]. This single comparison showed a non-statistically significant trend consistent with ED design bias;

- Classic RCT for indirect comparison; OR = 0.55; 95% CI 0.28-1.07
- ED designs OR = 0.37; 95% CI 0.33-0.42.

[8] Unpublished data from a fourth trial (Stahl 2000(323)) has been requested from the trial sponsor, Forest laboratories.

A single report was available an indirect comparison using a different intervention in phase 1 and phase 2 (Lepine et al 2004) (209). This trial is classed as an indirect comparison for completeness sake, but it could be argued that it is actually a crossover design. Patients with a minimum of 3 past episodes of depression and who had been compliant and 'responsive' to an SSRI for at least 6 months were enrolled in one of 83 centers. These patients, 75% of whom were receiving fluoxetine (half life approximately 10 days) were given open placebo treatment for 8 weeks followed by randomization to either ongoing fluoxetine or sertraline, another SSRI. Because of fluoxetine's long half-life, the placebo phase acted as a washout period prior to crossover to sertraline, resulting in a highly atypical designation of 'placebo responder'. We will therefore not comment further on this report.



Phase 1	Phase 2
SSRI	SSRI vs placebo
Placebo	Placebo vs SSRI

Indirect comparison type B above (Lepine et al)

Phase 1	Phase 2
SSRI	SSRI vs placebo
Classic RCT	Responders followed as originally allocated

Indirect comparison type A above

Figure 2. Summary of SSRI indirect comparisons. Type A comparisons show a trend consistent with an ED design bias favoring the open phase intervention.

4.3d Direct comparisons:

Two reports provided direct comparisons (Table 6). This approach has the advantage of comparing different treatment strategies using the same population (chapter 2). Results were mixed. Emslie 2004 (101) favored placebo in the classic arm and fluoxetine in the phase 1 discontinuation arm, consistent with either selection bias or medication withdrawal-rebound bias. The outcome of Montgomery et al 1993 (244) did not point to bias favoring the phase 1 intervention, with similar outcomes in the classic RCT and the discontinuation arm of their 'direct comparison' trial.

Trial	Odds Ratio: P-P vs SSRI-SSRI **'classic RCT'**	Odds Ratio: SSRI-SSRI vs SSRI-P **'ED arm'**
Montgomery '93	.30 (95% CI .13-.69) favors SSRI N=165	.26 (95% CI= .11-.64) favors SSRI N= 147
Emslie '04	2.07 (95% CI = .57-7.59) favors placebo N=55	.29 (95% CI = .08-1.06) favors SSRI N=40

Table 6. Summary of cumulative relapses under contrasting treatment strategies from direct comparison trials. Inconsistent evidence of bias.

4.3e Post-hoc exploratory analysis:

A number of factors are likely to impact phase 2 reported outcomes in ED trials and should be considered when using trial data to make comparisons. These factors include differences in natural illness course, total trial duration, duration of phase 1, number of phase 1 drop-outs, and the use of tapering after phase 2 randomization. In addition, it is possible that secular trends toward a relatively higher placebo response over the past

three decades as described for 8-week antidepressant trials (353), may also apply to maintenance trials.

Impact of illness course:

The chances of relapse in the year following recovery from an episode of depression increase with the number of prior lifetime episodes (321). Trials excluding patients with a first lifetime episode of depression would therefore be expected to show a greater effect size than trials including this group, since patients with recurrent depression are more at risk for future episodes.

As illustrated in figure A1 (Appendix), ED trials excluding first-episode cases have larger effect sizes than trials including these cases. This holds for TCA and MAOI trials (OR 0.39; 95% CI 0.24-0.63 in trials including first-episode depression vs OR 0.23; 95% CI 0.14-0.37) as well as in SSRI trials (OR 0.47; 95% CI 0.38-0.57 in trials including first episodes vs OR 0.29; 95% CI 0.24-0.35).

Impact of phase 1 duration and total trial duration:

Depression tends to spontaneously resolve over time. The duration of phase 1 would therefore be expected to relate inversely to trial effect size, since patients randomized later in their natural illness course would be expected to have a higher spontaneous response rate than those randomized earlier in their natural course. Differences in effect size were not observed when trials were stratified by the duration of phase 1 (phase 1 less than or greater than 16 weeks) (figure A2, Appendix). And a non-significant trend toward

increased effect with longer trials was seen when trials were stratified by total duration (figure A3, Appendix).

Impact of phase 1 drop-outs:

If ED trials preferentially increase the proportion of medication responders or tolerators to placebo responders in phase 1 (ie. if the assumption of phase 1 non-differential attrition is violated), trials with greater drop-outs in phase 1 may have higher phase 2 effect sizes than those with lower phase 1 drop-outs. This was not observed when trials were stratified by phase 1 drop-outs >51%, 31-50% and <30% (figure A4 Appendix).

Impact of tapering in phase 2:

Antidepressant withdrawal symptoms are believed to affect up to a third of patients abruptly discontinued from SSRI's (287). Since the symptoms of withdrawal-rebound are difficult to distinguish from relapse, we expected that the use of tapering would have been associated with a lower phase 2 reported effect size. This was not seen when SSRI trials were stratified according to their use of tapering (figure A5 Appendix).

Secular trends:

Among brief 6-8 week antidepressant trials, a trend toward an increasing placebo effect over the past three decades has been described (353). A similar trend appears to exist among antidepressant maintenance trials (figure A6 Appendix).

4.3f Assessment of publication bias:

A funnel plot (figure 3) provides a general estimate of publication bias among ED trials. Evidence of a marked publication bias might offer one alternative explanation for the apparently uniform outcome in published ED maintenance trials. The plot however, does not suggest an obvious imbalance of effect size between trials of large and small effects.

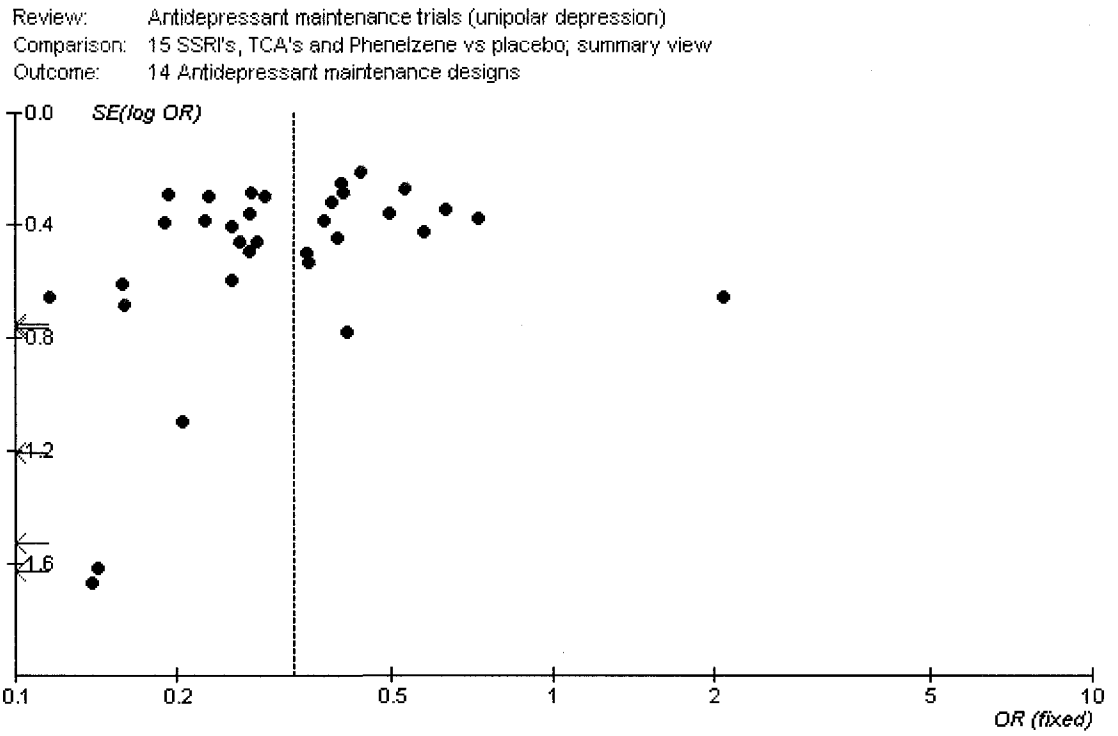


Figure 3. Funnel plot of SSRI and TCA maintenance trials. No obvious publication bias.

4.3g Impact of report quality

Since the quality of reporting is believed to correlate with reported effect sizes (173), we also stratified trials by Jadad scores (less than 3, equal or greater than 3) (figure 4). Lower quality scores are associated with higher reported effect sizes.

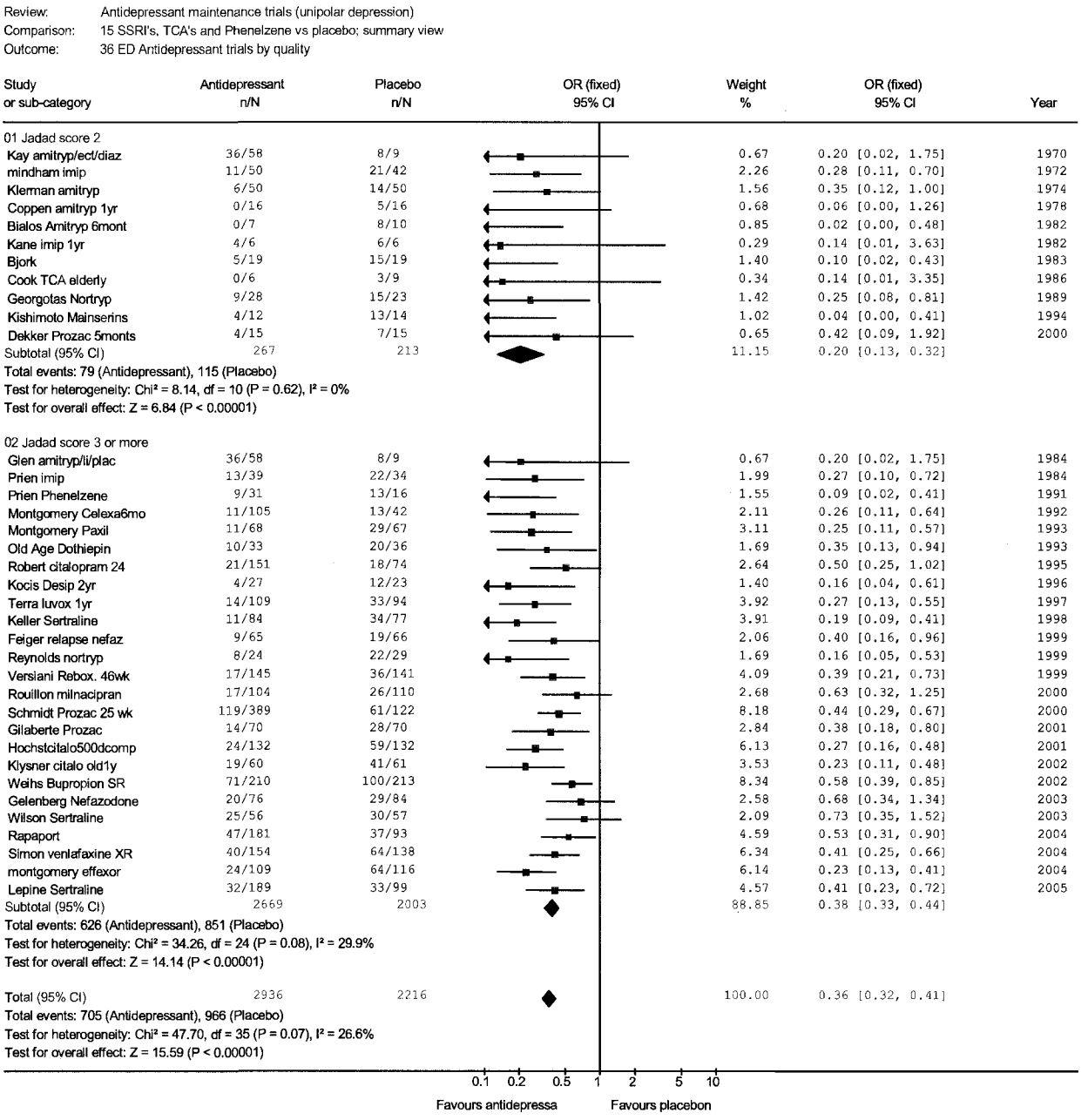


Figure 4. Trials stratified by Jadad quality score. Lower quality reports have higher effect sizes.

4.3h Monomethod bias

The object of the current study is to identify a design bias among antidepressant trials using the ED designs. Almost all identified trials over the past three decades have used the design, leading to a situation resembling Shadish's 'mono-method' bias (308). For

antidepressant maintenance trials, medication response is defined almost exclusively in relation to relapses among cases withdrawn from medication rather than in relation to spontaneous responders. Joffe has recently argued that spontaneous responders are the most relevant comparator group (176). For reasons outlined in chapter 3 for mood stabilizers, an observation of Shadish's 'mono-operational' bias also seems to apply to antidepressant maintenance trials.

4.4 Discussion

ED designs have dominated antidepressant maintenance trials for the past 30 years, so the issues common to this RCT variant, including concerns about phase 1 differential attrition, phase 2 withdrawal-rebound symptoms, and the feasibility of phase 2 blinding, apply to the overall evidence base for maintenance antidepressants.

Only five trials have used alternative treatment strategies to ED designs in defining treatment responders, severely limiting our method, which depends on finding a central effect estimate from among many studies. Large samples of trials conducted across a range of settings are important in sorting out the signal of a design bias from the noise inherent in a syndrome as heterogeneous as major depressive disorder. As a result of this limitation in our study, we will discuss the merits of the comparisons on a case-by-case basis.

The indirect comparison (figure 3), consisting of the classic RCT arms from Emslie 2004(101), Montgomery 1993 (244), and the Hypericum-Setraline study (169) suggests a weak trend toward a smaller effect size when responders from a classic RCT are followed forward than is observed in ED designs. Put another way, if patients get better

spontaneously, they are more likely to stay better without any active intervention than if they were first treated with an antidepressant and then later withdrawn.

Direct comparisons have the advantage of reducing the heterogeneity facing indirect comparisons. Essentially, responders from a classic RCT are followed forward in parallel to responders from an ED trial. This allows a comparison, from the same population, of placebo responders from an extended classic RCT with 'medication responders' who are randomized to placebo discontinuation in phase 2. The two available reports from 'direct comparisons' yield conflicting results. Montgomery et al 1993 (244) suggests that there is little difference between the groups, while Emslie et al 2004 (101), favors cases responding to placebo in the classic RCT. Two plausible rival hypotheses should be considered. First, these trials studied different populations. Montgomery studied adults while Emslie studied adolescents. The efficacy of antidepressants in adolescents has long been controversial(101), so it is possible that the conflicting results reflect only population differences. On the other hand, there are major differences in the quality of reporting between these reports. Montgomery presents scattered data, and it is not possible to locate the reported classic RCT from this large trial. Only parts of the cohort referred to can be found in the published literature across four separate reports with different authors (244-246, 282). The reports refer to one another, but despite our best efforts, it was not possible to trace the full inception cohort. This raised concerns about reporting quality, which were confirmed when we contacted the sponsor to attempt access to data. We were told that the data were difficult to access because of logistic problems and that there might be problems locating some of the data.

Emslie on the other hand, reported using a format compatible with the CONSORT statement (238). Patients can be traced, in a single report, from start to completion. Although the paper highlighted only on the discontinuation arm, which was favorable to the sponsor's product, the transparent reporting style allowed the acute RCT placebo responders to be followed forward, showing that if adolescents responded to placebo initially, they are no more likely to relapse than those who 'respond' to fluoxetine and are discontinued to placebo. The trial published survival curves for the discontinuation arm but not for the classic RCT arm and the sponsor has not released the placebo responder data despite two separate applications.

Returning to the indirect comparisons, analogous factors apply, since we have used data from both Emslie and Montgomery. In addition, the third trial (169), was conducted on adults. It reports transparently, although not in compliance with the CONSORT statement. This trial found that the odds of remaining well for 18 weeks given recovery on either placebo or SSRI were similar.

Taken together, the limited comparisons between treatment strategies, questions about heterogeneity in illness course and trial quality limit what can be said about potential ED design effects in antidepressant maintenance trials.

Post hoc analyses confirmed a relationship between lower trial quality and greater reported effect size. Similarly, as expected, we observed greater effect sizes in reports

excluding patients with a single episode of depression. And the secular trend over the past three decades showing an increasing placebo response in maintenance trials (figure A6), replicates a similar trend reported for 8-week antidepressant trials (353). This effect, at least in shorter trials, has been attributed to threshold shifts in the severity of cases enrolled in placebo-controlled trials (with contemporary trials enrolling milder cases).

Three post-hoc observations however are counterintuitive and require further consideration. The observations are as follows (figures A3-5):

1) There is no apparent difference in effect size between trials using antidepressant tapering and those that use abrupt discontinuation. In contrast, trials that used tapering actually showed a greater effect size than those that did not.
2) ED trial effect size does not differ by duration of phase 1.
3) ED trial effect size does not differ by duration of phase 2.

The first observation is counterintuitive, since antidepressant withdrawal syndromes severe enough to be misclassified as relapses affect up to a third of patients from whom medications are abruptly withdrawn (25, 287). Withdrawal is distinguished from recurrence or relapse primarily by the timing of symptom appearance although at least one scale has been developed in an attempt to differentiate the two(25). Withdrawal effects correlating with a mean increase of 7 points on the Hamilton Depression Inventory (HAM-D) score is described within two weeks of antidepressant

discontinuation (25, 260, 287)[9], and there is some indication that the intensity of effect varies from compound to compound (25). Withdrawal effects have been described across all classes of antidepressants (287). Approximately 40% of clinical relapses in ED trials are concentrated in the 8 weeks following placebo randomization (101, 200, 243). Concerns about misclassification of early relapses have been raised by several ED trial investigators (244, 276, 277, 279, 315), and a number of reports have given special consideration to the first month following phase 2 randomization (reviewed in chapter 5). Plausible explanations for our observation that medication tapering did not impact trial effect size include the inconsistent use of tapering procedures, inconsistent recording of symptoms believed to differentiate medication withdrawal from relapse and problems with blinding ('wish bias' favoring active medication (27)).

The two observations showing a lack of relationship between trial duration and outcome are puzzling (figures A2 and A3), since depression naturally remits over time. Longer studies should yield lower effects relative to placebo than shorter studies. Reimherr et al (277) showed this in a large fluoxetine ED trial, in which discontinuation phases were staggered, with the longest being 9 months (9 months in phase I prior to phase II discontinuation). At nine months, there was no difference in recurrence rates between placebo and fluoxetine. It is possible that our chosen outcome, cumulative relapses using LOCF is too coarse to pick up these time-dependent changes in recurrence rates. A lack of sensitivity may be magnified by attrition, a non-random process that affects maintenance trials with rates of 25-87% (tables 1 and 2). Similarly, this lack of sensitivity

[9] The threshold for clinical depression is a HAM-D of 13 and for recovery the threshold is 7 (287).

may have limited our ability to detect a relationship between phase 1 drop-outs and phase 2 relapses (figure A6).

Taken together, this chapter demonstrates that design effects are difficult to disentangle from heterogeneous report quality, probable classification issues and challenges with blinding patients who have been exposed to an open treatment for an extended period then withdrawn from it.

In chapter five, we explore the recognition of ED designs in the psychiatric literature as a source of uncertainty in the evidence base for maintenance medications.

Summary:

ED designs have been the predominant method for testing antidepressant maintenance efficacy over the past three decades. A limited number of alternative treatment strategies were available to assess bias favoring phase 1 interventions. Based on the limited available data, bias cannot be ruled out. Inconsistent quality of reports including questions about phase 2 blinding place further limitations on what can be gathered from the available comparisons. The dominance of a single method for testing the efficacy of antidepressant maintenance raises the likelihood of Shadish's monomethod bias and mono-operational bias.

Appendix

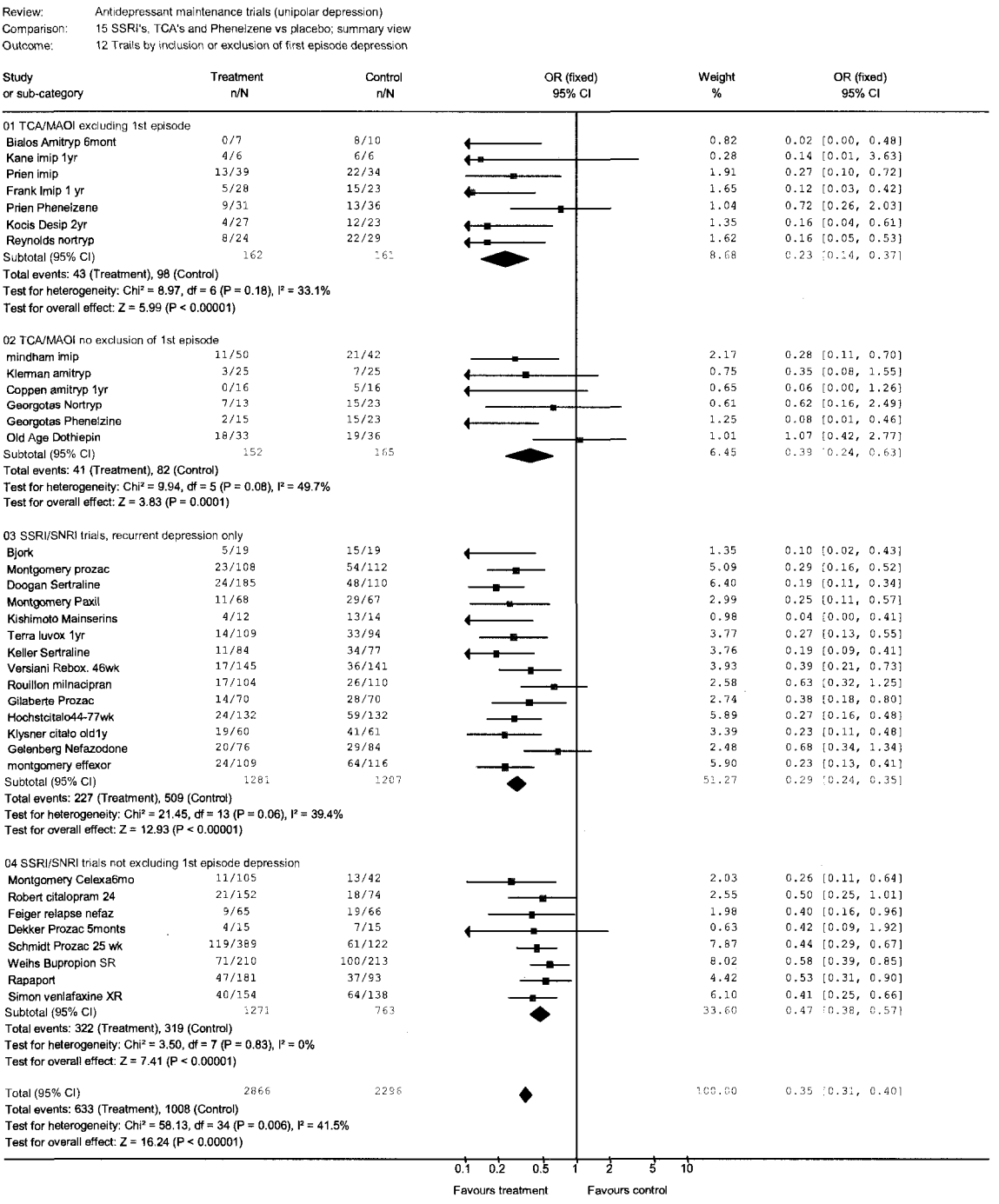


Figure A1: Post-hoc analysis showing trend towards greater effect size in both SSRI/SNRI trials and TCA/MAOI trials by exclusion of first episode depression.

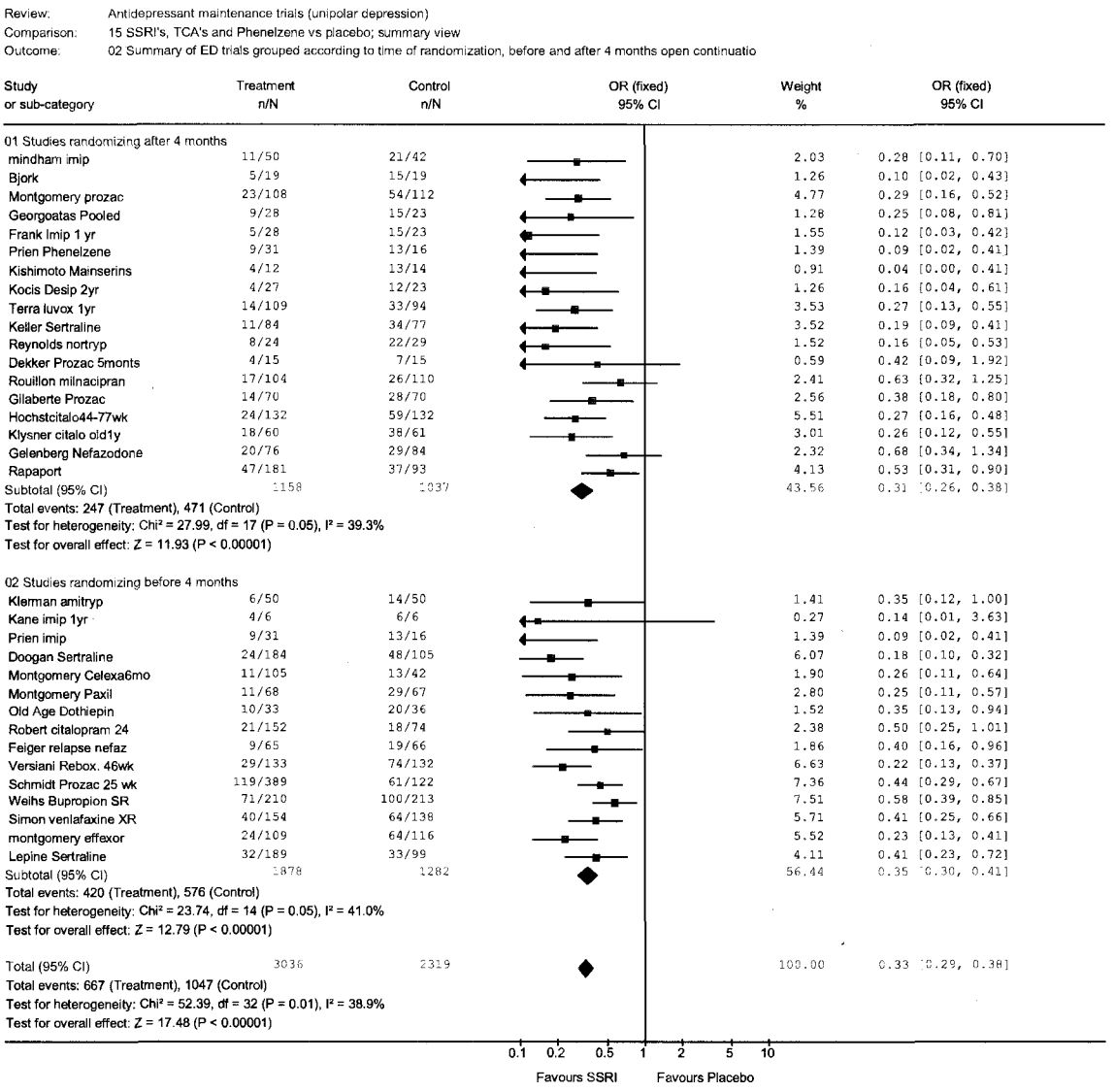


Figure A2: Trial effect size by duration of phase 1 (greater or less than 16 weeks).

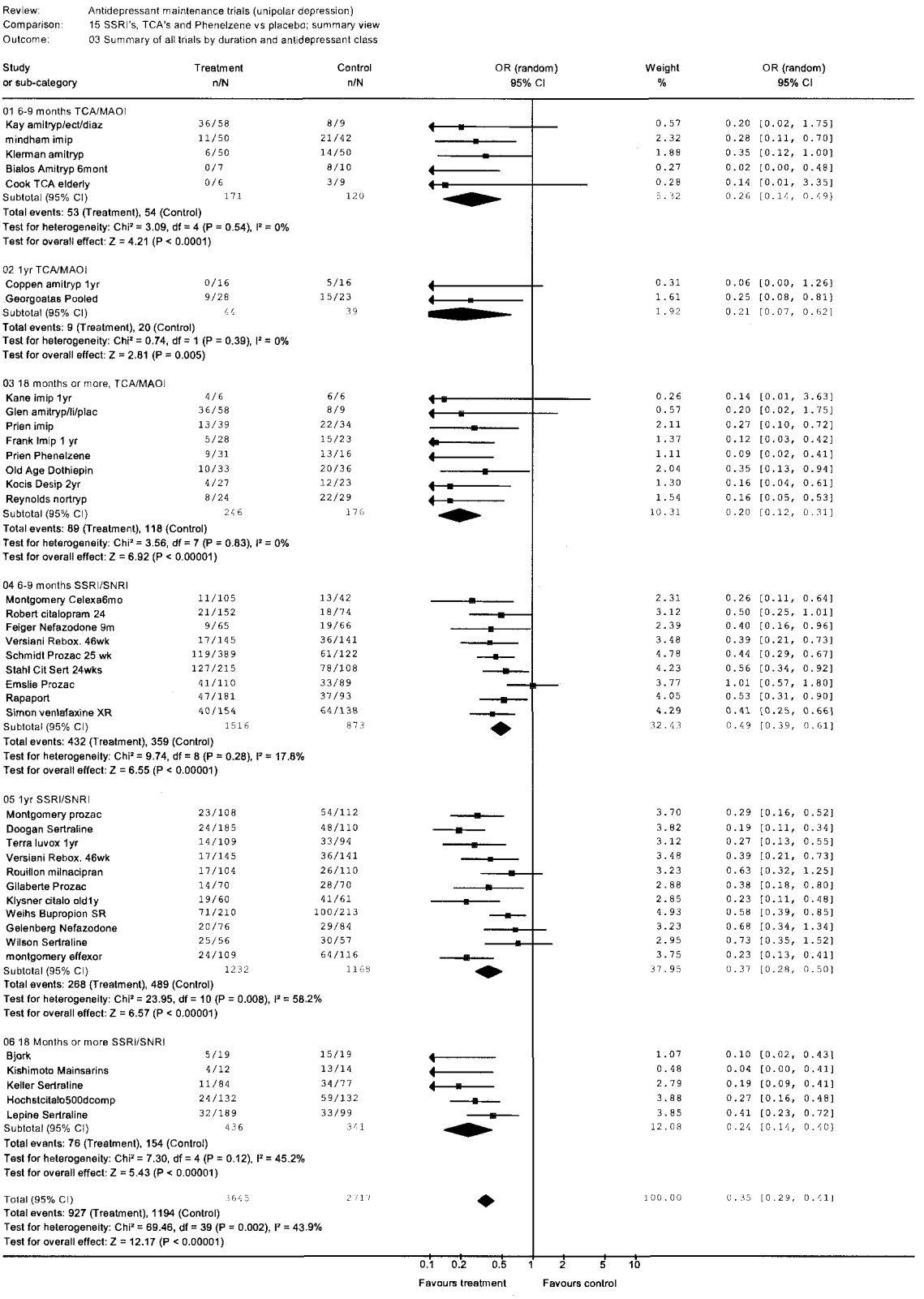


Figure A3: Trial effect size by duration. Contrary to expectations, there is a non-significant trend to higher effect with longer trials, both antidepressant classes.

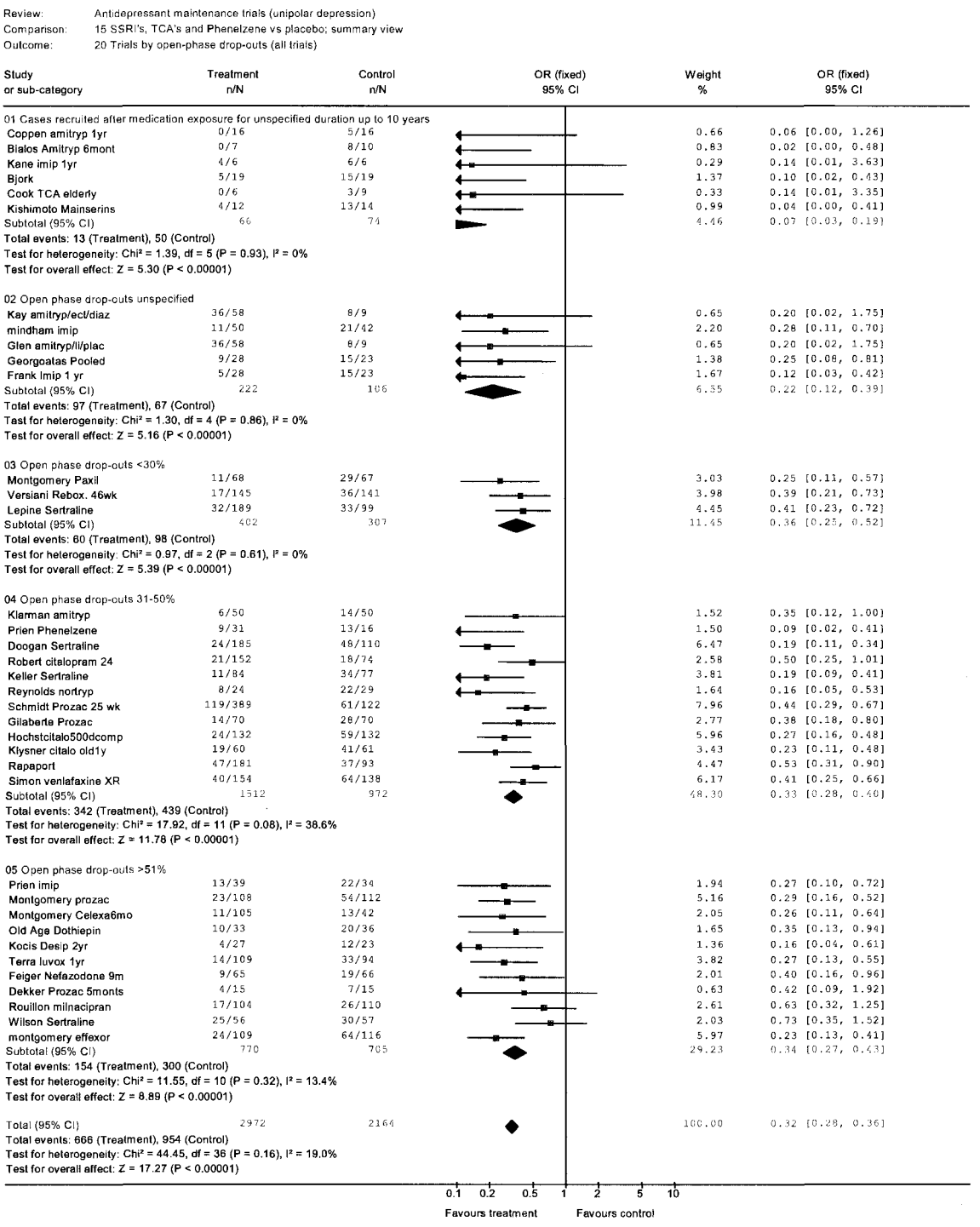


Figure A4: Trial effect size as function of phase 1 drop-outs. No relationship between phase 1 drop-outs and phase 2 effect size.

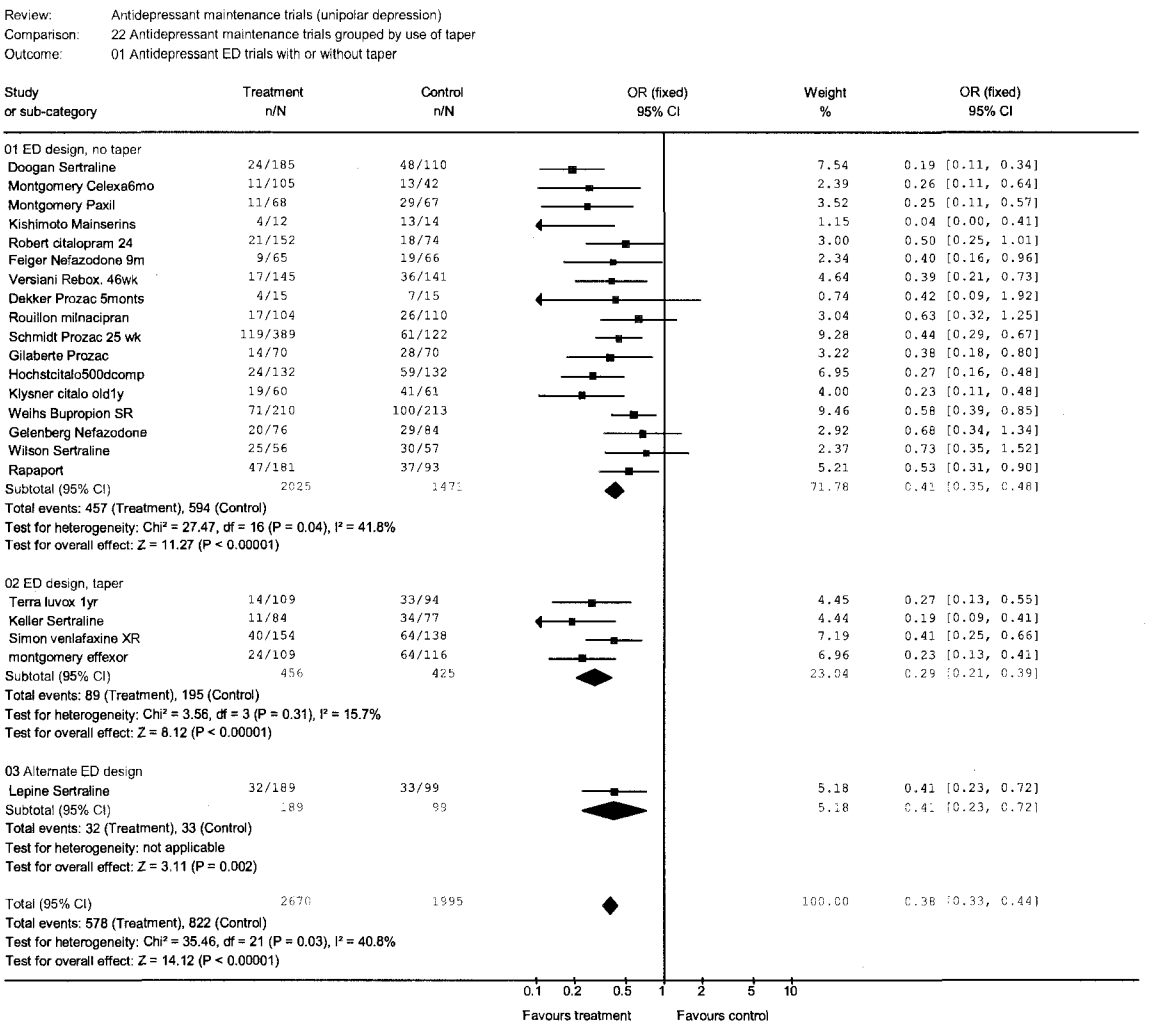


Figure A5: Trials according to use of taper, no taper.

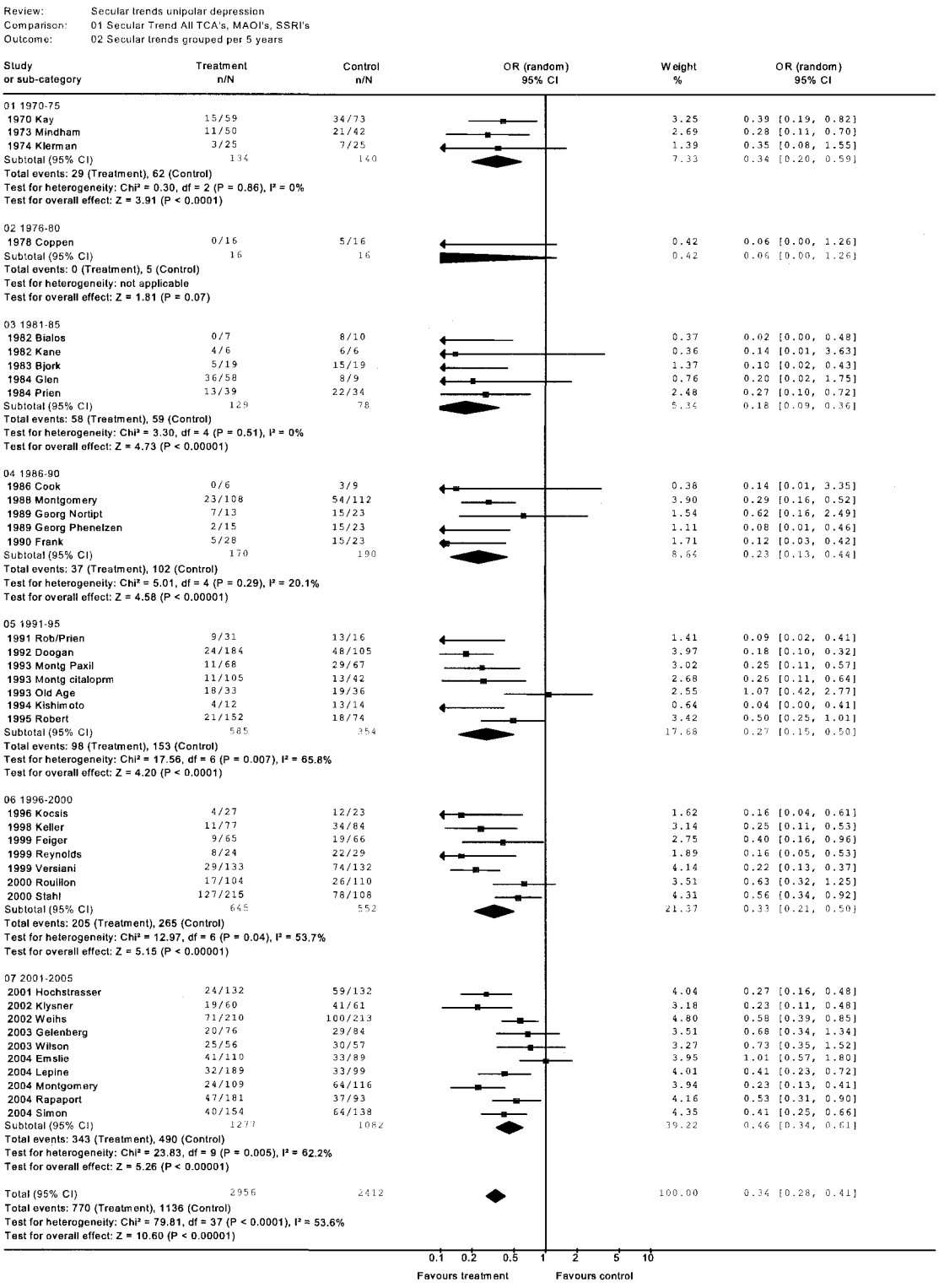


Figure A6: Secular trend toward lower effect size, 5-year increments.

Chapter 5

ED trials vs classic RCT's in psychiatric drug maintenance of mood disorders Does one size fit all?

"Most experiments are highly local but have general aspirations"

(Shadish et al p18 (308))

5.1 Introduction:

Clinical evidence is sometimes ranked hierarchically, with anecdotal case reports and expert opinion at the least definitive end of the spectrum and RCT's at the most definitive end (212). ED trials, with their vulnerability to withdrawal bias and questionable applicability, likely rate somewhere in between. Strict hierarchies can however, be misleading, and Shadish et al propose a more contextual approach to linking methods with strength of evidence(308). They suggest that the definition of an 'optimal' experimental design will differ according to the research context.

In the context of antidepressant and mood stabilizer trials, ED designs have been applied to the near exclusion of classic RCT's, informing an idiosyncratic definition of 'medication responders', defined in relation to 'non-relapse' under conditions of medication withdrawal. The resulting gap in the literature defining treatment response in relation to spontaneous response adds a layer of uncertainty to the rational use of psychiatric maintenance drugs.

In preparing this thesis, only five references to ED methodology were found in psychiatric journals (143, 176, 190, 270, 332), giving the impression that the prevalence and unique limits of this RCT variant are not widely recognized by the psychiatric community. This is unexpected given the high prevalence of the design in psychiatric maintenance trials. To further explore our impression about the recognition of ED designs in the psychiatric literature, we reviewed a sample of treatment guidelines, commonly used textbooks of psychiatry, systematic reviews, and published ED trials to identify whether ED trials are considered identical to or distinct from traditional RCT's, and whether the potential for withdrawal bias or limited applicability to source populations is specifically discussed. As a frame of reference, we also reviewed ED trials from general medicine cited in Kopec et al (197), to identify the extent to which design limits were discussed.

5.2 Methods

Clinical guidelines published between 2001 and May 2006 for drug maintenance in recurrent depression and bipolar disorder (109) were identified using the keywords 'depression' or 'bipolar disorder' in three dedicated guideline internet sites; guidelines.gov , cma/cpg infobase , and Guidelines International Network. General psychiatric textbooks considered 'classic' or 'core textbooks' for medical education published since 2001 and held in the University of Ottawa library network were reviewed

for discussion of maintenance drug therapy for indications of DSM-IV bipolar disorder, type I and type II and for DSM-IV major depressive disorder.

Systematic reviews of maintenance drug therapy for depression and bipolar disorder published between 2001 and May 2006 were identified in MEDLINE accessed through the OVID interface using keywords 'depression' and 'bipolar disorder' and limiting to 'EBM Reviews and English Language'. Finally, discussions from all maintenance trials included in chapters 2 and 3 of this project were reviewed to identify the frequency with which the role of ED designs in limiting trial generalizability was noted. As a control, all ED trails identified in Kopec et al (197) were similarly reviewed.

5.3 Results

Guidelines (table 1)

From Guidelines International Network, a total of 3 guidelines were identified, two on depression (250, 337) and one on bipolar disorder(304). The CMA infobase identified 13 guidelines using the keyword depression, and 1 for the keyword bipolar disorder. Based on screening of title and abstract, 1 was selected for relevance to maintenance drug therapy in depression(57) and 1 for bipolar disorder(365).

The National Guideline Clearinghouse identified 398 citations using the keyword 'depression' and 61 with the keyword ' bipolar disorder'. Of these, 11 covered an aspect of the management of depression, not restricted to post-partum depression(11-13, 16, 48, 59, 170, 171, 202, 235, 316, 344). Two referred to the management of bipolar

disorder(14, 304). After removing duplicates, the search yielded a total of 12 treatment guidelines for depression (11-13, 18, 47, 58, 59, 170, 235, 249, 316, 345)and 3 for bipolar disorder(14, 304, 365).

Of the guidelines identified, one (the American Medical Director's Association) was available for review for members only. Three were based entirely on other guidelines. Of the independently generated guidelines for maintenance therapy in depression, all but one used episode counting as the primary way of determining a need for prophylaxis. The exception was the NICE guidelines, which used the quality of recovery from the most recent episode. Kaiser Permanente guidelines were the most aggressive in recommending maintenance therapy after a single episode. As summarized in table 1, ED designs were used as evidence supporting maintenance. The design was referred to as 'RCT' in all but one guideline (NICE), which provided a section on methodological limitations to the evidence base. Of the 12 depression guidelines identifies, 10 declared funding sources. Of the ten, seven were funded by either a private health care firm or by pharmaceutical manufacturers. Three were either government funded or stated 'no conflict of interest'.

In the three identified guidelines for bipolar disorder, North American guidelines recommended maintenance after a single manic episode and Scottish guidelines suggested a more flexible approach to maintenance therapy as a result of diagnostic heterogeneity. Canadian guidelines were outliers in recommending a specific medication (lithium) based on specific symptom profiles, even though this recommendation was based on professional opinion. One guideline (American Psychiatric Association) noted

the use of ED designs as a potential limitation to the generalizability of recent maintenance trials in bipolar disorder.

General Psychiatric Textbooks (table 2)

Two general classic psychiatric texts published since 2001 were held in the University of Ottawa library network, so we also included texts from 1999, increasing the number to 4 (2-4, 6). (Only the most recent editions of a given text were included). In the four texts reviewed, the term RCT was used synonymously with ED designs for antidepressant and mood stabilizer maintenance trials. One text (Harvard Guide) (6) discussed potential withdrawal confounding in lithium trials. One text described concerns with applicability of mood stabilizer trials due to design issues but did not identify the ED design as the specific methodological challenge. The American Psychiatric Association textbook of psychiatry discussed general sampling problems in trials, the use of last observation carried forward and the role of regulatory approval as the prime motivator for clinical trials.

Systematic Reviews (table 3)

From MELINE, 228 systematic reviews were identified, of which 34 were selected for closer screening. 14 were excluded because they dealt only with acute management (26, 64, 121, 123-125, 150, 160, 218, 233, 240, 318, 334, 362), 5 because they were restricted to post-partum depression(88, 168, 210, 216, 253), 1 because it dealt only with sub-threshold depression(8), 1 included only treatment resistant depression(329) and 4 reviewed depression management only at the program level but did not deal specifically

with maintenance drug therapy (23, 174, 251, 303). This left a total of 8 reviews covering maintenance therapy in bipolar disorder(45, 84, 128, 133, 152, 221, 343) and 1 covering maintenance therapy in major depressive disorder(129). As summarized in table 3, five of the eight reviews relied exclusively on ED designs. Three of the five had some discussion related to the methodological limits of the 'RCT's, but did not discuss ED designs as RCT variants requiring a special approach to interpretation.

ED maintenance trials (table 4)

A total of 41 placebo controlled ED maintenance trials for DSM-IV major depressive disorder and 12 trials for DSM-IV bipolar disorder as identified in chapters 3 and 4 of this thesis were reviewed. In addition, all 21 general medical ED trials cited in Kopec et al were reviewed(10, 60, 91, 92, 95, 98, 103, 105, 119, 120, 135, 136, 167, 237, 263, 264, 291, 319, 320, 327, 328, 336, 338). In trial discussions, 59% of antidepressant trials and 25% of mood stabilizer trials discussed potential bias due to rebound/withdrawal symptoms in phase 2, while only 20% of antidepressant trials discussed problems with reduced applicability stemming from ED designs, while 67% of mood stabilizer trials discussed potentially limited applicability due to extended open run-ins.

General medical ED trials discussed the potential for withdrawal-rebound bias in 43% of reports, and the discussed concerns about reduced applicability stemming from open phase 1 treatment in 38% of the reports.

Table 1: Summary of identified clinical guidelines for DSM-IV major depressive disorders and bipolar disorder.

Year	Guideline Developer (total refs.)	Maintenance Recommended?	Level of evidence	ED Design limits Noted or cited?	NNT?	Specific Evidence Cited	Funding
	Depression						
2006	Brigham and Women's Hospital (47)	Yes, if family Hx of recurrent Depression or bipolar disorder, Hx past depressive recurrence within 1Year of discontinuing effective treatment Onset of depression <age 20 Severe, sudden or life-threatening episode (ex. suicide attempt)	Based on APA guidelines(13) (details in this table)	no	no	no	Brigham and Women's Hospital
2006	Institute for Clinical Systems Improvement (ICSI) (170)	First episode: 6-12 months Second episode: 3 years Second episode with complicating factors Ex. dysthymia, recurrence after stopping meds : Lifetime Third episode: Lifetime	Meta-analysis Non-randomized Trial, Consensus Statements ED Trials.	no	no	Yes, 17 Citations including Metaanalysis. No ED trials Trials (13, 100, 108, 114, 129, 142, 180, 181, 183, 257, 259, 288, 297, 312-314, 359) One trial questions ED trials, but this is not discussed in guideline(332)	Blue Cross and Blue Shield of Minnesota Medica, Metropolitan Health plan; ICSI members
2005	American Medical Director's Association (12)	Yes, Specific recommendations available to members only	Specific recommendations available to members only	Specific recommendations available to members only	n/a	Specific recommendations available to members only	Lilly
2005	University of Michigan Health System (344)	Refers to ICSI Guidelins					unclear
2004	Kaiser Permanente Care Management Institute (59)	First episode: 'optional discontinuation after 12 months Second episode: continue 15 months to five years Hx of dysthymia: 15-28 months	Consensus for episodic Depression. 'Evidence based' for patients with history of dysthymia	no	no	Citations unavailable to public. Search strategy in guideline clearinghouse suggests that at least 6 specific citations were included	Kaiser Permanente
2004	NICE depression Guidelines (UK) (249)	Maintenance recommended if recovery is incomplete	Meta-analysis of maintenance trials	Extensive discussion of general limits to all RCT's in psychiatry as well as issues of differential attrition and response thresholds. ED trials not specifically	no	Geddes et al (129) Meta-analysis	independent

				covered but withdrawal symptoms discussed as a general limitation			
2004	Michigan Quality Improvement Consortium	Refers to ICSI Guidelines					No 'conflict of interest'
2004	Singapore Ministry of Health (316)	After three or more episodes OR Two episodes plus family hx of Bipolar disorder, recurrence within a year of discontinuing meds, fam hx. Recurrent MDD, first episode <age 20, sudden or Severe episodes in past 3 years	Evidence graded as 'GPP', good practice point, based on clinical experience of the guideline development group	no	no	Two citations, APA guidelines(13) and Agency for Health Policy and Research Guideline (9)	Singapore Ministry of Health
2003	American Medical Director's Association (11)	Yes, Specific recommendations available To members only	Specific recommendations only available to members	Specific recommendations only available to members	n/a	Specific recommendations only available to members	Lilly
2001	Canadian Psychiatric Association (58)	All patients receive at least 6 months treatment Minimum 2 years if: older age, psychotic features, chronic Symptoms, recurrent episodes (3 or more lifetime), severe episodes	6 months based on 'level 1'(meta-analysis or repeat RCT evidence) 2 years based on at least 1 RCT with placebo or active comparator	no	no	3 supporting citations. 1 ED trial (277) 2 qualitative reviews (116, 349)	Brisol-Meyers Wyeth-Ayerst GlaxoSmithKline Lundbeck Organon Pharmacia
2000	British Association For Psychopharmacology(18)	Maintenance after 5 or more episodes Fewer than 5 episodes if risk factors	RCT evidence, meta-analysis; cited placebo controlled RCT's are ED designs	General methodological limits to antidepressant trials noted, including a section on differential response in milder cases and withdrawal symptoms for all antidepressants	*no*	Specifically cites Frank(117), chronic depression	Not stated
2000	American Psychiatric Association (325 refs) (13)	Maintenance loosely defined, possibly lifetime; use if risk of recurrence deemed high (residual inter-episode symptoms), severe episodes, consider preference, side-effects	RCT evidence (all included trials are ED Trials)	no	no	'over 20 trials in maintenance phase have generally demonstrated effectiveness' RCT's conducted' refers to ED trials, summarized in one review(322) and two summaries of a	Not stated

						nationally funded ED trial (117, 200)	

Table 1: continued

	Bipolar Disorder						
2005	Canadian Network for Mood and Anxiety Network (CANMAT)	Maintenance after one moderately severe manic episode.	Level 1 evidence including RCT's (ED designs) and meta-analyses of ED trials. States 'there are more and better predictors for response to lithium than for other meds'(p30), cites opinion Paper based on open series (148)	no	no	Three meta-analyses of ED trials, 6 ED trials cited (40, 43, 49, 54, 81, 128, 144, 157, 193)	unclear
2005	Scottish Intercollegiate Guidelines Network (SIGN) (182 refs)	Maintenance recommended, but flexibility needed due to heterogeneous illness course and patient acceptability.	Level 1 evidence, including ED trial evidence and systematic reviews	Indirectly, uncertainty as to which patients benefit and heterogeneity of patients is noted	no	8 ED Trials cited as evidence for lithium and anticonvulsant maintenance. One systematic review, Geddes et al for lithium evidence (41, 42, 54, 73, 86, 128, 144, 217, 254, 261)	unclear
2002	American Psychiatric Association (472 refs)	Recommended with lithium, divalproex or oxcarbazepine after a manic episode. Duration unclear	Recommended with 'substantial clinical confidence' (in bipolar type I), Moderate evidence for bipolar type II	Use of pre-randomization open phase described for two trials, but general interpretive challenges not identified(42, 54)	no	Four naturalistic lithium studies, seven ED lithium and anticonvulsant trials, two crossover trials (21, 41, 42, 54, 70, 72, 86, 96, 156, 211, 223, 268, 333)	unclear

Table 1: continued

Textbook	Maintenance Recommendations?	Discussion of methods and limits to evidence base?	Specific discussion of high prevalence and limits of ED designs?
Kaplan & Saddock (2005)(4)	Yes, for recurrent or chronic depression. 'When to discontinue treatment is unclear' (p1660)	Yes. One ED trial described in detail (50)with concerns raised about generalizability (p1687)	No
American Psychiatric Association Textbook of clinical psychiatry (2003)(3)	Yes, for recurrent or chronic depression. 'Of 23 prospective double blind randomized studies of continuation and maintenance pharmacotherapy for depression, 22 found more recurrence in the placebo group' (p513).	Yes. General discussion of sampling problems, high drop rates, use of last observation carried forward, role of regulatory approval driving trials. (p491)	No. ED trial = RCT
Textbook of Psychiatric Disorders (2000)(2)	Yes, for 'some patients'. (p1144) Four ED trials cited as support	No	No
Harvard Guide to Psychiatry (1999)(6)	Yes. Refers to widespread differences in opinion re: duration of treatment and selection of patients (p479)	Discusses potential confounding withdrawal effect	No

Table 2: ED trials identified in general psychiatric textbooks.

Year, Author	# RCT's cited	Proportion ED/total RCT's	ED design limits Discussed?
2004 Geddes(128)	5	100%	Partial discussion. Open treatment alluded to indirectly. Focus on potential withdrawal bias, but selection and generalizability challenges related to extended open phases not identified.
2004 Hadjipavlou(152)	7	4/7	ED named 'Double-blind RCT' and treated as same level of evidence
2003 Geddes(129)	31	100%	Potential for withdrawal bias discussed, but role ED designs per se as a limit to generalizability not addressed.
2003 Brambilla(45)	4	0	Review focuses only on atypical antipsychotics, for which RCT maintenance trials are only starting to be conducted. Review cites only open continuation/maintenance trials and 4 acutely randomized short term trials (<6 wks). No discussion of ED trials or their role in the evidence base for maintenance medication.
2003 Tondo(343)	6	100%	Meta-analysis combines ED trials and open trials. A general overview of the literature to look at long-term medication response as a function of medication type used. No specific discussion of ED designs or other specific design limits. General discussion of shortage of high quality maintenance trials. Concludes that literature is inadequate to make valid correlations between illness course or symptom patterns and response to long-term medications.
2001 Burgess (49)	9	7/9	This is the Cochrane review of lithium maintenance therapy. Randomized 'Discontinuation trials' were formally an exclusion criteria, and trials conducted prior to 1973 were therefore excluded. These trials randomized patients on long-term lithium to placebo discontinuation. In contrast, ED trials use 8-12 week active run-ins. A Comment registered in 2004 addressed this apparent contradiction. The comment was addressed by Geddes who stated that 8-12 week open run-ins should be treated more like placebo run-ins used to ensure compliance, but recognized that this position is a compromise.
2001 Ghaemi(133)	7	100%	ED trials labeled Double blind RCT's, but no discussion of maintenance design limits
2001 Macritchie(221)	1	100%	Cochrane review of divalproex maintenance. Cites trial with 3 month open run-in. Run-in allowed investigator choice of medications, divalproex or lithium prior to randomization to divalproex, lithium or placebo. Concerns raised about early drop-out of more severe cases, so trial may not reflect clinical population. Concerns raised about possible lithium rebound syndrome for cases discontinued from lithium after the open run-in.
2001 De Leon(84)	9	6/9	No. Qualitative review covered all anticonvulsant mood stabilizers, but did not deal specifically with design variants. Sampling issues and ethical challenges introducing placebo arms to chronically ill patients were noted as restrictions for future trials.

Table 3: Systematic Reviews of Maintenance therapy in DSM-IV major depressive disorder and bipolar disorder

	All ED trials in Chapters 2 and 3 (total N=53) (%)	Antidepressant ED Trials (N=41) (%)	Mood Stabilizer ED Trials (N=12) (%)	General medical ED trials in Kopec (mixed topics) N=21
Potential bias due to rebound/withdrawal Symptoms discussed. Includes special sub-group analysis to test for rebound/withdrawal bias	27 (51)	24 (59)	3 (25)	9 (43)
Potential impact of open run-in discussed as limit to applicability independent of diagnostic inclusion/exclusion criteria?	22 (42)	8 (20)	8 (67)	8(38)

Table 4: ED Trials included in Chapters 3 and 4 in which at least one sentence of the discussion or introduction refers to potential limits of extended open run-ins or withdrawal-related inflation of effect size (aggregate, antidepressants and mood stabilizers). Comparator group is a full sample of ED trials conducted for a range of medical conditions as cited by Kopec et al(197).

Discussion:

All psychiatric guidelines and textbooks referred to some form of maintenance therapy following recovery from an acute mood episode, but specific recommendations varied widely. At one end of the spectrum, Kaiser Permanente's depression guidelines recommended 'optional discontinuation' a year after the first episode and five years to life after the second episode. At the other end of the spectrum, the NICE guidelines (249) reserved maintenance therapy for the most highly recurrent or chronic cases with residual depressive symptoms. Recommendations for maintenance in bipolar disorder also differed between developers. Canadian (CANMAT) guidelines and US (APA) guidelines recommended indefinite maintenance following a singe manic episode, while caution regarding blanket recommendations for maintenance was suggested by the SIGN guidelines, due to diagnostic heterogeneity within DSM-IV 'bipolar disorder'. Canadian

guidelines were alone in their endorsement of a relationship between symptom sub-types and clinical response to lithium[10]. The relationship between effect sizes stemming from classic RCT's and discontinuation trials was not discussed. Both ED and acutely randomized RCT's were cited in support of maintenance therapy.

Guidelines assigning 'levels of evidence' to maintenance therapy for DSM-IV major depressive disorder equated ED trials with RCT's and systematic reviews of ED trials as the highest grade of evidence[11]. Only one guideline (developed in collaboration with NICE) dealt in some detail with the methodological problems facing antidepressant trials, dedicating two pages to sampling, diminishing response with diminishing illness severity, high drop rates and investigator conflict of interest(249). Special limits or controversy surrounding ED maintenance trials as applied to antidepressants were not addressed by any of the guidelines.

Only two identified systematic reviews extended open run-ins as a potential source of limited applicability. Of these, the Cochrane lithium review took a unique approach to ED designs, excluding trials with 1-2 year open run-ins conducted in the 1960's due to potential selection bias, while including trials with 16 week open run-ins. The apparently arbitrary cut-off for run-in duration (i.e., the duration considered likely to cause selection bias) was raised in a comment to the authors at the end of the review (49). The authors countered with the argument that open run-ins reflect clinical practice, since people must

[10] Supporting evidence for the predictors was at the expert opinion level. In contrast, APA guidelines made specific reference to open run-in phases of ED trials as a potential limit to predictive models for medication response in bipolar disorder.

[11] Singapore guidelines ranked the evidence for maintenance therapy below RCT levels without providing specific rationale for the modified confidence level.

first recover from acute episodes and tolerate acute phase treatments before starting maintenance. While the author's position is intuitively appealing for the most severe or highly recurrent cases, particularly those for whom acute placebo randomization may not be ethically acceptable, it remains difficult to answer the sceptic's challenge of differential phase 1 attrition and the resulting circular definition of 'responder'. Discussions of diminishing returns of long term medications with less severe illness forms appeared in the SIGN guidelines. A similar discussion of illness thresholds and diminishing medication effects appeared in two out of four textbook reviews of mood stabilizer maintenance therapy in bipolar disorder. Discussion regarding ED designs and their potential limitations appears in textbooks and systematic reviews covering mood stabilizers over the past five years, while ED designs appear to have received virtually no discussion in the context of antidepressant trials at the level of guidelines, systematic reviews or textbook chapters, where they are referred to as RCT's. Antidepressant trials account for approximately 80% of all published ED maintenance trials for DSM-IV mood disorders (table 5).

At the individual trial level, over half of the antidepressant maintenance trials reviewed in chapters 3 and 4 of this thesis contained discussions of potential rebound-withdrawal bias, but only 20% discussed potentially limited applicability. This contrasts with 25% discussing withdrawal bias and 67% discussing limited applicability in mood stabilizer trials. Similar proportions of general medical ED trials identified by Kopec et al discussed these limitations.

The limited recognition of ED trials as distinct RCT variants in the psychiatric literature is surprising in light of their widespread application and their potential to impact the evidence base as a whole. However, it should be noted that other methodological issues related to applicability, including sampling techniques and high drop-outs also received little attention.

Finally, this chapter is itself exploratory and limited in scope. The general theme is identifying the extent to which ED trials are recognized in the literature. Preliminary indications point to reasons for debating the advantages an disadvantages of ED trials, a process that could help raise awareness of the current 'one size fits all' approach to testing long-term psychotropic medications.

Chapter 6

Discussion

Enrichment-discontinuation designs for antidepressants and mood stabilizers: Efficient in theory but interpretable in practice?

> "If important alternative hypotheses are compatible with available evidence in investigations of causal relationships, the primary research question is unsettled, even if the observed association is established from a comparative randomized clinical trial." Cornfield (Cited in Greenhouse et al 1993) (143)

6.1 Statement of principal findings:

ED designs dominate randomized mood stabilizer and antidepressant trials of more than 20 weeks duration, accounting for approximately 90% of the published reports. Questions about bias, either from phase 1 differential attrition favoring open interventions or from withdrawal-rebound effects are therefore highly relevant to the application of this RCT variant in testing psychiatric maintenance drug therapy.

Mood stabilizers

Published reports in the mood stabilizer class for which comparisons were possible included trials of lithium vs placebo and trials of lithium vs anticonvulsants. Only 'indirect comparisons' were possible. Table 1 summarizes the results.

ED designs favored phase 1 interventions in placebo controlled lithium trials (table 1). This was not replicated in active comparator trials of lithium vs anticonvulsants, where

none of the indirect comparisons suggested an enrichment effect favoring the phase 1 intervention.

Comparison	**Trials used to generate comparison (n)**	Result summary
Lithium vs Placebo ***over placebo***		***larger OR favors lithium***
Direct Comparisons	0	n/a
Indirect Comparisons	5 (2 classic RCTs, 3 lithium enriched, 2 anticonvulsant enriched)	Favors phase 1 intervention Classic RCT OR 3.18 (95% CI 0.65-15.46) Lithium enriched OR 22.0 (95% CI 7.0-68.7) Anticonvulsant enriched OR 1.9 (95% CI 1.2-2.8)
Lithium vs Anticonvulsant ***over anticonvulsant***		***larger OR favors lithium***
Direct Comparisons	0	n/a
Indirect Comparisons	8 (3 classic RCTs, 2 lithium enriched, 3 anticonvulsant enriched)	Does not favor phase 1 intervention Classic RCT OR 0.80 (95% CI 0.36-1.79) **Lithium Enriched OR 1.27 (95% CI 0.52-3.10)** Anticonvulsant Enriched OR 1.07 (95% CI 0.78-1.39)

Table 1: Summary of mood stabilizer maintenance trial comparisons. An apparent 'enrichment effect' in placebo controlled lithium trials is not replicated among active comparator trials involving lithium and anticonvulsants.

Antidepressants

Reports of antidepressant trials offered few contrasting treatment strategies for inclusion in comparison arms of this study (summarized in table 2). All tricyclic antidepressant (TCA) trials used the same design with an extended open phase followed by placebo discontinuation. This single treatment strategy allowed no comparisons. SSRI reports provided alternative treatment strategies for indirect as well as direct comparisons. A total of four classic RCT's were available for indirect comparisons. As shown in table 2 below, these comparisons showed a trend toward ED design bias.

Comparison	Trials (n)	Result summary
SSRI vs Placebo		***Smaller OR favors SSRI over placebo***
Direct Comparisons	2	See table 3
Indirect Comparisons	27 (23 SSRI enriched, 3 classic RCTs, 1 placebo 'enriched')	**Does not favor phase 1 intervention** Classic RCT OR 0.55 (95% CI 0.28-1.07) SSRI enriched OR 0.37 (95% CI 0.33-0.42) Placebo 'enriched' OR 0.41 (95% CI 0.23-0.72)
TCA's: No comparisons available		n/a

Table 2: Summary of antidepressant comparisons. Indirect comparisons do not show evidence of bias, but only 4 alternative designs were reported.

An attempt was made to obtain unpublished data for cumulative relapses from a 24-week classic RCT design in which 306 depressed patients were randomized to placebo SSRI treatment (Stahl 2000) (323). A request from the trial sponsor was only partially granted (we were provided with endpoint responder data from which cumulative relapses could not be calculated), so we are unable to include this trial with our indirect comparisons at this time.

Results from two 'direct comparisons' are mixed. One report (Emslie, 2004 (101)) suggests that the design favors the phase 1 intervention, while the other report (Montgomery 1993(244)) does not (table 3).

Direct SSRI vs Placebo comparisons	Classic RCT P-P vs SSRI-SSRI	SSRI enriched SSRI-SSRI vs SSRI-P
Montgomery '93 (adult population, poor reporting quality)	.30 (95% CI .13-.69) Favors SSRI N=165	.26 (95% CI= .11-.64) **Favors SSRI** N= 147
Emslie '04 (adolescent population, excellent reporting quality)	2.07 (95% CI = .57-7.59) **Favors placebo** N=55	.29 (95% CI = .08-1.06) **Favors SSRI** N=40

Table 3: Summary of cumulative relapses under contrasting treatment strategies from direct comparison trials. Mixed evidence of bias.

Discussion of special limits of ED designs in the psychiatric literature

Fifty-nine percent of antidepressant maintenance trials using the ED design included discussions of potential withdrawal bias stemming from the discontinuation of phase 1 interventions vs 25% of ED mood stabilizer trials. In contrast, only 20% of ED antidepressant trials discussed potentially limited applicability while 67% of ED mood stabilizer trials discussed this limitation. A brief survey of ED trials in general medicine suggested that a similar proportion of reports specifically address potential withdrawal bias or limited applicability. Thirteen out of 15 treatment guidelines equated ED trials with classic RCT's. Two out of 9 systematic reviews of mood stabilizer or antidepressant drug maintenance therapy identified and discussed ED trials as distinct from classic RCT's. This suggests that the dominant role played by ED designs in psychiatric drug maintenance trials and the resulting uncertainty over the applicability of ED reports is broadly under-recognized (176).

6.2 Strengths and weaknesses of the study:

In a theoretical exercise, it is possible to separate the signal of design effects from the noise of explicitly named and modeled confounders. Theoretical modeling can be carried out in the absence of any real-world application, and the results later applied to situations in which key assumptions are met. A meta-epidemiological approach on the other hand, is qualitatively different. It attempts retrospectively, to separate a signal from the unavoidable noise of the experimental situation using a 'strength in numbers' approach. This method of gathering and interpreting data has advantages and disadvantages.

The major strength of our approach is its broad perspective, allowing relevant hypotheses to flow from unexpected observations. For example, in this thesis, our observation of an apparent 'mono-operational bias' (308) raises questions about the construct of 'medication responder', particularly for long-term antidepressants (176). Mono-operational bias is described by Shaddish et al as complicating scientific inference, since "any one construct both under represents the construct of interest and measures irrelevant constructs" (Shaddish et al p. 75) (308). Paradoxically, our finding of mono-operational bias, while provocative, goes hand in hand with a 'mono-method bias', which limited our ability to carry out planned comparisons, playing to the weak suit of our method. This made it difficult to sort the 'signal' of bias from the noise of patient heterogeneity. Referring to monomethod bias, Shaddish et al state ..., " when all operationalizations use the same method [in this case, ED designs], that method is part of the construct actually studied [in this case, response to drug maintenance therapy]" (Shaddish p.75).

As an observational study, this thesis is prone to selection bias, which we attempted to minimize by setting up rival hypotheses prospectively and by using a systematic approach to selecting and abstracting data (1). During the course of the thesis, it became clear that the practical issues involved in conducting the proposed series of systematic reviews was daunting. We addressed the issue by using trials identified by DF in his recent systematic review of SSRI's (107) and supplementing by identifying trials from published systematic reviews of antidepressants. This introduced an inevitable loss of control over the trial identification process, but was essential to ensure timely thesis completion. Also, while we had initially intended to have two reviewers abstracting and

rating trials (DD and JA), practical realities made this possible for only one medication (lithium), while the remainder of data abstraction and trial quality ratings had to be completed by a single reviewer (DD). This introduced the potential for observer bias. By using both direct and indirect comparisons, our approach attempted to avoid a mono-operational bias.

At the individual trial level, blinding of 'responders' to phase 2 allocation in ED trials may be incomplete, since experience with side effects in phase 1 (or prior to the trial) may effectively reveal to patients their allocation to placebo or competitor. This in turn may alter patient's interpretation of mood fluctuations, introducing a 'wish bias'(27) against allocation to placebo which could inflate ED trial outcomes. None of the reports attempted to correlate patients' ability to identify their treatment with outcome. As we proposed in chapter 3, trial quality, including questions about the quality of blinding, may explain the cohort effect among placebo controlled lithium maintenance trials, as can trial size, use of tapering, and year of publication (figure 1). Trial quality also can explain significant differences in reported effect size between antidepressant trials.

We did not formally explore the implications of dropout rates ranging from 50-85% at trial completion, which may overshadow the more subtle selection effects we were looking for. Rather, we assumed that the factors influencing long-term adherence were equally distributed across the treatment strategies.

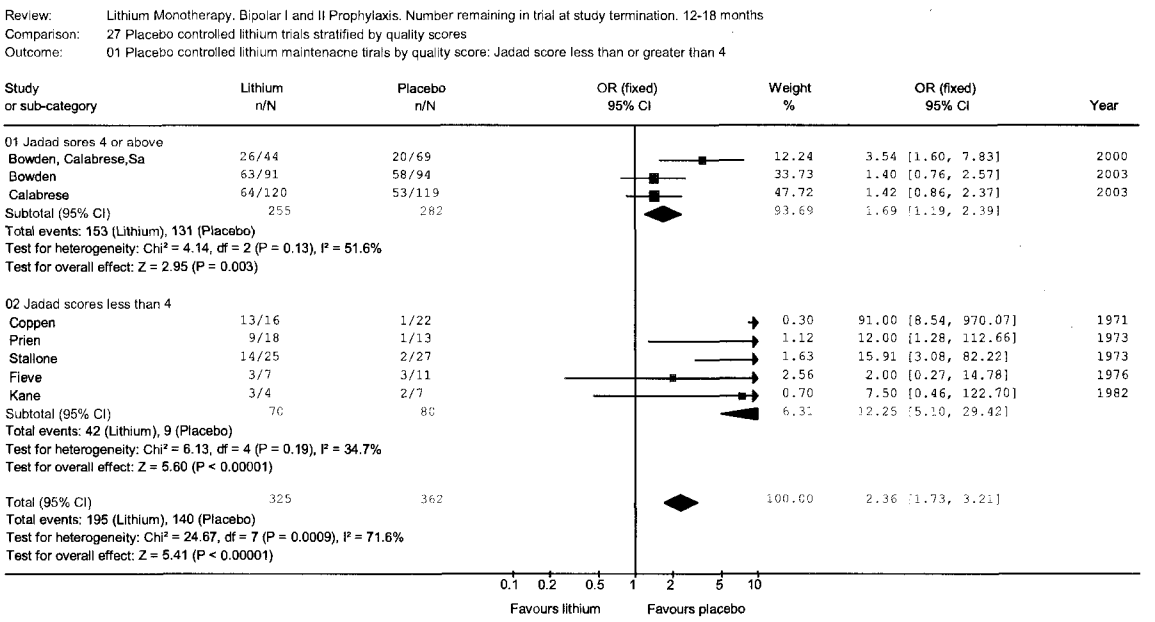


Figure 1: A cohort effect among placebo controlled lithium trials grouped according to Jadad quality scores, size, use of lithium taper, and year of publication. The cohort effect may be explained by any or all of these factors, in addition to preferential selection of lithium responders in phase 1.

6.3 Strengths and weaknesses in relation to other studies:

Kopec et al in their seminal methods paper cautioned that while ED designs are theoretically promising under strict assumptions, their value would ultimately have to be tested in the context of real-life chronic illness and real-life interventions. We have applied such a test to an entire body of clinical evidence.

Questions about medication-specific selection bias and patient classification bias in psychiatric ED trials go back at least as far as the Shepherd/Schou debate, and were picked up by Greenhouse et al, who added concerns that phase 2 withdrawal-related relapses were prone to misclassification as 'medication responders'. As a solution, they proposed a method using a series of comparisons between alternative treatment strategies for interpreting ED trials (summarized in chapter 2, figure 2). In this thesis, we

independently arrived at, expanded and tested an analogous approach, which we named 'direct and indirect comparisons'. In contrast to Greenhouse et al, who focused on a re-analysis of a single trial, our study looked at ED designs across two large classes of psychiatric illnesses and two broad classes of drug treatments.

The main drawback in relation to Greenhouse et al is our relative lack of data, resulting in unstable estimates. Greenhouse et al were able to avoid this problem by focusing on original data from a single trial.

Finally, we noted that all reports published since 1995 reflected industry-sponsored trials, whereas the majority of earlier trials were sponsored by trusts and government agencies. The implications for trial objectives, designs, conduct, reporting and relevance remain to be established (75). Inconsistent reporting styles made it difficult or at times impossible to follow patient cohorts from inception to trial termination. Examples of the latter include one large classic RCT referred to above (Stahl 2000) (323) and the large Montgomery 'direct comparison' trial, which is only partially reported across four separate reports in four separate journals and with differing authorship (full details in chapter 4 discussion). Only one of four trials used for our direct comparisons used a clear flow diagram consistent with the CONSORT statement (238).

6.4 Meaning of the study

At the outset of this thesis, we hypothesized that a cohort effect among placebo controlled lithium maintenance trials in bipolar disorder may be due to the self-selection

effects of 'lithium responders' in phase 1 (89). If replicated, this would have not only suggested that ED designs are biased in favor of phase 1 interventions, but it would have lent support to the hypothesis that responder self-selection effects can be observed at the level of published reports and eventually tested on individual interventions (148, 162). During the course of the study, it became evident that such selection effects, if present, are too subtle to be detected using published data (90). We reached this conclusion both because of a lack of replication and the plausibility of several alternative explanations for the cohort effect. This should not however, be taken as evidence that ED designs are unbiased when applied to psychiatric maintenance drug trials across a widely heterogeneous patient groups and different interventions.

Theoretically, ED designs offer a way to demonstrate efficacy while randomizing fewer patients than classic RCT's. When key assumptions about illness course, phase 1 non-differential attrition and phase 2 withdrawal-rebound effects are satisfied, they may be a preferred method for establishing efficacy in chronic conditions. A meta-epidemiological approach was intended to address the problem of heterogeneous illness courses between comparator trials. However, an unforeseen mono-operational bias (308) extending over three decades, in which the construct of 'medication response' is defined retrospectively in phase 2 of ED designs (176) has ironically returned us to a vestige of the Shepherd Schou debate (31) – the problem of misclassifying patients with intermittent natural illness courses as distinct from those for whom the natural course is better described as 'chronic'. In the former, as long as the duration between episodes is longer than the study duration, maintenance medication would not be expected to confer benefit over placebo. In the latter, 'maintenance' treatment can be viewed more like long-term treatment than

as prophylaxis, and treatment effects would be expected to be relatively consistent over the long term.

Because it is difficult or impossible to causally differentiate withdrawal-rebound symptoms from relapses due to the removal of an effective therapy, even if a large number of alternative treatment strategies were available to overcome patient heterogeneity, our approach could only conclude that some combination of responder selection (phase 1 differential attrition) and withdrawal-rebound explained any observed bias.

The plausibility of withdrawal confounding as a contributor to bias in SSRI ED trials is supported through narrative reports that 80% of relapses occur in the first 6-8 weeks of medication withdrawal (101, 244). In our review, we found two antidepressant reports from which it was possible to estimate the ratio of 6-month outcomes in phase 2 participants who had not relapsed in the first 8 weeks following randomization. As illustrated in figure 2, no statistically significant difference is observed between relapses in placebo arms vs relapses in medication arms.

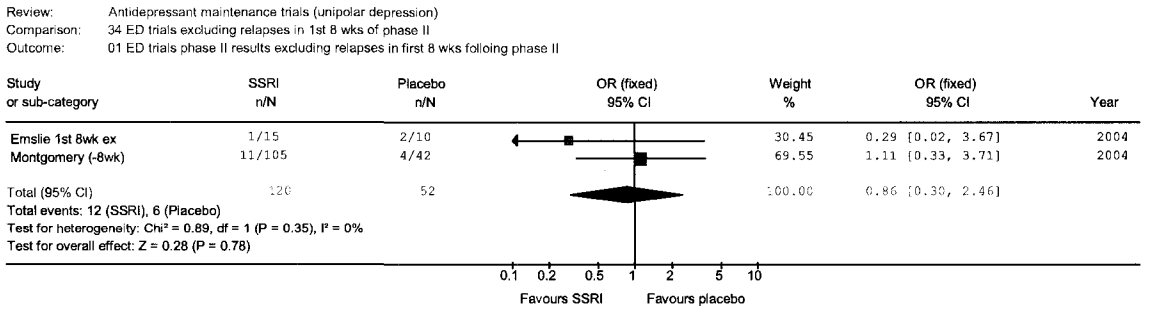


Figure 2: 6-month outcomes in ED antidepressant trials after relapses in the first 8 weeks are excluded.

Broadly, there are five ways to view ED trials as applied to psychiatric maintenance drugs:

1. The design is unbiased and provides a realistic reflection of outcomes in compliant patients recovering on medications.
2. The design favors phase 1 interventions due to an intervention-specific selection bias.
3. The design overstates efficacy due to phase 2 medication withdrawal-rebound effects.
4. Some combination of 2 and 3
5. There is inadequate information available to decide.

Based on our analysis, particularly for SSRI trials, acceptance of option 1 would require a leap of faith. For reasons outlined above, options 2 and 3 cannot be sorted out. Option 4, while attractive, ignores the unexplained but important possibility that problems with blinding ED trials results in a 'wish bias' favoring medications in phase 2 (27).

Finally, it is difficult to ignore the potential for conflict of interest (75) in the conduct and reporting of psychiatric ED trials, all of which have been industry sponsored since 1995. Specifically, if a disinterested understanding of drug maintenance was the sole aim of the trials included in this thesis, one would have expected greater attention to replicable sampling techniques and to the resolution of lingering questions about blinding in ED trials raised more than thirty years ago in the context of lithium.

In light of the limited data available we conclude that the question of bias in ED designs as applied to antidepressant and mood stabilizer trials, and therefore to the respective evidence bases, remains unanswered.

6.5 Unanswered questions and future research:

Because of the limited observations available, we were unable to 'control' the impact of inconsistent classification of patients with different illness courses across reported ED trials. Because of inconsistent reporting quality and heterogeneous patient groups (see discussion, chapter 4), conflicting reports are difficult to interpret.

One feasible strategy to reduce the uncertainty about design bias relative to SSRI trials would be to add data to our group of classic RCT's with continuation or maintenance phases, thus increasing the number of 'indirect comparisons' available. At the time of submission, we have re-requested cumulative relapse data from a large trial using a classic RCT design (323). Access to this data is important in light of the mixed outcome data from placebo responders in the three trials making up our SSRI indirect comparison.

A second strategy would be to re-analyze the two 'direct comparison' trials using individual patient-level data (66). This would allow us not only to look at 'time to relapse' but also to stratify individual cases by risk factors such as the previous number of episodes. To date however, access to this information has not been possible. Because uncertainty about patient heterogeneity is likely to impact the interpretation of indirect comparisons, the direct comparison approach is preferable for evaluating bias in ED designs as applied to psychiatric drug maintenance therapy.

In the SSRI review, two observations seemed counterintuitive. First, the consistency of ED trial effect sizes regardless of run-in duration and across different phase 1 dropout rates is puzzling. As discussed in chapter 4, we would have intuitively expected lower effect sizes with longer run-ins (since the natural course of depression is toward spontaneous resolution) and greater effects with higher phase 1 dropouts (since phase 1 differential attrition is more likely when drop-rates are high than when they are low). The fact that neither was observed raises the likelihood that issues other than illness course or responder selection are impacting ED trials in psychiatry. These issues potentially include withdrawal-rebound effects and blinding problems.

Finally, since most treatment guidelines now recommend treatment of any episode of depression for at least 6 months and a second episode for at least a year, it is evident that clinical equipoise has shifted over the past three decades. As a result, the current evidence base of ED trials, with open treatment phases of 8-16 weeks is becoming dated and is in need of recalibration to the new realities of clinical practice. If the evidence is to be re-calibrated in light of contemporary equipoise, it is likely that ED designs will again be called upon to test long-term outcomes, and an opportunity to test the questions arising in this thesis may emerge in the process.

References

1. Systematic Reviews in Health Care. London: BMJ Books, 2001.

2. Treatments of Psychiatric Disorders. 2000.

3. The American Psychiatric Association Textbook of Clinical Psychiatry. Washington DC: 2003.

4. Kaplan and Sadock's Comprehensive Textbook of Psychiatry. Philadelphia: Lippicott Williams and Wilkins, 2005.

5. Consensus guidelines outline drug selection and sequencing options for bipolar disorder. Report on Medical Guidelines & Outcomes Research.11(10):1-2, 5, 2000.

6. The Harvard Guide to Psychiatry. Cambridge, Massachusetts: The Bellknap Press of Harvard University Press, 1999.

7. Abou-Saleh M, Coppen A. Who responds to prophylactic lithium? Journal of Affective Disorders 1986;10:115-25.

8. Ackermann RT, Williams JW Jr. Rational treatment choices for non-major depressions in primary care: an evidence-based review. Journal of General Internal Medicine 2002;17:293-301.

9. Agency for Healthcare Ploicy and Research. Treatment of Major Depression. AHCPR Publications 1993;Vol. 2.

10. Ahern MJ et al. D-Penicillamine withdrawal in rheumatoid arthritis. Ann Rheum Dis 1984;43:213-7.

11. American Medical Directors Association. Depression. American Medical Directors Association 2003.

12. American Medical Directors Association. Pharmacotherapy companion to the depression. American Medical Directors Association 2005.

13. American Psychiatric Association. Practice guideline for the treatment of patinets with major depresive disorder (revision). American Journal of Psychiatry 2000;157:1-45.

14. American Psychiatric Association. Practice Guidelines for the treament of patients with bipolar disorder (revision). American Journal of Psychiatry 2002;159:4-51.

15. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition. Washington D.C.: American Psychiatric Association, 1994.

16. American Psychiatric Association practice guideline for the treatment of patients with major depressive disorder. Practice guideline for the treatment of patients with major depressive disorder. American Journal of Psychiatry 2000;April:1-45.

17. Amery W, Dony J. A clinical trial design avoiding undue placebo treatment. Journal of Clinical Pharmacology 1975;15:674-9.

18. Anderson IM et al. Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 1993 British Association for Psychopharmacology guidelines. Journal of Psychopharmacology 2000;14:3-20.

19. Angst J et al. Lithium prophylaxis in recurrent affective disorders. British Journal of Psychiatry 1970;116:604-14.

20. Arroll B et al. Efficacy and tolerability of tricyclic antidepressants and SSRI's compared with placebo for the treatment of depression in primary care: A Meta-Analysis. Annals of Family Medicine 2005;4:449-56.

21. Baastrup PC et al. Prophylactic Lithium: Double Blind Discontinuation in Manic-Depressive and Recurrent-Depressive Disorders. Lancet 1970;326-30.

22. Baastrup PC, Schou M. Lithium as a prophylactic agent. Archives of General Psychiatry 1967;16:162-72.

23. Badamgarav E et al. Effectiveness of disease management programs in depression. American Journal of Psychiatry 2003;160:2080-90.

24. Baldessarini RJ, Tohen M, Tondo L. Maintenance treatment in bipolar disorder.[comment]. Archives of General Psychiatry.57(5):490-2, 2000.

25. Baldwin DS et al. A double-blind, randomized, parallel-group, flexible-dose study to evaluate the tolerability, efficacy and effects of treatment discontinuation with escitalopram and paroxetine in patients with major depressive disorder. International Clinical Psychopharmacology 2006;21:159-69.

26. Barbui C, Guaiana G, Hotopf M. Amitriptyline for inpatients and SSRIs for outpatients with depression? Systematic review and meta-regression analysis. Pharmacopsychiatry 2004;37:93-7.

27. Barbui C et al. "Wish Bias" in antidepressant drug trials? Journal of Clinical Psychopharmacology 2004;24:126-30.

28. Bialos D. et al. Recurrence of Depression After Discontinuation of Long-term Amitriptyline Treatment. American Journal of Psychiatry 1982;139:325-8.

29. Bjork K. The efficacy of zimeldine in preventing depressive episodes in recurrent major depressive disorders -- a double blind placebo-controlled study. Acta Psychiatrica Scandinavica 1983;68:182-9.

30. Blackburn S. Oxford Dictionary of Philosophy. Oxford: Oxford Univerity Press, 1996.

31. Blackwell B, Shepherd M. Prophylactic lithium: Another therapeutic myth? Lancet 1968;May 4:968-71.

32. Blazer DG et al. The prevalence and distribution of major depression in a national community sample -- The National Comorbidity Survey. American Journal of Psychiatry 1994;151:979-86.

33. Blomhoff S. et al. Randomised controlled general practice trial of sertraline, exposure therapy and combined treatment in generalized social phobia. British Journal of Psychiatry 2001;179:23-30.

34. Blondal T. et al. The effects of fluoxetine combined with nicotine inhalers in smoking cessation -- a randomized trial. Addiction 1999;94:1007-15.

35. Bouman TK et al. The effectiveness of lithium prophylaxis in bipolar and unipolar depressions and schizoaffective disorders. Journal of Affective Disorders 1986;11:275-80.

36. Bowden CL, Singh V. Valproate in bipolar disorder: 2000 onwards. Acta Psychiatrica Scandinavica 2005;426:13-20.

37. Bowden CL. Acute and maintenance treatment with mood stabilizers. International Journal of Neuropsychopharmacology.6(3):269-75, 2003.

38. Bowden CL. Valproate. Bipolar Disorders 2003;5:189-202.

39. Bowden CL. Clinical correlates of therapeutic response in bipolar disorder. [Review] [84 refs]. Journal of Affective Disorders.67(1-3):257-65, 2001.

40. Bowden CL et al. Efficacy of divalproex vs lithium and placebo in the treatment of mania. The Depakote Mania Study Group.[see comment][erratum appears in JAMA 1994 Jun 15;271(23):1830]. JAMA.271(12):918-24, 1994;-30.

41. Bowden CL et al. A randomized, placebo-controlled 12-month trial of divalproex and lithium in treatment of outpatients with bipolar I disorder. Divalproex Maintenance Study Group.[see comment]. Archives of General Psychiatry.57(5):481-9, 2000.

42. Bowden CL et al. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I

disorder.[erratum appears in Arch Gen Psychiatry. 2004 Jul;61(7):680]. Archives of General Psychiatry.60(4):392-400, 2003.

43. Bowden CL et al. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently manic or hypomanic patients with bipolar I disorder.[erratum appears in Arch Gen Psychiatry. 2004 Jul;61(7):680]. Archives of General Psychiatry.60(4):392-400, 2003.

44. Bowden CL, McElroy SL. History of the development of valproate for treatment of bipolar disorder. Journal of Clinical Psychiatry.56 Suppl 3:3-5, 1995.

45. Brambilla P, Barale F, Soares JC. Atypical antipsychotics and mood stabilization in bipolar disorder. Psychopharmacology 2003;166:315-32.

46. Breum L. et al. Long-term effects of fluoxetine on glycemic control in obese patients with non-insulin-dependent diabetes mellitus or glucose intolerance: Influence on muscle glycogen synthase and insulin receptor kinase activity. Metabolism 1995;44:1570-6.

47. Brigham and Women's Hospital. Depression. A guide to diagnosis and treatment. Brigham and Women's Hospital 2006.

48. Brigham and Women's Hospital. Depression. A guide to diagnosis and treatment. Brigham and Women's Hospital 2001.

49. Burgess S et al. Lithium for maintenance treatment of mood disorders. Cochrane Database of Systematic Reviews 2001.

50. Calabrese JR et al. A placebo-controlled 18-month trial of lamotrigine and lithium maintenance treatment in recently depressed patients with bipolar I disorder. Journal of Clinical Psychiatry 2003;69:1013-24.

51. Calabrese JR, Delucchi GA. Spectrum of efficacy of valproate in 55 patients with rapid-cycling bipolar disorder. American Journal of Psychiatry.147(4):431-4, 1990.

52. Calabrese JR et al. Spectrum of efficacy of valproate in 78 rapid-cycling bipolar patients. Journal of Clinical Psychopharmacology.12(1 Suppl):53S-56S, 1992.

53. Calabrese JR, Rapport DJ. Mood stabilizers and the evolution of maintenance study designs in bipolar I disorder. [Review] [42 refs]. Journal of Clinical Psychiatry.60 Suppl 5:5-13; discussion 14-5, 1999.

54. Calabrese JR et al. Long-term treatment of bipolar disorder with lamotrigine. [Review] [27 refs]. Journal of Clinical Psychiatry.63 Suppl 10:18-22, 2002.

55. Calabrese J et al. A double-blind, placebo controlled, prophylactic study of lamotrigine in rapid cycling bipolar disorder. Journal of Clinical Psychiatry 2000;61:841-50.

56. Callahan C.M., Berrios G. Reinventing Depression: A history of the treatment of depression in primary care 1940-2004. New York: Oxford University Press, 2005.

57. Canadian Psychiatric Association. Prescribing antidepressants for depression in 2005: Recent concerns and recommendations. Canadian Journal of Psychiatry 2004;49.

58. Canadian Psychiatric Association and CANMAT Depression work-group. Clinical Guidelines for the Treatment of Depressive Disorders. Canadian Journal of Psychiatry 2001;46:1-91.

59. Care Management Institute KP. Adult primary care depression guidelines. Kaiser Permanente 2004.

60. Casaer P et al. Flunarizine in alternating hemiplegia in childhood. An international study in 12 children. Neuropediatrics 1987;18:191-5.

61. Centorrino F et al. Multiple versus single antipsychotic agents for hospitalized pschiatric patients: Case-control study of risks versus benefits. American Journal of Psychiatry 2004;161:700-6.

62. Chick J., Aschauer H., Hornick K. Efficacy of fluvoxamine in preventing relapse in alcohol dependence: a one-year, double-blind, placebo-controlled multicetre study with analysis by typology. Drug and Alcohol Dependence 2004;74:61-70.

63. Christodolulou GN, Lykouras EP. Abrupt lithium discontinuation in manic-depressive patients. Acta Psychiatrica Scandinavica 1982;65:310-4.

64. Cipriani A et al. Fluoxetine versus other types of pharmacotherapy for depression. Cochrane Database of Systematic Reviews 2005.

65. Cipriani A, Rendell J, Geddes J. Olanzapine in long-term treatment for bipolar disorder. Cochrane Database of Systematic Reviews.(1):CD004052, 2003;2.

66. Clarke M, Stewart L. Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. American Journal of Epidemiology 1998;148:102-3.

67. Coiffier B et al. LNH-84 Regimen: A multicenter study of intensive chemotherapy in 737 patients with aggressive malignant lymphoma. J Clin Oncol 1989;1018-26.

68. Connoly V., Gallagher A., Kesson C. A study of fluoxetine in obese elderly patients with type 2 diabetes. Diabetic Medicine 1995;12:416-8.

69. Cook B.L. et al. Unipolar depression in the elderly. Reoccurrence on discontinuation of tricyclic antidepressants. Journal of Affective Disorders 1986;10:91-4.

70. Coppen A et al. Prophylactic Lithium in Affective Disorders. Controlled Trial. Lancet 1971;275-9.

71. Coppen A. et al. Continuation therapy with amitriptyline in depression. British Journal of Psychiatry 1978;133:28-33.

72. Coryell W et al. Lithium and recurrence in a long-term follow-up of bipolar affective disorder. Psychological Medicine 1997;27:281-9.

73. Coxhead N, Silverstone T, Cookson J. Carbamazepine versus lithium in the prophylaxis of bipolar affective disorder. Acta Psychiatrica Scandinavica 1992;85:114-8.

74. Coyne J. et al. Emotional Disoders in Primary Care. Journal of Consulting and Clinical Psychology 2002;70:798-809.

75. Coyne J. Lessons in conflict of interest: The construction of the martyrdom of David Healy and the dilemma of bioethics. American Journal of Bioethics 2005;5:W3-W14.

76. d'Amato C. et al. Fluoxetine for migraine prophylaxis: A double-blind trial. Headache 1999;39:716-9.

77. Dardennes R et al. Comparison of carbamazepine and lithium in the prphylaxis of bipolar disorders: a meta-analysis. British Journal of Psychiatry 1995;166:378-81.

78. Darga L. et al. Fluoxetine's effect on weight loss in obese subjects. American Journal of Clinical Nutrition 1991;54:321-5.

79. Davidson J. et al. Efficacy of sertraline in preventing relapse of posttraumatic stress disorder: results of a 28 week double-blind, placebo-controlled study. American Journal of Psychiatry 2001;158:1974-81.

80. Davidson J., Raft D. Use of phenelzine in continuation therapy. Neuropsychobiology 1984;11:191-4.

81. Davis JM, Janicak PG, Hogan DM. Mood stabilizers in the prevention of recurrent affective disorders: a meta-analysis. Acta Psychiatrica Scandinavica.100(6):406-17, 1999.

82. Davis LL et al. Comprehensive review of the psychiatric uses of valproate. [Review] [204 refs]. Journal of Clinical Psychopharmacology. 2000;1S-17S.

83. Dawson EB, Moor TD, McGanity WJ. The mathematical relationship of drinking water lithium and rainfall to mental hospital admission. Diseases of the nervous system 1970;31:811-20.

84. De Leon OA. Antiepileptic drugs for the acute and maintenance treatment of bipolar disorder. Harvard Review of Psychiatry 2001;9:209-22.

85. Dekker J., de Jonghe F., Tuynman H. The use of anti-depressants after recovery from depression. European Journal of Psychiatry 2000;14:207-12.

86. Denicoff KD et al. Comparative prophylactic efficacy of lithium, carbamazepine and the combination in bipolar disorder. Journal of Clinical Psychiatry 1997;58:470-8.

87. Denicoff KD et al. Valproate prophylaxis in a prospective clinical trial of refractory bipolar disorder.[see comment]. American Journal of Psychiatry.154(10):1456-8, 1997.

88. Dennis CL, Stewart DE. treatment of postpartum depression, part 1: a critical review of biological interventions. Journal of Clinical Psychiatry 2004;65:1242-51.

89. Deshauer D et al. Re-evaluation of randomized control trials of lithium monotherapy: a cohort effect. Bipolar Disorders 2005;7:382-7.

90. Deshauer D, Fergusson D, Grof P. Reply to Severus, Kleindienst and Griel. Bipolar Disorders 2006;In press.

91. DiBianco R et al. Long-term efficacy of bepridil in patients with chronic stable angina pectoris: results of a multicenter, placebo-controlled study of extended bepridil use. Am J Cardiol 1985;55:50C-4C.

92. DiBianco R et al. Oral amrinone for the treatment of chronic congestive heart failure: reslts of a multicenter randomized double blind and placebo-controlled withdrawal study. J Am Coll Cardiol 1984;4:855-66.

93. DiCostanzo E, Schifano F. Lithium alone or in combination with carbamazepine for the treatment of rapid-cycling bipolar affective disorder. Acta Psychiatrica Scandinavica 1991;83:456-9.

94. Doogan D.P., Caillard V. Sertraline in the Prevention of Depression. British Journal of Psychiatry 1992;160:217-22.

95. Dunk AA et al. The safety and efficacy of tripotassium dicitrato bismuthate (De-Nol) maintenance therapy in patients with duodenal ulceration. Aliment Pharmacol Ther 1990;157-62.

96. Dunner DL, Stallone F, Fieve RR. Lithium carbonate and affective disorders V: a double-blind study of prophylaxis of depression in bipolar disorder. Archives of General Psychiatry 1976;33:117-20.

97. Eaton W.W. et al.
Natural history of Diagnostic Interview Schedule/DSM-IV depression. The Baltimore Epidemiological Catchment Area follow-up. Archives of General Psychiatry 1997;54:993-9.

98. Echt DS et al. Mortality and morbidity in patinets receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Eng J Med 1991;324:781-8.

99. Eklund K, Forsman A. Minimal effective dose and relapse -- double blind trial: haloperidol decanoate vs placebo. Clin Neuropharmacol 1991;14:S7-S12.

100. Ellis PM, Smith DAR. Treating depression: the beyondblue guidelines for treating depression in primary care. Medical Journal of Australia 2002;176:77-83.

101. Emslie G J et al. Fluoxetine Treatment for Prevention of Relapse of Depression in Children and Adolexcents: A Double-Blind, Placebo-Controlled Study. Journal of the American Academy of Child & Adolescent Psychiatry 2004;43:1397-405.

102. Ertugrul A, Meltzer HY. Antipsychotic drugs in bipolar disorder. Neuropsychopharmacology 2003;6:277-84.

103. Evans JR et al. Chonic oral amrinone therapy in congestive heart failure: a double-blind placebo-controlled withdrawal study. Int J Clin Pharm Res 1984;IV:9-18.

104. Even C, Siobud-Dorocant E, Dardennes RM. Critical approach to antidepressant trials. Blindness pretenction is necessary, feasible and measurable. British Journal of Psychiatry 2000;177:47-51.

105. Fabricias PG et al. Efficacy of one-a-day terazosin in benign prostatic hyperplasia. Prosate Suppl 1990;85-93.

106. Feiger A.D. et al. Double-blind, placebo-substitution study of nefazodone in the prevention of relapse during continuation treatment of outpatients with major depression. International Clinical Psychopharmacology 1999;14:19-28.

107. Fergusson D et al. Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. BMJ 2005;330:396.

108. Ferrier IN. Treatment of major depression: is improvement enough? Journal of Clinical Psychiatry 1999;60:10-4.

109. Field MJ, Lohr KN (Eds.). Clinical Practice Guidelines: Directions for a New Agency. Washington, DC: National Academy Press, 1990.

110. Fieve R, Kumbaraci T, Dunner D. Lithium prophylaxis of unipolar and bipolar II illness. American Journal of Psychiatry 1976;133:925-8.

111. Fieve RF et al. Lithium carbonate in affective disorders IV, A double-blind study of prophylaxis in unipolar recurrent depression. Archives of General Psychiatry 1975;32:1541-4.

112. Fieve RR, Platman SR, Plutchik RR. The use of lithium in affective disorders: I. Acute endogenous depression. American Journal of Psychiatry 1968;125:79-83.

113. Findling RL et al. Double-blind 18-month trial of lithium versus divalproex maintenance treatment in pediatric bipolar disorder. Journal of the American Academy of Child & Adolescent Psychiatry.44(5):409-17, 2005.

114. Finely PR et al. Impact of a collaborative pharmacy practice model on the treatment of depression in primary care. American Journal of Health-System Pharmacy 2002;59:1518-26.

115. Fletcher AE, Franks PJ, Bulpitt CJ. The effect of withdrawing antihypertensive therapy: a review. J Hypertension 1988;431-6.

116. Frank E et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. Archives of General Psychiatry 1991;48:851-5.

117. Frank E. et al. Three-year outcomes for maintenance therapies in recurrent depression. Archives of General Psychiatry 1990;47:1093-9.

118. Frankenburg F et al. Long-term response to carbamazepine: a retrospective study. Journal of Clinical Psychopharmacology 1988;8:130-2.

119. Frazier LM et al. Need for linsulin therapy in type II diabetes mellitus. A randomized trial. Arch Intern Med 1987;147:1085-9.

120. Freis ED et al. Effects of reduction in drugs or dosage after long-term control of systemic hypertension. Am J Cardiol 1989;702-8.

121. Freudenstein U et al. Treatments for late life depression in primary care -- a systematic review. Family Practice 2001;18:321-7.

122. Friedman LM, Furberg CD, DeMets DL. Fundamentals of Clinical Trials. Littleton, MA: PSG Publishing Company, 1985.

123. Furukawa T, McGuire H, Barbui C. Low dosage tricyclic antidepressants for depression. Cochrane Database of Systematic Reviews 2003.

124. Furukawa TA, McGuire H, Barbui C. Meta-analysis of effects and side effects of low dosage tricyclic antidepressants in depression: systematic review. BMJ 2002;325.

125. Furukawa TA, Streiner DL, Young LT. Antidepressant and benzodiazepine for major depression. Cochrane Database of Systematic Reviews 2002.

126. Gajwani P et al. Antiepileptic drugs in mood disordered patients. Epilepsia 2005;46:38-44.

127. Gaul A. et al. Sertraline for the prevention of relapse in detoxicated alcohol dependent patients with a comorbid depressive disorder: A randomized controlled trial. Alcohol and Alcoholism 2003;38:619-25.

128. Geddes JR et al. Long-term Lithium Therapy for Bipolar Disorder: Systematic Review and Meta-Analysis of Randomized Controlled Trials. American Journal of Psychiatry 2004;161:217-22.

129. Geddes JR et al. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. Lancet 2003;361:653-61.

130. Gelenberg A et al. Randomized, placebo-controlled trial of nefazodone maintenance treatment in preventing recurrence in chronic depression. Biological Psychiatry 2003;54:806-17.

131. Georgotas A., McCue R.E., Cooper T.B. A Placebo-Controlled Comparison of Nortriptyline and Phenelzine in Maintenance Therapy of Elderly Depressed Patients. Archives of General Psychiatry 1989;46:783-6.

132. Georgotas A, Gershon S. Historical Perspectives and Current Highlights on Lithium Treatment in Manic-Depressive Illness. Journal of Clinical Psychopharmacology 1981;1:27-31.

133. Ghaemi SN, Lenox MS, Baldessarini RJ. Effectiveness and safety of long-term antidepressant treatment in bipolar disorder. Journal of Clinical Psychiatry 2001;62:565-9.

134. Gilaberte I. et al. Fluoxetine in the Prevention of Depressive Recurrences: A Double-Blind Study. Journal of Clinical Psychiatry 2001;21:417-24.

135. Giles TD et al. Remission of mild to mederate hyperension after treatment with carteolol, a beta-adrenoceptor blocker with intrinsic sympathomimetic activity. Arch Intern Med 1988;148:1728.

136. Gisslinger H et al. Long-term interferon therapy for thrombocytosis in myeloproliferative diseases. Lancet 1989;634-7.

137. Gitlin MJ et al. Relapse and impairment in bipolar disoder. American Journal of Psychiatry 1995;152:1635-40.

138. Glen A., Johnson A., Shepherd M. Continuation therpay with lithium and amitriptyline in unipolar depressive illness: a randomized, double-blind, controlled trial. Psychological Medicine 1984;14:37-50.

139. Goldstein TR et al. Antidepressant discontinuation-related mania: critical prospective observation and theoretical implications in bipolar disorder. Journal of Clinical Psychiatry.60(8):563-7; quiz 568-9, 1999.

140. Goodwin FK, Murphy DL, Bunney WE. Lithium carbonate treatment in depression and mania. Archives of General Psychiatry 1969;21:486-96.

141. Gray D. et al. Fluoxetine treatment of the obese diabetic. International Journal of Obesity 1992;16:193-8.

142. Greden JF. Antidepressant maintenance medications: when to discontinue and how to stop. Journal of Clinical Psychiatry 1993;54:39-47.

143. Greenhouse J.B. et al. Methodologic Issues in Maintenance Therapy Clinical Trials. Archives of General Psychiatry 1991;48:313-8.

144. Greil W et al. Lithium versus carbamazepine in the maintenance treatment of bipolar disorders - a randomized controlled study. Journal of Affective Disorders 1997;43:151-61.

145. Greist J. et al. A 1 year double-blind placeb-controlled fixed dose study of sertraline in the treatment of obsessive-compulsive disorder. International Clinical Psychopharmacology 1995;10:57-65.

146. Grof P. Has the effectiveness of lithium changed? Impact of the variety of lithium's effects. Neuropsychopharmacology 1998;19:183-8.

147. Grof P. Lithium update: selected issues. Ayd F, Taylor JT, and Taylor BT. 1983. Baltimore, Ayd Medical Communications. Affective Disorders Reassessed. Ref Type: Serial (Book,Monograph)

148. Grof P et al. The challenges of predicting response to stabilizing lithium treatment. The importance of patient selection. British Journal of Psychiatry 1993;Supp:16-9.

149. Grof P et al. The Challenge of Predicting Response to Stabilising Lithium Treatment. British Journal of Psychiatry 1993;163:16-9.

150. Guaiana G, Barbui C, Hotopf M. Amitriptyline versus other types of pharmacotherapy for depression. Cochrane Database of Systematic Reviews 2003.

151. Gyulai L et al. Maintenance efficacy of divalproex in the prevention of bipolar depression. Neuropsychopharmacology.28(7):1374-82, 2003.

152. Hadjipaviou G, Mok H, Yatham LN. Pharmacotherapy of bipolar II disorder: a critical review of current evidence. Bipolar Disorders 2004;6:14-25.

153. Hahn CG et al. The current understanding of lamotrigine as a mood stabilizer. [Review] [119 refs]. Journal of Clinical Psychiatry.65(6):791-804, 2004.

154. Harris M et al. Mood-stabilizers: the archeology of the concept. Bipolar Disorders 2003;5:446-52.

155. Harrison W. et al. Phenelzine for chronic depression: a study of continuation treatment. Journal of Clinical Psychiatry 1986;47:346-9.

156. Harrow M et al. Outcome in manic disorders: a naturalistic follow-up study. Arch Gen Psychiatry 1990;47:665-71.

157. Hartong E et al. Prophylactic efficacy of lithium versus carbamazepine in treatment-naive bipolar patients. Journal of Clinical Psychiatry 2003;64:144-51.

158. Haug T.T. et al. Exposure therapy and sertraline in social phobia: 1-year follow-up of a randomised controlled trial. British Journal of Psychiatry 2003;182:312-8.

159. Hayes S.G. Long-term use of valproate in primary psychiatric disorders. Journal of Clinical Psychiatry 1989;5:35-9.

160. Hazell P et al. Tricyclic drugs for depression in children and adolescents. Cochrane Database of Systematic Reviews 2002.

161. Healey D. The Antidepressant Era. Cambridge: Harvard University Press, 1997.

162. Healy D. Lithium. The Psychopharmacologists II 1994;259-84.

163. Hirschfeld RMA. American helath care systems and depression: The past, present, and the future. Journal of Clinical Psychiatry 1998;59:5-10.

164. Hirschfeld RM, Kasper S. A review of the evidence for carbamazepine and oxcarbazepine in the treatment of bipolar disorder. [Review] [100 refs]. International Journal of Neuropsychopharmacology.7(4):507-22, 2004.

165. Hochstrasser B. et al. Prophylactic effect of citalopram in unipolar, recurrent depression. Placebo-controlled study of maintenance therapy. British Journal of Psychiatry 2001;178:304-10.

166. Hollon S et al. Presenting characteristics of depressed outpatients as a function of recurrence: Preliminary findings from the STAR*D clinical trial. Journal of Psychiatric Research 2006;40:59-69.

167. Horan RF, Sheffer AL, Austen KF. Cromolyn sodium in the management of systemic mastocytosis. J Allergy Clin Immunol 1990;852-5.

168. Howard LM et al. Antidepressant prevention of postnatal depression. Cochrane Database of Systematic Reviews 2005.

169. Hypericum Depression Trial Study Group. Effect of Hypericum perforatum (St.John's Wort) in Major Depressive Disorder; A Randomized Controlled Trial. JAMA 2002;287:1807-14.

170. Institute for Clinical Systems Impovement. Major depression in adults in primary care. Institute for Clinical Systems Improvement 2006.

171. Institute for Clinical Systems Improvement. Major depression in adults for mental health care. Institute for Clinical Systems Improvement 2004.

172. Jacobsen F.M. Low-dose valproate: a new treatment for cyclothymia, mild rapid cyclinc disorders, and premenstrual syndrome. Journal of Clinical Psychiatry 1993;54:229-34.

173. Jadad AR et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Controlled Clinical Trials 1996;17:1-12.

174. Jane-Llopis E et al. Predictors of efficacy in depression prevention programmes. British Journal of Psychiatry 2003;183:384-97.

175. Jefferson JW. Lamotrigine in psychiatry. CNS Spectrums 2005;10:224-32.

176. Joffe RT. Discontinuing treatment for psychiatric disorders. Journal of Psychiatry & Neuroscience 2006;31:11.

177. Johnson FN, Cochrane R. Drinking water lithium and sites of mental hospitals. Lancet 1973;976.

178. Kane J. et al. Lithium Carbonate and Imipramine in the Prphylaxis of Unipolar and Bipolar II illness. Archives of General Psychiatry 1982;39:1065-9.

179. Kane JM et al. Lithium carbonate and imipramine in the prophylaxis of unipolar and bipolar II illness. Archives of General Psychiatry 1982;39:1065-9.

180. Katon W et al. A multifaceted intervention to improve treatment of depression in primary care. Archives of General Psychiatry 1996.

181. Katon W, Von Korff M, Lin E. Stepped collaborative care for primary care patients with persistent symptoms of depression: a randomized trial. Archives of General Psychiatry 1999;56:1109-15.

182. Katon W et al. Adequacy and duration of antidepressant treatment in primary care. Medical Care 1992;30:67-76.

183. Katon W et al. Collaborative management to achieve depression treatment guidelines. Journal of Clinical Psychiatry 1997;58:20-3.

184. Kay E., Fahy T., Garside R. A seven-month double blind trial of amitriptyline and diazepam in ECT-treated depressed patients. British Journal of Psychiatry 1970;117:667-71.

185. Kaye W. et al. Double-blind placebo-controlled administration of fluoxetine in restricting and restricting-purging-type anorexia nervosa. Biological Psychiatry 2001;49:644-52.

186. Keck PE et al. Placebo effect in randomized controlled maintenance studies of patients with bipolar disorder. Biological Psychiatry 2000;47:756-61.

187. Keck PE Jr, McElroy SL. Divalproex in the treatment of bipolar disorder. Psychopharmacology Bulletin 2003;37:67-73.

188. Keck PE, Jr., McElroy SL. Carbamazepine and valproate in the maintenance treatment of bipolar disorder. [Review] [30 refs]. Journal of Clinical Psychiatry.63 Suppl 10:13-7, 2002.

189. Keck PE, Jr., McElroy SL, Nemeroff CB. Anticonvulsants in the treatment of bipolar disorder. [Review] [78 refs]. Journal of Neuropsychiatry & Clinical Neurosciences.4(4):395-405, 1992.

190. Keller M.B. et al. Maintenance Phase Efficacy of Sertraline for Chronic Depression: A Randomized Controlled Trial. JAMA 1998;280:1665-72.

191. Kessler RC et al. National Comorbidity Survey replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003;289:3095-105.

192. Kishimoto A. et al. Prophylactic effect of mainserin on recurrent depression. Acta Psychiatrica Scandinavica 1994;89:46-51.

193. Kleindienst N, Greil W. Differential efficacy of lithium and carbamazepine in the prophyalxis of bipolar disorder: Results of the MAP study. Neuropsychobiology 2000;42:S2-S10.

194. Klerman G. et al. Treatment of depression by drugs and psychotherapy. American Journal of Psychiatry 1974;131:186-91.

195. Klysner R. et al. Efficacy of citalopram in the prevention of recurrent depression in elderly patients: placebo-controlled study of maintenance therapy. British Journal of Psychiatry 2002;181:29-35.

196. Kocsis J. et al. Maintenance therapy for chronic depression. Archives of General Psychiatry 1996;53:769-74.

197. Kopec JA, Abrahamowicz M, Esdaile JM. Randomized Discontinuation Trials: Utility and Efficiency. Journal of Clinical Epidemiology 1993;46:959-71.

198. Kopec JA, Abrahamowicz M, Esdaile JM. Randomized Discontinuation Trials: Utility and Efficiency. Journal of Clinical Epidemiology 1993;46:959-71.

199. Koran L. et al. Efficacy of sertraline in the long-term treatment of obsessive-compulsive disorder. American Journal of Psychiatry 2002;159:88-95.

200. Kupfer D.J. et al. Five-year outcome for maintenance therapies in recurrent depression. Archives of General Psychiatry 1992;49:769-73.

201. Kupfer DJ et al. Demographic and clinical characteristics of individuals in a bipolar disoder case registry. Journal of Clinical Psychiatry 2002;63:120-5.

202. Kurlowicz LH. Depression in older adults. In: Mezey M et al., eds. Geriatric nusing protocols for best practice. 2003:185-206.

203. Lambert P.A. Acute and prophylactic therapies of patients with affective disorders using vapromide (dipropylacetamide). Amsterdam: 1984:33-44.

204. Landabaso M. et al. A randomized trial of adding fluoxetine to a naltrexone treatment programme for heroin addicts. Addiction 1998;93:739-44.

205. Langford HG et al. Dietary therapy slows the return of hypertension after stopping prolonged medicatoin. JAMA 1985;657-64.

206. Lee S et al. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation. Depression and Anxiety 2004;19:10-9.

207. Lennard-Jones JE et al. Prednisone as maintenance treatment for ulcerative colitis in remission. Lancet 1965;Jan 23:188-9.

208. Leon AC. Measuring onset of action in clinical trials: An overview of definitions and methdology. Journal of Clinical Psychiatry 2001;62:12-6.

209. Lepine J.P. et al. A Randomized, Placebo-Controlled Trial of Sertraline for Prophylactic Treatment of Highly Recurrent Major Depressive Disorder. American Journal of Psychiatry 2004;161:836-42.

210. Levitt C et al. Systematic review of the literature on postpartum care: methodology and literature search results. Birth 2004;31:196-202.

211. Licht RW et al. A lithium clinic for bipolar patients: 2-year outcome of the first 148 patients. Acta Psychiatrica Scandinavica 2001;104:387-90.

212. Lieberman J et al. Comparing the effects of antidepressants: Consensus guidelines for evaluating quantitative reviews of antidepressant efficacy. Neuropsychopharmacology (Nature) 2005;30:445-60.

213. Lieberman DZ, Goodwin FK. Separate and concomitant use of lamotrigine, lithium, and divalproex in bipolar disorders. [Review] [50 refs]. Current Psychiatry Reports.6(6):459-65, 2004.

214. Liebowitz MR et al. Pharmocotherapy of social phobia: an interim report of a placbo-controlled comparison of phenelzine and atenolol. J Clin Psychiatry 1988;49:252-7.

215. Lovett L., Watkins S.E., Shaw D.M. The use of alternative drug therapy in nine patients with recurrent affective disorder resistant to conventional prophyalxis. Biological Psychiatry 1986;21:1344-7.

216. Lumley J, Austin MP, Mitchell C. Intervening to reduce depression after birth: a systematic review of the randomized trials. International journal of technology assessment in health care 2004;20:128-44.

217. Lusznat RM, Murphy DP, Nunn CMH. Carbamazepine vs lithium in the treatment and prophylaxis of mania. British Journal of Psychiatry 1988;153:198-204.

218. MacGillivary S et al. Efficacy and tolerability of selective serotonin reuptake inhibitors compared with tricyclic antidepressants in depression treated in primary care: systematic review and meta-analysis. BMJ 2003;326 (7397):1014.

219. MacLehose RR et al. Hybrid study designs and RCT variants. Heatlh Technology Assessment 2000;4:57-80.

220. Macritchie KA et al. Lamotrigine in the maintenance treatment of bipolar disorder. Cochrane Database of Systematic Reviews.(1):CD004052, 2001;1.

221. Macritchie KA et al. Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder. [Review] [63 refs]. Cochrane Database of Systematic Reviews.(3):CD003196, 2001.

222. Maggs R. Treatment of manic illness with lithium carbonate. British Journal of Psychiatry 1963;109:56-65.

223. Marangell LB et al. Lamotrigine treatment of bipolar disorder: data from the first 500 patients in STEP-BD. Bipolar Disorders.6(2):139-43, 2004.

224. Marcus M. et al. A double-blind, placebo-controlled trial of fluoxetine plus behavior modification in the treatment of obese binge-eaters and non-binge-eaters. American Journal of Psychiatry 1990;147:876-81.

225. Martenyi F. et al. Fluoxetine v. placebo in prevention of relapse in post-traumatic stress disorder. British Journal of Psychiatry 2002;181:315-20.

226. Mavissakalian M, Perel JM. Clinical experiments in maintenance and discontinuation of imipramine therapy in panic disorder with agoraphobia. Arch Gen Psychiatry 1992;318-23.

227. McCormack PL, Wiseman LR. Olanzapine: a review of its use in the management of bipolar I disorder. Drugs 2004;64:2709-26.

228. McCreadie RG et al. The Scottish first episode schizophrenia study. VII. Two-year follow-up. Acta Psychiatr Scand 1989;80:597-602.

229. McElroy S.L., Keck P.E.Jr., Harrison G.P.Jr. Sodium Valproate: Its use in primary psychiatric disorders. Journal of Clinical Psychopharmacology 1987;7:16-24.

230. McElroy SL et al. Valproate in the treatment of rapid-cycling bipolar disorder.[see comment]. Journal of Clinical Psychopharmacology.8(4):275-9, 1988.

231. Melia PI. Prophylactic lithium: A double-blind trial in recurrent affective disorders. British Journal of Psychiatry 1970;116:621-4.

232. Mendels J, Secunda SK, Dyson WL. A controlled study of the antidepressant effects of lithium carbonate. Archives of General Psychiatry 1972;154-7.

233. Michael KD, Crowley SL. How effective are treatments for child and adolescent depression? A meta-analytic review. Clinical Psychology Review 2002;22:247-69.

234. Michelson D. et al. Outcome assessment and clinical improvement in panic disorder: Evidence from a randomized controlled trial of fluoxetine and placebo. American Journal of Psychiatry 1998;155:1570-7.

235. Michigan Quality Improvement Consortium. Management of adults with major depression. Michigan Quality Improvement Consortium 2004.

236. Mindahm R., Howland C., Shepherd M. An evaluation of continuation therapy with tricyclic antidepressants in depressive illness. Psychological Medicine 1973;3:5-17.

237. Misra SP et al. Long-term treatment of irritable bowel syndrome: results of a randomized controlled trial. Q J Med 1989;73:931-9.

238. Moher D et al. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001;357:1191-4.

239. Moncrieff J. Lithium: evidence reconsidered. British Journal of Psychiatry 1997;171:113-9.

240. Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. Cochrane Database of Systematic Reviews 2004.

241. Montgomery S.A. et al. The Prophylactic Efficacy of Fluoxetine in Unipolar Depression. British Journal of Psychiatry 1988;153:69-76.

242. Montgomery S.A., Dunbar G. Paroxetine is better than placebo in relapse prevention and the prphylaxis of recurrent depression. International Clinical Psychopharmacology 1993;8:189-95.

243. Montgomery S.A. et al. Venlafaxine versus placebo in the preventive treatment of recurrent major depression. Journal of Clinical Psychiatry 2004;65:328-36.

244. Montgomery S.A., Rasmussen J.G., Tanghoj P. A 24-week study of 20mg citalopram, 40mg citalopram, and placebo in the prevention of relapse of major depression. International Clinical Psychopharmacology 1993;8:181-8.

245. Montgomery S.A. et al. Dose Response Relationship of Citalopram 20mg, Citalopram 40 mg and Placebo in the Treatment of Moderate and Severe Depression. International Clinical Psychopharmacology 1992;6:65-70.

246. Montgomery SA, Rasmussen JGC. Citalopram 20 mg, Citalopram 40 mg and Placebo in the Prevention of Relpase of Major Depression. International Clinical Psychopharmacology 1992;6:71-3.

247. Montori VM et al. Optimal search strategies for retrieving systematic reviews for Medline: analytical survey. BMJ 2005;330.

248. Muzina DJ et al. Lamotrigine and antiepileptic drugs as mood stabilizers in bipolar disorder. Acta Psychiatrica Scandinavica, Supplementum.(426):21-8, 2005.

249. National Collaborating Center for Mental Health. Depression: Management of depression in primary and secondary care. National Institute for Clinical Excellence 2004.

250. National Collaborating Centre for Mental Health Royal College of Psychiatrists' Research and Training Unit. Depression in children and young people. Identification and management in primary, community and secondary care. The British Psychological Society 2005.

251. Neumeyer-Gromen A et al. Disease management programs for depression: a systematic review and meta-analysis of randomized controlled trials. Medical Care 2004;42:1211-21.

252. O'Kane M., Wiles P., Wales J. Fluoxetine in the treatment of obese type 2 diabetic patients. Diabetic Medicine 1994;11:105-10.

253. Ogrodniczuk JS, Piper WE. Preventing postnatal depression. Harvard Review of Psychiatry 2003;11:291-307.

254. Okuma R et al. A preliminary double-blind study on the efficacy of carbamazepine in prophylaxis of manic-depressive illness. Psychopharmacology 1981;73:95-6.

255. Old Age Depression Interest Group. How long should the elderly take antidepressants? A double-blind placebo-controlled study of continuation/prophylaxis therapy with Dothiepin. British Journal of Psychiatry 1993;162:175-82.

256. Oliver SL, Comstock GW, Helsing KJ. Mood and lithium in drinking water. Archives of Environmental Health 1976;31:92-5.

257. Oxman TE et al. A three component model fro reengineering systems for the treatment of depression in primary care. Psychosomatics 2002;43:441-50.

258. Pablos-Mendez A, Barr R G, Shea S. Run-in periods in randomized trials. Implications for the application of results in clinical practice. JAMA 1998;279:222-5.

259. Paykel ES et al. Residual symptoms after partial remission: an important outcome in depression. Psychological Medicine 1995;25:1171-80.

260. Perahia DG et al. Symptoms following abrupt discontinuation of duloxetine treatment in patients with major depressive disorder. Journal of Affective Disorders 2005;89:207-12.

261. Placidi GF et al. The comparative efficacy and safety of carbamazepine versus lithium: a randomized double-blind 3-year trial in 83 patients. Journal of Clinical Psychiatry 1986;47:490-4.

262. Pokorny AD, Sheehan D, Atkinson J. Drinking water, lithium and mental hospital admissions. Diseases of the nervous system 1972;33:649-52.

263. Ponten J et al. Beta-receptor blocker withdrawal. A preoperative problem in genral surgery? Acta Anaesth Scand 1982;Suppl. 76:32-7.

264. Ponten J et al. Beta-receptor blockade and spinal anaesthesia. Acta Anaesth Scand 1982;Suppl. 76:62-9.

265. Post RM et al. Carbamazepine prophylaxis in refractory affective disorders: a focus on long-term follow up. Journal of Clinical Psychopharmacology 1990;10:318-27.

266. Post RM et al. Prophylactic Efficacy of Carbamazepine in manic-depressive illness. American Journal of Psychiatry 1983;140:1602-4.

267. Prien R. et al. Drug therapy in the prevention of recurrences in unipolar and bipolar affective disorders. Archives of General Psychiatry 1984;41:1096-104.

268. Prien RF, Klett CJ, Caffey EM. Lithium carbonate and imipramine in prevention of affective episodes. Archives of General Psychiatry 1973;29:420-5.

269. Puzynski S, Klosiewcz L. Valproic acid amide in the treatment of affective and schizoaffective disorders. Journal of Affective Disorders. 1984;6:115-21.

270. Quitkin FM, Rabkin JG. Methodological problems in studies of depressive disorder: utility of the discontinuation design. Journal of Clinical Pharmacology 1981;1:283-8.

271. Rabkin JG et al. Effects of pill giving on maintenance of placebo response in patients with chronic mild depression. Am J Psychiatry 1990;1622-6.

272. Randlov C et al. The efficacy of St. John's Wort in patients with minor depressive symptoms or dysthymia -- a double-blind placebo-controlled study. Phytomedicine 2006;13:215-21.

273. Rapaport M., Endicott J., Clary C. Posttraumatic stress disorder and quality of life: Results across 64 weeks of sertraline treatment. Journal of Clinical Psychiatry 2002;63:59-65.

274. Rapaport M. et al. Sertraline treatment of panic disorder: results of a long-term study. Acta Psychiatrica Scandinavica 2001;104:289-98.

275. Rapaport M. et al. Sertraline treatment of panic disorder: results of a long-term study. Acta Psychiatrica Scandinavica 2001;104:289-98.

276. Rapaport M.H., Bose A., Zheng H. Escitalopram Continuation Treatment Prevents Relase of Depressive Episodes. Journal of Clinical Psychiatry 2004;65:44-9.

277. Reimherr F.W. et al. Optimal Length of Continuation Therapy in Depression: A Prospective Assessment During Long-Term Fluoxetine Treatment. American Journal of Psychiatry 1998;155:1247-53.

278. Rendell JM, Geddes JR. Risperidone in long-term treatment for bipolar disorder. Cochrane Database of Systematic Reviews.(1):CD004052, 2004;1.

279. Reynolds C. et al. Nortriptylikne and interpersonal psychotherapy as maintenance therapies for recurrent major depression. A randomizaed controlled trial in patients older than 59 years. JAMA 1999;281:39-45.

280. Ricca V. et al. Fluoxetine and Fluvoxamine Combined with individual cognitive-behaviour therapy in binge eating disorder: A one-year follow-up study. Psychotherapy and psychosomatics 2001;70:298-306.

281. Rickels K et al. Long-term diazepam therapy and clinical outcome. JAMA 1983;250:767-71.

282. Robert Ph., Montgomery S.A. Citalopram in doses of 20-60mg is effective in depression relapse prevention: a placebo-controlled 6 month study. International Clinical Psychopharmacology 1995;10.

283. Robinson D. et al. Maintenance therappies in trecurrent depression: New findings. Psychopharmacology Bulletin 1991;27:31-9.

284. Robinson DS et al. Continuation and maintenance treatment of major depression with the monoamine oxidase inhibitor phenelzine: a double blind placebo-controlled discontinuation study. Psychopharmacol Bull 1991;31-9.

285. Romano S. et al. Long-term treatment of obsessive-compulsive disorder after an acute response: a comparison of fluoxetine versus placebo. Journal of Clinical Psychopharmacology 2001;21:46-52.

286. Romano S. et al. A placebo controlled study of fluoxetine in continued treatment of bulimia nervosa after successful acute fluoxetine treatment. American Journal of Psychiatry 2002;159:96-102.

287. Rosenbaum J.F. et al. Selective Seotonin Reuptake Inhibitor Discontinuation Syndrome: A Randomized Clinical Trial. Biological Psychiatry 1998;44:77-87.

288. Rost K et al. Managing depression as a chronic disease: a randomised trial of ongoing treatment in primary care. BMJ 2002;325:934-7.

289. Rouillon F. et al. Milnacipran efficacy in the prevention of recurrent depression: a 12-month placebo-controlled study. International Clinical Psychopharmacology 2000;15:133-40.

290. Ruggiero G. et al. Nutritional management of anorexic patients with and without fluoxetine: 1-year follow-up. Progress in Neuro-Psychopharmacology & Biological Psychiatry 2003;27:425-30.

291. Ruoff G. Effect of withdrawal of terazosin therapy in patients with hypertension. Am J Med 1986;80:35-41.

292. Ruskin PE, Nyman G. Disontinuation of neuroleptic medication in older, outpatient schizophrenics. J Nerv Ment Dis 1991;212-4.

293. Saiz-Ruiz J. et al. Sertraline treatment of pathological gambling: A pilot study. Journal of Clinical Psychiatry 2005;66:28-33.

294. Schaff MR, Fawcett J, Zajecka JM. Divalproex sodium in the treatment of refractory affective disorders. Journal of Clinical Psychiatry.54(10):380-4, 1993.

295. Schmidt M.E. et al. The Efficacy and Safety of a New Enteric-Coated Formulation of Fluoxetine Given Once Wekly During the Continuation Treatment of Major Depressive Disorder. Journal of Clinical Psychiatry 2000;61:851-7.

296. Schneider A.L., Wilcox C.S. Divalproate augmentation in lithium-resistant rapid cycling mania in four geriatric patients. Journal of Affective Disorders 1998;47:201-5.

297. Schoenbaum M et al. The effects of primary care depression treatment on patient's clinical status and employment. Health Serv Res 2002;37:1145-58.

298. Schou M et al. The treatment of manic psychoses by the administration of lithium salts. Journal of Neurology Neurosurgery and Psychiatry 1954;16:250-60.

299. Schou M, Shaw DM. Lithium in Recurrent Manic-Depressive Disorder. The Practitioner 1972;105-11.

300. Schrauzer GN, Shreshtha KP. Lithium in drinking water and the incidences of crimes, suicides, and arrests related to drug addictions. Biological Trace Element Research 1990;25:105-13.

301. Schulberg HC et al. The 'usual care' of major depression in primary care practice. Archives of Family Medicine 1997;6:334-9.

302. Schulz KF et al. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408-12.

303. Scott J, Gutierrez MJ. The current status of psychological treatments in bipolar disorders: a systematic review of relapse prevention. Bipolar Disorders 2004;6:498-503.

304. Scottish intercollegiate guidelines network. Bipolar affective disorder. A national clinical guideline. NHS Quality Improvement Scotland 2005.

305. Sencan S. et al. A study to compare the therapeutic efficacy of aerobic excercise and paroxetine in fibromyalgia syndrome. Journal of Back and Musculoskeletal Rehabilitation 2004;17:57-61.

306. Senn S. Individual response to treatment: is it a valid assumption? BMJ 2004;329:966-8.

307. Serrano-Blanco A, Gabarron E, Garcia-Bayo I et al. Effectiveness and cost-effectiveness of antidepressant treatment in primary health care: A six-month randomised study comparing fluoxetine to imipramine. Journal of Affective Disorders 2006;91:153-63.

308. Shadish W., Cook T., Campbell D. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin Company, 2002.

309. Shattell M, Keltner NL. The case for atypical antipsychotics in bipolar disorder. Perspectives in Psychiatric Care 2006;40:34-8.

310. Shepherd M. The use and abuses of drugs in psychiatry. Lancet 1970;January 3:31-3.

311. Simhandl C, Denk E, Thau K. The comparative efficacy of carbamazepine low and high serum level and lithium carbonate in the prophylaxis of affective disorders. Journal of Affective Disorders 1993;28:221-31.

312. Simon GE. Long-term prognosis of depression in primary care. Bulletin of the World Health Organization 2000;78:439-45.

313. Simon GE et al. Cost-effectiveness of a program to prevent depression relapse in primary care. Medical Care 2002;40:941-50.

314. Simon GE et al. Randomised trial of monitoring, feedback and managment of care by telephone to improve treatment of depression in primary care. BMJ 2000;320:550-4.

315. Simon J. et al. Extended-release venlafaxine in relapse prevention for patients with major depressive disorder. Journal of Psychiatric Research 2004;38:249-57.

316. Singapore ministry of health. Depression. Singapor Ministry of Health 2004.

317. Small J et al. Carbamazepine compared with lithium in the treatment of mania. Archives of General Psychiatry 1991;48:915-21.

318. Smith D et al. Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: a meta-analysis. British Journal of Psychiatry 2002;180:396-404.

319. Snider DE et al. Six-month isoniazid-rifampin therapy for pulmonary tuberculosis. Am Rev Respir Dis 1984;573-9.

320. Sobel JD. Recurrent vulvovaginal candidiasis. A prospective study of the efficacy of maintenance ketoconazole therapy. N Engl J Med 1986;1455-8.

321. Solomon D.A. et al. Multiple Recurrences of Major Depressive Disorder. American Journal of Psychiatry 2000;157:229-33.

322. Solomon DA, Bauer MS. Continuation and maintenance pharmacotherapy for unipolar and bipolar mood disorders. Psychiatric Clinics of North America 1993;16:515-40.

323. Stahl S. Placebo-controlled comparison of the selective serotonin reuptake inhibitors citalopram and sertraline. Biological Psychiatry 2000;48:894-901.

324. Stallone F et al. The use of lithium in affective disorders, III: A double-blind study of prphylaxis in bipolar illness. American Journal of Psychiatry 1973;130:1006-10.

325. Stein D. et al. Efficacy of paroxetine for relapse prevention in social anxiety disorder. Archives of General Psychiatry 2002;59:1111-8.

326. Stein M., Rickels K., Weise C.C. Maintenance therapy with amitriptyline: A controlled trial. American Journal of Psychiatry 1980;173:370-1.

327. Stellon AJ et al. Maintenance of remission in autoimmune chronic active hepatitis with azathioprine after corticosteroid withdrawal. Hepatology 1988;781-4.

328. Stellon AJ et al. Randomized controlled trial of azathioprine withdrawal in autoimmune chronic active hepatitis. Lancet 1985;668-70.

329. Stimpson N, Agrawal N, Lewis G. Randomised controlled trials investigating pharmacological and psychological interventions for treatment-refractory depression. British Journal of Psychaitry 2002;181:294.

330. Stocchi F. et al. Efficacy and tolerability of paroxetine for the long-term treatment of generalized anxiety disorder. Journal of Clinical Psychiatry 2003;64:250-8.

331. Stoll KD et al. Carbamazepine versus haloperidol in manic syndromes: first report of a multicenter study. In: Shagass C et al., eds. New York: Elsevier, 1985:332-4.

332. Storosum JG et al. Relapse and recurrence prevention in major depression: a critical review of placebo-controlled efficacy studies with special emphasis on methodological issues. European Psychiatry 2001;16:327-35.

333. Stromgren LS. The combination of lithium and carbamazepine in treatment and prevention of manic-depressive disorder: a review and a case report. Comprehensive Psychiatry 1990;31:261-5.

334. Taylor WD, Doraiswamy PM. A systematic review of antidepressant placebo-controlled trials for geriatric depression: limitations of current data and directions for the future. Neuropsychopharmacology 2004;29:2285-99.

335. Terra J.L., Montgomery S.A. Fluvoxamine prevents recurrence of depression: results of a long-term, double-blind, placebo-controlled study. International Clinical Psychopharmacology 1998;13:55-62.

336. The Canadian Hydroxychloroquine Study Group. A randomized study of the effect of withdrawing hydroxychloroquine sulfate in systemic lupus erythematosus. New England Journal of Medicine 1991;324:150-4.

337. The Johns Hopkins University Evidence-based Practice Center. Post-Myocardial Infarction Depression. Agency for Healthcare Research and Quality 2005.

338. The sixty plus reinfarction study research group. A double blind trial to asess long-term oral anticoagulant therapy in elderly patients after myocardial infarction. Lancet 1980;990-4.

339. Tohen M et al. Relapse prevention in bipolar I disorder: 18-month comparison of olanzapine plus mood stabiliser v. mood stabiliser alone. British Journal of Psychiatry.184:337-45, 2004.

340. Tohen M et al. Efficacy of olanzapine in combination with valproate or lithium in the treatment of mania in patients partially nonresponsive to valproate or lithium monotherapy.[see comment]. Archives of General Psychiatry.59(1):62-9, 2002.

341. Tohen M et al. Olanzapine versus divalproex sodium for the treatment of acute mania and maintenance of remission: a 47-week study. American Journal of Psychiatry.160(7):1263-71, 2003.

342. Tollefson G. et al. Continuation treatment of OCD: double-blind and open-label experience with fluoxetine. Journal of Clinical Psychiatry 1994;55:69-76.

343. Tondo L, Hennen J, Baldessarini RJ. Rapid-cycling bipolar disorder: effects of long-term treatments. Acta Psychiatrica Scandinavica 2003;108:4-14.

344. University of Michigan Health System. Depression. University of Michigan Health System 2004.

345. University of Michigan Health System. Depression. University of Michigan Health System 2005.

346. Vencovsky E, Soucek K, Kabes J. Prophylactic effect of dipropylacetamide in patients with bipolar affective disorder. In: Emrich HM, Okuma T, Muller AA, eds. Anticonvulsants in affective disorders. Amsterdam: Elsevier Science Publishers B.V., 1984:66-7.

347. Verhaaeghe L. Treatment of angina pectoris with lidoflazine. Arzneim-Forsch 1969;1842-8.

348. Versiani M. et al. Reboxetine, a unique selective NRI, prevents relapse and recurrence in long-term treatment of major depressive disorder. Journal of Clinical Psychiatry 1999;60:400-6.

349. Viguera AC, Baldessarini RJ, Friedberg J. Discontinuing antidepressant treatment in major depression. Harvard Review of Psychiatry 1998;5:293-306.

350. Voors AW. Drinking-water lithium and mental hospital admission in North Carolina. North Carolina Medical Journal 1972;33:597-602.

351. Wadden T. et al. Sertraline and relpase prevention training following treament by very-low-calorie diet: A controlled clinical trial. Obesity Research 1995;3:549-57.

352. Walker J. et al. Prevention of relapse in generalized social phobia: results of a 24-week study in responders to 20 weeks of sertraline treatment. Journal of Clinical Psychopharmacology 2000;20:636-44.

353. Walsh T et al. Placebo response in studies of major depression. Variable, substantial and growing. JAMA 2002;287:1840-7.

354. Walsh BT et al. Long-term outcome of antidepressant treatment for bulimia nervosa. Am J Psychiatry 1991;1206-12.

355. Waraich P et al. Prevalence and incidence studies of mood disorders: a systematic review of the literature. Canadian Journal of Psychiatry 2004;49:124-38.

356. Watkins SE et al. The effect of carbamazepine and lithium on remission from affective illness. British Journal of Psychiatry 1987;150:180-2.

357. Wearden A. et al. Randomised, double-blind, placebo-controlled treatment trial of fluoxetine and graded exercise for chronic fatigue syndrome. British Journal of Psychiatry 1998;172:485-90.

358. Weihs KL et al. Continuation phase treatment with bupropion SR effectively decreases the risk for relapse of depression. Biological Psychiatry 2002;51:753-61.

359. Wells KB et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. JAMA 2000;283:212-20.

360. Wermuth L. et al. Depression in idiopathic Parkinson's disease treated with citalopram. A placebo controlled study. Nordic Journal of Psychiatry 1998;52:163-9.

361. Wharton RN, Fieve RR. The use of lithium in the affective psychoses. American Journal of Psychiatry 1966;123:706-12.

362. Whittington CJ et al. Selective serotonin reuptake inhibitors in childhood depression : systematic review of published and unpublished data. Lancet 2004;363:1341-5.

363. Wilson K.C.M. et al. Older community residents with depression: long-term treatment with sertraline. Randomised double-blind, placebo-controlled study. British Journal of Psychiatry 2003;182:492-7.

364. Yatham LN. Atypical antipsychotis for bipolar disorder. Psychiatric Clinics of North America 2005;28:325-47.

365. Yatham LN et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) guidelines for the management of patients with bipolar disorder: consensus and controversies. Bipolar Disorders 2005;7:5-69.

366. Zimmerman M, Mattia J., Posternak M. Are Subjects in Pharmacological Treatment Trials of Depression Representative of Patients in Routine Clinical Practice? American Journal of Psychiatry 2002;159:469-73.