POST-TREATMENT OUTCOMES FOR ADULTS TREATED FOR DEPRESSION

BY

JENNIFER VOLK

A Thesis in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Copyright © 2012 Jennifer Volk

Abstract

Depression has been cited as the most prevalent of the Axis I disorders affecting upward of 16% of American adults in their lifetimes (Kessler et al., 2005). The literature on effective treatments for depression is substantial, however the follow-up literature that speaks to what happens after treatment ends is much smaller. This thesis describes two studies. The first is an overview of reviews of post-treatment outcomes for adults treated for major depressive disorder (MDD). The second is a narrative systematic review of studies of long-term (at least 12 months) post-treatment outcomes after the completion of treatment for major depressive disorder (MDD). These studies synthesize the available evidence concerning post-treatment outcomes and discuss the limitations of these data. Relapse is a significant issue for many people who respond to treatment with upwards of 50% of people relapsing within a year of the end of treatment, depending on the type of treatment. Some treatments have significant relapse prevention effects, including continued medication treatment, acute and continuation phase CBT, and variations of CBT designed specifically for addressing residual symptoms after acute treatment or specifically aimed at relapse prevention. Given that the risk of relapse after treatment is significant, it should be discussed during acute treatment, as should approaches to reducing the risk of relapse. Recommendations for future research are discussed. Within the overview of reviews, there was considerable consistency across reviews which aided in the formulation of practical recommendations for clinicians and for patients. Examples include provision of education about the probability of relapse and planning for relapse prevention during acute phase treatment. Engaging in continuation and maintenance treatments that are aimed at reducing relapse, and

whenever possible, continuing treatment until patients are considered to be in recovery, and not just for a certain period of time, or until the point of remission of symptoms are also recommended to reduce rates of relapse.

Acknowledgements

The author gratefully acknowledges the dedication and support of Dr. John Walker, research supervisor, and Dr. Bruce Tefft, academic advisor. The thoughtful input of Dr. Steven Feldgaier, Dr. Rob Santos, Dr. Warrant Eaton, and Dr. David Clark, committee members, was much appreciated and contributed to the quality of this paper. This project was completed with support from a knowledge mobilization team grant awarded by the Canadian Institutes for Health Research and the Mental Health Commission of Canada to the Mobilizing Minds Research Group, of which Dr. John Walker is the principal investigator and the author is a research associate. The author wishes to acknowledge all members of the Mobilizing Minds research group for their tremendous work in the area of knowledge synthesis and translation. Thanks goes to James Ducas, Brooke Beatie, and Julie Erickson for their contribution as research assistants.

Dedication

This thesis is dedicated to those who suffer with depression, to the care providers who work to alleviate this suffering, and to researchers who continue to search for the answers. Keep up the work. Keep up the hope.

A personal dedication goes out to all those who have supported me through the writing of this paper, and through my graduate degrees. I am thankful for my parents; my father Karel, who taught me the value of always doing my best, and my mother Alice, who taught me to love learning for its own sake. I am thankful for the support of all of my family and friends who have been there through the years to cheer me on and celebrate the milestones. A special thanks to Shannon, my forever friend, who is truly always there. I am thankful for my fellow students, who have journeyed with me through this program. I am thankful for my employers and co-workers who always encouraged me in my studies. I am thankful for my yoga practice and my teachers there, especially Shauna, who have helped to keep me grounded and reminded me to savour even the briefest moments of stillness and the smallest of joys. Lastly, I am thankful for Brad who encouraged me to take on this project in the first place, who has been there through the ups and downs, and who always believed that I would succeed. Thank you everyone.

Table of Contents

Abstract ii

Acknowledgements ……iv

Dedication ……..……..……..v

Table of Contents vi

List of Tables …..……....xi

List of Figures ……..……..……...xii

Introduction.....1

Major Depressive Disorder ……..……..……...3

Why This Review? ……...……....9

Long-term Outcomes ……..……...……..… 11

Problems in Measurement 11

Gaps in Knowledge ……..……...18

Patient Preference and Need for Information ……..……20

Evidence-Based Treatments for Depression ……..……..…24

Pharmacological Treatments ……..…… 25

Limitations of Antidepressant Trials ……..…… 30

Psychological Interventions ……..…. 32

Cognitive Behavioural Therapy ……..…….... 33

Interpersonal Psychotherapy ……..……..…….. 37

Other Therapies ……..…… 38

Limitations of Psychological Intervention Trials ……...……... 39

Combined Treatment ……...…… 41

Comparative Studies ……. 44

Other Treatments ……..…….. 47

The Current Study ……..……..…… 48

Study 1: Overview of Reviews …….……..……52

Study 1 Method …….……..……..…….52

Criteria for Selecting Reviews for inclusion ……..……...…….…….52

Types of Reviews …….……...…….…….…..52

Types of Participants ……..…….52

Types of Interventions …….……..….52

Outcomes …….……...53

Search Methods for Identification of Reviews …….……..53

Data Collection and Analysis ……..……...…….54

Selection of Reviews ……..…….54

Data Extraction ……..……..…...54

Assessment of Methodological Quality of Included Reviews ……...…54

Data Synthesis ……...…….....55

Study 1 Results 55

Description of Included Reviews ……..…….…….57

Methodological Quality of Included Reviews ..……..66

Key Findings of the Reviews …….……71

Antidepressant Medications …….……...71

Cognitive Behavioural Therapy …….……73

Direct Comparison: Antidepressant Medication and CBT ……...…77

Behavioural Activation ……..…..82

Interpersonal Psychotherapy ……...…84

Psychodynamic Therapy ……..…84

Combined Psychotherapy and Medication Treatment …….……...86

Study 1 Discussion …….…..……...89

Summary of Main Results …….…….…….89

Quality of the Evidence …….……..……91

Potential Biases in the Overview Process …….……..92

Agreement and Disagreements with Other Studies or Reviews …….……93

Implications for Research …….……..94

Recommendations for Future Reviews …….……..94

Recommendations on Study Design …….……..94

Recommendations on Measurement …….……..96

Recommended Analyses …….……98

Implications for Practice ……..……...……99

Study 1 Conclusions …….…..……...100

Study 2: Narrative Systematic Review ……...…..…….101

Study 2 Method …….……...…..……101

Criteria for Selecting Studies for this Review …….…….……101

Study Design …….…...…….101

Types of Participants …….……...…….101

Types of Interventions …….…….……102

Outcomes …….…….……102

Search Methods for Identification of Studies ……102

Data Collection and Analysis ……103

Selection of Studies ……103

Data Extraction ……103

Assessment of Methodological Quality of Included Studies ……103

Data Synthesis ……104

Study 2 Results ……104

Description of Studies ……104

Risk of Bias/Quality of the Included Studies ……121

Clinical Outcomes at Follow-up ……132

Antidepressant Medications ……132

Direct Comparison: Antidepressant Medications and CBT ……133

CBT in Primary Care ……137

CBT for Residual Symptoms ……139

CBT for Relapse Prevention: Mindfulness Based Cognitive Therapy ……145

Behavioural Activation ……149

Interpersonal Psychotherapy ……150

Psychodynamic Therapy ……151

Study 2 Discussion ……153

Summary of the Main Results ……153

Attrition ……154

Future Studies ……155

Limitations ……156

Study 2 Conclusions ……157

General Discussion ……157

General Conclusions ……162

References ……163

Appendix A: Inclusion checklist ……195

List of Tables

Table 1. Key Characteristics and Results of Included Reviews 58

Table 2. Quality Ratings for Included Systematic Reviews …….……67

Table 3. Study Characteristics and Results ……...……106

Table 4. Quality Assessment of Included Studies …….122

List of Figures

Figure 1. Course of Depression Over Time …….……...…..…2

Figure 2. Flowchart of Eligibility Assessment and Inclusion 56

Figure 3. Progress Through the Stages of a Systematic Review for RCTs …..……105

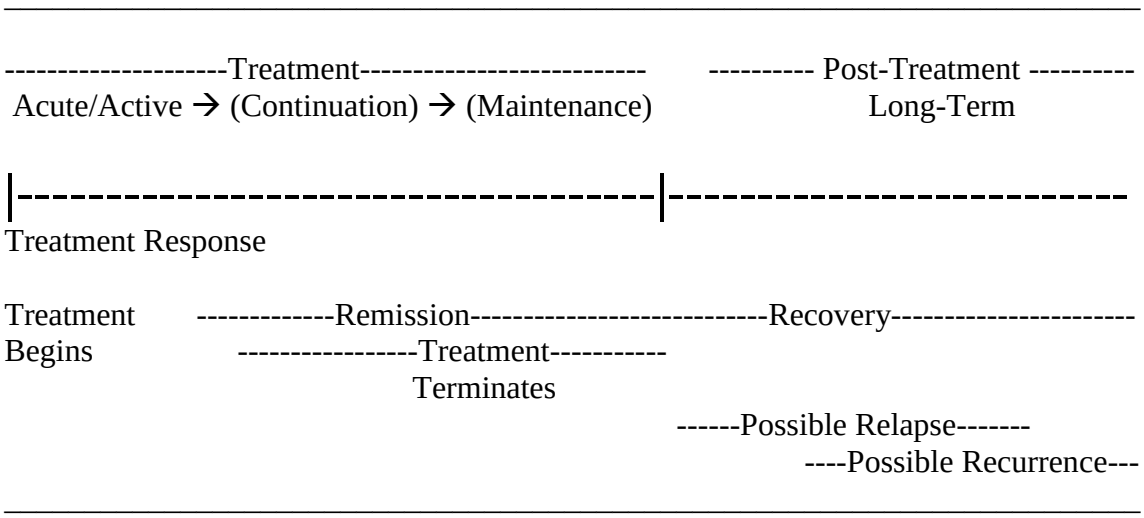
Post-Treatment Outcomes for Adults Treated for Depression

Depression is currently the most common Axis I disorder (Kessler, Berglund, Demler, Jin, Merikangas, & Walters, 2005) and, although there is a voluminous and growing evidence base on the short-term efficacy of psychological and pharmacological treatments of major depressive disorder (MDD), we are not well informed about the long-term, post-treatment (i.e. after treatment has been terminated) outcomes for individuals who respond to these treatments in the short term. In the following section I review the prevalence of MDD and outline the need for more synthesis of the evidence on long-term, post-treatment outcomes. Issues such as the limitations in available data, as well as evidence of patients' and health care providers' need and desire for this information are discussed. Challenges in evaluating long-term outcome are described. Further context is provided in a summary of the evidence base and current guidelines for both psychological and pharmacological treatment of MDD. Following this are methods and results for a review of both previous reviews and primary studies that contribute to our knowledge about the long-term, post-treatment outcomes for individuals treated for depression. These take the form of an overview of reviews, reported in Study 1, and a systematic narrative review, reported in Study 2.

Figure 1 below depicts some of the important time-related constructs discussed in this paper. Briefly, the top half of Figure 1 shows the phase of treatment. Treatment begins with an acute or active phase. For some, but not all, individuals this is followed by continuation treatment or maintenance treatment phase. The time frames for each of these phases are not well established, but this is their order of occurrence. When treatment ends, which can be after acute treatment, or after continuation and/or

maintenance treatment, the individual is said to be in the post-treatment period. The bottom half of Figure 1 shows the individual's status. Once an effective course of treatment has been received, the person may experience a remission of symptoms. When full remission is maintained for a period of time, a person is considered to be in recovery. At some point during remission or recovery, treatment is usually terminated. In the post-treatment phase, a person can (a) maintain their gains, remaining in remission or recovery, (b) move from remission back into a *relapse* of their *initial* episode of depression, or (c) move from recovery into a new, *distinct* episode of depression referred to as *recurrence*. The lack of consensus on the time frames for these constructs will be explored later, however this diagram depicts the general order of occurrence.

Figure 1. Course of treatment and response over time



Major Depressive Disorder

Depression presents with a collection of affective, cognitive, and somatic symptoms including sad or depressed mood, loss of interest in normal activities, feelings of guilt, worthlessness or hopelessness, difficulties with concentration, reduced energy, changes in appetite and sleep, and withdrawal from normal activities. When five or more of these symptoms have been present during the same 2-week period, and represent a change from previous functioning, with at least one of the symptoms being depressed mood, or loss of interest or pleasure, an individual is considered to have reached a clinical threshold and qualifies for a diagnosis of major depressive disorder (MDD; American Psychiatric Association, 2000).

Depression has been cited as the most prevalent of the Axis I disorders, affecting upward of 16% of American adults in their lifetimes (Kessler et al., 2005). That said, there is considerable variability in the reported prevalence of MDD. The World Mental Health Survey conducted in 17 countries and currently considered to be the authoritative source for prevalence of major psychiatric conditions globally, found that on average about 1 in 20 people reported having an episode of depression in the previous year (Marcus et al., 2012). According to Statistics Canada's 2002 Canadian Community Health Survey, Cycle 1.2, 4.8% of the Canadian population aged 15 years and over met criteria for major depression over a period of 12 months with 12.2% meeting criteria for major depression at some point during their lifetime (Public Health Agency of Canada, 2009). Other well-conducted prevalence studies such as the National Co-morbidity Survey – Replication (NCS-R; Kessler, Berglund, Demler, Jin, Koretz, & Merkingas, 2003) report similarly, with 1-year and lifetime prevalence rates (up to the age at which

the interview was conducted) of 6.6% (95% CI, 5.9-7.3) and 16.2% (95% CI, 15.1-17.3), respectively. This is equivalent to national population projections of 13.1 - 14.2 million US adults with MDD in any given year, and 32.6 - 35.1 million with lifetime MDD. The NCS-R also calculated a projected lifetime risk up to age 75 years of 23.2%, which is likely a conservative estimate due to potential sources of bias (Kessler et al., 2005).

Rates are comparable in Canada. The Joint Canada/United States Survey of Health, measuring Canadian and US residents' ratings of health and health care services, found similar 1-year prevalence rates of depression (8.2 and 8.7% respectively; Vasiliadis, Lesage, Adair, Wang, & Kessler, 2007). This study also found similar use of mental health care services in Canada and the US, with some disadvantage experienced by US citizens with no medical insurance.

A recent systematic review examined 1-year prevalence rates for MDD around the world and found rates ranging from a low of 0.64 per 100 in Taipei to a high of 22.5 per 100 in Udmurtia, a region in Russia. This is a difference of approximately 35 fold (Waraich, Goldner, Somers, & Hsu, 2004). The authors of this review noted that this variability is due in part to the use of different criteria for diagnosis, as well as the level of the training of the individual who diagnoses, which can vary from clinician to lay interviewer to computer algorithm depending on the study. The same review reported that lifetime prevalence rates for MDD range between a low of 0.88 per 100 in Taipei to a high of 29.6 per 100 in Montreal, Canada, again a difference of more than 33-fold (Waraich et al., 2004). It is difficult to say what these differences represent given that they often do not employ common measures. Waraich et al.'s review does suggest that

further, methodologically sound, comparisons are needed in order to determine the true variability in prevalence of depression around the globe. It is also important to consider that stigma concerning mental health problems varies across cultures, which may impact reported rates regardless of the methods used.

There are gender differences in the prevalence of MDD. Wariach et al.'s (2004) review reported pooled 1-year and lifetime prevalence of 4.1 per 100 (95% CI, 2.4 to 6.2 per 100) and 6.7 per 100 (95% CI, 4.2 to 10.1 per 100) respectively. Their heterogeneity analysis indicated a significant gender difference, namely that both one-year and life-time prevalence rates were 1.5 to 2.5 times higher for women. Waraich et al. noted that their best estimates of life-time prevalence rates for MDD are lower than the commonly reported rates of 5 – 12% for men and 10 – 25% for women (American Psychiatric Association, 1994), and suggest that the estimates provided by their systematic review are more accurate than the rates reported in standard medical texts because their review has taken important sources of variation into account. In an editorial response to this comment, Goldney and Fisher (2005) counter that although Waraich et al. may have accounted for many sources of variation in prevalence estimates, that does not necessarily enhance accuracy. Goldney and Fisher suggest that most of the studies reviewed by Waraich et al. used instruments that have the potential to overlook many cases of depression and so their rates are not more accurate. They explained that, for example, some instruments screen out individuals who do not perceive their symptoms to interfere "a lot" with their daily life and activities, individuals who have not taken medication for their symptoms on more than one occasion, and individuals whose symptoms could be related to medication, drugs, alcohol, physical illness, or injury.

Although some of these exclusions may be consistent with DSM-IV guidelines, Goldney and Fisher remind us that co-morbidity is common in clinical practice and that excluding these individuals when calculating prevalence rates creates an underestimate of depression. Goldney and Fisher also note that excluding those who do not rate their symptoms as serious risks omitting all of those individuals who may be suffering from depression and simply have poor mental health literacy. Finally, they caution that excluding those who have not taken medication more than once may inappropriately exclude many people with depression who have not pursued drug treatment, as well as those who have experienced problematic medication side-effects and so choose not to continue to pursue that treatment option.

During adulthood, rates of depression appear to vary with age. Waraich et al. (2004) examined prevalence over the life course and suggest that lifetime prevalence rates for MDD seem to be fairly stable through ages 18 – 64 years. However, their data depict a non-significant trend for decreasing lifetime prevalence rate from young- to old-adulthood. Kessler et al.'s (2005) report on the NCS-R data indicates variability in lifetime prevalence. Kessler et al. (2005) find that reports of lifetime prevalence increase from young adulthood (age 18-29, prevalence = 15.4%) to middle adulthood (age 30-44, prevalence = 19.8%; age 45-59, prevalence = 18.8%) and then decline to the lowest point in older adulthood (age 60+, prevalence = 10.6%). Regarding age at onset, Kessler et al. (2005) cite a median age of 32 years for MDD. In another report examining the same 2001 NCS-R data, Kessler et al. (2003) found a significant cohort effect for age at onset. Their analysis showed that risk was low in the early teens, rising in a linear fashion with increasingly steep slopes in more recent cohorts. This means that age of

onset appears to be becoming younger among adults in successive generations or cohorts.

The prevalence of MDD is predicted to be on the rise as prevalence rates appear to continue to increase even when controlling for age-related recall differences, openness to disclosing information, and other methodological factors (Klerman & Weissman, 1989; Murray & Lopez, 1996). This is concerning, given that the prevalence problem is further compounded by the significant disability associated with MDD for many individuals. The World Health Organization (WHO) ranks MDD as one of the most burdensome diseases in the world (WHO, 2002) and according to the Global Burden of Disease study, the disability associated with MDD is second only to that associated with ischemic heart disease (Murray & Lopez, 1996).

The NCS-R (Kessler et al., 2003) measured role impairment among depressed individuals in four domains including work, household, relationship, and social roles. Almost all respondents (96.6%) with 12-month MDD reported role impairment associated with their depression in at least one of the four domains measured. Most (87.4%) described this impairment as moderate or worse, with 28.1% in the moderate category, 40.2% in the severe category, and 19.1% in the very severe category. People with 12-month MDD reported a mean of 35.2 days in the past year when they were unable to work or carry out normal activities because of their depression. Depression has been labelled, not just as a major source, but as *the main* source of disability in the workplace, both in Canada (Dewa et al., 2004) and the U.S. (Wang et al., 2004).

Now, consider that in the U.S., in the year 2000, the economic burden of depressive disorders was an estimated $83.1 billion (Greenberg et al., 2003). Although

the social and economic impacts of depression are serious, population surveys indicate that many people do not receive treatment for their depression at all in the 12 months prior to their being interviewed (Kessler et al., 2003; Vasiliadis et al., 2005). In both Canada and the United States, just over half (55.7% in both countries) of individuals diagnosed with depression, who have access to health insurance, report accessing some type of mental health care service for their depression (Vasiliadis et al., 2005). The situation is worse for U.S. respondents without medical insurance, with only 36.5% of those with a diagnosis of MDD reporting use of mental health services within the last 12 months (Vasiliadis et al., 2005). The NCS-R (Kessler et al., 2003) finds a similar rate for the 12-month MDD sample, with 57.3% receiving some type of treatment in the 12 months before their interview. It is important to consider, however, that when Kessler et al. (2003) set criteria for treatment adequacy, only 21.6% of those receiving treatment actually received adequate treatment. Adequate treatment was defined as either (1) at least four outpatient visits with any type of physician for pharmacotherapy, including the use of an antidepressant or mood stabilizer for a minimum of 30 days, or (2) at least eight outpatient visits with a professional in the specialty mental health sector for psychotherapy lasting a mean of at least 30 minutes. This definition was based on the recommendations of evidence-based treatment guidelines. Notably, of those who received treatment from a general medical practitioner only (not from a mental health specialist), only 9.6% received adequate treatment. This is important knowledge for policy makers who manage health care systems in which general practitioners are the first, and often only, point of contact for mental health care concerns.

Why This Review?

As prevalence rates suggest, depression is a common diagnosis. Moreover, individuals seeking treatment for MDD are faced with numerous, complex options. They must choose between psychological, pharmacological, or combined treatment with a myriad of choices in each of these categories. It is important for individuals seeking treatment to have access to current information about treatment efficacy, adverse effects, and long-term prognosis, which will assist them in making treatment decisions. It is also important for the professionals offering treatment to be able to meet the information needs of their clients and to assist in the best possible shared decision-making process. This thesis will contribute to the Mobilizing Minds Research Group's knowledge translation work, part of which includes the creation and dissemination of an informational decision-aid about depression and its treatment, for consumers and practitioners. The Mobilizing Minds Research Group is a knowledge translation research program funded by a Canadian Institutes of Health Research Team Grant. The team grant focuses on developing knowledge translation approaches to meet the information needs of young adults (aged 18 to 25) concerning common mental health problems. The team has members from the University of Manitoba, York University, McMaster University, and Brock University. The nominated principal applicant for the project is John Walker, Ph.D. in the Department of Clinical Health Psychology at the University of Manitoba. The research team has thirteen academic members, three research associates, eight young adult members, and two community partners. The current study is funded by this grant. The team is active in presenting findings of the project at conferences and is preparing a number of publications on the research findings

of the project. This decision aid, which is currently being piloted by other members of the research group, will provide a rigorously researched tool for considering the treatment options for depression. It will be designed for use by the individual seeking treatment, or in collaboration with a health professional in the context of shared decision-making.

There is a large, growing evidence-base and consistent, evidence-informed guidelines for pharmacological and psychological treatment of depression, which will be described later. Of note, these guidelines focus almost exclusively on the active treatment phase when individuals are taking a drug or participating in a psychosocial treatment. In contrast, only brief mention is given to discontinuation of active treatment, where it is stated, for example, that "The precise timing and method of discontinuing psychotherapy and pharmacotherapy for depression have not been systematically studied" (APA, 2010). It is recommended that the same consideration be given to terminating treatment as is provided at the point of treatment initiation, including "probability of recurrence, the frequency and severity of past episodes, the persistence of dysthymic symptoms after recovery, the presence of comorbid disorders, and patient preferences" (APA, 2010). Although these guidelines are descriptive of what a patient might expect during the course of various treatments, there is no similar description for the post-treatment prognosis.

In the above APA recommendation are two important issues related to the current proposal. The first is that our knowledge about the discontinuation of active treatment and the prognosis for long-term outcome is limited. The second is that patient

preferences are important at all stages of treatment planning. The following is a more in-depth examination of these two key points.

Long-Term Outcomes

The long-term outcomes of the various treatment options are not as well researched as the short-term efficacy of these treatments. Most research has been focused on the first 8 to 20 weeks of treatment, whereas many people remain in treatment or are at risk of relapse or recurrence over much longer periods of time. Importantly, patients describe themselves as wanting but lacking information about long-term outcome (Pollock, Grime, Baker, & Mantala, 2004; Walker, Joyce, Furer, Vincent, & Kjernisted, 2000).

The long-term outcomes most commonly reported in the literature are rates of relapse and recurrence. Relapse is defined as a return of symptoms after a period of remission (Belsher & Costello, 1988). Recurrence on the other hand, refers to the development of a new episode of disorder, after full recovery, which does not represent a return of symptoms associated with a previous episode. The distinction is said to be important because relapse assumes that the returning symptoms have the same etiology as those previously experienced, whereas recurrence represents a distinct episode. It has been argued that unipolar depression can have very different etiology from one episode to another (Hays, 1984), so that an individual who experiences recurrent depression may not simply be relapsing. This may have implications for choice of treatment for distinct episodes of MDD.

Problems in Measurement. There are a number of conceptual and methodological issues that complicate the measurement of relapse and recurrence.

Relapse and recurrence are measured from the point of remission, and remission itself is not uniformly defined in the literature. The American Psychiatric Association advises that remission ought to be defined as "a return to the patient's baseline level of symptom severity and functioning and should not be confused with substantial but incomplete improvement" (APA, 2010). Symptoms that never resolved to the point of remission should not be considered as recurrent when measured over time, as they were never satisfactorily resolved in the first place.

This advice is not followed consistently and remission is variably defined. At times remission is defined by referring to the same criteria used to define onset. Remission is said to occur when the number and severity of symptoms fall below the threshold for defining onset, and remain there for a specified duration of time (Belsher & Costello, 1988). Other studies consider patients to be in remission if their scale scores are reduced by a certain magnitude or percentage of the initial score (Targum, 1984). In some cases remission is reported as being "determined by the clinician" with no more specific criteria included (e.g. Hooley, Orley, & Teasdale, 1986).

In addition to a change in symptoms, definitions of remission refer to varying points of time in the recovery process. Some definitions require that the state of symptom remission have endured for a given period of time, though this is not standardized, while others do not mention this at all. This point of determination is key to the relapse rates that will be reported, because if several weeks of remission of symptoms are required for a person to be considered "in remission", there will be a shorter time to relapse, if it occurs, than if being "in remission" is defined as the point in time that remission of symptoms is first achieved. On the other hand, if remission is said

to occur as soon as sufficient symptom change has occurred, relapse rates may be artificially high because some of those considered in remission may not have been in a stable enough remission to qualify by more stringent standards. It may be more accurate to describe a person as having neither a recurrence nor a relapse, but as chronically depressed and not in remission at all (Belsher & Costello, 1988).

To further complicate matters, some studies rely on patients' retrospective accounts of their own psychological well-being to determine remission, recovery, relapse, and recurrence. For example, the National Institute of Mental Health (NIMH) Collaborative Depression Study (e.g. Keller, Lavori, Endicott, Coryell, & Klerman, 1983; Keller, Lavori, Lewis, & Klerman, 1983) used the Longitudinal Interval Follow-up Evaluation (LIFE) which asks patients to report on their week-to-week psychiatric well-being over the previous 6 months, the validity of which has been questioned (Aneshensel, Estrada, Hansell, & Clark, 1987).

There is also a lack of consensus in defining the time-line that distinguishes relapse from recurrence. There is no standard cut-off time at which return of symptoms no longer represents relapse and is considered to represent a new episode or recurrence (Belsher & Costello, 1988). A comparison of relapse and recurrence across studies requires vigilance around the definitions used by each study being considered. Take, for example, Keller and Shapiro's (1981) report that relapse has been used to describe return of depressive episodes during time-frames that range from a few weeks to as much as five years. These extremes are likely not comparable in many ways.

Another problem in defining relapse is the number, severity, and duration of symptoms required to constitute a relapse. Authors do not always report these criteria

and, when they are reported, there is disparity in defining the thresholds required for this definition. Measures of symptom-return vary. Most commonly used is the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960), with the Montgomery-Asperg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979) being second most common (Zimmerman, Chelminski, & Posternak, 2004). Each of these scales has longer and shorter formats and the cut-off scores representing relapse vary between studies. Frank et al. (1991) published recommendations around the use of HDRS scores which have led to the use of a more consistent cut-off score. However, other thresholds are still used by some (Schweizer et al., 2001; Trivedi, Rush, Pan, & Carmody, 2001). No such standards have been recommended for the MADRS and cut-off scores vary widely (Forlenza, Almeida, Stoppe, Hirata, & Ferreira, 2001; Guilfi, Anseau, Timmerman, & Korsgaard, 2001; Kyle, Petersen, & Overo, 1998; Levine, Deo, & Mahadevan, 1989; Nierenberg, Feighner, Rudolph, Cole, & Sullivan, 1994; Schweitzer et al., 2001).

A recent review of studies of the MADRS in healthy control subjects, sought to clarify the MADRS score used to define remission in treatment studies of depression (Zimmerman et al., 2004). The authors identified 10 studies of 14 samples including data on the MADRS for 569 healthy controls. They concluded that the narrow definition of remission as a complete absence of clinically significant symptoms required a cut-off score of 4 or less, whereas when remission was based on a broader definition, the optimal cut-off score was 9 or less. The authors cite significant differences between individuals scoring 4 and those scoring between 5-9 with regards to their level of psychosocial impairment and advocated for use of the lower cut-off. Their review did not find differences in scores based on gender or age group. There is no empirical work

to support a difference in HDRS scores that could be considered important or clinically significant (Montgomery et al., 1994), and there is little work relating scores of rating scales such as the HDRS and the MADRS to actual behavioural observation (Montcrieff, 2001). One study found the correlation between scale scores and behavioural observation ratings to be below 0.5 (Mazure, Nelson, & Price, 1986). Zimmerman and colleagues caution that this lack of agreement on the definition of remission allows researchers to examine different cut-off levels in a "search for significance," basing their published cut-off score on the most favourable findings in the data rather than an a priori decision around what constitutes meaningful remission.

Zimmerman et al. (2004) commented on the common practice in treatment studies of comparing mean improvement (change in MADRS or HDRS scores) in the treatment group to improvement in the control group. They cautioned that although this may tell us something about the efficacy of treatment, it does not tell us whether patients actually achieved remission of symptoms or not. They stressed that statistically significant change does not necessarily equate with clinically significant change. Therefore, the likelihood that patients will respond to treatment by reaching remission is of much greater clinical interest than simply knowing whether patients will likely experience some change which may not have practical implications for their daily functioning. A typical requirement for remission is a 50% change in score. This is problematic because a severely depressed individual who enters the study with an HDRS score of 40 could improve to 20, still be clinically depressed, but be considered in remission for the purposes of outcome measurement. Additionally, the commonly accepted 50% reduction is not based on any empirical findings (Moncrieff, 2001).

Definitions of remission that include a minimum % change *and* a final score below some pre-determined threshold are more useful because they tell us both that an individual has had a response to treatment (change in symptoms), and that this response has been sufficient to meet a pre-determined level (complete remission, not partial remission which may or may not be clinically meaningful). Ideally, the definition of remission should also include a period of time over which symptoms remain in remission, as well as a patient's return to regular functioning, but this is not often the case (Zimmerman et al., 2004). It is also important to consider the validity of the measures themselves. The HDRS, for example, contains a large number of items that relate to anxiety and six items that relate to sleep. This means that any medication with sedating properties may show favourable results when outcome is measured with the HDRS (Murray, 1989).

Note that reports of the total number of patient relapses within a sample are not particularly meaningful, as they will vary with the period of follow-up. Persons with frequent relapses can account for a large percentage of the total if multiple relapses are not distinguished from individuals with a single relapse. Further, occurrence of relapse may influence whether a participant drops out prior to the end of the follow-up. Also, reporting the mean duration between remission and relapse as the key outcome is flawed in that it excludes those who do not relapse, potentially significantly underestimating the duration of recovery. Studies that report the median time to relapse provide a clearer estimate of this outcome. Finally, variation in the length of follow-up conducted makes comparison of results across studies less than straightforward.

Survival analysis, which involves determining the probability that people will remain well at selected time intervals after recovery, is a preferred method of presenting relapse data (Lavori, Keller, & Klerman, 1984). This method produces a continuous picture of the probability over time that patients will be well or will have relapsed.

In addition to clarity around the definition of outcome measures, it is also important to exercise caution when interpreting comparisons between the treatment group and so-called 'normal' controls (Zimmerman et al., 2004). Reviewers point out that control groups often are not pre-screened for psychopathology or for the presence of medical illnesses that are known to inflate depression scores. It is thus important to consider the meaning of these comparisons: what does it mean to compare the treatment group to normal controls if the controls are not free of depression themselves? Statistical normality suggests that abnormality is defined as deviation from the population average. The medical model of normality, however, equates it with health and not in reference to deviation from a group mean. The statistical model would include everyone in the comparison group, whereas the medical model excludes individuals with disorder from the control group. Zimmerman et al. (2004) promote the use of the medical model and the exclusion of individuals with depressive symptoms from the control group. They provide the analogy of determining normal body temperature and suggest that you would not include individuals currently experiencing an infectious disease in your normal sample. Similarly, they suggest it is more meaningful and a cleaner definition of remission to compare to a normal control group that is free of depressive symptomatology. Finally, Zimmerman's review was unable to locate any large-scale epidemiological study of the MDRS among healthy controls. Therefore, the authors

caution that it may not be possible to generalize from the individual studies included in their review to the larger population.

The present studies consider the complexities of interpreting long-term outcome data and account for the length of the follow-up period and the presence and monitoring of treatment during the follow-up period. Treatment during follow-up is an important factor in interpreting long-term outcomes. During naturalistic follow-up extra-protocol treatment may be permitted or not. Where it is permitted, it may be monitored by a variety of methods. Other studies intentionally continue some form of intervention over the course of the follow-up period. The present studies refer primarily to relapse when discussing the return of symptoms after treatment. The definition of relapse is important, and the distinction between relapse and recurrence may also be meaningful. This distinction, however, is often based on an arbitrary time point, and there is no evident divide between relapse and recurrence when one observes these outcomes on a survival curve. For this reason, the current studies refer to relapse, unless citing a study in which the authors provided a clear definition and rationale for distinguishing relapse and recurrence as separate outcomes.

Gaps in Knowledge

Particularly within the pharmacological treatment literature, there is a relative paucity of follow-up data reported. Most intervention trials are very short. This is likely due to factors such as the feasibility of running longer-term studies, as well as the challenges related to attrition over longer time frames. One recent review reported that over 90% of the clinical trials they reviewed were between 6 – 8 weeks in duration (Frank, Revicki, Sorensen, & Shih, 2001). Psycho-social intervention studies often

include a follow-up. However, these studies vary greatly in duration from commonly reported follow-ups of several months duration (e.g. Santor & Segal, 2001) to very limited numbers of reports covering more than one year (e.g., Dobson et al., 2008; Fava, Ruini, Rafanelli, Finos, Conti, & Grandi, 2004). A further potential confound is that not all long-term follow-ups are created equal. Continuation studies of medication treatments may follow participants over the long-term, but these participants remain engaged in treatment. Follow-up studies can also occur after all planned treatment has been terminated, which provides information about the outcome post-treatment. This is important because many people will not continue treatment indefinitely. Discontinuation studies also provide a slightly different type of post-treatment outcome information because it is inherent in the design of these studies that all participants begin the treatment, and then are randomized to be withdrawn from treatment to form the control group. It is important to be aware of study design and the treatment status of individuals followed up over the long-term. This provides important information for generalizability of results, namely whether the results apply to individuals who are still engaging in treatment, or who have stopped treatment altogether. The treatment status of the comparison group is also important, in that we want to know about a treatment's long-term outcomes relative to different conditions (eg. no treatment, withdrawal).

Current guidelines for the treatment of depression (APA, 2010; Anderson et al., 2008) clearly reflect the uncertainty that exists around long-term, post-treatment outcomes. Recommendations for length of treatment are not based on empirical evidence, but on case-by-case clinical judgment (APA, 2010). Guidelines also note that it is not known if or when a patient should be withdrawn from medication, or what the

schedule of "withdrawal" should be from psychological treatment (APA, 2010). Given that it would be very helpful for practitioners and patients to have this information at the outset of treatment, as the following section of this thesis describes, it is imperative that the existing knowledge be accessible in a meaningfully synthesized way. It is also important to identify gaps in our knowledge and endeavor to more thoroughly examine outcomes over the long-term.

Patient Preference and Need for Information

Patient treatment preference has received minimal attention in the literature. In summary, it appears that when patients receive their treatment of choice, they may show earlier improvement and less switching between treatments. However, there is no evidence for any difference in eventual outcome (Chilvers et al., 2001; Lin et al., 2005; Peveler et al., 2005). Despite the potentially limited or non-existent effect on outcomes, patients' preferences and values are important, as they can play a role in their choice of treatment for depression and their willingness to initiate, continue, and comply with treatment.

The traditional, paternalistic model of health-care decision-making, in which the expert professional assesses the patient's needs and decides what form of treatment to implement (Charles, Gafni, & Whelan, 1999; Emanuel & Emanuel, 1992), is seen less favourably as attitudes shift with regards to the involvement of the person receiving the health care in making this decision. Informed decision-making, in which the expert provides information and the patient makes their own decision (Charles et al., 1999), or shared decision-making, in which the exchange of information and consideration of choices involves a collaborative effort between the patient and the health-care provider

(Brock & Wartman, 1990), have become common. Shared decision-making, in mental health care in particular, has been associated with empowerment, autonomy, and quality of life (Patel, Bakken, & Rulan, 2008). A study involving community mental health consumers found that 70% of these individuals had a moderate desire to be involved in sharing the decision-making process with their mental health care provider (Hill & Laugharne, 2006).

Given this shift, it is necessary for patients and health-care providers alike to be well informed about the aspects of treatment that will influence these decisions. It appears that this is not currently the status quo, as 65% of people in one study indicated both that (a) their GP was an essential source of information about depression and its treatment options and that (b) they experienced a lack of information about depression and its treatment options (Patel et al., 2008). These participants reported a desire for information on various topics including determining the presence of a mental disorder, the necessity of treatment, social stigma concerns, and the different types of treatment, including pharmacotherapy and psychotherapy.

Ludman et al. (2003) conducted a study in which the intervention group received a low-intensity 12-month intervention comprised of education about depression and shared decision-making regarding the use of medications and cognitive-behavioural therapy for self-management. The intervention was delivered via a book, a video tape, two in person sessions with a depression prevention specialist (psychologist or psychiatrist trained for this particular intervention), three telephone sessions and four personalized mail-outs for monitoring symptoms and treatment adherence. They found that, compared to usual care control patients, the intervention group showed greater self-

efficacy (measured with a previously validated 6-item scale; Bush et al., 2001) in managing their depression. Those in the intervention group were also more likely to track and monitor their symptoms and to plan for high risk situations.

While more research is required to establish the relationship between patient involvement and mental health outcomes, it is important to remember that, regardless of the magnitude of the impact that patient education and involvement may have on such outcomes, these things are necessary for ethically sound practice. Medical and psychological associations in Canada and the U.S. have ethical codes requiring practitioners to provide patients with the relevant information they require to make decisions about their care and to engage in shared decision-making (American Medical Association, 2001; Canadian Medical Association, 2004; Canadian Psychological Association, 2000). Unfortunately, in the real world, this does not always occur and discussion around decision-making has been found to be poor (Braddock, Edwards, Hasenberg, Laidley, & Levinson, 1999) with patients often wanting more information than they are given (Coulter, Entwistle, & Gilbert, 1999).

With regards to patients' desire for information, research has examined the information needs of people with various health conditions and found that they have a strong desire to be informed about their condition, as well as the risks and benefits of treatment (Deber, Kraetschemer, & Irvine, 1996; Hill & Laugharne, 2006; Stiggelbout & Kiebert, 1997). Patients also have a preference for detailed information and feel that they are not provided with enough information about their health concern and the treatment options (Cassileth, Zupkis, Sutton-Smith, & March, 1980; Simon, Loh, Wills, & Harter, 2006). Desire for information has been found to vary with age and education,

with younger and more educated individuals having a stronger desire for information (Cassileth et al., 1980).

Fewer studies have examined the information needs of patients with *mental health* problems but research is beginning in this field. One study found that inpatients with mental health diagnoses reported a desire for, and a lack of, information about their diagnosis, the names and dosage of their medications, as well as the side-effects, long-term effects, and implications of not taking medication (Pollock et al., 2004). Another study by Garfield, Francis, and Smith (2004) asked patients, at the beginning of pharmacological treatment for depression about their information needs. These patients reported that they lacked information on side-effects, the duration of treatment, concerns about drug dependency and dosage. A study of information needs among people with anxiety also found that the majority of patients reported having received "none" or "little" information about treatment (psychological or pharmacological) when they had received it (Walker et al., 2000). These patients indicated that information on the types of treatments, the efficacy of different treatments, the patient's role in treatment, and the effects of discontinuing treatment were very important. They also indicated that various sources would be acceptable for obtaining this information including information in a booklet, discussions with a health care provider, and discussions on videotape. Most participants in this study indicated that, in their previous decision-making experiences, the information they had received had come from discussions with their health care provider; a finding that is consistent with other research on mental health decision-making (Simon et al., 2006). In response to this need for information, this thesis contributes to the Mobilizing Minds Research Group's current work to create accurate

and accessible informational resources and a decision-aid tool for patients and health care providers concerned with the treatment of depression.

Evidence-Based Treatments for Depression

In a review of the long-term, post-treatment outcomes for individuals treated for depression, it is important to understand the range of evidence-based treatments currently being employed. These are treatments that have been found to be effective in one or more randomized controlled trials (RCTs), with the active treatment being compared to a condition that does not include the treatment and with participants being randomly assigned to these groups. Preferably, studies of evidence-based treatments have also been replicated by researchers independent of those who designed or created the treatment. There is a substantial literature on evidence-based treatments for depression in adults, with numerous well-researched pharmacological and psychological intervention options showing broad applicability. Guidelines for the treatment of major depressive disorder, including summaries of these evidence-based practices, have been published by several influential bodies including the American Psychiatric Association (APA, 2010), the Royal Australian and New Zealand College of Psychiatrists (2004), the British Association for Psychopharmacology (Anderson et al., 2008), and the National Institute for Health and Clinical Excellence in the United Kingdom (NICE, 2009), with the aim of assisting health care providers and patients in choosing an appropriate treatment based on the evidence and specific individual situations. Following is a brief synopsis of current evidence-based pharmacological and psychological interventions for depression, as well as the current evidence comparing these two treatment approaches and the state of knowledge to date on combining these

treatments. This brief synopsis mainly covers the short-term efficacy of these treatments, and is necessary context for the larger aim of this paper which is to investigate the long-term post-treatment outcomes of the same interventions.

Pharmacological Treatments

There are several classes of anti-depressant medications, with numerous drugs in each class. The older classes, including tricyclic antidepressants (TCAs) and monoamine oxidase inhibitors (MAOIs), are not currently favoured due to their side effect profiles (Anderson, 2001; Montgomery et al., 1994; Mulrow et al., 2000) and the higher risk of harm due to overdose or combination with other medications (Hansen, Gartlehener, Lohr, & Gaynes, 2005). The newer, "second generation" antidepressants have a relatively favourable side effect and safety profile. They include selective serotonin reuptake inhibitors (SSRIs), selective norepinephrine reuptake inhibitors (SNRIs), and other drugs that selectively affect the activity of neurotransmitters, such as, serotonin, norepinephrine, and dopamine (Hansen et al., 2005). These newer antidepressants currently play a prominent role in the treatment of MDD (Hansen et al., 2005) and the guidelines for their use and evidence for their effectiveness is presented here.

Medication treatment, for a period of 6 – 8 months, is recommended as the standard of care for treatment of depression in the current psychiatric guidelines (APA, 2010; Anderson et al., 2008; NICE, 2009). These guidelines recommend that patients who have responded to pharmacotherapy for depression continue to take the same medication at the same dosage during continuation and maintenance phases. Lower-dose maintenance treatment with anti-depressants has not been proven effective in preventing

relapse (APA, 2010). Length of recommended treatment is indefinite for those with chronic or recurrent depression (APA, 2010).

While use of antidepressant medication is recommended as the first line treatment for moderate to severe, acute depression and for sub-threshold chronic depression (Anderson et al., 2008), the research on the efficacy of these drugs is not as strong as one might assume. As an example, a review by Barbui, Furukawa, and Cipriani (2008), looked at the efficacy of paroxetine, an SSRI, versus placebo in treating MDD, including 29 published and 11 unpublished trials. Barbui et al. found that, although more participants in the paroxetine group reached at least 50% improvement in symptoms, paroxetine was not more effective than placebo with regards to overall treatment efficacy or acceptability. In another review, Deshauer et al. (2008) synthesized the results of 6 classic RCTs of SSRIs and reported similar findings. Deshauer et al. found that patients who continued treatment for 6-8 months were significantly improved, as measured by response to treatment, but they did not have significantly better rates of remission or ratings of treatment acceptability when compared to placebo groups. These two reviews highlight the importance of carefully considering the positive outcomes reported in treatment studies. The outcome of "at least 50% improvement" or "significantly improved" do not translate directly to remission rates.

Another recent meta-analysis of 75 short-term RCTs showed a 50% response rate in the antidepressant group, with 55-65% of those on antidepressants continuing to have significant symptoms (Walsh, Seidman, Sysko, & Gould, 2002). These reviewers also found a 30% response to placebo. It is important to compare treatment response to placebo response when determining the specific effect of the treatment because it is

acknowledged that placebo is likely better than no treatment at all, though the data on this are scarce. It is also important to consider and control for the severity of depression in both groups because there is evidence that the difference between treatment and placebo increases as the severity of the depression increases (Angst, Scheidegger, & Stabl, 1993; Khan, Brodhead, Kolts, & Brown, 2005; Kirsch, Scorboria, & Moore, 2002; Ottevanger, 1991). Hamilton Depression Rating Scale scores (HDRS) above 24 are commonly accepted as the dividing line between moderate and severe MDD, and are associated with the most consistent and clinically significant difference between drug and placebo (Khan et al., 2005; Muller, Himmerich, Kienzle, & Szegedi, 2003). These studies suggest that medication treatment, although it may be a first line treatment for depression, certainly cannot guarantee remission.

In addition to reports of varying levels of efficacy in the literature, there appears to be a lack of consensus on the *comparative* efficacy of second-generation antidepressants. That is, it is not clear, among the treatments believed to be effective, which are more effective than others or if differences in efficacy can even be found consistently. It is challenging to determine whether the apparently contradictory findings represent real differences in efficacy, or whether they reflect reviewer's inclusion of different studies, focus on different drugs, and variability in the meta-analytic methods employed. Because these differences exist, the contrary findings that follow must be carefully considered before drawing conclusions about the comparative efficacy of second-generation antidepressants.

First, consider the numerous recent systematic reviews and meta-analyses that indicate that second-generation antidepressants *are* comparably efficacious for the

majority of patients (Anderson, 2001; Gartlehner et al., 2008; Hansen et al., 2005). Hansen et al. (2005), for example, found that second-generation antidepressants are not significantly different in their efficaciousness, reporting that 88% of comparative efficacy studies found no statistically significant differences in any outcome measures at the end of the study. Hansen et al. do not recommend one drug over another, suggesting that choosing the right medication for the right patient is difficult if not impossible.

A more recent review by Gartlehner et al. (2008) supports the finding of comparative efficacy. Gartlehner et al. reviewed 203 studies comparing the benefits and harms of second-generation antidepressants and found that there were no significant differences in efficacy for the treatment of MDD. They also found no difference in the spectrum of adverse events experienced, although they did find differences in the incidence of specific adverse effects and in the onset of action. These authors suggested that these differences may be used to inform the choice of which medication to prescribe in the absence of differential efficacy.

In contrast, are the findings of a systematic review by Cipriani, Barbui, Furukawa, Hotopf, & Geddes (2006) that looked at the acceptability (defined as continued adherence to treatment) and efficacy of fluoxetine, the most widely studied of the newer antidepressants, in comparison to other anti-depressants, both new and old. Blinded and non-blinded studies were included for a total of 131 studies reviewed. Results indicated that fluoxetine was more *acceptable* than amitryptyline and pramipexole, but that sertraline and venlafaxine were more *effective* than fluoxetine. The authors of this review concluded that there *are* differences between fluoxetine and other antidepressants. Of note, Cipriani et al. (2006) looked at fluoxetine as a common

comparator to other antidepressants which, it could be argued, might produce more sensitive comparisons with these other options than if no common comparator is used.

More recently, a review of comparative RCTs found that there are clinically significant differences among the most commonly prescribed antidepressants in terms of both efficacy and acceptability (Cipriani et al., 2009). Specifically, these authors found that escitalopram, venlafaxine, and sertraline were significantly more efficacious than duloxetine, fluoxetine, fluvoxamine, paroxetine, and reboxetine. Moreover, escitalopram and sertraline were most acceptable, leading to fewer discontinuations than the others. Cipriani et al. (2009) further stated that sertraline might be the best choice for treating moderate to severe MDD because it has the best overall profile when accounting for efficacy, acceptability, and cost.

Some evidence suggests that the newer SNRIs, such as venlafaxine and duloxetine, as well as other serotonergic-noradrenergic antidepressants, are more efficacious than SSRIs. A review of the effects of serotonergic-noradrenergic (SNRI) antidepressants in comparison to SSRIs in treatment of MDD found that SNRIs, which simultaneously enhance serotonin and norepinephrine uptake, are statistically more efficacious than SSRIs and more likely to result in clinically significant response (Papakostas, Thase, Fava, Nelson, & Shelton, 2007). There was no difference among the various SNRIs examined (including venlafaxine, duloxetine, milnacipram, mirtazapine, miasnserin, and moclobemide), except that duloxetine was significantly less effective. This particular finding was unexpected and the authors recommend taking it with caution.

Both Hansen et al. (2005) and Gartlehner et al. (2008), who found no difference in comparative efficacy for second-generation antidepressants, conducted broader searches than Cipriani et al. (2006), and Cipriani et al. (2009), who did report differences in efficacy. Hansen et al. included not only comparative studies, but also observational studies in which treatments were not directly compared and placebo-controlled trials in which one of the comparisons was a placebo condition that did not involve an active treatment. Both Cipriani et al. (2006) and Cipriani et al. (2009) included only comparative studies in their reviews and, in the case of Cipriani et al. (2006), the scope was further refined, with fluoxetine acting as a common comparator. Gartlehner et al.'s search extended back to 1980, including an initial 2323 titles, 115 of which were retained for review. Cipriani et al. (2009) only went back to 1991, including an initial 345 titles, 117 of which were retained for review. Examining only more recent studies may lead to conclusions that are more relevant, as they will include trials with the currently prescribed medications, whereas older studies may be reporting on medications that are no longer commonly prescribed. The breadth of the search conducted in each review is important, as the apparent contradictions between one review and another may be at least partly due to the effect of different inclusion criteria clouding conclusions about actual differential efficacy.

Limitations of Antidepressant Trials

Research always has limitations and antidepressant trials have been critiqued on several methodological grounds. Recall, for example, the review of classic, long-term RCTs conducted by Deshauer et al. (2008) of which they located only six. They conducted a review of the use of SSRIs in treatment of unipolar depression. They

excluded discontinuation trials in which all participants begin the trial on medication and then the control group are withdrawn in a randomized fashion. These reviewers cite the discontinuation trial as problematic because the results apply only to those who respond to medication and not to those who experience spontaneous recovery, for which rates are potentially high. A second problem with discontinuation trials is that those who become the comparison group must be withdrawn from active medication treatment. Discontinuation symptoms can mimic depression and may lead to the withdrawn group appearing more symptomatic, thus providing an overestimate of the true effect of the medication (Baldwin, Cooper, Huusom, & Hindmarch, 2006; Greenhouse et al., 1991; Kopec, Abrahamowicz, & Esdaile, 1993).

Moncrieff (2001) conducted a review of the methodological problems in antidepressant trials and provided several cautions. She found that unblinding, due for example to the differential physiological effects of drug and placebo ingestion, can inflate response in the medication group. This unblinding would mean that people would know if they were taking the active medication or were receiving the placebo, and would defeat the purpose of the placebo group, which is to control for an individual's knowledge that they are taking an active medication. Placebo conditions are used to ensure that the act of "taking something" is not responsible for improvements in a person's condition. If, however, people can tell whether they are taking the placebo or the medication, this knowledge may bias their expectations and reports of improvement. Studies like that of Cipriani et al. (2006) have addressed this potential weakness by using an active treatment as a comparator, rather than a placebo.

With regards to inclusion criteria, Moncrieff (2001) observes that excluding some subjects from analysis (e.g., those who withdraw after randomization), may inflate the treatment results. She suggests that the problem of publication bias may be even more significant in trials of SSRIs because the majority of these trials are funded by pharmaceutical companies who are not obligated to publish negative results. This bias is a source of concern to other reviewers as well (Cipriani et al., 2006). Finally, with regards to methodological concerns, Moncrieff, like Deshauer et al. (2008), notes the potential for discontinuation-related effects to confound efficacy findings in the commonly employed design of a discontinuation trial.

Psychological Interventions

There are also numerous psychological interventions for the treatment of depression that have amassed a significant evidence base. Current treatment guidelines recommend cognitive behavioural therapy (CBT) and interpersonal psychotherapy (IPT) as the psychological interventions with the best documented efficacy in the treatment of major depression (APA, 2010; Anderson et al., 2008; NICE, 2009). The BAP guidelines recommend CBT, IPT and behavioural activation (BA) as alternatives to antidepressants in the acute treatment of mild to moderate MDD, with CBT named as the treatment of choice if psychological treatment is the monotherapy (Anderson et al., 2008). Other psychological treatments such as brief psychodynamic psychotherapy, problem solving therapy, and various client-centred approaches continue to be examined empirically, but are not recommended as first-line treatments for depression.

While the above guidelines are meant to be based on a summary of the best evidence to date, a closer examination of the literature and of the variety of knowledge

synthesis efforts reveals that the evidence base is not clear-cut. Following is a summary of the evidence to date on the state of research on CBT, IPT, and other evidence-based psychological interventions for depression.

Cognitive Behavioural Therapy (CBT). CBT is the most extensively researched psychological treatment for unipolar depression (Scott, 1996), with the first key treatment manual published more than thirty years ago (Beck, Rush, Shaw, & Emery, 1979). CBT involves identifying maladaptive thoughts and behaviours and engaging in strategies to correct them.

In an overview of current meta-analyses on treatment outcomes of CBT for various psychiatric disorders Butler, Chapman, Forman and Beck (2006) found large effect sizes ($d > 0.8$; as defined by Cohen, 1992) for unipolar depression. They concluded that CBT was equal to behavioural activation and superior to antidepressants in treating adult depression. Reviews of CBT assume that it is a standardized treatment across studies. CBT, as is the case for any treatment, manualized or not, is not delivered in a standardized manner and details are rarely, if ever, provided that would allow for replication of the precise treatment that was delivered in a given trial. This does not mean that we cannot or should not compare or synthesize studies of CBT, only that we must be aware of this significant limitation.

Other recent work supports Butler et al.'s finding that CBT is effective in the treatment of depression. For example, Dobson et al. (2008) found that brief CBT is as effective over the long-term as continued medication treatment, that CBT has more enduring effects than medication post-treatment, and that CBT is less expensive over the long-term than medication.

The recommendation of current guidelines for CBT as an effective treatment may be the best advice that can be offered at present, but they do not inform the reader of the inconsistencies in the literature. It is true that the best reviews to date agree that CBT is effective. They remain in disagreement, however, about the size of this effect. The most recent meta-analysis of CBT for the treatment of depression in adults (Lynch, Laws, & McKenna, 2010) suggests that the commonly accepted conclusion that CBT is an effective treatment is dependent, not on the results of individual trials (where the evidence is mixed), but on the results of meta-analyses that have methodological limitations that may inflate effect size. Past benchmark reviews (e.g. Churchill et al., 2001; Glauguen, Cottraux, Cucherat, & Blackburn, 1988) hail CBT as effective and superior to wait-list control, medication, and other therapies in achieving symptom remission and in prevention of relapse in the long-term. However, these reviews are criticized by Lynch et al., primarily for failing to take lack of blinding into account, which these authors suggest stands to confound their results. That said, Lynch et al. (2010), in their own meta-analysis, controlled for blindness and found that it did not significantly moderate effect size. It seems reasonable, then, to question how different the results of previous meta-analyses would have been had they too accounted for blinding in their reviews.

Lynch et al. (2010) identified several well-conducted studies that actually reported negative findings for CBT as a treatment for depression. For example the National Institute of Mental Health (NIMH) conducted a study of brief psychotherapeutic interventions and found no evidence for the efficacy of CBT and only marginal evidence for the efficacy of IPT (Elkin et al., 1989). However, despite these

findings and their rigorous control for blinding, Lynch et al. concluded that CBT does have a modest (d = .28) therapeutic benefit in the treatment of symptoms of depression and in prevention of relapse. Effect sizes were significantly larger in studies comparing CBT to medication than in studies comparing CBT to another psychological intervention (Lynch et al.). They count as an important function of their review the finding that a large, methodologically rigorous trial comparing CBT to a non-specific control intervention in the treatment of depression has yet to be carried out.

Finally, consider a review by Parker, Roy and Eyers (2003). They suggest that CBT is not superior to other psychotherapies or to antidepressants across the board, and cite a dearth of evidence for CBT as an augmentation of medication therapy. Parker et al. suggest that part of the confusion regarding the use of CBT for depression is that most studies of the efficacy of CBT are flawed in some important aspects. They report that many studies do not control well enough for the severity of depression experienced and that CBT is likely more effective for less severe depression. They do not report strong empirical support for this suggestion, however, and other reviewers report that CBT is effective *regardless* of initial Beck Depression Inventory (BDI) score (Glauguen et al., 1998). Some trials also find CBT to be as effective as medication for acute treatment (Dimidjian et al., 2006), and superior to medication for long-term relapse prevention, in those diagnosed with moderate to severe depression (Glauguen et al., 1998).

Parker et al. (2003) also suggest that CBT is not comparably effective for all depressive subtypes (e.g. psychotic, melancholic, and heterogeneous residue of nonmelancholic disorders), noting that this is seldom accounted for. Again, there is no

significant evidence to argue for or against this suggestion, and future studies might consider this as a variable to investigate. Parker et al. also point out that there are numerous variables to account for in CBT trials that do not play the same role in medication trials. These include therapist characteristics, adherence to a treatment manual, and the therapeutic relationship. Again, these variables are often not documented, and so their influence on the results of CBT research cannot be determined with confidence.

Parker et al. (2003) offer a conservative summary. First they suggest that it is fair to say that CBT is equally as effective as other psychotherapies or basic clinical management (which is not defined and likely varies considerably in practice). They attribute CBT's current reputation to the extensive scientific evaluation it has received relative to other therapies, and to the credibility it enjoys among patients and practitioners due to its logical and rational approach. Secondly, they suggest that CBT is similar to all other treatments for depression in its varying levels of benefit experienced across heterogeneous groups. Although their cautions ought to be considered, the lack of empirical support for some of them implies that these variables should be included in future research, and may not limit the conclusions from previous research.

Haby, Donnelly, Corry, and Vos (2005) also suggest important variables for consideration when interpreting and combining results of CBT trials. They found that heterogeneity in effect size was explained by treatment, duration of therapy, inclusion of severely depressed patients in the trial, the type of control group, and the number of drop-outs from the control group. Other variables that explained heterogeneity in effect size included the year of the study, the country the study was conducted in, and the

language. This latter group of variables likely represents a group of proxy variables for things like improvements in methodology over time (year of the study), and the qualifications of the therapists delivering the intervention, which may vary by geographical location. Many of these variables can and should be controlled for, as they cannot be eliminated. One exception, and an interesting approach to the confounding variables associated with therapist characteristics, is the use of computer technology for delivery of a CBT intervention. The literature on computerized CBT (CCBT) is in its infancy, but thus far some positive results shows evidence for CBT's efficacy in the absence of a therapist and the associated bias and confounds (Kaltenthaler, Parry, & Catherine, 2004; Kaltenhaler et al., 2006; 2008).

The cognitive aspects of CBT have received more attention than the behavioural aspects of this type of treatment, and the terms CBT and CT (cognitive therapy) are, in fact, often used interchangeably. Noteably, Hollon, Stewart, and Strunk, (2006) suggest in their review that more purely behavioural interventions may be just as effective as more purely cognitive interventions, and as effective as medication treatments in the short term.

Interpersonal Psychotherapy (IPT). IPT is also considered an evidence-based treatment for MDD (APA, 2010; Anderson et al., 2008; NICE, 2009). IPT was developed in the 1970's (Klerman, Weissman, & Rounsaville, 1984) to treat major depression and has since been applied to a variety of disorders (Klerman et al., 1984; Weissman, Markowitz, & Klerman, 2000). IPT deals with current relationships and focuses on the patient's immediate social context. It aims to intervene with symptom formation and the social dysfunction associated with depression. The literature on IPT is

compelling, though much smaller than the body of research that exists for CBT. Several reviews have been published citing the efficacy of IPT for treatment of depression (Feijo de Mello, Jesus Mari, Bacaltchuk, Verdeli, & Neugebauer, 2005; Jarrett & Rush, 1994; Klerman, Weissman, & Markowitz, 1994). The review by Feijo de Mello et al. was most thorough in that the authors examined not only the efficacy of IPT as a treatment for depression, but they conducted four separate meta-analyses comparing IPT vs. medication, combined IPT and medication vs. medication alone, IPT vs. placebo, and IPT vs. CBT. Their findings indicated that IPT was superior in efficacy and acceptability to placebo. There was no difference in efficacy or acceptability between IPT and medication. Also, when IPT and medication were given together, there was no difference compared to when medication was given alone, indicating that IPT did not add significantly to the outcome of medication treatment alone. Finally, IPT was more effective, but not more acceptable than CBT. This last comparison is based on only three studies and should be interpreted with caution.

Other Therapies. Other therapies that have received research attention and are mentioned here as less well-supported options in current practice guidelines include problem solving therapy, various forms of client centred therapy, and short term psychodynamic psychotherapy (STPP). These treatments require further study to accurately assess their efficacy as treatments for MDD. STPP has received enough research attention to have produced one meta-analysis (Abbas, Hancock, Henderson, & Kisely, 2006), the results of which indicate mixed findings. Of those studies included in Abbas et al.'s review, some indicate that STPP is superior to no treatment and to minimal treatment controls, and equal to other treatments (Anderson, 1995; Crits-

Christoph, 1992). Others, using a largely different group of studies, find STPP to be inferior to other treatments, and equal to minimal treatments, but this effect is lost at follow-up (Svartber & Stiles, 1991). A review by Hollon and Ponniah (2006) reports that, to date, no well-conducted studies have found dynamic psychotherapy to be superior to a nonspecific control or alternative treatment in the acute phase, and they conclude that there is insufficient data to conclude that it is an efficacious treatment. Proponents of psychodynamic approaches argue that there is sufficient evidence to consider these interventions evidence-based, and that one of the reasons that this literature is so small is that accepted methodologies such as the RCT do not fit well with psychodynamic approaches which are often longer than a typical intervention trial (Taylor, 2008).

Limitations of Psychological Intervention Trials

There does not appear to be a body of evidence to suggest the optimal duration of treatment or tapering schedule for termination for any of the currently accepted psychological treatments for depression. The recommendation of the American Psychiatric Association (2010) for patients who have responded to any form of psychotherapy for depression is that they continue this form of psychotherapy for a 16 – 20 week continuation phase after remission to prevent relapse. The optimal frequency of continuation visits is not suggested, and APA's guideline recommends basing this on the patient's needs and the practitioner's clinical judgement. Following the continuation phase, APA recommends that treatment continue during what is called the maintenance phase. During this phase, APA notes that sessions usually decrease in frequency. However, they do not cite any specific studies that have demonstrated this to be an

effective practice and, unlike most of their guidelines, they do not classify the degree to which this has been established as good practice.

In addition to this gap in the literature, there are several methodological concerns that are important to consider when interpreting the evidence base. While pill placebo is a standard control condition in studies of medication treatment and treating clinicians in these studies can be blind to the patient's treatment condition, it has been difficult to establish a practical psychological treatment placebo. Also, because clinicians are delivering the psychosocial intervention they are not blind to the patient's treatment condition. Placebo may not be an appropriate option for a control group condition in studies of psychological therapies (Parloff, 1986) and these studies often do include appropriate comparisons such as wait list control, treatment as usual control, or some other active treatment such as educational support as a control condition. It is important to consider both the presence and the nature of the control group. In addition, many studies are small and depression is often broadly defined, leading to problematic heterogeneity that is not accounted for in the analyses (Anderson et al., 2008). This increases the likelihood of bias and confounding, as reported in Lynch et al.'s (2010) review.

Recall another important and potentially confounding variable in psychotherapy studies, that of therapist effects. A meta-analysis of therapist effects in psychotherapy outcome studies (Crits-Christoph et al., 1991) found that use of a treatment manual and greater therapist experience was associated with small differences in effect between therapists. Less experienced therapists and no treatment manual were associated with larger therapist effects. This review found that type of treatment (cognitive versus

psychodynamic) was related to outcome, but that when the overlap between type of treatment and use of a manual were controlled, type of treatment was no longer a significant predictor of outcome.

Finally, recall an earlier point regarding the adequacy of treatment received from general practitioners. This is an area that begs for further research, as treatment commonly occurs in a GP's office and the findings for this type of delivery are mixed. The NCS-R study reports less than 10% of those treated for depression by their GPs receive adequate treatment (Kessler et al., 2005). However, a recent review of treatment of depression in primary care settings (which certainly encompasses the work of GPs) reports that psychological interventions delivered in the primary care setting achieve significant effect sizes (d = .31), though smaller than for psychological interventions delivered in other settings (d = .67), such as clinical samples from specialized mental health care (Cuijpers, van Straten, van Schaik & Andersson, 2009). Cuijpers et al.'s (2009) review of psychotherapy in primary care compared CBT, problem solving therapy (PST), and other forms of psychotherapy. They found a medium effect for CBT (d = .42), as well as a smaller effect for PST (d = .19) and other therapies (d = .27). Of note, this review did not involve a meta-analysis, as the studies they included proved too heterogeneous.

Combined Treatment

Having reviewed mono-therapies (both psychological and pharmacological), it is relevant to consider the state of research on combining these treatments. The evidence is mixed on the benefits of combined treatment. Although current guidelines (Anderson et al., 2008) recommend psychological treatment as an adjunct to medication for treatment

of severe MDD, a recent mega-analysis (a recently developed data synthesis method which combines individual data) reports equal outcomes for combined drug and psychotherapy versus IPT or CBT alone (Hollon et al., 2005a). Mega analysis allows for more refined and broader analyses than typical meta-analyses which do not combine individual data. Mega analysis, however, is not well established and not frequently employed.

Other meta-analyses find that combined treatment produces slightly better outcomes and, importantly, improves retention rate (Keller et al., 2000; Pampallona, Bollini, Tibaldi, Kupelnick, & Munizza, 2004). In studies where the focus of combined therapy is to enhance adherence to antidepressant medication, reviewers find that those in combined therapy had significant improvement over those on drug treatment alone (Pampallona et al., 2004). They also find that adherence rates are affected only for trials longer than 12 weeks, with the combined group having better adherence than the drug only group (Pampallona, et al., 2004). This finding is important given that guidelines recommend drug treatment duration significantly longer than 12 weeks. It may be that psychological treatment does not enhance the efficacy of medications in the short term, but does improve adherence over the long-term, thus increasing the probability that patients will reap the potential benefits of ongoing medication treatment. An assessment is provided by Otto, Smits, and Reese (2005), who report that, although combined therapy is not of such benefit that it ought to be considered the default treatment for MDD, it may be beneficial in cases of chronic depression and for prevention of relapse.

Sequential treatment, in which one treatment follows another, is a particular variant of combined treatment. The Sequenced Treatment Alternatives to Relieve

Depression (STAR*D) Trial, is the largest pragmatic trial on the treatment of depression that has been run to date (Sinyor, Schaffer, & Levitt, 2010). Participants were individuals presenting to psychiatry or family practice seeking treatment for depression (*n* = 2876). Each stage of treatment lasted from 12 – 16 weeks and was referred to as a "Level". At Level 1, all patients began on the antidepressant citalopram. For those not reaching remission, various augmentation or switching options were made available at Level 2, including other antidepressants (bupropion SR, buspirone, sertraline, venlafaxine XR) and/or CBT. Level 3 offered additional medication augmentation (lithium or T3) or switching (mirtazapine or nortriptyline). Level 4 involved randomization to either tranylcypromine or a combination of venlaxafine XR and mirtazapine. Patients who failed to reach remission in the time allotted at each level (ranging from 12 – 16 weeks) could move on to the next level of treatment. No significant differences were found when comparing outcomes across treatments at any one level. However, rates of remission and response were both higher at Levels 1 and 2 than at Levels 3 and 4 suggesting that if response to treatment is not achieved within the first 24 weeks of therapy, future response to treatment, even to changes and augmentations, becomes less likely as time goes on. The authors cite the duration of each level as a strength of this trial. They note that, whereas many people may not reach remission or even respond within several weeks of initiating their first therapy, many of these will remit within 14 weeks of their first therapy and even more will remit after a second treatment step involving augmentation or switching treatments. Results of the STAR*D trial expand on Otto et al.'s (2006) conclusion that combined treatment, though not the obvious default for treatment of MDD, can be beneficial. STAR*D

results indicate that it is not only those who suffer from chronic or recurrent depression that may benefit from combined treatment.

Comparative Studies

Finally, with regards to comparison of psychological and pharmacological treatments, the literature is not unanimous on the relative benefits. Numerous reviews have been conducted to compare these two treatment modalities and the mixed findings are summarized here.

Numerous studies comparing CBT to other therapies find little difference in efficacy. Specifically, authors report no significant difference between the efficacy of CBT and behavioural activation (Cuijpers, van Straten, & Warmerdam, 2007; Ekers, Richards, & Gilbody, 2008), IPT (Casacalenda, Perry, & Looper, 2002; Ekers et al., 2008; Luty et al., 2007), or antidepressants (Casacalenda et al., 2002; Dimidjian et al., 2006; Ekers et al., 2008). While potentially equal in efficacy, these treatments appear to be more effective than no treatment, as evidenced by Casacalenda et al.'s (2002) meta-analysis of 6 RCTs with control treatment arms, which found equal remission for antidepressants, CBT, and IPT (46%) over control conditions (26%).

A review of psychosocial and pharmacological interventions for MDD in primary care also reported comparable efficacy for psychotherapy and pharmacotherapy and concluded that both are favourable in comparison to usual care, though no meta-analysis was conducted due to the heterogeneity of the studies reviewed (Wolf & Hopko, 2007). A review of comparison studies in treating later-life depression reached similar conclusions regarding the comparable efficacy of pharmacological and psychological treatments but cautioned that medication trials are more likely to use a

credible active placebo, which may lead to smaller adjusted effect sizes (Pinquart, Duberstein, & Lyness, 2006).

Contrary to these findings, a recent review of studies comparing the efficacy of pharmacotherapy and psychotherapy concluded that SSRIs were *more* effective in treating MDD than was psychotherapy, and that treatment with TCAs and other antidepressants was similar to psychotherapy in efficacy (Cuijpers, van Straten, Andersson, & van Oppen, 2008). Although the difference was significant, the authors note that the contradiction to previous findings may not be of consequence, as the effect size was small (d = - 0.16) rendering its clinical significance debateable. Cuijpers et al.'s (2008) review was rigorous, including only studies that directly compared a pharmacotherapy to a psychotherapy group in an RCT. Comparisons included CBT (15), IPT (7), PST (5), other psychological treatments (12), SSRIs (15), TCAs (16), and other medications (6). While finding an efficacy advantage for SSRIs, the authors also reported that drop-out rates were lower in psychological treatments than in pharmacological treatments, which may have an impact on efficacy over the long-term. Cuijpers et al. (2008) conclude their review with a discussion of its limitations, including the fact that there may be important differences between psychotherapy and pharmacotherapy over the long-term that are not observed in short term outcomes studies. It is also worth noting that not all individuals are willing to take antidepressant medications given their potential side-effects (APA, 2010). Therefore, comparative trials that require participants to be willing to be randomized into a psychological or a pharmacological treatment may exclude all of those individuals who are not willing to be assigned to the pharmacological group (Hoffman et al., 1998; Huppert, Franklin, Foa,

& Davidson, 2002; Zoellner, Feeny, Cochran, & Pruitt, 2003). This issue of preference for treatment could cause the sample to be biased in favour of those who are more comfortable with pharmacological treatment.

With regards to comparison of mono- versus combined-therapies, Hollon, Jarret, Nierenberg, Thase, Trivedi, and Rush. (2005b) conducted a review and concluded that antidepressants alone are an effective treatment for MDD for as long as they are continued but do little to reduce risk once their use is terminated. They also concluded that IPT and CBT are as effective as medication, and that CBT in particular appears to reduce risk subsequent to treatment termination. They report that ongoing IPT or CBT appear to further reduce risk and that combining medication with IPT or CBT retains the specific benefits of each. It may also improve the probability of response over either treatment alone, especially in those who are chronically depressed.

It appears then, that with regards to comparative efficacy of psychological and pharmacological treatments there may not be a clear advantage for one over the other during the acute phase. Psychological interventions, particularly CBT and IPT however, may have more enduring effects and do not entail the withdrawal and increased risk of relapse associated with discontinuation of medication treatment. In considering studies that do find an advantage for one acute treatment over the other, it is important to discern the clinical significance of these differences, and to consider potential bias in the sample given patient preferences, and the selection bias introduced when participants must be willing to accept both psychological and pharmacological treatment in order to participate in a trial.

Other Treatments

The psychological and pharmacological treatments just reviewed are the most commonly employed, but not the only treatment approaches for depression. For some, self-help is an attractive option. However, there is little evidence and a lack of rigorous evaluation of self-help therapies, including self-administered computerized CBT (Kaltenthaler et al., 2006; Spek, Cuijpers, Nyklicek, Riper, Keyzer, & Pop, 2007), CBT-based bibliotherapy (Anderson, Bergstrom, Hollandare, Carlbring, Kaldo, & Ekselius, 2005), and other forms of guided self-help, either alone or as an addition to drug treatment (Mead et al., 2005; Salkovskis, Rimes, Stephenson, Sacks, & Scott, 2006).

Exercise is commonly recommended to combat depression, but is not proven to be an effective mono-therapy. One meta-analysis on exercise for depression (Lawlor & Hopker, 2001) found a large effect size of 1.1 against no treatment, and no difference when compared to CBT or sertraline. However, most studies included in this analysis had major methodological flaws, including poorly defined diagnosis of depression and use of non-clinical populations.

Finally, a well-researched treatment for the most severe cases of depression is electroconvulsive therapy (ECT). ECT is recommended as the first-line treatment for severe MDD in emergency situations, including patients who are not eating or drinking, or who exhibit depressive stupor, extreme distress, or suicidality (Anderson et al., 2008). A systematic review (UK ECT Review Group, 2003) reports that ECT is more effective in acute treatment than placebo (effect size .91) and drug treatment (effect size .80), with bilateral ECT being slightly more effective than unilateral (effect size 0.32) and higher dose being more effective than lower dose in the short term (effect size 0.58).

Unfortunately, both bilateral ECT and higher dosage are also associated with greater cognitive impairment in the short and possibly the long-term (UK ECT Review Group, 2003). Moreover, it is suggested that these impairments may be underestimated by clinicians (Rose, Fleischmann, Wykes, Leese, & Bindman, 2003). The use of ECT during the maintenance phase is not recommended (APA, 2010), as there is no evidence to support this application and, for those who improve following ECT, continuation treatment with medication is recommended to reduce relapse rates (Anderson et al., 2008), though evidence that this result is achieved is lacking.

Recall that the intent of this description of current evidence-based treatments for depression is to provide background on the interventions for which long-term post-treatment outcomes will be examined in the current study. The literature on acute phase treatment for depression is voluminous and is only described here briefly, to provide context for the outcomes of interest in the research questions to be described shortly.

The Current Study

The objective of the current thesis is to synthesize the literature on the post-treatment outcomes for adults who have received a range of treatments for depression. This is presented in two parts: Study 1 is an overview of reviews synthesizing the results of previous reviews that provide important information about long-term post-treatment outcomes. Study 2 is a systematic review of studies that include at least one year post-treatment follow-up and fit within the inclusion criteria laid out below. This thesis represents a unique contribution, in that previous reviews have examined (a) relapse and recurrence during *continuation and maintenance phase* treatments with second-generation antidepressants (Hansen et al., 2008; Zimmerman & Thongy, 2007), (b)

differential relapse rates for patients treated for a *first episode versus recurrent* depression (Kaymaz, van Os, Loonen, & Nolen, 2008), (c) relapse after discontinuation of *augmentation* treatment (Ross, 2008), (d) *discontinuation syndrome* which occurs when anti-depressant medication is stopped too abruptly (Tamam & Ozpoyraz, 2002), and (e) *paradoxical effects* during long-term treatment with anti-depressant drugs (Fava, 2002). They have not, however, broadly reviewed *long-term post-treatment* outcomes across *psychological and pharmacological* treatments. As we will see in the overview of reviews, even the existing reviews that do provide analysis of long-term outcomes often provide only limited analysis on these outcomes and focus primarily on outcomes immediately post-treatment. At the time of writing, the author is aware that protocols have been published for reviews that are underway of interventions for (a) early recognition of symptoms and (b) prevention of recurrence (Allen, Hetrick, Yap, & Simmons, 2008; Morriss, Bolton, McCarthy, Marshal, Williamson, & Jones, 2004) as well as (c) continuation and maintenance treatments for depression (Wilkinson & Izmeth, 2007). However, no protocol exists for a review like the current study, to the best of the author's knowledge.

The goal of these studies is to synthesize the existing evidence-base on post-treatment outcomes for adults treated for depression across treatments and study designs. These studies will also contribute to the translation of this synthesis into a format that is accessible to both practitioners and consumers as a part of the work of the Mobilizing Minds Research Group's knowledge translation project, funded by CIHR.

As is the case throughout the health and social sciences, the existing body of research on long-term post-treatment outcomes for treatment for depression is

heterogeneous (Petticrew & Roberts, 2009). Therefore, the current thesis is multi-layered, whereby (i) the reviews' and studies' descriptions are organized into logical categories such as type of intervention and study design, (ii) the findings are discussed within these categories, and (iii) the findings are synthesized in narrative format across categories (including all of the studies reviewed), taking into account biases, heterogeneous designs, and other measurement issues.

Previous reviews, including systematic meta-analyses, and systematic and non-systematic narrative reviews are synthesized in Study 1: Overview of Reviews. Many of these previous reviews focus mainly on immediate effects of acute treatment and report briefly on long-term follow-up data as a secondary focus. Reviews were considered for inclusion if they reported post-treatment outcomes for included interventions in comparison to a no-treatment control group, treatment-as-usual, wait-list control, placebo control, or an active-treatment comparison group. Reviews of continuation or maintenance studies in which outcomes were reported both for the treatment group and for the discontinued, or withdrawn to placebo group were also included.

Study 2: Narrative Systematic Review is a synthesis of long-term follow-up of a variety of treatment types, and study designs. Study 2 includes primarily post-treatment follow-up of acute or continuation treatment RCTs, in which both intervention and comparison groups are followed, for at least one year, after the treatment period has ended. Often these follow-ups include only those who responded to the acute treatment, so we know less about individuals who do not achieve significant response within the typical duration of a treatment trial (about 8-16 weeks). These follow-up studies include comparison to wait-list control, treatment as usual, or other active treatment groups.

One study design that is included in Study 1 but not Study 2 is the maintenance or continuation study design often used to examine the effects of ongoing medication treatment in comparison to medication withdrawal. These studies inform our knowledge about what happens after treatment in that they provide relapse rates for those who discontinue their medication treatment. This type of study has been the focus of high quality reviews including a large number of studies with large sample sizes (Kaymaz et al., 2008; Viguera, Baldessarini, & Friedberg, 1998). These reviews have reached clear and consistent conclusions about the outcomes of continuation or maintenance antidepressant treatment versus discontinuation. For these reasons, previous reviews that synthesize antidepressant continuation and maintenance studies are reviewed in Study 1, but these studies are not included in the primary reports of long-term post-treatment follow-up reviewed in Study 2. The post-treatment follow-up of acute and continuation trials included in Study 2, including at least one year of post-treatment follow-up on *all* groups, have not been previously reviewed in this way.

The questions addressed in this overview of reviews, and narrative systematic review include:

Question 1. What are the rates of relapse over time for adults who have completed psychological, pharmacological, or combined/sequential treatment for depression?

Question 2. When treatment gains are measured on a continuous measure, to what degree are treatment gains maintained over time during long-term follow-up?

Study 1: Overview of Reviews - Method

Criteria for Considering Reviews for Inclusion

Types of reviews. Systematic and non-systematic reviews including meta-analyses and narrative reviews were included. Reviews eligible for inclusion were those of randomised controlled trials (RCTs) with planned post-treatment follow-up including naturalistic follow-ups, maintenance, continuation, and discontinuation studies. Only reviews published in English were included.

Types of participants. Reviews were included that focused on adults (aged 18 years or older) who had a diagnosis of unipolar depression and who had completed treatment for this disorder. Reviews including studies of participants with co-morbid anxiety were included given the high rate of co-morbidity of anxiety and depression, provided that participants were not also receiving a separate intervention for their anxiety. Reviews focusing exclusively on participants with unipolar depression and co-morbid disorders such as substance abuse disorders, personality disorders, psychosis, chronic pain, and medical diagnoses such as cancer were excluded. Reviews focusing on participants with bipolar affective disorder or dysthymia were also excluded. Reviews that looked only at specific age groups such as the elderly were excluded.

Types of interventions. Interventions provided by a health professional for the treatment of unipolar depression, and established as an effective acute treatment, were considered. Included interventions were: first generation antidepressant drugs including tricyclics (TCAs) and monoamine oxidase inhibitors (MAOIs); second generation antidepressants including selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), other drugs that selectively affect the

activity of serotonin, norepinephrine, and dopamine; cognitive behavioural therapy (CBT); behavioural activation (BA); interpersonal psychotherapy (IPT) and psychodynamic therapies. Interventions used primarily in cases of treatment resistant depression, such as electroconvulsive therapy (ECT), or trans-cranial magnetic stimulation were not included, as treatment resistant depression is beyond the scope of this review. Interventions such as exercise, yoga, herbal remedies, computer or internet based interventions, and self-administered interventions such as bibliotherapy were excluded.

Outcomes. Clinical outcomes, including relapse and maintenance of gains were examined.

Search Methods for Identification of Reviews

Terms searched included: depression, treatment, follow-up, relapse, maintenance, continuation, discontinuation, and "return of symptoms." The following databases were searched: PsychINFO, PubMed, the Cochrane Library including Systematic Reviews and the Controlled Trials Register (CCTR), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts and Reviews of Effects (DARE), and Cochrane Methodology Register (CMR), CINAHL, Social Sciences Citation Index, SCOPUS, and Web of Science. Selected journals deemed relevant in the review were hand-searched. Reference lists of the selected reviews were also searched. Databases were searched from their earliest date (with PsychINFO reach back as far as the 1800s) up to October 2011. Hand-searching was completed in August 2012.

Data Collection and Analysis

Selection of reviews. The titles and abstracts obtained in the searches were screened against the inclusion and exclusion criteria to identify potentially relevant reviews. This was followed by a more detailed screening of the full papers identified as possibly relevant in the initial screening. Selection of reviews was carried out by two reviewers for reliability (the author and another member of the Mobilizing Minds research team – a graduate student in psychology who was trained by the author and the author's research supervisor on the application of the inclusion and exclusion criteria). Any disagreements (there were very few) were resolved by discussion, and when necessary through arbitration with a third team member (the author's research supervisor).

Data extraction. One author extracted the data from the identified articles. This included key characteristics of the review, quality of review methodology, and relevant results. This data can be found in Tables 1 and 2.

Assessment of methodological quality of included reviews. Evaluation of the methodological quality of each review was carried out according to the Oxman criteria (Oxman & Guyatt, 1991) using an application developed by Posadzki, Watson, & Ernst (2012). This application of the Oxman criteria includes 5 criteria including reporting of search methods, comprehensiveness of search, reporting of inclusion criteria, avoidance of bias, and support for conclusions. The sum of a review's score on these criteria suggest whether the review has significant, minimal, or no methodological flaws. Specifically, each question is scored as 1, 0, or -1. For each item respectively, 1 indicates that: (a) the review states the databases used, date of most recent searches, and

some mention of search terms; (b) the review searches at least 2 databases and looks at other sources; (c) the review states the criteria used for deciding which studies to include in the overview; (d) the review reports how many studies were identified by searches, numbers excluded, and appropriate reasons for excluding them; (e) the conclusions made by the author(s) are supported by the data and/or analysis reported in the review. For each item, 0 indicates that the criteria were partially fulfilled. For each item, -1 indicates that none of the criteria were fulfilled. A total score of 0 or below means the review has major flaws, 1-2 indicates minor flaws, and 3-5 indicates minimal or no flaws. The reproducibility and construct validity of quality measures such as the Oxman criteria require further research support. Their use, however, is recommended as the best measure for quality assessment currently available (Cochrane, 2011; Shea et al., 2007).

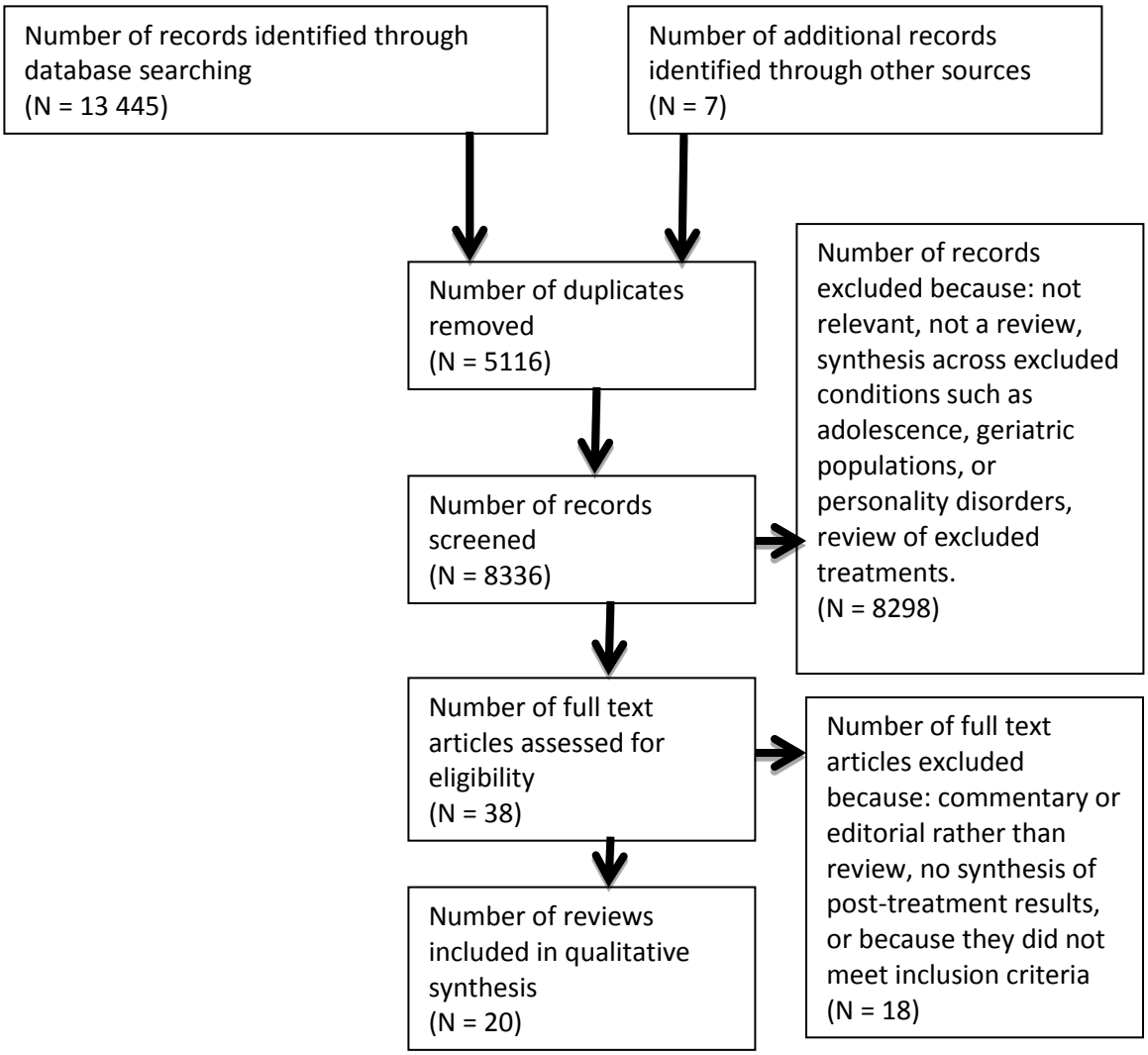
Data synthesis. A narrative summary of the characteristics and results of the reviews for each of the primary outcomes is provided. These are presented using tables and discussion in the text.

Study 1 Results

A total of 38 papers were identified for consideration. Twenty of these met the criteria for inclusion. Of these, 12 were systematic meta-analyses, 3 were systematic narrative reviews, and 5 were non-systematic narrative reviews. The remaining 18 papers were excluded because they were commentaries or editorials rather than reviews, they synthesized results across groups that were not included in the current overview such as adolescents and geriatric populations, they included studies focused on personality disorders, or they reviewed treatments not included in this overview. Figure 2 gives a flow diagram outlining the selection process and review numbers at each stage.

Note that, as is common in the review process, the greatest number of exclusions were due to removal of duplicates, and titles and abstracts that were identified in the search because they included key words, but were easily identified as irrelevant to the current review.

Figure 2. Flowchart of eligibility assessment and inclusion (PRISMA, Moher et al., 2009)



Description of Included Reviews

A summary of the characteristics and key findings of the 20 included reviews appears in Table 1 where they are organized by type: systematic meta-analysis, and systematic or non-systematic narrative review. Table 2 includes quality review of each review. All reviews (n = 20) examine long-term, post-treatment follow-up of adults treated for depression. The number of included studies ranged from 3 - 32 in each review. These reviews varied in their scope, including interventions, comparisons, and outcomes of interest. Eleven reviews included 1st and 2nd generation antidepressant medications (ADM), 2 included only 1st generation ADM, 1 included only 2nd generation ADM, 15 included CBT, 7 included BA, 6 included IPT, 3 included psychodynamic therapies, 3 included other psychological treatments, and 3 included combined treatment. Fourteen reviews examined studies that compared one intervention to another, including 7 that compared psychotherapy with ADM and 7 that compared one psychotherapy to another.

Relapse was reported in 15 reviews and maintenance of gains was reported in 6 reviews. Moderating variables were reported in 3 reviews (Kaymaz et al., 2008; Viguerra et al., 1998; Vittengle, Clark, Dunn, & Jarrett, 2007).

Table 1. Key characteristics and results of included reviews

Study	Number of studies (number of participants) [designs included]	Treatments	Outcomes reported	Duration of treatment (Duration of follow-up)	Results*
Systematic Review and Meta-Analysis					
Cuijpers, 2007	9 (154) [RCTs]	Behavioural Activation (BA), cognitive behavioural therapy (CBT), other psycho-therapies	Maintenance of gains*	6-20 sessions (1-24 months)	No significant change from post-test to follow-up for BA (*d* = 0.18 at 1-3 months follow-up, 0.03 at 4-6 months follow-up, both *ns*). Change from post-test to follow-up not reported for other treatment groups. Only the significance of the difference between BA and other treatment groups is reported. Outcomes are not significantly different for BA than CBT at follow-up (*d* = 0.02 at 1-3 months, -0.13 at 4-6 months, both *ns*).
Cuijpers, 2009	5 (359) [RCTs]	1st or 2nd generation antidepressant medication (ADM) + psychotherapy	Maintenance of gains	6-24 sessions (3-12 months)	No difference between combined medication and psychological treatment versus psychological treatment only, at follow-up** (*d* = -0.15 at 3-6 months, 0.00 at 12 months, both *ns*).

De Maat, 2006	6 (546) [RCTs]	1st or 2nd generation ADM, CBT	Relapse, recurrence*	8-20 weeks (1-2 years)	At follow-up relapse was higher in ADM treatment conditions (57%) than in CBT treatment conditions (27%)
Driessen, 2010	6 (150) [not specified – case studies excluded]	Short term psycho-dynamic psychotherapy (STPP), CBT, BA, ADM	Maintenance of gains	3-80 sessions (3,6, and 12 months)	There was no significant change from post-treatment to follow-up (3 months: $d = 0.03$, 6 months: $d = 0.21$, 12 months: $d = 0.02$, all *ns*). There was no difference between follow-up outcomes for STPP compared to other psychotherapies, though the trend favoured other psychotherapies ($d = -0.29$, $p < 0.10$)
Ekers, 2008	8‡ (271) [RCTs]	BA, CBT	Relapse, recurrence	5-24 sessions (Mean = 4 months)	BA is as effective as CBT in preventing relapse over short-term follow-up (Standardized mean difference = 0.25, $p < 0.30$). Note the difference favours BA but is non-significant (percent relapse not reported)
Friedman, 2004	17 (455) [RCTs]	Combined treatment including 1st or 2nd generation ADMs and one of CBT or IPT	Relapse, recurrence	Not consistently reported (not consistently reported)	Combined treatment leads to lower relapse rates than medication alone ($d = 0.68$, $\chi^2 = 4.06$), especially when treatment is sequential: medication followed by CBT (percent relapse not reported)
Gloaguen, 1998	8 (261) [RCTs]	1st generation ADM, CBT	Relapse, recurrence	4-79 weeks (1-2 years)	Responders withdrawn from CBT are about half as likely to relapse over follow-up as responders withdrawn from medication (30% versus 60% respectively)

Imel, 2008	11 (602) [RCTs]	ADM, psychotherapy (neither are reported by type)	Effect size at follow-up	10-52 weeks (Mean = 15 months)	When follow-up occurred post-treatment for both psychotherapy and pharmacotherapy, there was a significant advantage for psychotherapy at follow-up ($\lambda(13) = 0.37$, $p < 0.01$). When ADM was continued, over follow-up, but psychotherapy was terminated after the acute phase, there was no difference at follow-up ($\lambda(10) = -0.03 p > 0.50$)
Kaymaz, 2008	30 (4890) [RCTs]	1st or 2nd generation ADM	Relapse, recurrence	3 – 14 weeks acute treatment 0 – 192 weeks contin-uation (6-36 months)	Withdrawal from continuation antidepressant treatment leads to increased relapse, especially for those with recurrent depression (Odds Ratio = 0.25 $p < 0.01$ at follow-up; percent relapse not reported)
Leichsen-ring, 2001	3 (416) [RCTs]	STPP, CBT, BA	Relapse	8 – 20 sessions (1-2 years)	At follow-up, there is no significant difference in percent relapsing between STPP (56%) and CBT/BA (46%)

Spielmans, 2011	5 (527) [RCTs]	2nd generation ADM, bona fide psychological treatments[ł] - CBT, BA, cognitive behavioural analysis system of psychotherapy (CBASP), interpersonal psychotherapy (IPT); non-bonna fide psychological treatments	Treatment effect size	8-26 weeks (18-40 weeks)	In the analysis of acute phase results, bona fide psychological treatments were equivalent to ADM. Other treatment (usually studies with therapists with limited or unspecified training) were not. At follow-up, bona-fide psychotherapies were superior to discontinued medication by a small effect size (d = 0.29, $p < 0.01$) The authors point out that previous reviews have often grouped bona fide and non-bona fide psychological treatments together, suggesting lower effectiveness of psychotherapy.
Vittengle, 2007	28 (1880) [RCTs]	CBT	Relapse, recurrence	Not reported (20-332 weeks)	Of responders to Acute-CBT, 29% relapse within one year and 54% within two years. Continuation-CBT reduced relapse-recurrence further with C-CBT resulting in 40% relapse over three years as compared to73% for the non-active control groups (groups who received no continuation treatment, and no treatment over follow-up).

Systematic Narrative Reviews					
Hensley, 2004	5, only 3 reported data appropriate for re-analysis. For these 3 studies n = 234 [RCTs]	1st generation ADM, CBT	Relapse recurrence	8-20 weeks (1-2 years)	CBT leads to better long-term outcomes than antidepressants alone (at follow-up, fewer people in CBT group relapsed[53-60%] than in antidepressant group [80-93%], $p < 0.05$). Outcomes may be even stronger in studies that include maintenance phase CBT (This conclusion was based on only one study).
Hollon & Ponniah, 2010	7 (not reported) [RCTs]	CBT, Mindfulness Based Cognitive Therapy (MBCT), IPT, Psycho-dynamic treatment	Relapse, recurrence	6 – 46 sessions (1-2 years)	(Descriptive results provided, no report on indicators of central tendency or range/percentage) CBT has relapse-prevention effects post-treatment that are similar to continued medication treatment, and superior to discontinuation of medication treatment MBCT has prophylactic effects post-treatment, especially for those with 3+ previous episodes IPT shows no evidence of prophylaxis post-treatment Psychodynamic treatment shows no post-treatment advantage over comparison groups.

Viguera, 1998	27 (3037) [RCTs]	1st or 2nd generation ADM	Relapse, recurrence	0 - 48 months, *M* = 5.78 mo. (5 - 66 months, *M* =16.6 mo)	Discontinuation is associated with higher relapse rates (6%/month) than continued medication treatment (2%/month)
Non-Systematic Narrative Reviews					
DeRubeis, 1998	CBT: 4 (not reported) BA: 2 (not reported) IPT: 1 (not reported) [RCTs]	ADM, CBT, BA, IPT	Relapse, recurrence, maintenance of gains	CBT: 3-4 months (1 year) BA: Not reported (3-6 months) IPT: Not reported (18 months)	Responders withdrawn from CBT are about half as likely to relapse over follow-up as responders withdrawn from medication (26% versus 64% respectively) BA shows similar gains to CBT and superior gains to non-specific psychotherapy. These gains are maintained, or slightly decreased at follow-up Of the initial treatment group, 26% of IPT patients recovered and stayed well over follow-up, compared to 30% for CBT, 19% for ADM, and 20% for placebo. NOTE: these results are based only on the Treatment of Depression Collaborative Research Program (TDCRP) study (Shea et al., 1992)

Hollon, 2006	Not reported (not reported) [not consistently reported]	CBT, BA, IPT	Relapse, recurrence	Not reported (not reported)	(Narrative summaries and study specific statistics were provided – no index of central tendency or range/%) CBT has superior relapse-prevention effects when compared to active and non-active controls BA produces similar relapse-prevention effects to CBT/CT post-treatment No evidence that IPT prevents relapse post-treatment
Paykel, 2001	32 (Not reported) [RCTs]	1st or 2nd generation ADM	Relapse, recurrence	2-3 months (4 months – 2+ years)	Relapse rates are approximately twice as high for those who discontinue (range = 16 – 100%) versus those who continue treatment (Range = 0 – 32%) Differential rates of recurrence are less marked during longer-term maintenance treatment (discontinuation = 24 – 87% relapse, maintenance = 6 – 54% relapse)
Paykel, 2007	7 (802) [RCTs]	ADM (not specified), CBT, MBCT	Relapse	Not consistent-ly reported (8 months – years)	Compared to medication withdrawal, CBT for residual symptoms reduces relapse and recurrence (no quantitative synthesis provided). This reduction is not due to medication adherence as it is found regardless of presence or absence of ongoing medication, and regardless of level of adherence when this is measured. The authors suggest that CBT provides a coping framework that medication treatment does not.

Segal, 2002	12 (1247) [RCTs]	1st or 2nd generation ADM, CBT, IPT	Relapse, recurrence	8-25 weeks (6 months – 6 years)	Findings are mixed for combined treatment that is concurrent (2 treatments at the same time) or crossover (one follows discontinuation of another). Sequential treatment, when one is added to another, was reported in only 1 study in which combined treatment was superior to either monotherapy, however this is insufficient evidence to form conclusions. Combined treatment may be more successful with some populations, such as those with more severe depression, but more evidence is required to support this.

Note . ns – not significant, RCT – randomized controlled trial, ADM – antidepressant medication, BA – behavioural activation, CBASP – cognitive behavioural analysis system of psychotherapy, CBT – cognitive behavioural therapy, CT – cognitive therapy, IPT – interpersonal psychotherapy, MBCT – mindfulness based cognitive therapy, STPP – short term psychodynamic therapy.

* When reviews report on continuous measures – referred to here as maintenance of gains, effect sizes are reported. When reviews report dichotomous measures such as relapse or recurrence, percent relapse/recurrence is reported.

** All follow-up results refer to post-treatment follow-up, in which treatment was not continued during the follow-up period, unless otherwise stated. For example, cADM refers to continuation antidepressant medication and indicates that medication treatment was continued throughout the follow-up period.

ǂ Criteria applied by Spielmans (2011) to identify bona fide psychological treatments involve adequately trained therapists, face to face contact with a therapist, and treatment based on psychologically valid components.

ǂWhen a review includes studies of immediate effects of acute outcomes as well as studies of long-term post-treatment effects, only the number of included long-term post-treatment studies is reported in this table. For example, Ekers (2008) include 17 articles in their review, however, only 8 of these include follow-up data, so Table 1 indicates that 8 articles were reviewed as it is only the post-treatment outcomes that are relevant to this overview.

Methodological Quality of Included Reviews

The quality of each review is reported in Table 2. A 5-item operationalization of the Oxman (1991) criteria, as adapted by Posadzki et al. (2012), was used to rate the quality of the systematic reviews included in this overview. This included 12 meta-analyses and 3 systematic narrative reviews. Scores ranged from 2-5 of a total score of 5. Fourteen reviews scored 3-5 indicating that they had minimal or no flaws, 1 review scored 1-2 suggesting minor flaws, and no reviews scored 0 or lower suggesting major flaws. Ten reviews reported their search methods, 13 conducted comprehensive searches, 15 reported their inclusion criteria, 7 avoided selection bias, and in 15 the conclusions of the authors were supported by the reported data.

Quality assessment tools such as the Oxman criteria only account for the quality of specific elements of the design and comprehensive reporting of a given review. The comments column in Table 2 includes information not accounted for in this tool, but that affects the quality of the review itself, or the results that it reports.

Table 2. Quality ratings for included systematic reviews

Study (year)	Search Methods (a)	Search Compre-hensive (b)	Inclusion Criteria (c)	Selection Bias Avoided (d)	Conclusions Supported (e)	Sum	Comments
Systematic Meta-Analyses							
Cuijpers, 2007	1	1	1	0	1	4	Exclusions and reasons for exclusions not reported. Many included studies are not of optimal quality.
Cuijpers, 2009	1	1	1	0	1	4	Exclusions and reasons for exclusions not reported. Too few studies to report sub-group analyses by type of therapy.
De Maat, 2006	1	0	1	1	1	4	Only uses two databases, does not review other sources.
Driessen, 2010	0	1	1	1	1	4	Date of searches not reported. Many of the included studies were of poor quality.
Ekers, 2008	1	1	1	1	1	5	Average follow-up time only 4 months.
Friedman, 2004	1	1	1	-1	1	3	Number of hits, exclusions, and reasons for exclusion not reported.
Gloaguen, 1998	0	1	1	1	1	4	Search terms not reported.

Imel, 2008	1	1	1	0	1	4	Reasons for exclusion not reported. Very general description of treatments, so that results do not speak to specific treatments.
Kaymaz, 2008	0	1	1	1	1	4	Search dates not reported.
Leichsenring, 2001	1	1	1	1	1	5	Only 3 studies included follow-up data. Author advises against generalizing beyond specific variants of psychodynamic treatment reviewed, but does not describe these variants.
Spielmans, 2011	1	1	1	-1	1	3	Number of hits, exclusions, and reasons for exclusion not reported. Too few studies to report outcomes by specific treatment.
Vittengle, 2007	1	1	1	-1	1	3	Number of hits, exclusions, and reasons for exclusion not reported.

Systematic Narrative Reviews							
Hensley, 2004	0	1	1	1	1	4	Did not report date of searches. Very narrow inclusion criteria led to inclusion of only 3 studies. Defined medication discontinuation as equivalent to no treatment control – this is not an established practice. Defined pursuit of treatment during follow-up as relapse – this inflates relapse rates because some people pursue "booster" treatment for sub-threshold symptoms that do not constitute a relapse.
Hollon & Ponniah, 2010	0	1	1	-1	1	3	Search terms not listed. Number of hits, exclusions, and reasons for exclusion not reported. No description of design or quantitative results of included studies.
Viguera, 1998	1	0	1	-1	1	2	Searched only 1 database. Number of hits, exclusions, and reasons for exclusion not reported.

Note. This is an operationalization of the Oxman criteria, Oxman & Guyatt, 1991; as adapted by Posadzki, Watson, & Ernst, 2012. Each question is scored as 1, 0, or -1. For each item respectively, 1 indicates that: (a) the review states the databases used, date of most recent searches, and some mention of search terms; (b) the review searches at least 2 databases and looks at other sources; (c) the review states the criteria used for deciding which studies to include in the overview; (d) the review reports how many studies were identified by searches, numbers excluded, and appropriate reasons for excluding them; (e) the conclusions made by the author(s) are

supported by the data and/or analysis reported in the review. 0 indicates that the criteria were partially fulfilled. -1 indicates that none of the criteria were fulfilled. A total score of 0 or below means the review has major flaws, 1-2 minor flaws, 3-5 minimal or no flaws.

Key Findings of the Reviews

Antidepressant medications. Three reviews examined the post-treatment effects of antidepressant medications. Although TCAs and MAOIs (first generation antidepressants) are no longer commonly prescribed, with second generation antidepressants such as SSRIs and SNRIs taking their place, the first generation antidepressants more commonly appear in the follow-up literature and are included here. Of note, one review of discontinuation studies (Viguera et al.'s, 1998) that primarily included studies of TCAs, did make comparisons between TCAs, MAOIs, and SSRIs and found no differences between TCAs and MAOIs, or between these older antidepressants and SSRIs. Kaymaz et al.'s (2008) more recent review included second generation antidepressant (SSRI) and first generation antidepressant (TCA) studies and found no difference between the two in comparison to placebo withdrawal. Neither Kaymaz et al. and Viguera et al. found any studies that directly compared SSRIs to TCAs and placebo, so comparisons are indirect rather than direct.

In a review of 27 discontinuation studies Viguera et al. (1998) report relapse rates per month (simply calculating the % of participants becoming depressed per month for each study +/- standard error) They reported average relapse rates of 6.24 +/- 5.34 % per month for those who terminate treatment, versus 1.85 +/- 1.51 % per month for those who continue medication treatment. Viguera et al. conducted survival analysis based on the data from 19 studies with an average follow-up of 19.4 +/- 14.1 (Range = 6-60) months and found that time to 50% chance of relapse was significantly different for those who stopped treatment (14.2 +/- 0.5 months, n = 952) compared to those who continued treatment (48.0 +/- 4.7 months, n = 1663). With regards to possible

mediators, Viguera et al. (1998) report that longer treatment was not associated with lower relapse rates post-treatment, nor was gradual discontinuation (ranging from 2-12 weeks).

Paykel (2001) conducted a non-systematic narrative review of continuation and maintenance therapy and reached similar conclusions concerning relapse. They reported on the percentage of participants who relapsed in discontinuation arms, as compared to continuation or maintenance arms and found that relapse rates were at least double in the discontinuation arm than in the continuation arm of most studies. The authors also examined maintenance studies, which involve a longer period (2 years or more) of ongoing medication treatment than continuation studies (less than 2 years), and reported that maintenance treatment showed a similar advantage over discontinuation, but was less marked than the advantage found for continuation treatment. This suggests that the relapse-prevention effects of ongoing antidepressant medication may diminish over time.

Most recently, Kaymaz et al. (2008) conducted a meta-analysis examining both relapse/recurrence rates and potential mediators. They included only continuation studies with a comparison group which was randomly assigned to medication withdrawal and placed on placebo for the duration of the follow-up period. The 30 RCTs included in their review provided information on 4890 patients, with 2749 randomized to receive maintenance antidepressant treatment and 2141 assigned to placebo. Those who were withdrawn to placebo did have a greater chance of relapse than those who were maintained on antidepressants (odds ratio (OR) = 0.3, 95% CI 0.25-0.35, $p < 0.01$). Of the mediating variables they examined, Kaymaz et al. found

that duration of continuation treatment prior to withdrawal (3, 6, 9, or 12 months), and abrupt versus gradual withdrawal (less than one week versus more than one week) did not appear to affect rates of relapse after withdrawal. History of depression was a significant mediator. Patients with a history of recurrent depressive episodes were at greater risk for relapse after withdrawal than those who had had only one episode (OR = 0.12, 95% CI 0.06-0.26, $p < 0.01$). In addition, those who had recurrent episodes were at greater risk for relapse if withdrawn abruptly (OR = 0.32, 95% CI 0.27-0.38) than if withdrawn gradually (OR = 0.11, 95% CI 0.06-0.21). This finding should be interpreted with caution given the definition of abrupt (less than one week) versus gradual (more than one week – range = 2-12 weeks) withdrawal. The authors conclude that maintenance treatment has relapse-prevention effects, but that the benefits are less for those who have had recurrent episodes.

These reviews clearly converge on the conclusion that within 2 years follow-up, for antidepressant medications, relapse rates are higher during post-treatment follow-up (approximately 60%) than when treatment is continued (approximately 30%). What is less clear is the magnitude of this difference over time, and the possibility that treatment-related variables such as duration of treatment and gradual versus abrupt withdrawal can moderate these rates.

Cognitive behavioural therapy. Four reviews reported on the post-treatment outcomes for adults who had received CBT for depression. Gloaguen et al. (1998) conducted a hallmark review on the effectiveness of CBT, in which they reported that only 30% of CBT patients relapsed over follow-up as compared to 60% of patients treated with antidepressants. Gloaguen et al.'s review included 8 studies, 5 of which

demonstrated a relapse-prevention effect for CBT. At about the same time, DeRubeis and Crits-Christoph (1998) published their review of four large scale clinical trials which reported that CBT patients had a lower relapse rate over one year follow-up than did pharmacotherapy patients (26% versus 64% respectively). Their findings were, not surprisingly, congruent with the Gloaguen et al. review, as all 4 of the studies reviewed by DeRubeis and Crits-Christoph were also included in the Gloaguen et al. review.

More recently, in a high quality systematic review and meta-analysis of relapse rates following CBT, Vittengle et al., (2007) reviewed 28 studies, including 1880 participants, and reported on rates of relapse and recurrence after discontinuation of acute-phase CBT (A-CBT) and continuation-phase CBT (C-CBT). A-CBT is defined as CBT "applied during a major depressive episode with the goal of reducing depressive symptoms and producing initial remission" (p. 475). C-CBT is defined as CBT "applied to sustain remission of a major depressive episode and reduce the probability of relapse-recurrence" (p. 475). The authors do not further define the two phases of CBT with regards to differential intensity or frequency. They compared the rates of relapse-recurrence after A-CBT and C-CBT to other active treatments as well as to non-active control conditions.

Vittengle et al., (2007) found that in the A-CBT group, 29% relapsed within one year and 54% within two years. Proportions varied significantly between studies so they tested for moderating variables. They found higher relapse rates in studies with longer follow-up periods, studies reporting survival analysis instead of simple proportions, studies reporting on CBT therapists' competence (conducting CBT skillfully; e.g., scores > 39 on the Cognitive Therapy Scale; Young & Beck, 1988), and studies using

major depressive episode (MDE) diagnostic criteria rather than cut-points on scales such as the HDRS in relapse-recurrence definitions. They found lower relapse rates in studies that reported CBT therapists adherence (conducting treatment identifiable as CBT and consistent with the cognitive model of depression; e.g., measured with the Collaborative Study Psychotherapy Rating Scale; Hollon et al., 1988), studies with gaps in time in the follow-up assessment, and studies using instrument cut-point rather than MDE diagnosis in their relapse-recurrence definitions. Note that the authors do not report on the difference in relapse rates based on variation on these variables, for example, we do not know whether assessment of greater therapist competence or adherence was associated with lower relapse rates, only that relapse rates varied based on whether or not these variables were reported. It would have provided a valuable addition to this review had the authors reported on the association of relapse with variation on these moderating variables, and not just on the presence of their measurement.

Vittengle et al. (2007) found that for individuals who had achieved remission in acute phase treatments, C-CBT reduced relapse significantly compared to non-active controls. C-CBT is delivered in many ways and can involve weekly sessions over a few months, either individually or in a groups setting, or it can extend for up to a year, often including monthly sessions. Vittengle et al. synthesize results across C-CBT studies, but do not describe the duration or frequency of sessions in these C-CBT studies. In their analysis, C-CBT resulted in 40% relapse over three years of post-treatment follow-up as compared to 73% for groups not receiving C-CBT. C-CBT also reduced relapse-recurrence compared to active controls (including continuing medication and treatment as usual). Over an average of 114 weeks of post-treatment follow-up, relapse rates were

42% in the C-CT group and 61% in the active control groups. In sum, Vittengle et al.'s (2007) review concludes that A-CBT and C-CBT have greater relapse prevention effects than discontinued antidepressant medication, but that about half of A-CBT responders and about 40% of C-CBT responders will still experience relapse or recurrence within two to three years post-treatment.

In a systematic narrative review, Hollon and Ponniah (2010) synthesized the literature on efficacious psychological treatments for depression, focusing mainly on the much larger body of literature on immediate effects of acute treatment (125 studies were included). They also reviewed the smaller number of long-term follow-up studies that met their inclusion criteria (only 7 studies). The review is systematic, which is a strength, but it is less informative than it could be in that the authors do not provide descriptions of the designs or quantify the results of the included studies.

Hollon and Ponniah (2010) report that patients who are treated with CBT to remission are less likely to relapse after termination of treatment than are patients who were treated to remission and then withdrawn from antidepressant medication. They also report that patients who receive CBT are no more likely to have a relapse or recurrence after discontinuation than patients who are maintained on antidepressant medication throughout follow-up. They suggest that mindfulness based cognitive therapy (MBCT), a specific, manualized variation of continuation CBT, is efficacious in the prevention of relapse/recurrence. It appears to be most effective with patients who have had three or more prior episodes of depression and it is suggested that the mechanism of action may be different from that in traditional CBT. Hollon and Ponniah appear to base this statement on the findings of one multi-site trial with a one-year post-treatment follow-

up. This result is promising, but should be interpreted with caution pending replication. Their review supports the conclusions of previous reviews, and the seven long-term follow-up studies they review have all been included in one of the previous reviews included in this overview (DeRubeis & Crits-Christoph, 1998; Gloaguen et al., 1998, Hollon et al., 2006; Vittengle et al., 2007).

These reviews all find significant relapse prevention effects for CBT. Promising results are also reported for continuation phase CBT which appears to offer an additional relapse-prevention advantage when delivered after acute phase CBT. Similar findings are reported for Mindfulness based cognitive therapy which is typically delivered as a maintenance phase treatment. These findings for improved relapse-prevention with continuation and maintenance variants of CBT are based on a smaller evidence base that will benefit from replication.

Direct Comparison: Antidepressant Medication and CBT

Six reviews include only studies that *directly* compare the post-treatment outcomes for antidepressant medication and psychological treatment. Hensley, Nadiga, and Uhlenhuth (2004) conducted a systematic review of studies directly comparing the long-term post-treatment effects of CBT versus antidepressant medications. In their review of the literature published between 1966 - 2003, they were surprised to find only 5 studies that met their inclusion criteria, only three of which provided sufficient data for their proposed intent to treat (ITT) analysis.

Hensley et al. (2004) found that the way long-term outcomes are typically examined is by follow-up of only those patients who responded to acute treatment. They

used intent-to-treat (ITT) analysis in their review in an attempt to include all patients initially randomly assigned to either treatment or control groups, taking account of dropouts by carrying forward their last available observation, and by considering dropouts to be treatment failures. This decision was based on the assertion that most dropouts do so due to non-response or adverse events (Rickels, Noyes, Robinson, Schweizer, & Uhlenhuth, 1994, p.384). Treatment failure was distinguished from relapse so as not to artificially inflate relapse rates by counting as relapsers, those who had never reached remission in the first place. Hensley et al. (2004) suggest that ITT analysis prevents the differential sieve effect, in which patients who respond to acute treatment are necessarily those who benefit from it, and following up only with these treatment responders serves to narrow the contrast between the treatment and control groups. With regard to control groups, Hensley et al. (2004) deemed it essential for studies to include long-term follow-up of the control group in order to control for spontaneous improvement over longer periods. Control groups in their included studies were discontinued antidepressant medication arms (5 studies) and pill placebo plus clinical management (1 study). Note that one study had both of these conditions, thus the total number of control groups totals 6 even though only 5 studies were included. Hensley et al. (2004) also required at least 6 months of follow-up data as less than this was deemed not clinically significant. Finally they excluded follow-up studies that looked at only cross-sectional measurement at long time intervals, as this would miss occurrences of relapse and additional treatment in the interim period and would thus under report on relapse rates.

Hensley et al. (2004) assigned patients one of two outcomes: (1) achieved and maintained remission, or (2) failed to achieve or maintain remission. Patients who received additional treatment during the follow-up period were placed in the latter category. Hensley et al. (2004) found that in 4 of the 5 studies included in their review, those patients who received CBT had better outcomes at long-term follow-up (53-78% had not reached, or had not maintained remission over 15-27 months) than those who received medication only (80-93% had not reached or had not maintained remission over 15-27 months). Three of these 4 studies provided sufficient data for ITT analysis.

The authors note that their ITT analysis did not produce different conclusions than the original study authors' analysis. However, this refers to only 3 studies, which is too few to conclude that ITT is not a useful approach. Hensley et al. (2004) conclude that, on balance, the evidence favours a longer-term effect for CBT over tricyclic (first generation) antidepressants alone. Recall that Viguera et al. (1998) and Kaymaz et al (2008) noted that there is likely no difference in post-treatment outcomes for newer and older generation antidepressants, but were unable to find examples of direct comparisons. It is possible, but not certain, that this finding would generalize to the second generation antidepressant medications more commonly prescribed today.

Hollon et al. (2006) report in a non-systematic narrative review that patients treated to remission with CBT are about half as likely to relapse after treatment is terminated as those who are treated with medication. They provide examples from the work of their own group to support this statement, referring to DeRubeis et al., (2005) and Hollon et al., (2005c) who found that patients treated with CBT or with medication made the same gains by 6 weeks, but that after termination, at 12 months follow-up,

those who had prior CBT had a relapse rate of 31% as compared to those who had been treated with medication and withdrawn to placebo who relapsed at a rate of 76%, and those who continued medication throughout the follow-up period who relapsed at a rate of 47%. Hollon et al.'s (2006) review is illustrative and not systematic, so it is possible that the results reported are biased by the trials that the authors chose to describe. It is also possible that the articles they cite are of high quality, in which case this bias may not lead to erroneous conclusions. According to the studies they do cite, Hollon et al. (2006) conclude that when directly compared, post-treatment effects for CBT indicate relapse rates about half that of discontinued medication treatment.

De Maat, Dekker, Schoevers, and De Johghe (2006) conducted a systematic review and included 6 studies that compared ADM and CBT groups. Their findings were consistent with Hollon et al. (2006) as they found that relapse rates were higher in the ADM groups (57%) than in the CBT groups (27%).

In another non-systematic narrative review Paykel (2007) review 7 long-term follow-up studies examining CBT, specifically CBT for residual or recurrent depression. He included comparisons to ongoing medication treatment or comparisons in which CBT is added to ongoing medication treatment versus ongoing medication treatment lone. He also included CBT for residual or recurrent depression in comparison to medication withdrawal. All 7 studies found reduced relapse rates for the groups who received CBT compared to the groups who did not receive CBT and were either discontinued or continued on their antidepressant medications. The author noted that the difference between the groups is smaller when the medication group continues treatment than when it is discontinued. Paykel, (2007) noted that the difference in relapse rates

does not appear to be due to enhanced medication compliance because a significant difference was present in studies when medication was withdrawn, when patients were never on medication, and in studies when medication compliance was good throughout the follow-up period. No indication of search strategy or inclusion and exclusion criteria were provided. The authors state that these are the only 7 follow-up studies of this nature that had been conducted to date but do not report the process followed to arrive at this conclusion.

In another meta-analysis of studies directly comparing acute phase psychotherapy (any psychotherapy "intended to be therapeutic") and medication treatment (any medication accepted for treatment of depression), Imel, Malterer, McKay, and Wampold (2008) included 11 studies with long-term follow-up data (average of 15 months), and found a significant advantage, with moderate effects size, for psychotherapy over discontinued medication treatment at follow-up ($\lambda(13) = 0.37$, $p < 0.01$). They found no difference at follow-up between acute phase psychotherapy (not continued during the follow-up period), compared to continued medication treatment over the follow-up period ($\lambda(10) = -0.03$, $p > 0.5$). The lack of description of the psychotherapies included in this review is a limitation as outcomes may differ by treatment, and the authors do not provide information to identify which specific therapies were included in their comparisons. This review reported only aggregate effect sizes and did not provide categorical outcome data such as relapse and recurrence, or continuous outcome data such as depression scores or symptom ratings.

Spielmans, Berman, and Usitalo (2011) conducted a similar review to that of Imel et al. (2008), but with slightly different inclusion criteria. They included studies in

which any psychological treatment was compared to a second generation antidepressant medication. They coded psychotherapies as “bona-fide” or “non-bona-fide”. Bona-fide treatments were treatments delivered by a trained therapist with at least a master’s degree, in a face-to-face individualized context, and that contained psychologically valid components such as a basis in an established approach, identification of active ingredients and citations to support these, or a manualized format. The studies for which follow-up data were available examined 7 psychotherapy interventions, including CBT (4 studies), rational emotive therapy (1 study), supportive-expressive (1 study), GP delivered problem solving therapy (1 study), and Bellak’s therapy (1 study; this type of therapy is named but not described). Six of these were deemed bona-fide, but it was not specified which 6 these were. It is likely that the 2 non-bona fide interventions were the supportive-expressive treatment and one of the other non-CBT interventions. Spielmans et al. found that bona-fide psychotherapies were superior to medications on continuous outcome measures at follow-up. The effect size was small ($d = 0.29$, $p < 0.01$). There was no effect for non-bona-fide psychotherapies ($d = 0.15$, $p > 0.40$).

Overall, these reviews conclude that during the post-treatment follow-up phase, relapse rates are significantly lower for those who received CBT than for those who received ADM that has been discontinued. It also appears that prior CBT has a relapse-prevention effect similar to that of continued antidepressant medication over follow-up.

Behavioural activation. Cognitive behavioural therapy has received significant research attention. There also a small, but growing body of evidence that the behavioural components of CBT can be applied independently of the cognitive aspects, and this treatment referred to as behavioural activation, is equally efficacious when

compared to CBT (Hollon et al., 2006). Cuijpers, van Straten, and Smit (2007) conducted a meta-analysis of 9 studies including follow-up data on behavioural activation (BA), and found that there was no significant change from post-treatment to follow-up suggesting that patients had maintained their gains over the follow-up periods. Specifically, pooled effect sizes were reported separately for the change between post-test and follow-up occurring 1-3 months post-treatment (5 studies), 4-6 months post-treatment (5 studies), and 7-12 moths post-treatment (2 studies), (d = 0.18, 0.03, and 0.53 respectively, all *ns*). Note that although the effect size at 7-12 months is not statistically significant, this may be due to a lack of power as this analysis was based on only 2 studies. There was also no significant difference between BA and CBT at post-treatment follow-up with effect sizes reported separately for comparison between BA and CBT at 1-3 months post-treatment (4 studies) and 4-5 months post-treatment (4 studies), (d = 0.02, - 0.13 respectively, both *ns*). No other comparisons were reported at follow-up.

Another systematic meta-analysis by Ekers et al. (2008) looked at the effectiveness of BA against comparison groups including TAU, non-treatment control, CBT, brief psychotherapy, and supportive counseling. No long-term follow-up data was available to compare behavioural interventions to waitlist or placebo control groups or to supportive therapy. Comparisons of BA to CBT found no difference at long-term follow-up (*SMD* = 0.25, 95% *CI* -0.21 to 0.70. $p > 0.20$). Eight studies contributed to that analysis and the average follow-up period was only 4 months. Only one of these studies overlaps with the Cuijpers et al. (2007) review. Ekers et al. also compare BA to brief psychotherapy. They include two studies, with an average follow-up of 4.5 months

and report a medium sized positive effect in favour of behavioural interventions (*SMD* = -0.50, 95% *CI* -0.90 to -0.09, $p < 0.05$).

These two reviews converge on the finding that BA has relapse-prevention effects similar to CBT. It also appears that the gains achieved during acute phase BA are maintained over follow-up. The duration of follow-up, and comparison to control groups over the long-term was limited in many of the reviewed studies. These findings are promising and warrant additional, longer-term follow-up.

Interpersonal psychotherapy. There are no reviews that focus on the post-treatment outcomes for IPT. Hollon et al. (2006) reports that although IPT appears to prevent relapse and recurrence while it is continued, there is no evidence as to whether these effects endure beyond the termination of treatment. They cite two studies with reference to continued treatment and two studies with reference to relapse-prevention post-treatment, but they do not provide any description of these studies or their results.

Psychodynamic therapy. Three reviews examine psychodynamic treatments. Leichsenring (2001) conducted a meta-analysis on the comparative effects of short-term psychodynamic psychotherapy (STPP) and CBT in depression and reported that at 1 – 2 years post-treatment follow-up there was no significant difference between the percent of the population that relapsed (56% for SSTP, 46% for CBT), though they advise interpreting this finding with caution as their meta-analysis included only 3 studies with follow-up data. Leichsenring also warns that these findings apply only to the specific variants of STPP delivered in the included studies (these variants are not described in the review) and should not be generalized to other forms of STPP.

Hollon and Ponniah (2010) review the literature on brief dynamic therapy (BDT) and report on only one long-term follow-up study in which patients who had BDT were less likely to relapse over a 48 month follow-up than were patients who had received brief supportive psychotherapy (BSP). In this study, both BDT and BSP were added to medications as the medications were about to be discontinued and the 48 month follow-up was treatment-free (Maina, Rosso, &Bogetto, 2009).

A meta-analysis by Driessen, Cuijpers, de Maat, Abbass, de Johnghe, and Dekker (2010), which has no overlap with the Leichsenring (2001) review, found that the effects of STPP were large (d = 1.34, pre- to post-treatment) and that these gains were maintained at 3, 6, and 12 months post-treatment follow-up (d = 0.03, 0.21, and 0.02, all *ns*). Compared to other psychotherapies, STPP was not significantly different at 3 and 12 months follow-up (d = -0.09, -0.29, both *ns*). The effect size at 12 months follow-up represents a non-significant trend favouring other psychotherapies ($p < 0.10$). The authors assessed for publication bias and reported that their results were not significantly altered. This review should be interpreted with caution as the authors report that some of the included studies had very small sample sizes, and the quality of the included studies was not optimal. In fact, a number of studies did not include a treatment fidelity check or a treatment manual, permitted use of antidepressants in addition to psychotherapy, did not train the therapists delivering treatment, or did not include a control group. The authors also note that they included a variety of different STPP methods and that the number of studies using the same variant of STPP was too small to perform subgroup analyses.

These reviews suggest that the long-term relapse-prevention effects of psychodynamic therapies are not well established. The reviews just cited include a small number of studies, and these studies were frequently under-powered. In addition, psychodynamic therapies are diverse and rarely described in sufficient detail to permit confident synthesis of results across studies.

Combined psychotherapy and medication treatment. Three reviews examined the post-treatment outcomes for combined treatment. The first was a non-systematic narrative review by Segal, Vincent, and Levitt (2002) that included relapse rates for 12 follow-up studies. In each case, the monotherapy was compared to a combination of the same monotherapy plus another treatment type. Segal et al. distinguished between *combination* treatment in which both treatments are offered at the same time, *sequential* treatment in which one therapy is added to another, and *crossover* treatment in which one treatment is discontinued and then followed by another. They reported mixed findings for combination and crossover treatments and only one study in their review reported on sequential treatment. The authors suggest that combined treatment may be most appropriate for patients with severe or chronic or recurrent depression, but cite a lack of power in studies that had aimed to investigate these variables. They also suggest that combined treatment may offer an advantage with regards to relapse-prevention but acknowledge a similar lack of power and note that larger studies are required to answer this question. With regard to sequential treatment in particular, the authors suggest that it may be most beneficial for patients who do not achieve full remission on the first treatment. This statement is made with reference to response to treatment, and not to enhanced relapse prevention.

Friedman, Detweiler-Bedell, Leventhal, Horne, Keitner, and Miller (2004) conducted a meta-analytic and narrative review of combined treatment. Their review included RCTs in which a monotherapy (pharmacological or psychological) was compared to a combined pharmacological and psychological treatment, and the monotherapy had to be the same as one of the elements of the combined therapy. They included acute and maintenance studies in their review. Friedman et al. (2004) found that adding CBT to medication treatment appears to prevent relapse, particularly among individuals who terminate their medication treatment. With regards to relapse rates during the follow-up period, the authors report that these studies demonstrate a significant effect for combined CBT plus medication versus medication alone (d = 0.68, χ^2= 4.06), and only a small effect for combined CBT plus medication compared to CBT alone (d = 0.12, χ^2 = 1.48). The authors report only effect sizes and not relapse rates, and do not include an indication of statistical significance for these effects.

Cuijpers, van straten, Warmerdam, and Anderson (2009) conducted another meta-analysis of 7 studies that directly compared psychotherapy to a combination of the same type of psychotherapy plus pharmacotherapy. Analysis by type of therapy was not possible at long-term follow-up, and studies of CBT (3 studies), IPT (1 study), problem solving therapy (1 study), and self-control therapy (1 study) were all included in the overall effects. The authors found no difference between combined treatment and psychological mono-therapy at 3 – 6 months (d = - 0.15, *ns*), or 12 months (d < 0.00, *ns*) follow-up. This finding should be taken with caution for several reasons. The main focus of Cuijpers et al.'s (2009) review was acute treatment effects. Because of the small number of long-term follow-up comparisons available (7 studies, 10 comparisons)

and the wide range of follow-up duration (1 month to 2 years – both of these extremes being excluded from analysis) limited analyses were conducted with this follow-up data. Although Cuijpers et al.'s (2007) review is of good quality, the authors note that many of the studies they reviewed were not of optimal quality. Their meta-analysis is thus limited by the quality of the studies that they were able to include.

Only three studies overlap between these two systematic reviews (Cuijpers et al., 2009; Friedman et al., 2004) due to their differing inclusion criteria (type of mono-therapy included, and inclusion or exclusion of maintenance studies). Friedman et al.'s (2004) inclusion of maintenance studies may partly account for their significant findings at long-term follow-up, given that there is some evidence that combined treatment may have the greatest relapse prevention effects when CBT is added at the end of medication treatment, during discontinuation of medication or immediately thereafter (Fava, Grandi, Zielezny, Rafanelli, & Canestrari, 1996; Fava, Rafanelli, Grandi, Canestrari, & Morphy, 1998b). All three reviews presented here note that more long-term follow-up studies of combined treatment are required. Given the likelihood that not all combined treatments are equivalent, it will also be important to be able to conduct separate analyses for different types of therapies, for different phases of treatment (acute, maintenance), and even for different methods of combining treatment (eg. combined, sequential, crossover).

Study 1 Discussion

Summary of Main Results

This overview includes 12 meta-analyses, 3 systematic narrative reviews, and 5 non-systematic narrative reviews assessing the long-term, post-treatment effects of a range of interventions for the treatment of major depressive disorder. Overall, it appears that relapse over the long-term is a concern regardless of treatment modality, with relapse rates commonly falling between 30-60% within 2 years after the end of treatment. It is important that both clinicians and patients consider this significant risk, as well as the measures that can be taken to reduce it, or reduce the impact of relapse when it does occur.

While antidepressant medications have been shown to be effective during acute treatment, and to have relapse-prevention effects when they are taken during continuation and maintenance periods, they have relatively high relapse rates after treatment is terminated. CBT is similar in effectiveness to antidepressant medication during the acute treatment phase, and also appears to reduce relapse rates by about half compared to medication treatment alone, for at least a couple of years after treatment is terminated. The body of literature on the long-term post-treatment effects for combined versus mono-therapy is small, but suggests that combined medication and CBT during acute phase treatment improves long-term outcomes when compared to medication discontinuation alone. The advantage for combined CBT and medication over medication alone may be most effective in preventing relapse when administered sequentially, i.e. medication → discontinuation →CBT but this finding requires further

study. Combined treatment with CBT and medication does not appear to offer better long-term outcomes than CBT alone.

Behavioural treatment, particularly behavioural activation, may be as effective as CBT in decreasing rates of relapse. This finding is promising, but requires further replication as it is supported by only two reviews to date (Cuijpers et al., 2007; Ekers et al., 2008). Behavioural activation may be a cost effective option if it proves to have good long-term outcomes because it is simpler to deliver and it is easier to train practitioners to deliver it effectively (Centre for Economic Performance's Mental Health Policy Group, 2006; Jacobson et al. 1996; Jacobson, Martell, & Dimijan, 2001). No reviews provided evidence for relapse-prevention effects during the post-treatment period for IPT or psychodynamic therapies.

Briefly, this overview finds that:

- Return of symptoms is common after treatment is terminated.
- Continued antidepressant treatment over the follow-up period reduces relapse significantly compared to discontinuation of antidepressant medication.
- Acute phase CBT reduces relapse relative to medication treatment alone even after it is terminated, at least for a period of 1 - 2 years. Continuation phase CBT further reduces relapse during follow-up, even after it is discontinued.
- Even with continued antidepressant medication or continuation CBT, relapse rates still reach about 40 percent within 1 – 2 years. It is important that we continue to examine ways to provide ongoing support to promote maintenance of

gains and decrease the impact of recurrent episodes of depression. MBCT provides one example of a potential maintenance strategy to reduce both the frequency and intensity of relapse/recurrence, with the benefit seen primarily for those with a history of repeated episodes of depression (three or more).

- Combined CBT and antidepressant medication may provide post-treatment relapse-prevention effects for those who reach remission on antidepressants when CBT is added during discontinuation of antidepressant medication.

Quality of the evidence

The methodological quality of the included systematic meta-analyses and narrative reviews was very good, with almost all reviews falling into the category of "minimal or no flaws" on an application of the Oxman Criteria. One review (Viguera et al., 1998) was identified as having "minor flaws". Inherent in the non-systematic narrative reviews, is a likelihood of bias, and because of the lack of a defined methodology in these reviews, they cannot be evaluated for quality in the same way that systematic reviews can be. This suggests that a small number of the reviews included in this overview have limitations in design and/or execution that may influence the results when considered both individually and collectively.

While the quality of the reviews themselves is high, according to the Oxman criteria employed, many of the included reviews note methodological limitations of the individual studies they included. In addition, the problems created when reviewers synthesize outcome data that is not collected on a common metric, as described above for measurement of relapse rates, present a compromise in the quality of the analyses

performed in these reviews. The limitations in the quality of the evidence have been taken into account in this overview by interpreting the results and drawing conclusions in light of the quality of the included reviews. However methodological limitations at both the included study and review levels require that the results of this overview be interpreted with caution.

Potential Biases in the Overview Process

This overview followed rigorous methods to minimize the impact of bias contributed to the overview by the review process itself. The present analysis does have several limitations. First, although searches were broad and a large number of databases were included, the author cannot be certain that all relevant articles were located. Second, all systematic reviews are prone to publication bias. Also, although two reviewers screened titles and abstracts for inclusion, the fact that only one reviewer extracted the data and performed quality assessment might have caused additional bias. Finally, there were some cases in which individual studies contributed to more than one review. This was taken into account and the results were interpreted in light of this limitation, but it does remain to be a limitation of the overview. It is not possible to eliminate the risk of bias in an overview of reviews and the synthesis of evidence involves judgement and is not an exact science. It is thus important to be aware of potential sources of bias and to interpret overviews in light of these limitations.

Agreements and Disagreements with Other Studies or Reviews

This is the first overview of reviews that examines long-term post-treatment outcomes for adults treated for depression, across multiple treatments. Overviews of reviews exist for acute treatment outcomes for individual treatments (e.g. Butler et al., 2006) but not for the post-treatment period, and not across interventions. In their overview of existing meta-analyses on the effects of CBT, Butler et al. (2006) note that, although the literature on treatment effectiveness is vast and growing, there are very few reports on post-treatment follow-up. In fact, in their overview, they found only two previous reviews that spoke to this issue (DesRubies & Critschristoph, 1998; Glauoguen et al., 1998).

The findings of the current overview are consistent with other research on the long-term course of major depressive disorder. This overview finds that despite effective acute treatment, major depressive disorder is highly recurrent. Lewis Judd and colleagues have published a series of studies using data from the National Institute of Mental Health (NIMH) Collaboration Depression Study (CDS), which is a long-term prospective naturalistic study of patients with mood disorders. This series of studies is unique in that it follows patients for up to 31 years using rigorous periodic evaluations including weekly symptoms severity measures that allow for a continuous description of the course of MDD for this sample. Judd and his colleagues also report high rates of relapse and recurrence, even after treatment, and they go as far as to recommend viewing MDD as a chronic condition (Judd, 2012).

Implications for Research.

There are numerous ways in which the quality of long-term, post-treatment follow-up studies could be improved. The following recommendations represent a synthesis of those previously stated in the literature, as well as priorities identified in the process of conducting the current overview:

Recommendations for future reviews.

- Documentation of study selection including reasons for exclusion.
- Thorough description of search strategies and inclusion and exclusion criteria.
- Inclusion of quality analysis for included studies using a standardized tool (Cochrane, 2011).

Reviews that do not report all of the necessary details to determine how included studies were selected cannot be evaluated for their accuracy in representing the body of literature they review. Poor reporting results in an assessment of low quality, and requires caution in interpretation of their results. In addition, an overview's quality is limited if it must include reviews whose methodology is less than optimal or is not well reported.

Recommendations on study design.

- Studies should be designed at the outset to include a long-term follow-up, including non-responders and comparison groups (Friedman et al., 2004; Hensley et al., 2004; Hollon et al., 2006; Kaymaz et al., 2008; Scott, 1996).

- Studies should be designed to treat to criteria (remission or recovery), rather than for a set period of time (Hollon et al., 2006; Judd, 2012).

Planned follow-up of the control group would move us toward understanding, not just the comparative efficacy of one carefully controlled intervention to another, but the absolute efficacy of an intervention compared to the natural course of untreated, or treated depression (as in the case of a TAU control group). In a reality where various treatment options are increasingly available, it may become less feasible, and certainly less ethical to require a no-treatment control group to abstain from pursuing treatment over a long-term follow-up period. A TAU condition, in which treatment is continuously measured and reported, may be the most realistic comparison group to include in a study design.

While treating to criteria would reflect the real world aim to treat patients until they are well, it is also the case that real world patients often engage in time limited treatment for various reasons, and may not achieve remission by the end of treatment. For this reason, it is important to have both kinds of data, reflecting what happens to those who reach remission and those who do not. Studies that treat to remission would also allow us to more accurately assess factors such as optimal duration of treatment. This would avoid overestimating relapse by stopping treatment before an individual is in stable remission, and then counting them as a case of relapse, when in reality they had reached remission. In addition, we know from the decades of work by Judd and colleagues on the NIMH-CDS data that symptoms below threshold are associated with significant impairment, and that patients who reach full recovery remain free of symptoms more than 5 times longer than those who terminate treatment with residual

symptoms (Judd, 2012). This group's research has also shown that among patients recovering from their first episode of depression, residual symptoms at recovery are associated with significantly faster relapse, and a significantly more severe and chronic course of lifetime depressive disorder (Judd et al., 2000).

Recommendations on measurement.

- Consistent use of consensus definitions of relapse and recurrence (Frank et al., 1991; Rush et al., 2006; Scott, 1996; Vittengle et al., 2007).
- Continuous, and not cross-sectional, longitudinal outcome measurement during follow-up (Hollon et al., 2006; Judd, 1998; Vittengle et al., 2007).
- Use of a common metric such as % relapse per month, to permit clearer comparison between studies with varying lengths of follow-up.
- For psychological interventions, consistent measurement and reporting of therapist characteristics such as affiliation with a particular intervention, level of training and experience, and adherence/fidelity and competence for which standardized measures exist (Beck, 2005; Vittengle, 2007).
- Reporting of adherence to medication treatment for pharmacological studies, and of actual dose received for psychological studies (Vittengle et al., 2007).
- Systematic, continuous measurement and description of treatment received by the TAU group, and of extra-protocol treatment received by the intervention group during treatment and follow-up phases (Hollon et al., 2006; Vittengle et al., 2007).

Given the data available, reviewers must use the definitions of relapse and recurrence employed by the original study authors. The use of consensus definitions and a common metric would allow for much more accurate comparison between, and synthesis across studies. Continuous outcome measures would reduce the number of relapses that are missed when we use cross-sectional measures. Although longitudinal continuous measures would be ideal, they have been difficult to achieve. Continuous retrospective measures such as the LIFE interview would be helpful in many situations. The LIFE interview relies on patients' use of chronological prompts to remember their weekly functioning over a period of time, commonly ranging from 1-6 months (Judd et al., 1998). The LIFE interview is thorough in that it asks patients to report on symptoms week by week, but a weakness is that information is typically collected infrequently and consequently retrospective report covers a long time period (such as the semi-annual schedule in the NIMH-CDS study). Inclusive follow-up of non-responders and those initially assigned to the control condition would minimize the inaccuracies inherent in obtaining follow-up data from only a portion, often half or less (Hollon et al., 2005a; Hollon et al., 2006) of the original study sample.

The recommendations to more accurately report adherence/dosage, and therapist characteristics are in service of better documenting the variations in treatment delivered and received. One of the major limitations of current meta-analyses is that they assume uniformity of treatment. We know that, for example, CBT is not one uniform treatment (Beck, 2005), either in regards to content, or therapists' abilities. It is likely unrealistic to expect that we will ever achieve this uniformity in the literature. It would be helpful

to document as much of the variation as possible and control for this variation, or test for the potential of these variables to mediate outcomes.

With regards to extra-protocol treatment, it is not only important to report on continuous measurements to enhance our ability to control for variation. It is also necessary to document patient status and/or reason for pursuing these treatments. Currently, some studies do not report this variable at all, and some simply monitor for treatment received and control for this variable. Others consider seeking out additional treatment as an indication of relapse, which is not necessarily true. Some individuals may seek additional treatment for prodromal symptoms, in an effort to prevent a full relapse, and in fact, this may be a prescribed action in their relapse-prevention plan. Receipt of additional treatment should not be used as a measure of relapse unless it is established that an individual meets criteria for relapse at the time of that additional treatment. Finally, measurement of treatment received by the TAU group is important for making informed comparisons as described above.

Recommended analyses.

- Reporting of intent to treat (ITT) analyses in addition to completer analyses (Hensley, 2004).
- Use of time-to-event, or "survival" analysis rather than reporting the proportion who relapse/recur (Vittengle et al., 2007).
- Consistent reporting of, and subgroup analysis on, potential mediators such as number of previous episodes, and duration and severity of index episode (Kaymaz et al., 2008).

Conducting ITT analyses that are inclusive of the complete sample, and reporting time-to-event ("survival") analyses which provide more accurate estimates of relapse when studies include participants who do not relapse or withdraw by the end of follow-up, both serve to improve the accuracy of outcome reporting. Subgroup analyses would help to answer questions about mediating variables, which are frequently proposed, and as yet, infrequently tested.

Implications for Practice.

This overview suggests that regardless of the treatment applied, relapse is likely for nearly half of all those who respond to treatment for an episode of MDD. Taking this into consideration, the following are recommendations for practice in the treatment of adult depression.

- At the acute treatment stage, practitioners can advise patients of the likelihood of relapse, and incorporate a relapse-prevention plan into their treatment of the index episode (Judd, 2012).

- Currently, continued antidepressant medication, or a course of continuation CBT or mindfulness based cognitive therapy for relapse prevention all appear to be effective in reducing relapse rates.

- Sequential treatment with acute phase ADM and continuation phase CBT may be the best approach to combined treatment with regards to reducing relapse rates.

As we come to understand that, for most people who experience depression, it is a recurrent or chronic disorder (Judd, 2012; Keller et al., 1992; Mueller, et al., 1999), we must continue to explore ways to optimize maintenance of the gains achieved in acute treatment, and to minimize the frequency, duration, and intensity of relapses when they do occur. It is important that patients are accurately informed about the potential for relapse, and that they are able to make informed choices about how best to minimize its occurrence.

Study 1 Conclusions

In conclusion, while relapse is a significant issue after treatment ends, continued ADM, and prior CBT and BA have been found to have relapse-prevention effects. Continued use of antidepressant medications has about the same relapse-prevention effect as prior CBT or BA, and continuation phase CBT appears to have increased relapse-prevention effects regardless of whether acute treatment was CBT or ADM. Additional research, including the refinements to design and reporting described here will enhance our understanding of the post-treatment period. The consistency of results across the reviews included in this overview provide support for recommendations, for both clinicians and patients, to assist in planning for relapse prevention. Given the recurrent nature of depression, discussion of and planning for relapse-prevention should be included in all forms of acute treatment. An increased research focus and implementation of primary prevention interventions is also important as we do not currently have, what might be considered, a “cure”.

Study 2: Narrative Systematic Review - Method

Criteria for Selecting Studies for this Review

Study design. Studies with post-treatment follow-up of randomized controlled trials (RCTs) were included if there was a comparison group in place during the follow-up – either an active treatment or a control group (e.g. treatment as usual, clinical management, placebo). A minimum of 12 months of post-treatment follow-up was required for inclusion. Studies published in English that met the selection criteria were included. Continuation and maintenance studies, for which the focus was on the group that continued treatment, were excluded. These studies are relevant to our understanding of post-treatment outcomes, as they report relapse rates for the group that is randomized to placebo withdrawal. They have, however, been well reviewed elsewhere (e.g. Kaymaz et al., 2008; Viguera et al., 1998) in reviews covering a large number of studies with a large number of participants (reported in Study 1).

Types of participants. Studies were included that enrolled adults (aged 18 years or older) who had a diagnosis of unipolar depression and who had completed/terminated treatment for this disorder. Studies including participants with co-morbid anxiety were included given the high rate of co-morbidity of anxiety and depression, provided that participants were not also receiving a separate intervention for their anxiety. Studies focusing on participants with unipolar depression and co-morbid substance abuse disorders, personality disorders, psychosis, chronic pain, and medical diagnoses such as cancer were excluded. Studies including participants with bipolar affective disorder or dysthymia were also excluded. When participation was limited to a restricted age range (e.g., 65 and over) the study was excluded.

Types of interventions. Interventions or combinations of interventions given by a health professional for the treatment of unipolar depression, and established as an effective acute treatment, were included. Medication treatments included first generation anti-depressant drugs, namely tricyclics (TCAs) and monoamine oxidase inhibitors (MAOIs), and second generation antidepressants including selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), other drugs that selectively affect the activity of serotonin, norepinephrine, and dopamine. Psychological treatments included cognitive therapy or cognitive behavioural therapy (CBT), behavioural activation (BA), interpersonal psychotherapy (IPT) and other psychotherapies. Internet based interventions were not included, as the literature on these interventions is limited and they may reach participants with widely different characteristics.

Outcomes. Clinical outcomes, including relapse and maintenance of gains, were examined. These are the most commonly reported outcomes after successful treatment of depression. They are also outcomes that consumers report wanting to be informed about in choosing a course of treatment (Walker et al., 2000).

Search Methods for Identification of Studies

Terms searched included: depression, treatment, follow-up, relapse, maintenance, continuation, discontinuation, and "return of symptoms." The following databases were searched: PsychINFO; PubMed; the Cochrane Library including Systematic Reviews and the Controlled Trials Register (CCTR), Cochrane Database of Systematic Reviews (CDSR), Database of Abstracts and Reviews of Effects (DARE), and Cochrane Methodology Register (CMR); CINAHL; Social Sciences Citation Index;

SCOPUS; and Web of Science. Selected journals which the author deemed relevant in the review were hand-searched This included The Archives of General Psychiatry, Journal of Clinical and Consulting Psychology and the Journal of Affective Disorders. Reference lists of the selected studies and reviews were also searched.

Data Collection and Analysis

Selection of studies. The titles and abstracts obtained in the searches were screened against the inclusion criteria to identify potentially relevant studies. This was followed by a more detailed screening of the full papers identified as possibly relevant in the initial screening. Selection of studies was carried out by two reviewers for reliability (the author and another member of the Mobilizing Minds research team – a graduate student in psychology who was trained by the author and the author's research supervisor on the application of the inclusion and exclusion criteria). Any disagreements (there were very few) were resolved by discussion, and when necessary through arbitration with a third team member (the author's research supervisor).

Data extraction. A checklist for inclusion was completed by two reviewers which extracted basic information about study design and characteristics (see Appendix A). One reviewer extracted the data from the included articles according to predefined criteria, including key characteristics and results (see Table 3).

Assessment of methodological quality of included studies. There is evidence that the quality of study design can have an effect on the outcomes of a trial (Moher et al., 1998). For example, lack of double-blinding or concealment of allocation can result in exaggeration of the effect size of a trial (Juni, 2001). It is recommended that quality assessment be used in a systematic review in order to minimise bias (Cochrane, 2011;

Lynch et al., 2010). Quality assessment was used in this review to determine the robustness of the results of an included study. Quality assessment was not used to weight studies, as there is no evidence for this approach (Juni, 2001). Information extracted for quality assessment is presented in Table 4.

Data synthesis. Narrative synthesis was conducted by extracting and critically analysing descriptions of studies' methods, results, and conclusions, as well as by contrasting and synthesizing results across studies. Quantitative results such as the proportion relapsing, or effect sizes, were compared when appropriate. Planned subgroup synthesis was conducted by type of intervention.

Study 2 Results

Description of Studies

A total of 13 445 titles and abstracts were obtained. Of these, 265 abstracts were identified as possibly meeting inclusion criteria and selected for additional review. Review of abstracts resulted in exclusion of another 181 based on lack of follow-up data, treatment continued throughout follow-up, or no random assignment. For the remaining 84 studies, full text articles were reviewed and 30 of these were included. The other 54 studies were excluded because the interventions were not well enough described, did not qualify as one of the eligible evidence based treatments, or when treatment was continued throughout the follow-up period. The results of this systematic screening process can be seen in Figure 3.

All inclusion criteria were met by 30 studies, with a total of 2624 subjects. One study did not report sample size (Conradi, de Jonge, & Ormel, 2008). Selected

characteristics and results of these studies are described in Table 3. Eight studies included treatment arms with 1st generation antidepressant medications (ADMs), 3 with 2nd generation ADMs, 3 with unspecified ADMs, 26 with CBT, 3 with BA, 1 with IPT, and 7 with other forms of psychotherapy.

Figure 3. Progress through the stages of a systematic review for RCTs (PRISMA; Moher et al., 2009)

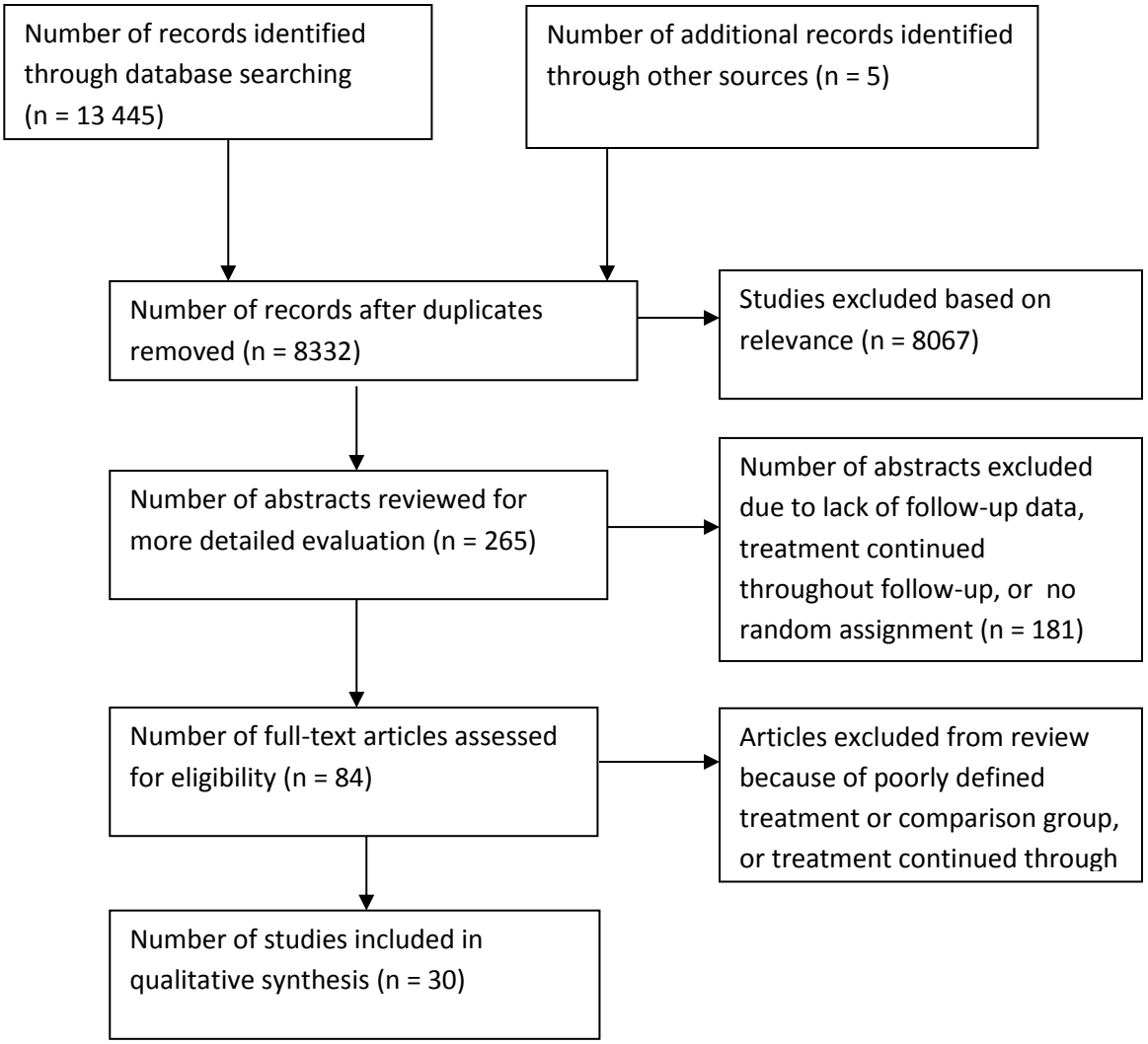


Table 3. Study characteristics and results.

Study	Treatments, N [Inactive control, N]	Duration: treatment/ follow-up	Follow-up N/ % attrition	Outcomes	Results
Direct Comparison: Antidepressant Medication vs CBT					
Blatt, 2000	CBT, Imipramine + clinical management, Interpersonal Psychotherapy (IPT), [Placebo + clinical management] All conditions terminated after 16 weeks N = 239 were randomized, not reported by treatment group.	16 weeks/ 18 months	162/42% Not reported by treatment group Completer analysis and supplementary analyses with n = 204 including partial completers, no ITT	Symptom reduction (aggregate of several depression screens), Current clinical condition (continuous, patient ratings)	Symptom reduction: No significant difference among 4 groups, F = 1.17, df = 129, *ns* Current clinical condition: No significant difference among 4 groups, F = 0.26, df = 138, *ns* Results not reported by treatment group.
Dobson, 2008	CBT, 30 BA, 27 Paroxetine withdrawn to placebo, 21 Paroxetine continued for 1st year of follow-up, 28	16 weeks/ 1 and 2 years	CBT, 26/10% BA, 21/19% Placebo, 17/19% cADM, 26/7% Survival analysis	Relapse (year 1), recurrence (year 2)	Relapse at 1 year: CBT – 39%, BA – 50%, Paroxetine – 53%, placebo – 59%; as a set, the three active treatments were superior to placebo, $\chi2(1, N= 106) = 4.07$, $p < 0.05$; separately, CBT was superior to placebo $\chi2(1, N= 51) = 5.30$, $p < 0.10$, BA showed a non-significant

			trend over placebo $\chi 2(1, N= 51) – 2.81, p < 0.10$, and continued Paroxetine was not significantly different from withdrawal to placebo $\chi 2(1, N= 49) = 0.97, p > 0.30$ Recurrence at 2 years (for those who survived year 1 without relapse): CBT – 24%, BA – 26%, prior continuation Paroxetine – 52%, taken together, CBT and BA were superior to prior continuation Paroxetine, $\chi 2(1, N= 46) = 3.58, p < 0.07$; separately, CBT was not superior to prior continuation Paroxetine $\chi 2(1, N= 34) = 2.71, p < 0.10$, and BA was not superior to prior continuation Paroxetine $\chi 2(1, N= 29) = 1.40, p > 0.20$ Overall effect for full 2 year follow-up: taken together, CBT and BA were superior to withdrawal after continued Paroxetine, $\chi 2(1, N= 85) = 6.59, p < 0.05$; separately, CBT was superior to withdrawal after continued Paroxetine $\chi 2(1, N= 58) = 5.97, p < 0.05$, BA showed a non-significant trend over

					withdrawal after continued Paroxetine $\chi 2(1, N= 55) = 3.09, p < 0.10$
Evans, 1992	CBT, Imipramine continued (12 weeks + 1 year), Imipramine discontinued (12 weeks only), Combined CBT and Imipramine (12 weeks only) Acute phase n = 107, not reported by group, Treatment responders eligible for follow-up = 50, not reported by group	12 weeks/ 2 years	CBT, 10 cADM, 11 ADM discontinued, 10 CBT + ADM, 13 Attrition, n = 12, 24% of follow-up sample (not reported by group)	Relapse	Relapse: CBT – 21.0%, cADM – 32%, ADM discontinued – 50%, CBT+ADM discontinued – 15% The difference between CBT(alone and in combination with ADM - pooled) and ADM discontinued was significant, $\chi 2(1) = 4.05, p < 0.05$
Hollon, 2005	CBT, 35 ADM continued, 34 ADM withdrawn to placebo, 35 (ADM not specified)	16 weeks/ 1 and 2 years	CBT, 32/9% cADM, 29/15% Placebo, 27/23% Survival analysis	Relapse	Relapse at 1 year: CBT - 31%, cADM – 47%, Placebo – 76%; CBT significantly lower than Placebo ($p < 0.01$) and not significantly different than cADM ($p > 0.10$) Recurrence at 2 years for those surviving year 1 without relapse: CBT – 17%, prior cADM (withdrawn) – 54%; Significantly different, $\chi 2(2, 1) = 6.81, p < 0.01$

Kovacs, 1981	CBT, 19 Imipramine (12 weeks), 25	12 weeks, maximum of 20 sessions, 50 minutes for CBT, 20 minutes for ADM/ 1 year	CBT, 18/5% ADM, 17/32%	Relapse, Depressive symptoms (BDI)	Relapse: CBT – 39%, ADM discontinued – 65%; $\chi 2 = 2.33$, significance not reported Depressive symptoms (BDI scores): No significant difference from termination to 1 year follow-up for either group ($t = -0.11$, $p > 0.90$). CBT group had lower BDI scores than ADM group ($F(1,28) = 4.96$, $p < 0.03$). CBT group had less pessimistic and hopeless thinking than ADM group ($F = 6.37$, $df = 1$, $p < 0.02$).
Schramm, 2011	CBASP, 15 IPT, 15	16 weeks/ 1 year	CBASP, 7/53% IPT, 12/20% ITT analysis	BDI score	Mean BDI score: CBASP – 12.92, IPT – 18.66, not significantly different, $F = 0.67$, $p > 0.40$ Note that power is limited by small sample size and high rate of attrition.

Simons, 1986	CBT, Nortriptyline, CBT + Nortriptyline, CBT + placebo, 87 patients entered the trial, not reported by group 70 patients completed treatment and were followed for one year	12 weeks, maximum of 20 sessions, 50 minutes for CBT, 20 minutes for ADM, 60 minutes for combined/ 1 year	CBT, 19/ Nortriptyline, 16/ CBT + Nortriptyline, 18 CBT + placebo, 17/ 11% attrition overall, not reported by group	Relapse	Relapse rates: Groups who received ADM (the nortriptyline group and the CBT + nortriptyline group), 52%, Groups who had not received ADM (CBT and CBT + placebo), 19%, $p < 0.05$. Groups who had received CBT (the CBT group, the CBT + placebo group, and the CBT + nortriptyline group), 28%, Groups who had not received CBT (the nortriptyline group), 66%, $p < 0.05$.
CBT in Primary Care					
Conradi, 2008	CBT + depression education, Depression education(low intensity, 6 sessions), [TAU] N at randomization not reported	12 weeks/ 2 years	CBT + depression education, 41 Depression education – low intensity – 6 sessions, 104 TAU, 63 Attrition not reported Completer analysis	Beck Depression Inventory (BDI) reported as average of 8 repeated measures over follow-up, adapted CIDI depression module (adaptation	Patients with 4 or more previous episodes: Mean BDI scores (s.e.): CBT + depression education = 9.01(1.54), depression education alone = 13.08(1.11), TAU = 13.03(1.25); CBT + depression education significantly lower than TAU (effect size = 0.50, $p < 0.05$), depression education alone not significantly different than TAU CIDI depression – % reporting

				not described)	cognitive problems: CBT + depression education– 15%, TAU – 47% (not reported for depression education alone group) For patients with less than 4 previous episodes: Mean BDI scores(s.e.): CBT + depression education = 9.81(1.44), depression education alone = 10.62(0.82), TAU = 11.48(1.14); no significant differences
Scott, 1997	Brief CBT + TAU,24 [TAU, 24]	6 weeks/ 58 weeks	Brief CBT, 16/33% TAU, 8/67% No ITT	HRSD	Mean HDRS score: Brief CBT = 6.10, TAU = 10.70, significantly different, $F = 7.4$, $p < 0.01$
CBT for Residual Symptoms					
Bockington, 2005	Post-remission group CBT, 88 [Treatment as usual (TAU), 84]	8 weeks/ 2 years Acute treatment not identified	CBT, 81/8% TAU, 84/0% ITT analysis	Relapse	For patients with 5 or more previous episodes (41% of the sample): Relapse: CBT–46%, TAU –72%, significantly different, $p < .01$) For patients with fewer than 5

					previous episodes: Relapse: CBT – 63%, TAU – 59%, not significantly different, $p > 0.70$
Bockington, 2009 (Follow-up to Bockington 2005)	Post-remission group CBT, 88 [TAU, 84]	8 weeks/ 5.5 years Acute treatment not identified	CBT, 81/8% TAU, 84/0% ITT analysis	Relapse	For patients with 4 or more previous episodes (52% of the sample) Relapse: CBT –75%, TAU – 95%, significantly different, $p < .01$ For patients with fewer than 4 previous episodes: Relapse: CBT – 82%, TAU – 79%, not significantly different, $p < 0.60$
Fava, 1994	Amitriptyline + CBT for residual symptoms, 21 [Amitriptyline + Clinical management, 22]	20 weeks/ 2 years	ADM + CBT, 20/5% ADM + CM, 20/9% Survival analysis	Relapse	Relapse: ADM + CBT– 15%, ADM + clinical management – 35%, not significant, $p > 0.30$, Fisher's exact test
Fava, 1996 (Follow-up to Fava, 1994)	Amitriptyline + CBT for residual symptoms, 23 [Amitriptyline + Clinical management, 22]	20 weeks/ 4 years	ADM + CBT, 20/13% ADM + CM, 20/9% Survival analysis	Relapse	Relapse: ADM + CBT –35%, ADM + clinical management – 75%

Fava, 1998b (Follow-up to Fava, 1994)	Amitriptyline + CBT for residual symptoms, 23 [Amitriptyline + Clinical management, 22]	20 weeks/ 6 years	ADM + CBT, 20/13% ADM + CM, 20/9% Survival analysis	Relapse	Relapse: ADM + CBT –50%, ADM + CM –75%; not significantly different Mean number of depressive episodes: ADM + CBT = 0.80, ADM + CM = 1.70, significant difference, $t = 2.50$, $df = 38$, $p < 0.05$
Fava, 1998a NOT THE SAME SAMPLE AS OTHER FAVA PUBLIC-ATIONS	Amitriptyline + Enhanced CBT for residual symptoms, 23 [Amitriptyline + Clinical management, 22] Enhanced CBT included standard CBT + Lifestyle management + Wellbeing therapy	20 weeks/ 2 years	ADM + Enhanced CBT, 20/13% ADM + CM, 20/9% Survival analysis	Relapse	Relapse: ADM + Enhanced CBT– 25%, ADM + clinical management– 80%

Paykel, 1999	ADM + CBT for residual symptoms (20 wks + 2 boosters), 80 [ADM + clinical management (once every 8 wks throughout follow-up year), 78]	20 weeks/ 1 year (Acute ADM - Amytripty line or Fluoxetine continued through-out)	ADM + CBT,61 /24% ADM + CM, 66/15% ITT analysis	Relapse	Relapse: ADM + CBT – 29%, ADM + CM – 47%, significantly different, Hazard ratio for relapse (95% CI) = 0.54 (0.32 – 0.93), $p < 0.05$
Vittengle, 2009	A-CBT + C-CBT, 41 [A-CBT + clinical management, 43]	12-14 wks A-CBT, 10 sessions over 8 months C-CBT/ 16 months	C-CBT, 35/15% CM, 34/21%	Remission (>/= 6 continuous weeks no or few mild symptoms) Recovery(>/= 35 continuous weeks no or few mild symptoms), Relapse (>/= 2 weeks meeting MDD criteria)	Remission (post A-CBT): C-CBT – 97%, CM – 88%, not significantly different, $\chi2(1, N = 84) = 2.52$, $p < 0.20$ Recovery (post A-CBT): C-CBT – 84%, CM – 62%, significantly different, $\chi2(1, N = 84) = 4.20$, $p < 0.05$ All patients without remission or recovery relapsed. Of those who remitted, 40% relapsed. Of those who recovered, 25% relapsed. Relapse rates not reported by treatment group.

CBT for Relapse Prevention					
Bondolfi, 2010	Mindfulness Based Cognitive Therapy (MBCT)+ TAU, 31 [TAU, 29]	8 weeks/ 60 weeks Acute treatment not identified	MBCT + TAU, 27/13% TAU, 28/3% ITT analysis	Relapse, time to relapse	Relapse: MBCT + TAU – 29%, TAU – 34%; not significantly different, Fisher's exact test, $p > 0.70$ Median time to relapse (range), MBCT + TAU – 204(35 – 330) days, TAU – 69(15-191) days; significantly longer for MBCT + TAU, Mann-Whitney U-test, $p < 0.01$
Godfrin, 2010	MBCT + TAU, 52 TAU, 54 Sample currently in remission or recovery, having at least 3 previous episodes. No description of previous treatment.	8 weeks/ 12 months	MBCT + TAU, 34/35% TAU, 42/22% ITT analysis	Relapse, time to relapse	Relapse: MBCT + TAU – 30%, TAU – 68%; significantly different, $\chi2 = 12.5$, $df = 1$, $p < 0.01$ Mean time to relapse, MBCT + TAU – 53.7 weeks, TAU – 39.5 weeks; significantly different, $\chi2$ (n = 66) = 12.81, df = 1, $p < 0.01$
Kuyken, 2008	MBCT (prior, non-specified ADM withdrawn) withdrawal, 61 ADM (various, not specified) continued, 62	MBCT: 8 weeks + 4 booster sessions over the following year/ 15 months	MBCT, 52/15% cADM, 52/16% ITT analysis	Recurrence	Recurrence: MBCT – 47%, cADM – 60%, significantly different, $\chi2(1) = 3.32$, $p < 0.08$

Ma, 2004	MBCT, 37 TAU, 38 Sample currently in remission or recovery, having at least 2 previous episodes. No description of previous treatment.	8 weeks/ 1 year	MBCT, 36/3% TAU, 37/3% ITT analysis	Relapse	Relapse: For those with 3 or more previous episodes: MBCT – 36%, TAU – 78%, significantly different, $\chi 2(1, n = 55) = 9.89, p < 0.05$ For those with 2 previous episodes: MBCT – 50%, TAU – 20%, not significantly different, Fisher's exact test $p > 0.30$
Manicavas-gar, 2011	MBCT, 30 Group CBT, 39	8 weeks/ 12 months	MBCT, 9/70% CBT, 14/ 64% No ITT	Maintenanc e of gains (BDI)	Maintenance of gains: No loss of for MBCT or CBT ($p > 0.15$ for all comparisons) Mean BDI scores: MBCT = 18.56, CBT = 18.93, not significantly different
Segal, 2010	Citalopram discontinued + MBCT, 26 Citalopram continued, 28 Citalopram withdrawn to placebo, 30	8 months acute ADM, + 8 weeks for MBCT group/ 18 months	Citalopram discontinued + MBCT, 21/8% Citalopram continued, 21/8% Citalopram withdrawn to placebo, 24/8% ITT analysis	Relapse	Relapse: Overall – MBCT – 38%, cADM – 46%, placebo – 60%, not significantly different, $\chi 2(2,2) = 0.84, p > 0.70$ Grouped by stability of remission: Unstable - MBCT – 27%, cADM – 28%, placebo – 71%, MBCT and cADM not significantly different, $\chi 2(2,1) = 1.07, p > 0.95$ MBCT and placebo were significantly different, $\chi 2(2,1) = 6.01, p < 0.01$

					Stable - MBCT – 62%, cADM – 59%, placebo – 50%, MBCT and cADM not significantly different, $\chi2(2,1) = 0.47$, $p > 0.40$ MBCT and placebo not significantly different, $\chi2(2,1) = 0.73$, $p > 0.30$
Teasdale, 2000	MBCT, 76 TAU, 69	8 weeks/ 1 year	MBCT, 63/17% TAU, 69/0% ITT analysis	Relapse	Relapse: For those with 3 or more previous episodes: MBCT – 40%, TAU – 66%, significantly different, $\chi2(1, n = 105) = 7.10$, $p < 0.01$ For those with 2 previous episodes: MBCT – 56%, TAU – 31%, not significantly different, $\chi2(1, n = 32) = 2.03$, $p > 0.10$
Behavioural Activation					
Gortner, 1998	BA, 50 Enhanced BA (BA + automatic thought modification), 39 CBT, 48	16 weeks/ 2 years	BA, 29/ 42% Enhanced BA, 30/ 23% CBT, 38/21% Survival analysis	Relapse	Relapse: BA - 32%, Enhanced BA – 50%, CBT - 46%, not significantly different, χ^2 (2, n = 68) = 0.57

McLean, 1990	BA, nondirective psycho-therapy, Amitriptyline (11 wks), relaxation therapy, [healthy comparison group, 55] N = 196 admitted to study and randomized to 4 treatment groups, n not provide by treatment	10 weeks/ 2.25 years	121 from treatment groups available for follow-up, 38% attrition BA,33 Non-directive, 28 Relaxation, 35 Amitriptyline, 25 Healthy comparison, 45/18% No ITT	Mood variable on non-standard-ized depression questionnair e; average score across 6 follow-up points reported	Percent scoring within upper region of the healthy comparison group: BA - 63%, non-directive – 36%, relaxation – 26%, Amitriptyline – 28%, BA was significantly higher than the other three treatment conditions $\chi(3) = 13.48$, $p < 0.01$
Psychodynamic Treatment					
Barkham, 1996	8 week CBT, 16 week CBT, 8 week Psychodynamic-Interpersonal (PI), 16 week Psychodynamic-Interpersonal, 20 patients received 8 sessions, 16 received 16 sessions, 18 received CBT and 18 received PI, not	8 weeks or 16 weeks/ 1 year	Drop out not reported. Authors report missing data and different sample sizes for different analyses, but no indication of the number of completers,	BDI	All groups improved from intake to post-treatment, and intake to follow-up with t values ranging from 3.07 – 7.64 ($p < 0.01$ in all cases). No comparison of post-treatment and follow-up reported. No significant treatment by duration interaction (p values > 0.05)

	reported specifically for each of the 4 treatment groups.		or drop-out by treatment group.		
Koppers, 2011	Psycho-dynamic therapy (PDT) + ADM, PDT alone, N = 140, not reported by treatment	16 sessions/ 5 years	PDT + ADM, 25 PDT alone, , 27 63% overall attrition: 29% refused, 34% untraceable Not reported by treatment No ITT	Recurrence	Recurrence: PDT + ADM – 44%, PTD alone– 37%, not significantly different, $\chi 2 = 0.216$, $p > 0.60$
Maina, 2009	Brief dynamic therapy (BDT – 6 months) + ADM – 6 months acute, 6 months continuation, 41 ADM alone – 6 months acute, 6 months continuation, 51 Only those who had remitted at the end of 6 months were included in this study.	1 year/ 4 years	BDT, 40/2% cADM, 49/4% ITT analysis	Relapse	Relapse: BDT + ADM - 28%, ADM - 47%; significantly different, $\chi^2 = 3.53$ $df = 1$, $p < 0.05$)

Shapiro, 1995	8 week CBT, 29 16 week CBT, 30 8 week Psychodynamic-Interpersonal (PI), 30 16 week Psychodynamic-Interpersonal, 27	8 weeks or 16 weeks/ 1 year	8 wk CBT, 26/10%, 16 wk CBT, 25/17% 8 wk PI, 28/7% 16 wk PI, 24/11%	BDI	Mean BDI score: CBT = 7.15, PI = 8.25. not significantly different, $F<1$). Mean BDI score, by treatment duration: 8 wk CBT = 6.72, 16 wk CBT = 7.55, 8 wk PI = 11.39, 16 wk PI = 5.21 8 wk PI did significantly less well than the other three groups, for simple effect of treatment $p < 0.05$ for 8 wk treatment and $p > 0.30$ for 16 wk treatment. For simple effect of duration, $p > 0.80$ for CBT and, $p < 0.05$ for PI

Note . ns – not significant, RCT – randomized controlled trial, ADM – antidepressant medication, cADM – continuation antidepressant medication, BA – behavioural activation, CBASP – cognitive behavioural analysis system of psychotherapy, CBT – cognitive behavioural therapy, A-CBT – acute cognitive behavioural therapy, C-CBT – continuation cognitive behavioural therapy, IPT – interpersonal psychotherapy, MBCT – mindfulness based cognitive therapy, STPP – short term psychodynamic therapy, PDT – psycho-dynamic therapy, PI – psychodynamic-interpersonal therapy, TAU – treatment as usual, CM – clinical management, BDI – Beck Depression Inventory, BAI – Beck Anxiety Inventory, HDRS – Hamilton Depression Rating Scale, CIDI – composite international diagnostic interview, ITT – intent to treat analysis.

Risk of Bias/Quality of the Included Studies

The quality of the studies varied. Of the 30 included studies, only 5 studies reported that allocation to conditions was conducted by an independent party. Concealment of random allocation to respondents was not possible or not reported in any of the studies, whereas blinding of assessors was reported in 20 studies. Attrition ranged from 0 – 70%. See Table 4 for detailed information on the quality of the included studies. Attrition rates are reported in Table 3.

Table 4. Quality assessment of included studies

Study	Randomization	Blinding/ independent assessors	Monitoring outcomes and treatment received during follow-up	Comments
Direct Comparison: ADM vs CBT				
Blatt, 2000	Method not described. Reader referred to previous study.	Medication and placebo arms double blind	Clinical interview, and self-report HDRS, BDI, GAS, SCL-90, conducted at 6, 12, and 18 months – appear cross-sectional. Assess patients' need for further treatment at time of assessment, but do not report on receipt of treatment during follow-up	Three sites, part of the large scale Treatment of Depression Collaborative Research Program (TDCRP) study sponsored by the National Institute of Mental Health (NIMH).
Dobson, 2008	Computer generated randomization list	Evaluators were blind, medication arm was triple blind (patient, pharmaco-therapist, and evaluator were blind to ADM versus placebo)	BDI, HRSD, LIFE, and PSR completed bi-weekly for first 2 months, and at months 3, 6, 12, 13, 14, 18, and 24. LIFE used to track other treatment received (retrospectively). Patients treated or referred to treatment upon relapse.	Single centre. GlaxoSmithKline provided medications and placebo.

Evans, 1992	Method not described, reader referred to a previous publication for details.	Assessment by blind clinical evaluator (every 6 months) and by mail in self-report (monthly).	Relapse monitored by monthly BDI and HDRS every 6 months. Treatment during follow-up monitored monthly, retrospective self-report.	None.
Hollon, 2005	Method not described	Assessment by blind clinical evaluator.	Relapse monitored by HDRS, weekly for first 2 weeks, bi-weekly for next 2 months, monthly thereafter. Treatment during follow-up not permitted, monitoring not reported.	Two centres. GlaxoSmithKline provided medications and placebo.
Kovacs, 1981	Random assignment, with last three participants assigned to medication condition to offset the higher attrition expected in this group.	Assessment by blind clinical evaluator.	BDI self-report, monthly mail in survey. Treatment monitored by monthly self-report mail in survey.	Single site. Planned face-to-face data collection but found this was not feasible, so changed to self-report measures via mail out surveys.
Schramm, 2011	Computer randomization, 1:1 treatment allocation ratio. Stratified by early trauma, in blocks or variable size, to guarantee concealment.	Evaluator was blind.	HDRS (interview) and BDI (self-administered) post-treatment and at 12 month follow-upTreatment over the follow-up period is reported but measurement intervals and instrument are not reported.	Single site.

Simons, 1986	Method not described	Not described	Relapse monitored by BDI and HDRS at 1, 6, and 12 months Treatment during follow-up not reported/monitored.	This may not be a true one-year post-treatment follow-up as the ADM groups were tapered and discontinued during that one year, and the tapering schedule was not reported.
CBT in Primary Care				
Conradi, 2008	Method not described	Not described	BDI self-report and adapted CIDI depression interview by telephone every 3 months. Treatment received during follow-up not reported.	Sites not reported.
Scott, 1997	Randomization stratified by gender, severity (BDI < 30 or > 30), and chronicity (duration > 12 months or < 12 months)	"independent assessment"	DSM-III-R and HDRS by interview, BDI self-report at 3, 6, and 12 months. Treatment during follow-up not reported	Multi-site: 11 general practices

CBT for Residual Symptoms				
Bockington, 2005	Blocked randomization, stratified by study location and type of aftercare (family physician, psychiatric centre, no aftercare). Concealed assignment by independent researcher.	Interviewers and evaluators were blind.	SCID-DSM-IV(assessed by psychiatrist), HRSD(administered by research assistants) at 3, 12, and 24 months. Treatment during follow-up (type and dosage) reported retrospectively, every 3 months on Tribos/Institute for Medical Technology Assessment Self Report Questionnaire for Costs Associated with Psychiatric Illness (IMTA)*.	Multi-site trial.
Bockington, 2009	As above	As Above	As above, with additional follow-up at 36 and 66 months.	As above. This is a follow-up of Bockington 2005.
Fava, 1994	Method not described	Evaluator was blind	Paykel Clinical Interview for Depression (CID) by clinical psychologist at 3, 6, 9, 12, 15, 18, 21, and 24 months. Treatment during follow-up was monitored retrospectively at each follow-up interview (no instrument described)	Single site. All treatments delivered by one psychiatrist.

Fava, 1996	As above	As above	As above, with additional follow-up at 30, 32, 36, 42, and 48 months.	As above. This is a follow-up to Fava, 1998a
Fava, 1998b	As above	As above	As above, with additional follow-up every 6 months up to 72 months.	As above. This is a follow-up to Fava, 1998a
Fava, 1998a	Method not described	Evaluator was blind	Paykel Clinical Interview for Depression (CID) by clinical psychologist at 3, 6, 9, 12, 15, 18, 21, and 24 months. Treatment during follow-up was monitored retrospectively at each follow-up interview (no instrument described)	Single site. All treatments delivered by one psychiatrist. This is NOT the same sample as the other Fava publications above.
Paykel, 1999	Randomization stratified by centre, # of previous episodes (>/= 2 vs.< 2), length of present illness, including both index major depression and residual symptoms (>/= 1 year vs.< 1 year), and severity (global ratings of mild or moderate vs. severe or psychotic)	Evaluators were blind.	Assessment by study psychiatrist, HDRS and BDI every 8 weeks Treatment received recorded at each psychiatrist rating.	Two sites.

Vittengle, 2009	Method not reported	Evaluators were blind up to 12 months, but not at 16 and 20 months due to high cost of blind evaluations.	Weekly PSR retrospective ratings and LIFE - to assess DSM-IV Axis I psychopathology and extra-protocol treatment at 4, 8, 12, 16, 20, and 24 months after A-CBT. Administered by "highly trained clinicians"	Single site.
CBT for Relapse Prevention				
Bondolfi, 2010	Block randomization by independent researcher. Stratified by site, number of previous episodes (3 or less vs > 3), and duration since remission from last episode (0-12 months, vs 13-24 months)	Evaluators were blind	SCID-DSM-IV administered every 3 months by trained research assistants. Treatment during follow-up monitored retrospectively every 3 months, instrument not identified.	Single research team, two sites.
Godfrin, 2010	Computer-generated randomization procedure.	Evaluators were blind	HDRS, BDI, SCID-DSM-IV-TR at termination and 6 and 12 months follow-up Treatment during follow-up monitored retrospectively at 6 and 12 months follow-up interviews	Single site. Appears to be the first published report on MBCT in which the intervention was delivered by someone other than the developers – this may add to the significance of their replication.

Kuyken, 2008	Block randomization by independent statistician using computer-generated quasi-random numbers. Stratified by status (asymptomatic = HDRS < 8, partially symptomatic = HDRS>/= 8)	Evaluators were blind.	SCID-DSM-IV, retrospectively every 3 months. HDRS and BDI at 15 months follow-up – not reported as repeated measures throughout follow-up. Treatment during follow-up assessed on the Adult Service Use Schedule (AD-SUS), recording number and duration of contacts with service providers retrospectively every 3 months.	Multi-site.
Ma, 2003	Randomization stratified by severity of last episode and number of previous episodes	Evaluators were blind clinical psychologists	HDRS, BDI, SCID – DSM – III – R, every 3 months Treatment during follow-up was monitored in the tri-monthly interviews (retrospectively)	One site.
Manicavasgar, 2011	Computerized randomization by independent researcher.	Not described	BDI and BAI (self-administered) at 6 and 12 months Monitoring of treatment during follow-up not reported	Multi-site.

Segal, 2010	Block randomization, block size = 4, by independent statistician using computer-generated quasi-random numbers	Clinical evaluators were blind	HDRS (interview) bi-monthly, SCID – DSM IV if HRSD was 16 or higher Treatment over follow-up not reported.	Two sites
Teasdale, 2000	Randomized by a central independent allocator, stratified by recency of recovery from last episode and number of previous episodes. Randomized by strata within each site.	Clinical evaluators were blind	HDRS, BDI, SCID-DSM-III-R, bi-monthly Treatment during follow-up period monitored retrospectively at bi-monthly interviews	Three sites
Behavioural Activation				
Gortner, 1998	Matched randomization, variables not reported.	Not reported	LIFE to assess DSM-III-R psychopathology over previous 6 months, HDRS, and BDI at 6, 12, 18, and 24 months Report that return to treatment was considered a relapse, but do not report instrument.	None.

McLean, 1990	Not described	Not described: evaluation was self-report	Non-standardized depression questionnaire administered at 6 time points during follow-up (interval, and method of administration not indicated). The questionnaire assessed only the last 7-10 days retrospectively. Report monitoring frequency of seeking additional help, and hospitalization, but do not report instrument or intervals.	None.
Psychodynamic Treatment				
Barkam, 1996	Method not described	Evaluation by self-report	BDI, SCL-90 at 3 months and 1 year Treatment during follow-up not reported.	Three sites
Koppers, 2011	Method not described.	Not described, evaluation conducted by study author, no mention of blind.	CIDI, HDRS reported only at 5 years. Treatment during follow-up not reported.	None.

Maina, 2009	Randomized by study recruiter who drew one of two coloured balls from a bag, each ball was assigned to one of the treatment options.	Evaluators were blind.	HDRS 24 and 48 months. LIFE every 6 months Treatment during follow-up not reported.	Single site.
Shapiro, 1995	Method not reported	Not reported. Evaluation by self-report.	BDI, SCL-90 at 3 months and 1 year. No restrictions on treatment during follow-up, treatment not monitored.	None.

*TICP, Hakkart-van Roijen, van Straten, Donker, &Tiemens, 2002

Note . ADM – antidepressant medication, CBT – cognitive behavioural therapy, BDI – Beck Depression Inventory, BAI – Beck Anxiety Inventory, HDRS – Hamilton Depression Rating Scale, CIDI – composite international diagnostic interview, SCL-90 – Symptom Checklist-90 item, LIFE – Longitudinal Interval Follow-up Evaluation, DSM – Diagnostic and Statistical manual, SCID – Structured Clinical Interview for DSM Disorders, PSR – Psychiatric Status Rating, GAS – Global Adjustment Scale.

Clinical Outcomes at Follow-up

The narrative review that follows is organized by treatment modality including antidepressant medications (ADM), cognitive behavioural therapy (CBT) including applications in primary care and for the specific purposes of post-remission and relapse-prevention treatment, behavioural activation (BA), interpersonal psychotherapy (IPT), and psychodynamic therapies.

Antidepressant Medications

Antidepressant medications are most often studied over the long-term in continuation and maintenance studies, and less often in post-treatment follow-up designs. The continuation and maintenance literature has been reviewed elsewhere (e.g. Kaymaz et al., 2008) and was presented in the overview of reviews in Study 1. Consequently, studies of this design are not included in Study 2. This review did not find any post-treatment follow-up studies (other than continuation and maintenance studies) focused solely on ADM. There are a number of comparative studies looking at the relative relapse rates between ADM and CBT (e.g. Blatt, Zuroff, Bondi, & Sanislow, 2000; Dobson et al., 2008; Hollon et al., 2005a) or psychodynamic therapies (Koppers, Peen, Niekerken, Van, & Dekker, 2011; Maina et al., 2009), which are reported below. These studies suggest that relapse rates are relatively high after antidepressant medication treatment is terminated. Relapse post- ADM treatment is roughly 50% within one year. More detail about the long-term outcomes post- ADM treatment is provided next, in the section on comparison studies.

Direct Comparison: Antidepressant Medications and CBT

Six studies directly compared CBT to antidepressant medications (Blatt et al., 2000; Dobson et al., 2008; Evans et al., 1992; Hollon et al., 2005a; Kovacs, Rush, Beck, & Hollon, 1981; Simons, Murphy, Levine, & Wetzel, 1986). These studies indicate that regardless of treatment type, relapse occurs for a large number of people once they stop treatment. Five studies found that CBT results in lower relapse rates over the long-term than does discontinued antidepressant medication (Dobson et al., 2008; Evans et al., 1992; Hollon et al., 2005; Kovacs et al., 1981; Simons et al., 1986) and one found no difference (Blatt et al., 2000).

Kovacs and colleagues (1981) published what appears to be the first long-term post-treatment follow-up for CBT. They included a one-year naturalistic follow-up following their acute phase trial of 12 weeks of CBT compared to 12 weeks of Imipramine (10 weeks at maximum dose, 2 weeks taper and discontinuation). Treatment was delivered by trained therapists and quality was monitored. They also controlled for extra-protocol treatment during the follow-up year. All of their follow-up measurement was by self-report mail-in survey, as their initial plan for face-to-face interviews turned out not to be feasible. They found that 39% of the CBT group relapsed as compared to 65% of the ADM group.

Simons et al. (1986) compared CBT to nortriptyline and included CBT plus placebo and CBT plus nortriptyline comparison groups. Treatment lasted 12 weeks at which time CBT was terminated and medications were tapered and discontinued. The tapering schedule is not reported. At one-year follow-up, the authors did not find a

significant difference across the four groups ($\chi 2 = 6.55$, $p < 0.08$). The authors then pooled the groups in two ways. In one comparison, they looked at relapse rates for groups who had received ADM (the nortriptyline group and the CBT + nortriptyline group) versus groups who had not received ADM (CBT and CBT + placebo). They found that the group who had received ADM had a higher relapse rate (52%) than those who did not receive ADM (19%), $p < 0.05$. They then compared patients who had received CBT (the CBT group, the CBT + placebo group, and the CBT + nortriptyline group) versus those who had not received CBT (the nortriptyline group). Simons et al. found that those who had received CBT had a significantly lower relapse rate (28%) than those who had not received CBT (66%), $p < 0.05$.

Evans et al. (1992) report similar findings comparing CBT to imipramine. They included comparison groups with a continued ADM arm and a combined CBT + ADM arm. Their study reported relapse rates for CBT (21%), ADM discontinued (50%), ADM continued (32%), and CBT + ADM discontinued (15%). They found no significant difference between CBT alone and CBT + ADM, which led them to report the remainder of their results pooling these two groups. They found that the pooled CBT group had significantly lower relapse rates than the discontinued ADM group ($\chi 2(1) = 4.05$, $p < 0.05$). They also found that the continued ADM group did not differ significantly from either the discontinued ADM group or the pooled CBT group ($p > 0.15$).

Blatt et al. (2000) report an 18-month follow-up to the NIMH sponsored Treatment for Depression Collaborative Research Program (TDCRP; Elkin, 1994) study. Theirs is the one follow-up that found no difference between CBT and ADM over

the long-term. Specifically, they found no significant difference between placebo withdrawn ADM, continued ADM, CBT, and IPT groups on their aggregate measure, at 18 months follow-up. They do not report relapse or recurrence, nor do they report scores on standardized measures of depression (e.g. HDRS, BDI, MDRAS). These measures were collected, but were then combined to create an aggregate maladjustment score. They do not report this aggregate score by treatment group, only stating that the differences were not significant. This makes it difficult to compare their results with those of other studies. Blatt et al. did find differences in satisfaction with treatment and life adjustment. At 18-month follow-up, those in the CBT and IPT conditions were more satisfied with their treatment, and reported that the treatment had had a more positive impact on their life adjustment than the two ADM groups.

Hollon et al. (2005) conducted a 3-year study with follow-up of adults treated with CBT or ADM. They followed patients who responded to acute treatment in year 1, with either CBT (16 weeks) or medication (continued through years 1 and 2), and who remained in remission at the beginning of year 2. At the beginning of year 3, they withdrew the ADM group from their medication and continued to follow both groups for another 12 months. The authors report that in the 12 months following acute treatment, patients withdrawn from CBT were less likely to relapse than were patients withdrawn from medication (31% versus 76%). Further, patients withdrawn from CBT were not more likely to relapse than patients who continued taking medication over one-year follow-up (31% versus 47%). During the final year of the naturalistic follow-up, when continuation medication had been withdrawn, Hollon et al. report on the 40 patients who began this period free of relapse. This included 20 who had received prior

CBT, 14 who had received continued medication (now discontinued), and 6 from the placebo group. Twenty-five percent, (5/20) of the recovered patients who had received 16 weeks of CBT had a recurrence during year 3, whereas 50% (7/14) of the recovered patients withdrawn from continuation medication had a recurrence during this time.

Dobson et al. (2008) conducted a 2-year follow-up of a study by Dimidjian et al. (2006), in which participants were randomly assigned to 16 weeks of ADM, CBT, or behavioural activation (BA). Dimidjian et al. included a pill placebo arm in their acute treatment study, but Dobson et al. do not report on this group in their follow-up – probably because they would have been offered treatment by the time this follow-up was conducted. Of the 188 participants assigned to treatment in the original study, 106 remitted by the end of acute phase treatment and comprise the follow-up sample. All treatment groups terminated treatment after the acute phase (16 weeks), except for the ADM group. Half of this group continued ADM treatment for the first year of follow-up, and were then withdrawn from treatment for the second year of follow-up. The other half of the ADM group was withdrawn to placebo at the beginning of the second year. They were withdrawn from this placebo at the end of the second year. Dobson et al. use the consensus definitions (Frank et al., 1991) for relapse and recurrence, with relapse being defined as an HDRS score of 14 or greater or Psychiatric Status Ratings (PSRs) of 5 or greater (1 = *absence of symptomatology*, 5 = *meets MDD criteria*, 6 = *definite and severe presence of symptomatology*) for 2 successive weeks during the 1st year of follow-up. Recurrence defined by the same criteria during the second year of follow-up. During the first year of follow-up, prior CBT was significantly superior to the placebo withdrawal group $\chi 2(1, n = 51) = 5.30, p < 0.05$. BA showed a non-significant trend in

the same direction $\chi 2(1, n = 51) = 2.81, p < 0.10$. There was no significant difference in relapse rates between cADM and placebo withdrawal $\chi 2(1, n = 49) = 0.97, p > 0.30$. During the second year of follow-up, recurrence rates were 24% for prior CBT, 26% for prior BA, and 52% for prior ADM. Prior psychotherapy (CBT and BA) showed a non-significant trend compared to prior cADM $\chi 2(1, n = 46) = 3.58, p < 0.07$. The authors conclude that both CBT and BA have significant relapse prevention effects, that there is no significant difference between the effects of these two types of therapy, and that the relapse prevention effects of these therapies is at least as powerful as the effects of continued antidepressant treatment.

CBT in Primary Care

CBT may be delivered in primary care settings and two trials included in this review take place in that setting. Unfortunately, both of these studies are of less than optimal quality. Scott et al. (1995) report a 12-month follow-up to their trial of brief CBT compared to treatment as usual (TAU) in primary care. All but one participant in each of the CBT and the TAU groups were on ADM at some point in the treatment and/or follow-up and no more detail is provided on this issue. At 12-month follow-up, only 8 of 24 participants remained in the control group and 16 of 24 remained in the intervention group. Of these, 4 in the control group relapsed as compared to only 1 in the intervention group. Mean HDRS score was significantly lower at follow-up in the intervention group ($M = 6.1, SD = 4.3$) than in the control group ($M = 10.7, SD = 6.5$), $F = 7.4, p < 0.01$. This result should be interpreted with caution due to small sample size, high attrition rates, and the cross-sectional nature of the follow-up measurement, which

could have excluded relapse that occurred and possibly remitted between measurement points.

In another primary care study, Conradi et al. (2008) randomly assigned patients to TAU, depression education, or depression education plus CBT. This study, also, appears to be of poor quality, failing to report on most of the criteria required to assess study quality. This study reports only self-report BDI scores, as well as the presence or absence of discreet symptoms as indicated in the Composite International Diagnostic Interview (CIDI). It does not report relapse rates. With regards to BDI scores at 2-year follow-up, the authors found no significant difference between the TAU and psycho-education groups. However, they did find lower scores for the CBT group compared to the TAU group for all patients (effect size = 0.41, $p < 0.05$). When they divided their sample into two groups based on number of previous episodes, they found no difference at two-year follow-up for the group with three or fewer episodes (effect size = 0.31, $p > 0.30$, but they did find a difference for the group who had had four or more previous episodes (effect size = 0.50, $p < 0.05$). They also found that, for the group who had had four or more episodes, those in the TAU group reported having cognitive problems (on the CIDI) during 47% of the 2-year follow-up period (equivalent of 11.3 months). The CBT group with four or more previous episodes reported having cognitive problems during only 15% of the 2-year follow-up period (equivalent of 3.6 months). This difference was significant ($z = 2.328$, $p < 0.05$). The authors suggest that it is valuable to establish the effectiveness of CBT in primary care settings where the vast majority of depressed patients are treated. They advise that their findings require replication.

CBT for Residual Symptoms

CBT has also been studied as a treatment for residual symptoms (sub-threshold symptoms that were not resolved during acute phase treatment), with the specific aim of relapse-prevention among patients who have already reached remission during acute treatment. Eight included studies examine this variant of CBT, six with individual and two with group treatment.

Fava, Grandi, Zielezny, Canestrari, and Morphy (1994) randomly assigned 40 patients with recurrent depression, who had been treated to remission with ADM to receive either CBT for residual symptoms or clinical management for 20 weeks. Both groups received 10 bi-weekly 30-minute sessions with the CBT group receiving the CBT intervention and the clinical management group attending appointments with their doctor but receiving no additional intervention. The content of these clinical management sessions is not described. During these 20 weeks, medications were withdrawn on a tapered schedule from both groups. Fava et al. (1994) report on follow-up at 2 years, during which time no ADM were administered unless relapse occurred. At follow-up, the CBT group had a lower relapse rate (15%) than the clinical management group (35%). However, this difference was not statistically significant, Fisher's exact test; log-rank test, $\chi 2 = 2.27$, $df = 1$, $p > 0.13$.

The above sample was followed-up repeatedly and outcomes were reported at 4 years (Fava, 1996) and 6 years (Fava et al., 1998b) post-treatment. At 4 years, 35% of the CBT group had relapsed as compared to70% of the clinical management group. This difference was statistically significant, $\chi 2 = 4.69$, $df = 1$, $p < 0.05$. At 6 years, 50% of the

CBT group and 75% of the clinical management group had relapsed and the difference between the two groups was no longer significant, though it did continue to approach significance (log rank test, $\chi 2 = 3.45$, $df = 1$, $p < 0.06$). At this time, the authors also examined frequency of relapse during the follow-up period, and here the difference was significant. Individuals in the CBT group had a significantly lower number of relapses during the follow-up period ($M = 0.80$, $SD = 0.95$) than did those in the clinical management group ($M = 1.70$, $SD = 1.30$). Fava et al. (1998b) conclude that their study provides support for sequential treatment, in which ADM to remission is followed by CBT for residual symptoms. They caution that their study is preliminary, the sample is small, and that replication is required.

Fava, Rafanelli, Grandi, Conti, and Belluardo (1998a) conducted another study following the same methods as their 1994 study, in which they randomly assigned 40 patients with recurrent depression, who had been treated to remission with ADM, to receive either enhanced CBT for residual symptoms or clinical management for 20 weeks. The enhanced CBT condition included standard CBT with additional lifestyle management and wellbeing therapy components. Lifestyle management techniques were derived from approaches found to be effective in clinical cardiological studies and aimed to make changes to accommodate such factors as an individual's life stress, interpersonal friction, excessive work, and inadequate rest. Wellbeing therapy components were based on Ryff and Singer's (1996) conceptual model of well-being and aimed to change beliefs and attitudes detrimental to well-being, stimulate awareness of personal growth and recovery from affective illness, and reinforce behaviour promoting wellbeing. Medications were withdrawn on a tapered schedule from both

groups. Fava et al. (1998a) report on follow-up at 2 years, during which time no ADM were administered unless relapse occurred. At follow-up, the CBT group had a significantly lower relapse rate (25%) than the clinical management group (80%). Mean survival times, i.e. the average time to relapse, were 91.8 (22.4) weeks for the CBT group and 62.2 (26.6) weeks for the clinical management group ($t_{38} = 3.81$, $p < 0.01$). Though most studies report median time to first relapse, this study reports the mean and does not describe how it was calculated. It may be that it is the mean time to first relapse for those who did relapse in each group, but this is not reported.

Paykel et al. (1999) conducted a study similar to Fava et al. (1998a, 1998b). Their population was different, in that they included 158 patients who had only partially remitted at the end of a course of ADM treatment and who had continued to exhibit residual symptoms over a range of 2-18 months. These participants were randomized to receive clinical management alone, or clinical management plus CBT for 16 sessions over the course of 20 weeks. Paykel et al. report on follow-up at one-year post-treatment. The cumulative relapse rate at follow-up was significantly lower for the CBT group (29%) than for the clinical management group (47%).

Bockington et al. (2005) provided group CBT or TAU to 187 patients in remission after various types of acute treatment. They found an interaction for relapse rates between treatment group and number of previous episodes of depression. For those in TAU, relapse rates increased with number of previous episodes. CBT, however, had the greatest effect for those with 5 or more previous episodes. For these individuals (41% of the sample), relapse rates for the CBT group were significantly lower (46%) than in the TAU group (72%; $p < 0.01$). This difference in relapse rates was evident by 3

months post-treatment and remained steady over 2 years of follow-up. In addition to considering whether or not people relapsed, these authors also examined the severity of relapse (defined by mild, moderate, and severe ranges on the HDRS), and the number of times individuals relapsed over the 2 year follow-up. The interaction of treatment condition and previous episodes was significant for severity, $F(1, 168) = 3.79$, $MSE = 1.37$, $p < 0.05$. In the group who had five or more previous episodes, those who received CBT not only had a lower relapse rate than those who received TAU, they also had less severe relapse when it did occur. The interaction of treatment condition and previous episodes was also significant for the number of relapses experienced over the follow-up period, $F(1, 166) = 3.94$, $MSE = 1.04$, $p < 0.05$. Among those who had five or more episodes, those who received CBT experienced fewer episodes of relapse over the follow-up period compared to those with five or more episodes who received TAU.

Bockington et al. (2009) conducted a further follow-up with the above sample at 5.5 years post-treatment and reported on relapse rates and time to first relapse for the CBT compared to the TAU group. In this analysis, they found that for those who had four or more previous episodes, the CBT group had a lower relapse rate (75%) than the TAU group (95%). Their survival analysis showed that CBT had an increasing protective effect over time for those with 4 or more episodes. That is to say, for individuals with four or more previous episodes, those who received CBT had lower relapse rates than those who received TAU, and this difference increased over time. Bockington et al. (2009) also re-examined severity of relapse, number of relapses during the follow-up period, and percentage of time free of depression during follow-up in patients who did relapse. There was no significant difference, either overall or when the

treatment condition by number of previous episodes interaction was considered. For these interactions, the differences present at 2 years follow-up (Bockington et al., 2005) were no longer evident at 5.5 years follow-up (Bockington et al., 2009).

Most recently, Vittengle, Clark, and Jarrett (2009) studied the effects of continuation-phase CBT for 84 individuals who had responded (but not necessarily remitted) to acute-phase CBT (20 individual sessions, 50-60 minutes each). These participants were randomized to receive 8 months of continuation CBT (10 sessions, 60-90 minutes each) or clinical management (10 evaluation sessions, same schedule as C-CBT). The authors report follow-up at 16 months post-treatment. Weekly psychiatric status ratings (PSRs) of DSM-IV MDD (on a scale of 1-6) defined remission and recovery, respectively as >/= 6 and >/= 35 continuous weeks of a PSR of 1 (*no symptoms*) or 2 (*one or two mild symptoms*). Relapse was defined as >/= 2 weeks of a PSR of 5 (*meets MDD criteria*) or 6 (*meets MDD criteria with severe impairment and/or psychosis*). Patients who had not reached full remission or recovery during acute-phase treatment all relapsed during follow-up. However, most of those who had remitted and recovered did not relapse over the follow-up period. Of this group, 40% of those who had remitted had a relapse and 25% of those who had reached recovery had a relapse during the follow-up period. There were no differences between the C-CBT and the clinical management groups on rates of relapse/recurrence over the follow-up period. Recall the review on the relapse-prevention effects of C-CBT, by the same first author (Vittengle et al., 2007), presented previously in the overview of reviews, that reported significant relapse prevention effects for C-CBT. Although this study appears to contradict the previous review, it may be that the population included here is

significantly different, in that it includes those who did not reach remission. The aim of this study is, in fact, to determine not only whether those who reached remission on A-CBT are better able to maintain it with a course of C-CBT, but also to determine whether those who did not achieve remission on A-CBT can reach remission and maintain it if they also receive C-CBT. In addition, the Vittengle et al. (2007) review did not describe the duration or frequency of the C-CBT delivered in the studies they reviewed. This may be another way that the Vittengle et al. (2009) study differs from those in the review. Given Vittengle et al.'s (2009) finding regarding differential relapse rates for those who had reached recovery and those who had not, this study speaks to the importance of treating to recovery and not just reduction of symptoms, or even remission.

One specific variant of CBT is the Cognitive-Behavioural Analysis System of psychotherapy (CBASP). The treatment is manualized, highly structured, and integrates behavioural, cognitive, and interpersonal strategies to teach patients interpersonal problem solving skills. It is tailored for the treatment of chronic forms of depression, particularly early-onset depression and it focuses on the problems that result from inhibition of maturation during early childhood by using the therapeutic relationship in a personal disciplined way. Although CBASP is effective in acute, continuation, and maintenance phases of treatment, there has been little well designed research on its relapse-prevention effects post-treatment (Klien et al., 2004). One long-term post-treatment follow-up study of CBASP was located. In this study, comparing CBASP to IPT, Schramm et al. (2011) found no significant difference in mean depression score at one-year follow-up (Mean BDI(*SD*) = 12.92(11.83) and 18.66(14.53) respectively, *F* =

0.67, *ns*). The authors did not report on relapse. There is not enough evidence to draw conclusions about the post-treatment relapse prevention effects of CBASP or IPT.

CBT for Relapse Prevention: Mindfulness Based Cognitive Therapy

Mindfulness Based Cognitive Therapy (MBCT) is a manualized variant of CBT. It includes explicit relapse-prevention components, and it is suggested that the mechanism of action may be different from that in traditional CBT (Hollon & Ponniah, 2010). It is discussed here, as a specific relapse-prevention variant of CBT, separate from but related to CBT for residual symptoms. Seven studies report on long-term post-treatment outcomes for MBCT in comparison to a variety of other treatment and control groups.

Teasdale, Segal, Williams, Ridgeway, Soulsby, and Lau (2000) compared MBCT to TAU and found a significant interaction between number of previous episodes and treatment condition, Wald (1) = 4.23, $p < 0.05$. For those with 3 or more previous episodes, 40% of the MBCT group experienced relapse over one year, as compared to 66% of the TAU group. This difference was significant, $\chi 2(1, n = 105) = 7.10$, $p < 0.01$. For the group with 2 episodes, 56% of the MBCT group experienced relapse compared to 31% of the TAU group. This difference was not significant $\chi 2(1, n = 32) = 2.03$, $p > 0.10$. When comparing these two groups, Teasdale et al. found that those with 3 or more episodes were significantly younger when they experienced their first episode of depression ($M = 25.0$, $SD = 9.84$) than those who had only 2 episodes ($M = 33.38$, $SD = 8.65$). The authors suggest that the two groups may represent distinct patient populations.

Ma and Teasdale (2004) set out to replicate the findings of Teasdale et al. (2000). Their findings were consistent, with MBCT having significant relapse prevention effects for those with 3 or more previous episodes (MBCT = 36% relapse, TAU = 78% relapse, $p < 0.05$) but not for those with 2 previous episodes (MBCT = 50% relapse, TAU = 20% relapse, $p > 0.30$). Ma and Teasdale investigated 3 proposed explanations for this difference. First, they found that MBCT was most beneficial for individuals whose onset of depression was not associated with significant life events (Fisher's exact test $p < 0.05$). Life events were assessed at follow-up interviews. They were rated by the interviewer, who was blind to treatment condition, and by an independent rater as "a significant life event that is more than likely to bring about serious depressed feelings in an average person" or as an event of "borderline significance". These ratings were validated by correlation with the Social Readjustment Rating Scale (SRRS; Holmes & Rahe, 1967). There was excellent agreement between raters and between ratings and SRRS scores. As a second explanation, they found that individuals who had 3 or more previous episodes were significantly more likely to have experienced more adverse parenting styles during their early years (assessed using the Measure of Parenting Style; Parker, Roussos, Hadzi-Pavlovic, Mitchel, Wilhelm, & Austing, 1997) than those who had 2 previous episodes. Third, consistent with Teasdale et al. (2000), Ma and Teasdale found that in patients with 2 previous episodes, mean age of onset (37.5, $SD = 8.0$) was significantly older than in those with 3 or more previous episodes (28.4, $SD = 12.9$). The authors concluded that MBCT may be less effective when depression is associated with significant life events and, in addition, the groups that show differential response to MBCT based on number of previous episodes may

come from different base populations with regards to age of first episode and adverse early life experiences.

Kuyken et al., (2008) compared MBCT to maintenance antidepressant medication and found a significant difference between the two groups at 15 months post-treatment follow-up. Specifically, the MBCT group relapsed at a lower rate (47%) than the maintenance ADM group (60%). The MBCT group also experienced significantly better resolution of residual symptoms and higher quality of life at 15 months follow-up.

Another study by Bondolfi et al. (2010) reported post-treatment follow-up over 14 months for MBCT compared to TAU. Bondolfi et al. found that relapse rates at the end of 14 months were not significantly different (MBCT = 29%, TAU = 34%, $p > .70$). However, they did find an advantage for the MBCT group in terms of time to relapse, with the MBCT group having a median of 204 days to relapse and the TAU group having a median of 69 days to relapse. Although not discussed by the study authors, it is important to consider that the ability to delay the occurrence of relapse may also translate into fewer episodes of depression. Given the chronicity of MDD, how long it takes before a person relapses, or how many relapses a person experiences may, in the long-term, be a more clinically meaningful outcome than simply whether or not a person experiences relapse.

Segal et al., (2010) compared 8 week MBCT to ongoing maintenance antidepressant medication, with a placebo group (switched from ADM to placebo). They found, similar to Bondolfi et al. (2010), that over 18 months of follow-up there were no overall group differences. Segal et al. went on to provide sub-group analysis. They

compared stable versus unstable remitters, with stable remitters being those who had an HDRS score of 7 or lower across the 5 months between remission and the beginning of the follow-up period whereas unstable remitters being those who met the same threshold, but had occasional elevated scores across this time interval. For the stable group, MBCT (59%) was not significantly different from continued medication (62%) or from placebo (50%) For the unstable remitters, the active treatment groups were at a significant advantage over the placebo withdrawn group. MBCT and continued ADM did not produce significantly different relapse rates (27% and 28% respectively), and these two treatment groups had significantly lower relapse rates than the placebo group (71%).

Godfrin and van Heeringen (2010) appear to be the first authors to publish a report on the efficacy of MBCT in which the intervention was *not* delivered by the developers of the program. This is a significant contribution to the previously reported replication studies. Godfrin and van Heeringen only included patients who had three or more previous episodes of depression and found results consistent with the existing literature. They followed patients for one year after treatment termination and found that 30% of the MBCT group relapsed as compared to 68% of the TAU group ($\chi 2 = 12.5$, $df = 1$, $p < 0.01$). They also found that MBCT increases time to first relapse. The MBCT group had a mean of 53.7 weeks to first relapse and the TAU group had a mean of 39.5 weeks to first relapse, $\chi 2$ ($n = 66$) = 12.81, $df = 1$, $p < 0.01$.

Finally, in a study comparing group MBCT with group CBT, Manicavasgar, Parker, and Perich (2011) found that both MBCT and traditional CBT for relapse prevention produced similar gains (measured using self-administered BDI) and that

these gains did not diminish over 6 or 12 month follow-up. They also reported that, immediately post-treatment, as well as at 6 and 12 months post-treatment, there was no significant difference in mood rating (using the BDI and BAI) for the two treatment groups. It appears that MBCT does have a significant relapse prevention effect, but it is not clear whether it offers an advantage over more traditional CBT approaches. Additional well-designed trials are necessary to address this question.

Behavioural Activation

Three studies examined relapse-prevention effects for behavioural activation (BA). McLean and Hakstian (1990) conducted a 2.25 year post-treatment follow-up to an RCT, in which 196 participants were randomized to BA, relaxation therapy, non-directive psychotherapy, or pharmacotherapy. This study included a non-depressed healthy comparison group at all measurement points. The authors did not report on relapse or recurrence rates, but included measures of cognitive functioning, coping, personal activity, social activity, somatic indicators, average satisfaction, and positive mood. The BA group's mood ratings were significantly better than the other three treatment conditions, $\chi 2(3) = 13.48$, $p < 0.01$. The means were as follows: BA group ($M = 72.2$), non-directive psychotherapy ($M = 67.1$), relaxation therapy ($M = 65.4$), and pharmacotherapy ($M = 67.6$). The BA group was close to the healthy comparison group for mean mood ratings (healthy comparison group: $M = 78.2$). The authors also report that BA means were higher than the other treatment groups, indicating better social functioning and personal activity. A weakness of this study is that the outcome measure is not well described, and not easily compared to other studies, or to more commonly reported outcomes like relapse rates or cut off scores on standardized measures of

depression. Fifty-seven percent of participants pursued further treatment. The study authors' analysis of rates of additional professional help and hospitalization found that these rates did not differ by treatment group, suggesting that this variable should not have affected the comparative results.

In another study, Gortner, Gollan, Dobson, and Jacobson (1998) found that BA was just as effective as CBT post-treatment, and at 6 months and 2 years follow-up. Gortner et al.'s follow-up data indicate that there was no significant difference in relapse rates between BA (behavioural activation only), enhanced BA (behavioural activation plus automatic thought modification), and complete CBT at 6, 12, 18, or 24 months post-treatment. This was the case for three different measures of relapse including cut off scores on the BDI, the HDRS, and the LIFE-II. At the two-year follow-up, 35% of the BA group had relapsed compared to 47% of the CBT group. This difference was not significant, $\chi 2$ (2, n = 68) = 0.57, *ns*.

Recall that Dobson et al.'s (2008) more recent, and methodologically high quality study, found an advantage for BA over discontinued medication, but no difference between BA and CBT with regards to relapse rates over 3 years. These results were reported previously under comparative studies.

Interpersonal Psychotherapy

IPT was included in two follow-up studies of the comparative effects between treatments (Blatt et al., 2000; Schramm et al., 2011). As previously reported, neither of these studies indicate that IPT prevents relapse after treatment is terminated.

Psychodynamic Therapy

Psychodynamic therapy describes a wide range of treatments. Studies often lack sufficient detail about the therapy offered for the reader to make an informed comparison between them. The long-term post-treatment outcome data on psychodynamic treatments are also not of high quality, overall, so this literature should be interpreted with caution. Four follow-up studies of various forms of psychodynamic treatments are included in this review.

Shapiro, Rees, Barkham, Hardy, Reynolds, and Startup (1995) compared psychodynamic-interpersonal treatment (PI) to CBT and examined long-term outcomes for 8 week and 16 week variants of each. The only effect they found was a treatment x duration effect, with those who received 8 sessions of PI doing less well than the other three groups at follow-up. For example, on the BDI, the 8 session PI group mean (11.39) was significantly lower than the 8 session CBT group mean (6.71; $F(1, 46) = 3.97$, $p < 0.05$). This was only evident in a measure of how well gains were maintained, and was not evident in relapse rates, which showed no difference. Shapiro et al. report that only 6/103, or 5.8% of their sample, including all 4 treatment groups, relapsed over the course of one year. This is surprising, given what we have learned to expect at one-year follow-up, with well-designed studies finding rates of 30 – 40% relapse at one-year follow-up to CBT (Dobson et al., 2008; Hollon et al., 2005). Shapiro et al. do not provide a definition of relapse, only reporting that BDI and SCL-90 measures were administered. Therefore, it is difficult to interpret this finding and to compare it to the rates reported in other studies.

Barkham et al. (1996) conducted a replication study of Shapiro et al.'s (1995) study, with the same interventions and the same duration of treatment and follow-up. They found that their patients' improvement was equally as large as those in the Shapiro et al. study at treatment termination, but that these gains were not as well maintained at 1-year follow-up. Barkam et al.'s reporting is less than thorough, however, and there is no specific analysis reported to support this statement about maintenance of gains.

Maina et al. (2009) conducted a 4-year follow-up study in which they found that patients who had brief dynamic treatment (BDT) plus ADM were less likely to relapse over a 48 month follow-up than were patients who had received ADM only. Active treatment occurred over one year and the 48 month follow-up was treatment-free. Among those who remitted during treatment, 53% of the BDT + ADM group relapsed over the 4-year follow-up, compared to 72% of the ADM only group ($\chi 2 = 3.525$, $df = 1$; $p < 0.05$). This finding is promising but requires replication.

In another recent study, Koppers et al. (2011) provide 5-year follow-up on an RCT comparing 16 sessions of psychodynamic therapy (PDT) alone versus PDT + ADM. They report no difference in relapse rates at 5 years (PDT = 37%, PDT + ADM = 44%). However, the authors were able to follow-up with only 37% of the treatment responders, as 29% refused to complete follow-up and 34% could not be located. In addition, receipt of other treatments during the follow-up period was not reported in this study.

Study 2 Discussion

Summary of the Main Results

This study comprises a systematic narrative review examining long-term, post-treatment outcomes of effective acute treatments for adult unipolar depression. In contrast to existing reviews, this review examined relapse-prevention effects, or maintenance of gains, over the long-term and across treatments. Previous reviews have focused mainly on outcomes immediately post-treatment, or on the relapse prevention effects of one specific treatment or comparison.

The evidence indicates that relapse is a common occurrence after treatment is terminated. Depending on the treatment and on the length of follow-up, half or more of those who terminate treatment can expect to experience at least one relapse. It appears that CBT, as an acute treatment, or a continuation treatment for residual symptoms or for relapse prevention (specifically MBCT), provides a significant relapse-prevention effect post-treatment. The Hollon et al. (2005) study is a well-designed, strong example of the relapse-prevention effects of acute phase CBT, and is cited as an exemplar in other reviews and overviews (Butler et al. 2006; Hollon et al., 2006). BA appears to have a similar relapse-prevention effect to CBT, but the body of evidence is small at this time. ADM has a relatively high relapse rate post-treatment, with higher rates of relapse for those who discontinue medication than for those who remain on continuation antidepressant treatment. There is insufficient evidence to support the relapse-prevention effects of IPT and psychodynamic therapies.

With regards to potential moderating variables, the results indicate that the chronicity of depression appears to influence the effectiveness of CBT over the long-term. Persons having 3 or more previous episodes experience the greatest benefit from this treatment and, in the case of MBCT, those experiencing fewer than 3 previous episodes appear not to benefit more from this intervention compared to TAU. Two studies suggest that these two groups may actually represent distinct clinical populations. Therefore, it is possible that in choosing a maintenance, or relapse-prevention treatment, practitioners might consider variables including age of onset and adverse early life experiences, with earlier age and presence of adverse early life events suggesting the best fit for MBCT. It is possible that other variables such as severity or duration of index episode also influence the effectiveness of treatment over the long-term. CBT for residual symptoms appears to prevent relapse similarly whether the acute treatment was CBT or ADM. We know that many individuals first seek treatment for depression from their primary health care providers and that the first line approach to treating depression is often antidepressant medication. Given that ADM has a higher relapse rate than CBT post-treatment, it is encouraging to see that individuals treated to remission with ADM stand to benefit just as well as those who reached remission on CBT, when they go on to receive continuation CBT or MBCT for relapse prevention.

Attrition

Drop-out rates varied widely across studies, ranging from 0% to almost 70%. Partly due to the high number of between-treatment comparisons reviewed here, it is unclear whether drop-out rates are higher for some treatments than for others. Another potentially confusing factor is that, when reporting intent to treat (ITT) sample size,

most studies report the number of acute phase responders who were randomised to a continuation treatment. However, this is not always the case. Some studies report their initial *n* as the number of participants enrolled in the acute phase trial. On occasion studies simply report 0% drop-out, which is unlikely, and may mean that they are not reporting drop-outs and in fact conducting ITT analysis. Alternatively, this may mean that they are not reporting drop-outs and are reporting only the number of completers who were randomized to each condition.

Future Studies

Future studies should consider a number of methodological and analytical practices to improve our understanding of what happens after treatment ends. Moderators that may interact with treatment efficacy, or may indicate one treatment over another for individuals from different populations, should be measured and controlled for. A common metric, such as % relapse per month, should be reported to aid in comparisons of relapse rates across studies with follow-up of varying duration. Vittengle et al. (2009) provide an example of this method of reporting. Similarly, survival curves such as those reported in the well-designed Dobson et al. (2005) study show us that relapse is fairly orderly over time, with no apparent asymptote, which would support the usefulness of a % per month metric. This review found inadequate numbers of sufficiently homogeneous studies to conduct any meta-analyses. This speaks both to the limited number of long-term follow-ups, and to the lack of consistency in design and measurement that are required for meaningful quantitative synthesis of this kind.

Limitations

This review has several limitations. First, reviews that draw conclusions based only on the results of RCTs are inherently limited, most obviously by selection bias. RCTs and, in fact, the criteria for inclusion in this review, exclude patients with serious comorbidity, some of which are common in the population seeking treatment for depression (e.g. drug dependence, suicidality, personality disorders). This limits the generalizability of the results and conclusions drawn. Second, reviews are limited by the quality of the studies that they include. Table 4 provides details on the variability in the quality of the included studies and suggests that the quality of several of them is not optimal. Although it is not possible for studies of psychological treatments to conceal to subjects the condition to which they are assigned, many studies did not meet, or even report on, other major quality criteria such as assignment to conditions by an independent person and, in some cases, blinding of evaluators. Although there is no evidence that quality assessment can be effectively used to weight the results of studies included in a review, quality can be viewed as a reflection of a study's robustness. Those studies with less robust methodology and reporting should be interpreted with caution. Lack of accounting for variables such as therapists' allegiance to the treatment they are delivering contributes to publication bias in otherwise well-conducted RCTs (Cuijpers, Smit, Bohlmeijer, Hollon, & Andersson, 2010) and similarly introduces bias into the results of a review. Another limitation of this, and all reviews examining relapse and recurrence, is that researchers do not employ consistent definitions of relapse and recurrence. It would require re-analyzing the raw data for each individual study to present findings across studies based on a uniform definition of relapse and recurrence.

Nonetheless, some studies do adhere to the consensus definitions of relapse and recurrence set out by Frank et al. (1991), making comparison of results clearer and more meaningful.

Study 2 Conclusions

The results of this review support the findings of the overview in Study 1. Depression appears to be a chronic condition for many individuals. Regardless of treatment type, a significant proportion of people will relapse after treatment is terminated. It is thus important for practitioners to discuss this reality with patients seeking treatment for depression, and to include planning for maintenance and relapse-prevention as a component of acute treatment. While relapse occurs across treatments, patients appear to derive the greatest relapse-prevention effects from CBT, either as an acute or continuation treatment. Variants of CBT, including BA and MBCT, appear to be as effective in relapse-prevention as standard CBT. However, the evidence for these variants is smaller and would benefit from replication. Antidepressant medication has similar relapse prevention effects to CBT as long as it is continued, but it has higher rates of relapse after it is stopped. Finally, patients who reach remission on ADM benefit just as much from the relapse-prevention effects of continuation phase CBT or MBCT as those who reach remission in acute phase CBT.

General Discussion

This thesis has reviewed the evidence on long-term post-treatment outcomes for adults treated for depression and concluded that regardless of treatment type, relapse remains a significant risk for many who respond to treatment, even for those who reach

remission or recovery. For this reason, it is important that future studies continue to explore not only probability of relapse over a given time period, but also important variables such as time to relapse, frequency of relapse over time, and severity of relapse when it does occur.

It is also important that patients and professionals view depression as a condition which is often chronic. Treatment should include interventions targeted at relapse-prevention and discussion of the potential for relapse in the acute treatment phase. Patients should be prepared to manage their potential for repeat occurrences of depression in the future and not believe that remission equals cure. This requires that practitioners be educated about the potential for relapse, and be prepared to relay and discuss this information with their patients. Communicating to patients that depression is often recurrent must be done in a careful manner so as not to convey a sense of hopelessness to patients, while at the same time in a forthright manner that engenders an appropriate level of caution about the resiliency of gains made during treatment, and the importance of planning ahead in order to prevent relapse or minimize its impact. As discussed earlier in this paper, the current studies are a part of the work of the Mobilizing Minds research group, whose work includes development of an information and decision aid for patients and practitioners to assist with just this type of discussion and education process. The results of the current studies have been incorporated into an online information aid which is currently being piloted in another study. The manner in which these findings about long-term post-treatment outcomes will be evaluated for their clarity and for the degree to which they constitute provision of helpful information about depression and its treatment. The results of that pilot will be used to further refine

the manner in which that information is communicated so that it can be made as accessible and useful as possible to the public, where the decision aid will eventually be made available.

Guidelines for depression treatment should incorporate this knowledge by, for example, including recommendations for the inclusion of relapse prevention planning during acute treatment. Such guidelines could also include recommendations around treatment to recovery, rather than partial remission, as the former is associated with significantly lower relapse rates (Judd, 2012).

With regards to the body of evidence on post-treatment outcomes, we should consider the clinical significance, and practical utility, of the knowledge that we aim to create with this kind of research. Although the literature calls repeatedly for more rigorous follow-up of controls groups and less reliance on naturalistic follow-up, some technical challenges make it unlikely that this call will be answered. First, it is not feasible to follow-up over the long-term the types of control groups that are most widely used in acute treatment studies (placebo, waiting list). These groups are often promised treatment at the end of the acute treatment study. Therefore, it would be unethical to ask these groups to wait for treatment for the duration of follow-up that can last from one to many years. It is more likely that we will continue to see long-term follow-up of comparison groups who are offered treatment as usual (TAU) or clinical management. Although these comparison groups cannot show us how effective a given treatment is in comparison to no treatment, still they are of great value because, in the real world, the population that will stand to benefit from the treatments we study are those who are seeking treatment. That is to say, TAU or clinical management groups are likely a good

representation of the comparison we most need to investigate: Do we have treatments that are more effective than the options that people are already pursuing? These comparisons will be more revealing if the treatments received and symptoms of depression are monitored in all groups over the same follow-up period.

Another challenge in conducting long-term follow-up relates to sample size and its impact on statistical power. Follow-up studies usually include only those who responded to acute treatment. This is usually about half of the original sample, leading to small samples even in initially large studies. Another complication concerns funding of these follow-up studies. Funding for treatment research is often focused on the recruitment and acute treatment phases, rather than for follow-up studies. This results in fewer follow-ups, and potentially in follow-up of fewer participants as it can be resource-intensive to track down and obtain data from participants who have moved or lost contact with the study for various reasons. Less than adequate funding for follow-up studies can also compromise the quality of data collected, as face-to-face assessment at frequent intervals is much more costly than less-frequent data collection and/or self-report mail in surveys. With all of these issues, the longer the follow-up period, the more of a challenge each of these problems becomes. Attrition due to death, inability to locate, or loss of interest in participation also become more likely over time, leading to smaller and potentially more biased samples in longer-term follow-up studies. It would be helpful for large granting agencies to provide targeted funding for long-term follow-up studies – similar to the Collaboration Depression Study (CDS; Judd, 2012) and the STAR*D study (Fava et al., 2003; Rush et al., 2004; Trivedi et al., 2006), both funded by the National Institutes of Mental Health in the United States. There is the potential to

use large-scale, or population-based administrative data sets to link receipt of certain interventions to outcomes over the long term. It would be well worth investing in an examination of the indicators currently available in such data sets and considering where improvements can be made to that data to better answer our questions in this domain.

Both the present overview of reviews (Study 1), and the systematic narrative review (Study 2) provide recommendations for practice and for future research. The overview of reviews includes numerous reviews whose focus is primarily on immediate effects of acute or continuation treatment, and whose secondary focus is on the long-term relapse-prevention effects of these interventions. As such, this overview was able to synthesize not only the original review authors' findings, but also their discussions of the issues and challenges inherent in long-term follow-up studies. The narrative review in Study 2 focused only on these long-term follow-up studies and was able to discuss, in more depth, the methodological strengths and weaknesses of the existing evidence base. With regards to the results reported in Study 1 and Study 2, the overall findings are consistent. However, Study 2 provides additional depth and clarity as it focuses solely on the results of follow-ups at least one year in duration. Moreover, the detailed characteristics and quality assessment of each study allows the reader to consider the meaning of the results with these variables in mind. It is also important to note that, to date, there were not sufficient numbers of homogeneous, long-term outcome studies available to perform any meta-analyses. This speaks to the need for more follow-up research and, importantly, for more consistency in treatment delivery, study design, and outcome measurement which will enable us to synthesize data across studies in the future. A final priority, but one of great importance given the chronic nature of

depression, is ongoing research and implementation of interventions that prevent the occurrence of depression. Evidence based interventions exist, which have demonstrated the ability to reduce the incidence of new episodes of major depressive disorder by 25 – 50% (Cuijpers, Beekman, & Reynolds, 2012). A review of these interventions is beyond the scope of this paper, but examples of such reviews include the 2009 Institute of Medicine report on prevention of mental, emotional, and behavioural disorders (National Research Council & Institute of Medicine, 2009) and Munoz, Beardslee, and Leykin's (2012) very recent discussion of the available interventions, which includes a proposed roadmap for the next decade of work in the prevention of depression.

General Conclusions

In practice, we must make use of the existing knowledge, however limited, about what happens after treatment stops, and ensure that this knowledge is accessible to practitioners and those seeking treatment. From a research perspective, we need more information about these post-treatment effects in the form of pre-planned, well-designed studies that compare active treatments and their alternatives to answer our most clinically important questions. Depression is a common and recurrent condition. Acute treatment is crucial, but we must increase our focus on the long-term course, treatments, and outcomes if we are to adequately assist individuals, and indeed society, in managing what is fast becoming one of the top health concerns of our time.

References

Abbass, A. A., Hancock, J. T., Henderson, J., & Kisely, S. R. (2006). Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database of Systematic Reviews,* Issue 4. Art. No.: DC004687.

Allen, N., Hetrick, S. E., Yap, M. B. H., & Simmons, J. G. (2008). Interventions for preventing relapse and recurrence after the first episode of a depressive disorder in young people. *Cochrane Database of Systematic Reviews, 2008*(4), Art. No.: CD007504.

American Medical Association (2001). *Principles of medical ethics*. Retrieved April 23, 2005, from http://www.ama-assn.org/ama/pub/category/4256.html.pdf.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2010). Practice guidelines for the treatment of patients with major depressive disorder, third edition. Retrieved December 15, 2012 from http://psychiatryonline.org/content.aspx?bookid=28&sectionid=1667485 .

Anderson, R. M. (1995). Revisiting the behavioural model and access to medical care: Does it matter? *Journal of Health and Social Behaviour, 6,* 1-10.

Anderson, I. M., (2001). Meta-analytical studies on new antidepressants. *British Medical Bulletin, 57,* 161-178.

Anderson, G., Bergstrom, J., Hollandare, F., Carlbring, P., Kaldo, V., Ekselius, L. (2005). Internet-based self-help for depression: Randomized controlled trial. *British Journal of Psychiatry, 187,* 456-461.

Anderson, I. M., Ferrier, I. N., Baldwin, R. C., Cowen, P. J., Howard, L., Lewis, G., et al. (2008). Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2000 British Association for Psychopharmacology guidelines. *Journal of psychopharmacology (Oxford, England)*, *22*(4), 343-96.

Aneshensel, C. S., Estrada, A. L., Hansell, M. J., & Clark, V. A. (1987). Social psychological aspects of reporting behaviour: Lifetime depressive episode reports. *Journal of Health and Social Behaviour. 28*(3), 232-246.

Angst, J., Scheidegger, P., & Stabl, M. (1993). Efficacy of moclobemide in different patient groups. Results of new subscales of the Hamilton Depression Rating Scale. *Clinical Neuropharmacology, 16*(suppl 2), S55-S62.

Baldwin, D. S., Cooper, J. A., Huusom, A. K. T., & Hindmarch, I. (2006). 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, 21*(3), 159-169.

Barbui, C., Furukawa, T. a., & Cipriani, A. (2008). Effectiveness of paroxetine in the treatment of acute major depression in adults: a systematic re-examination of published and unpublished data from randomized trials. *CMAJ : Canadian Medical Association journal, 178*(3), 296-305.

Barkham, M., Rees, A., Shapiro, D. A., Stiles, W. G., Agnew, R. M., Halstead, J., Culverwell, A., & Harrington, V. M. G. (1996). Outcomes of time-limited psychotherapy in applied settings: Replicating the Second Sheffield

Psychotherapy Project. *Journal of Consulting and Clinical Psychology, 64*(5), 1079-1085.

Beck, A. T. (2005). The current state of cognitive therapy: A 40 year retrospective. *Archives of General Psychiatry, 62,* 953-959.

Beck, A. T., Rush, A. J., Shaw, B. M., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Belsher, G., & Costello, C. G. (1988) Relapse after recovery from unipolar depression: A critical review. *Psychological Bulletin, 104,* 84-96.

Blatt, S. J., Zuroff, D. C., Bondi, C. M., & Sanislow, C. A. (2000). Short- and long-term effects of medication and psychotherapy in the brief treatment of depression: Further analyses of data from the NIMH TDCRP. *Psychotherapy Research, 10*(2), 215-234.

Bockington, C. L. H., Spinhoven, P., Koeter, M. W. J., Wouters, L. F., Huyser, J., Kamphuis, J. G., & The DELTA Study Group. (2005). Preventing relapse/recurrence in recurrent depression with cognitive therapy: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 73*(4), 647-657.

Bockington, C. L. H., ten Doesschate, M. C., Spijker, J., Spinhoven, P., Koeter, M. W. J., Schene, A. H., & The DELTA study group. (2009).Continuation and maintenance use of antidepressants in recurrent depression. *Psychotherapy and Psychosomatics, 77,* 17-26.

Bondolfi, G., Jermann, F., Van der Linden, M., Gex-Fabry, M., Bizzini, L., Weber Rouget, B., et al. (2010). Depression relapse prophylaxis with Mindfulness-Based

Cognitive Therapy: Replication and extension in the Swiss health care system. *Journal of Affective Disorders, 122,* 224-231.

Braddock, C. H., Edwards, K. A., Hasenberg, N. M., Laidley, T. L., & Levinson, W. (1999). Information decision-making in outpatient practice: Time to get back to basics. *Journal of the American Medical Association, 282,* 2313-2320.

Brock, D. W., & Wartman, S. A. (1990). When competent patients make irrational choices. The *New England Journal of Medicine, 322,* 1595-1599.

Bush, T., Russo, J., Ludman, E., Lin, E., Von Korff, M., Simon, G., Katon, W., & Walker, E. (2001). *Perceived self-efficacy for depression self-management. A reliable and valid self-report measure with predictive validity.* Poster presentation at the American Psychological Society Meeting, June 2001, Toronto Canada.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clinical psychology review, 26*(1), 17-31.

Canadian Medical Association (2004). *CMA Code of ethics*. Retrieved April 23, 2007, from http://www.cma.ca/index.cfm/ci_id/43892/la_id/1.htm.pdf.

Canadian Psychological Association (2000). *Canadian code of ethics for psychologists* (3rd ed.). Ottawa, ON: Canadian Psychological Association.

Casacalenda, N., Perry, J. C., & Looper, K. (2002). Remission in major depressive disorder: A comparison of pharmacotherapy, psychotherapy, and control conditions. *American Journal of Psychiatry, 159,* 1354-1360.

Cassileth, B. R., Zupkis, R. V., Sutton-Smith, K., & March, V. (1980). Information and participation preferences among cancer patients. *Annals of Internal Medicine, 92*, 832-836.

Centre for Economic Performance's Mental Health Policy Group. (2006). *The depression report. A new deal for depression and anxiety disorders.* London School of Economics: London.

Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician-patient encounter: Revisiting the shared treatment decision-making model. *Social Science & Medicine, 49,* 651-661.

Chilvers, C., Dewey, M., Fielding, K., Gretton, V., Miller, P., Palmer, B., et al. (2001). Antidepressant drugs and generic counselling for treatment of major depression in primary care: Randomized trial with patient preference arms. *British Medical Journal, 322,* 1-5.

Churchill, R., Hunot, V., Corney, R., Knapp, M., McGuire, H., Tylee, A., & Wessely, S. (2001). A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment 5*, 1–173.

Cipriani, A., Barbui, C., Furukawa, T. A., Hotopf, M., Geddes, & J. R. (2006). Are all antidepressants really the same? The case of fluoxetine: A systematic review. *Journal of Clinical Psychiatry, 67,* 850-864.

Cipriani, A., Furukawa, T. a., Salanti, G., Geddes, J. R., Higgins, J. P., Churchill, R., et al. (2009). Comparative efficacy and acceptability of 12 new-generation

antidepressants: a multiple-treatments meta-analysis. *The Lancet, 373*(9665), 746-758.

Cochrane Collaboration. (2001). *Cochrane handbook for systematic reviews of interventions.* West Sussex, England: John Wiley & Sons.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Conradi, H. J., de Jonge, P., & Ormel, J. (2008). Cognitive behavioural therapy versus usual care in recurrent depression. *The British Journal of Psychiatry, 193,* 505-506.

Coulter, A., Entwistle, V., & Gilbert, D. (1999). Sharing decisions with patients: Is the information good enough? *British Medical Journal, 318*(7179), 318-322.

Crits-christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Kurcias, J. S., et al. (1991). Meta-Analysis of Therapist Effects in Psychotherapy Outcome Studies. *Psychotherapy Research,* 1(2), 81-91.

Crits-Christoph, P. (1992). The efficacy of brief dynamic psychotherapy: A meta-analysis. *American Journal of Psychiatry, 149*(2), 151-158.

Cuijpers, P., Beekman, A. T. F., & Reynolds, C. F. (2012). Preventing depression: A global priority. *Journal of the American Medical Association, 307*(10), 1033-1034.

Cuijpers, P., Smit, F., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). Efficacy of cognitive behavioural therapy and other psychological treatments for adult depression: Meta-analytic study of publication bias. *The British Journal of Psychiatry, 196,* 173-178.

Cuijpers, P., van Straten, A., & Smit, F. (2007). Psychological treatments of subthreshold depression: A meta-analytic review. Acta Psychiatrica Scandinavica, 115, 434-441.

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcomes tudies. *Journal of Consulting and Clinical Psychology, 76,* 909-922.

Cuijpers, P., van Straten, A., van Schaik, A., & Andersson, A. (2009). Psychological treatment of depression in primary care. *British Journal of General Practice, 59,* 120-127.

Cuijpers, P., van Straten, A., & Warmerdam, L. (2007). Behavioural activation treatments of depression: A meta-analysis. *Clinical Psychology Review, 27,* 318-326.

Cuijpers, P., van Straten, A., Warmerdam, L., & Anderson, G. (2009). Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Depression and Anxiety, 26,* 279-288.

Deber, R. B., Kraetschemer, N., & Irvine, J. (1996). What role do patients wish to play in treatment decision-making? *Archives of Internal Medicine, 156,* 1414-1420.

De Maat, S., Dekker, J., Schoevers, R., & De Jonghe, F. (2006). Relative efficacy of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Psychotherapy Research, 16*(5), 566-578.

DeRubeis, R. J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology, 66,* 37-52.

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., et al. (2005). Cognitive therapy versus medications in the treatment of moderate to severe depression. *Archives of General Psychiatry, 62,* 409-416.

Dewa, C. S., Lesage, A. D., Goering, P., et al. (2004). Nature and prevalence of mental illness in the workplace. *Healthcare Papers, 5,* 12-25.

Deshauer, D., Moher, D., Fergusson, D., Moher, E., Sampson, M., Grimshaw, J., et al. (2008). Selective serotonin reuptake inhibitors for unipolar depression: a systematic review of classic long-term randomized controlled trials. *CMAJ : Canadian Medical Association journal*, *178*(10), 1293-301.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M .E., et al. (2006). Randomized trial of behavioural activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. Journal *of Consulting and Clinical Psychology, 74,* 658-670.

Dobson, K. S., Hollon, S. D., Dimidjian, S., Schmaling, K. B., Kohlenberg, R. J., Gallop, R. J., et al. (2008). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *Journal of consulting and clinical psychology*, *76*(3), 468-77.

Driessen, E., Cuijpers, P., de Maat, S. C. M., Abbass, A. A., de Johnghe, F., & Dekker, J. J. M. (2010). The efficacy of short-term psychodynamic psychotherapy for depression: A meta-analysis. *Clinical Psychology Review, 30,* 25-36.

Ekers, D., Richards, D., & Gilbody, S. (2008). A meta-analysis of randomized trials of behavioural treatment of depression. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences, 38*(5), 611-623.

Elkin, I., Shea, M. T., Watkins, S. D., Imber, S. M., Sotsky, S. M., Collins, D. R., et al. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry, 46,* 971-982.

Emanuel, E. J., & Emanuel, L. L. (1992). Four models of the physician-patient relationship. *Journal of the American Medical Association, 267,* 2221-2226.

Evans, M. D., Hollon, S. D., DeRubeis, R. J., Piasecki, J. M., Grove, W. M., Garvey, M. J., & Tuason, V. B. (1992). Differential relapse following cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 49,* 802-808.

Fava, G. A. (2002). Long-term treatment with antidepressant drugs. *Psychotherapy and Psychosomatics, 71,* 127-132.

Fava, G. A., Grandi, S., Zielezny, M., Rafanelli, C., & Canestrari, R. (1996). Four-year outcome for cognitive behavioural treatment of residual symptoms in major depression. *American Journal of Psychiatry, 153*(7), 945-947.

Fava, G., A., Grandi, S., Zielezny, M., Canestrari, R., & Morphy, M. A. (1994). Cognitive behavioural treatment of residual symptoms in primarily major depressive disorder. *American Journal of Psychiatry, 151*(9), 1295-1299.

Fava. G., Rafanelli, C., Grandi, S., Canestrari, R., & Morphy, M. A. (1998b). Six-year outcome for cognitive behavioural treatment of residual symptoms in major depression. *American Journal of Psychiatry, 155*(10), 1443-1445.

Fava, G., Rafanelli, C., Grandi, S., Conti, S., &Belluardo, P. (1998a).Prevention of recurrent depression with cognitive behavioural therapy. *Archives of General Psychiatry, 55,* 816-820.

Fava, G. A., Ruini, C., Rafanelli, C., Finos, L., Conit, S., & Grandi, S. (2004). Six-year outcome of cognitive behaviour therapy for prevention of recurrent depression. *American Journal of Psychiatry, 161*, 1872-1876.

Feijo de Mello, M., Jesus Mari, J., Bacaltchuk, J., Verdeli, H., & Neugebauer, R. (2005). A systematic review of research findings on the efficacy of interpersonal thearpy for depressive disorders. *European Archives of Psychiatry and Clincial Neuropscience, 255,* 75-82.

Forlenza, O. V., Almeida, O. P., Stoppe, A. Jr., Hirata, E. S., & Ferreira, R. C. R. (2001). Antidepressant efficacy and safety of low-dose sertraline and standard-dose imipramine for the treatment of depression in older adults: Results from a double-blind, randomized, controlled clinical trial. *International Psychogeriatrics, 13,* 75-84.

Frank, E., Prien, R. F., Jarrett, J. B., Keller, M. B., Kupfer, D. J., Lavori, P., et al. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Response, remission, recovery, relapse, and recurrence. *Archives of General Psychiatry, 48,* 851-855.

Frank, L., Revicki, D. A., Sorensen, S. V., & Shih, Y. C. T. (2001). The economics of selective serotonin reuptake inhibitors in depression: A critical reviews. *CNS Drugs*, *15*(1), 59-83.

Friedman, M. A., Detweiler-Bedell, J. B., Leventhal, H. E., Horne, R., Keitner, G. I., & Miller, I. W. (2004). Combined psychotherapy and pharmacotherapy for the treatment of major depressive disorder. *Clinical Psychology: Science and Practice*, *11*(1), 47-68.

Garfield, S., Francis, S-A., & Smith, F. J. (2004). Building concordant relationships with patients starting antidepressant medication. *Patient Education and Counseling*, *55*, 241-246.

Gartlehner, G., Thieda, P., Hansen, R. a., Gaynes, B. N., Deveaugh-Geiss, A., Krebs, E. E., et al. (2008). Comparative risk for harms of second-generation antidepressants : A systematic review and meta-analysis. *Drug Safety: An International Journal of Medical Toxicology and Drug Experience*, *31*(10), 851-65.

Glauguen, V., Cottraux, J., Cucherat, M., & Blackburn, I. (1988). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders*, *49*, 59-72.

Godfrin, K. A., & van Heeringen, C. (2010). The effects of mindfulness-based cognitive therapy on recurrence of depressive episodes, mental health and quality of life: A randomized controlled study. *Behaviour Research and Therapy*, *48*, 738-746.

Goldney, R. D., & Fisher, L. J. (2005). Canadian depression prevalence. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, *50*(9), 569-70.

Gortner, E. T., Gollan, J. K., Dobson, K. S., & Jacobson, N. S. (1998). Cognitive-behavioural treatment for depression: Relapse prevention. *Journal of Consulting and Clinical Psychology, 66*(2), 377-384.

Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., et al. (2003). The economic burden of depression in the United States: How did it change between 1990-2000? *Journal of Clinical Psychiatry, 64,* 1465-1475.

Greenhouse, J. B., Stangl, D., Kupfer, D. J., et al. (1991) Methodologic issues in maintenance therapy clinical trials. *Archives of General Psychiatry, 48,* 313-318.

Guilfi, J. D., Ansseau, M., Timmerman, L., & Korsgaard, S. (2001). Mirtazapine versus venlafaxine in hospitalized severely depressed patients with melancholic features. *Journal of Clinical Psychopharmacology, 21,* 425-431.

Haby, M. M., Donnelly, M., Corry, J., & Vos, T. (2006). Cognitive behavioural therapy for depression, panic disorder and generalized anxiety disorder: a meta-regression of factors that may predict outcome. *The Australian and New Zealand journal of psychiatry, 40*(1), 9-19.

Hakkaart-van Roijen, L., van Straten, A., Donker, M., & Tiemens, B. (2002). *Manual Trimbos/IMTA questionnaire for Costs associated with Psychiatric Illness (TIC-P),* Institute for Medical Technology Assessment, Erasmus University Rotterdam.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23,* 56-62.

Hansen, R., Gaynes, B., Thieda, P., Gartlehner, G., Deveaugh-Geiss, A., Krebs, E., & Lohr, K. (2008). Meta-analysis of major depressive disorder relapse and recurrence with second-generation antidepressants. *Psychiatric Services, 59*(10), 1121-1130.

Hansen, R. A., Gartlehner, G., Lohr, K. N., & Gaynes, B. N. (2005). Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. *Annals of Internal Medicine, 143,* 415-426

Hays, P. (1984). The nosological status of schizophrenia, *Lancet, 1,* 1342-1345.

Hensley, P. L., Nadiga, D., & Uhlenhuth, E. H. (2004).Long-term effectiveness of cognitive therapy in major depressive disorder. *Depression and Anxiety, 20,* 1-7.

Hill, S. A., & Laugharne, R. (2006). Decision-making and information seeking preferences among psychiatric patients. *Journal of Mental Health, 15,* 75-84.

Hofmann, S. G., Barlow, D. H., Papp, L. A., Detweiler, M. F., Ray, S. E., Shear, M. K., Woods, S. W., & Gorman, J. M. (1998). Pretreatment attrition in a comparative treatment outcome study on panic disorder. *American Journal of Psychiatry, 155*(1), 43-47.

Hollon, S. D., DeRubeis, R., J., Shelton, R. C., Amsterdam, J. D., Salomon, R. M., et al. (2005c). Prevention of relapse following cognitive therapy versus medications in moderate to severe depression. *Archives of General Psychiatry, 62,* 417-422.

Hollon, S. D., Evans, M. D., Auerbach, A., DeRubeis, R. J., Elikin, I., Lowery, A., et al. (1988). *Development of a system for rating therapies for depression:*

Differentiating cognitive therapy, interpersonal psychotherapy, and clinical management pharmacotherapy. Unpublished manuscript.

Hollon, S. D., Jarrett, R. B., Nierenberg, A. A., Thase, M. E., Trivedi, M., et al. (2005a). Treatment of Adult and Geriatric Depression. *Psychiatry: Interpersonal and Biological Processes,* (April), 455-468.

Hollon, S. D., Jarrett, R. B., Nierenberg, A. A., Thase, M. E., Trivedi, M., & Rush, A. J. (2005b). Psychotherapy and medication in the treatment of adult and geriatric depression: Which monotherapy or combined treatment? *Journal of Clinical Psychiatry, 66*(4), 455-468.

Hollon, S. D., & Ponniah, K. (2010).A review of empirically supported psychological therapies for mood disorders in adults. *Depression and Anxiety, 27,* 891-932.

Hollon, S. D., Stewart, M. O., & Strunk, D. (2006).Enduring effects for cognitive behavioural therapy in the treatment of depression and anxiety. *Annual Review of Psychology, 57,* 285-315.

Holmes, T. H., & Rahe, R. H. (1967). The Social Readjustment rating Scale. *Journal of Psychosomatic Research, 11,* 213-218.

Hooley, J. M., Orley, J., & Teasdale, J. D. (1986). Levels of expressed emotion and relapse in depressed patients. *British Journal of Psychiatry, 148,* 642-647.

Huppert, J. D., Franklin, M. E., Foa, E. B., & Davidson, J. R. T. (2002). Study refusal and exclusion from a randomized treatment study of generalized social phobia. *Anxiety Disorders, 17,* 683-693.

Imel, Z. E., Malterer, M. B., McKay, K. M., & Wampold, B. E. (2008).A meta-analysis of psychotherapy and medication in unipolar depression and dysthymia. *Journal of Affective Disorders, 110,* 197-206.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., et al. (1996). A component analysis of cognitive-behavioural treatment for depression. *Journal of Consulting and Clinical Psychology, 64*(2), 295-304.

Jacobson, N., Martell, C., & Dimijan, S. (2001). Behavioural activation treatment for depression: Returning to contextual roots. *Clinical Psychology: Science and Practice, 8,* 255-270.

Jarrett, R. B., & Rush, A. J. (1994). Short-term psychotherapy of depressive disorders: Current status and future directions. *Psychiatry, 57*(2), 115-132.

Judd, L. (2012). Dimensional paradigm of the long-term course of unipolar major depressive disorder. *Depression and Anxiety, 29,* 167-171.

Judd, L. L., Akiskal, H. S., Maser, J. D., et al. (1998). Major depressive disorder: A prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. *Journal of Affective Disorders, 50,* 97-108.

Judd, L. L., Paulus, M. P., Zeller, P. J., et al. (2000). Does incomplete recovery from first lifetime major depressive episode herald a chronic course of illness? *American Journal of Psychiatry, 157,* 1501-1504.

Juni, P. (2001). Assessing the quality of randomised controlled trials. In: M. Egger, G. Davey-Smith, & D. G. Altman (Eds.), *Systematic reviews in healthcare.* London: BMJ publications.

Kaltenthaler, E., Parry, G., & Catherine, B. (2004). Computerized cognitive behaviour therapy: A systematic review. *Behavioural and Cognitive Psychotherapy, 32*(1), 31-55.

Kaltenthaler, E., Brazier, J., De Nigris, E., Tumur, I., Ferriter, M., Beverley, C., Parry, G., Rooney, G., & Sutcliffe, P. (2006). Computerised cognitive behaviour therapy for depression and anxiety update: A systematic review and economic evaluation. *Health Technology Assessment, 10,* 1-186.

Kaltenthaler, E., Sutcliffe, P., Parry, G., Beverley, C., Rees, a., Ferriter, M., et al. (2008). The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. *Psychological medicine*, *38*(11), 1521-30.

Kaymaz, N., van Os, J., Loonen, A. J. M., & Nolen, W. A. (2008). Evidence that patients with single versus recurrent depressive episodes are differentially sensitive to treatment discontinuation: A meta-analysis of placebo-controlled randomized trials. *Journal of Clinical Psychiatry, 69*(9), 1423-1436.

Keller, M. B., Lavori, P. W., Meuler, T. I., et al. (1992). Time to recovery, chronicity, and levels of psychopathology in major depression. *Archives of General Psychiatry, 49,* 809-816.

Keller, M. B., McCullough, J. P., Klein, D. M., Arnow, B., Dunner, D. L., Gelenberg, A. J., Markowitz, J. C., Nemeroff, C. B., Russell, J. M., Thase, M. E., Trivedi, M.H., & Zajecka, J. (2000). A comparison of nefazodone, the cognitive-behavioural analysis system or psychotherapy and their combination for the

treatment of chronic depression. *New England Journal of Medicine, 342,* 1462-1470.

Keller, M. B., Lavori, P. W., Endicott, J., Coryell, W., & Klerman, G. L. (1983). "Double depression": Two-year follow-up. *American Journal of Psychiatry, 140,* 689-694.

Keller, M. B., Lavori, P., Lewis, C. E., & Klerman, G. L. (1983). Predictors of relapse in major depressive disorder. *Journal of the American Medical Association, 250,* 3299-3304.

Keller, M. B., & Shapiro, R. W. (1981). Major depressive disorder: Initial results from a one-year prospective naturalistic follow-up study. *Journal of Nervous and Mental Disease, 169,* 911-915.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA : the journal of the American Medical Association, 289*(23), 3095-105.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62,* 593-502.

Khan, A., Brodhead, A. E., Kolts, R. L., & Brown ,W. A. (2005). Severity of depressive symptoms and response to antidepressants and placebo in antidepressant trials. *Journal of Psychiatric Research, 39,* 145-150.

Kirsch, I., Scoboria, A., & Moore, T. J. (2002). Antidepressants and placebos: Secrets, reveleations, and unanswered questions. *Prevention and Treatment, 5*: Article 33. http://journals.apa.org/prevention/volume5/toc-jul15-02.htm.

Klein, D. N., Santiago, N. J., Vivian, D., Arnow, B. A., Blalock, J. A., Dunner, D. L., et al. (2004). Cognitive-behavioural analysis system of psychotherapy as a maintenance treatment for chronic depression. *Journal of Consulting and Clinical Psychology, 72*(4), 681-688.

Klerman, G. L., Weissman, M. M., & Rounsaville, B. J. (1984). *Interpersonal Psychotherapy of Depression.* New York: Basic Books.

Klerman, G. L., Weissman, M. M., & Markowitz, J. C. (1994). Medication and psychotherapy. In: Bergin, A. E, & Garfield, S. L. (Eds.), *Handbook of Psychotherapy and Behaviour Change* (pp 734-782), New York: Wiley.

Klerman, G. L. & Weissman, M. M. (1989). Increasing rates of depression. *Journal of the American Medical Association, 261,* 2229-2235.

Kopec, J. A., Abrahamowicz, M., & Esdaile, J. M. (1993). Randomized discontinuation trials: utility and efficiency. *Journal of Clinical Epidemiology,46*, 959-71.

Koppers, D., Peen, J., Niekerken, S., Van, R., & Dekker, J. (2011).Prevalence and risk factors for recurrence of depression five years after short term psychodynamic therapy. *Journal of Affective Disorders, 134,* 468-472.

Kovacs, M., Rush, A. J., Beck, A. T., & Hollon, S. D. (1981). Depressed outpatients treated with cognitive therapy or pharmacotherapy. Archives of General Psychitary, 38, 33-39.

Kuyken, W., Byford, S., Taylor, R. S., Watkins, E., Holden, E, White, K., et al. (2008).Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *Journal of Consulting and Clinical Psychology, 76*(6), 966-978.

Kyle, C. J., Petersen, H. E. H., & Overo, K. F. (1998). Comparison of the tolerability and efficacy of citalopram and amitriptyline in elderly depressed patients treated in general practice. *Depression and Anxiety, 8,* 147-153.

Lavori, P. W., Keller, M. B., & Klerman, G. L. (1984). Relapse in affective disorders: A reanalysis of the literature using life table methods. *Journal of Psychiatric Research, 18,* 13-25.

Lawlor, D. A., & Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the managements of depression: Systematic review and meta-regression analysis of randomised controlled trials. *British Medical Journal, 322,* 763-767.

Leichsenring, F. (2001). Comparative effects of short-term psychodynamic psychotherapy and cognitive-behavioural activation in depression: A meta-analytic approach. *Clinical Psychology Review, 21*(3), 401-419.

Levine, S., Deo, R., & Mahadevan, K. (1989). A comparative trial of a new antidepressant, fluoxetine. *International Journal of Clinical Psychopharmacology, 4,* 41-45.

Lin, P., Campbell, D., Chaney, E. F., Liu, C. F., Heagerty, P., Felker, B. L., et al. (2005). The influence of patient preference on depression treatment in primary care. *Annals of Behavioural Medicine, 30,* 164-173.

Ludman, E., Katon, W., Bush, T., Rutter, C., Lin, E., Simon, G., Von Korff, M., & Walker, E. (2003). Behavioral factors associated with symptom outcomes in a

primary care-based depression prevention intervention trial. *Psychological Medicine, 33,* 1061-1070.

Luty, S. E., Carter, J. D., McKenzie, J. M., Rae, A. M., Frampton, C. M., Mulder, R. T., & Joyce, T. R. (2007). Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression. *British Journal of Psychiatry, 190,* 496-502.

Lynch, D., Laws, K. R., & McKenna, P. J. (2010). Cognitive behavioural therapy for major psychiatric disorder: Does it really work? A meta-analytical review of well-controlled trials. *Psychological Medicine, 40,* 9-24.

Ma, S. H., & Teasdale, J. D. (2004). Mindfulness-based cognitive therapy for depression: Replication and exploration of differential relapse prevention effects. *Journal of Consulting and Clinical Psychology, 72*(1), 31-40.

Maina, G., Rosso, G., & Bogetto, F. (2009). Brief dynamic therapy combined with pharmacotherapy in the treatment of major depressive disorder: Long-term results. *Journal of Affective Disorders, 114,* 200-207.

Manicavasgar, V., Parker, G., & Perich, T. (2011).Mindfulness-based cognitive therapy vscognitive behavioural therapy as a treatment for non-melancholic depression. *Journal of Affective Disorders, 130,* 138-144.

Marcus, M., Yasamy, M. T., van Onneren, M., Chisholm, D., Saxena, S., & the World Health Organization department of mental health and substance abuse. (2012). *Depression: A global public health concern.* Retrieved December 15, 2012 from http://www.who.int/mental_health/management/depression/en/index.html

Mazure, C., Nelson, J. C., & Price, L. H. (1986). Reliability and validity of the symptoms of major depressive illness. *Archives of General Psychiatry,* 43, 451-456.

McLean, P. D., & Hakstian, A. R. (1990). Relative endurance of unipolar depression treatment effects: Longitudinal follow-up. *Journal of Consulting and Clinical Psychology, 58*(4), 482-488.

Mead, N., MacDonald, W., Bower, P., Lovell, K., Richards, D., Roberts, C., & Bucknall, A. (2005). The clinical effectiveness of guided self-help versus waiting-list control in the management of anxiety and depression: A randomized controlled trial. *Psychological Medicine, 35,* 1633-1643.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-anlyses: The PRISMA Statement. *PLoS Medicine, 6*(7), 1-6.

Moher, D., Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., Tugwell, P., & Klassen, T. P. (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *The Lancet, 352,* 609-613.

Moncrieff, J. (2001). Are antidepressants overrated? A review of methodological problems in antidepressant trials. *The Journal of Nervous and Mental Disease, 189*(5), 288-295.

Montgomery, S. A., & Asberg, A. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry, 134,* 382-389.

Montgomery, S. A., Henry, J., McDonald, G., Dinan, T., Lader, M., Hindmarch, I., et al. (1994). Selective serotonin reuptake inhibitors: Meta-analysis of discontinuation rates. *International Journal of Clinical Psychopharmacology, 9,* 47-53.

Morriss, R., Bolton, C. A., McCarthy, J. P., Marshall, M., Williamson, P. R., & Jones, A. P. (2004). Interventions for helping people recognise early signs of recurrence in depression. *Cochrane Database of Systematic Reviews, 2004*(3), Art. No.: CD004855.

Mueller, T. I., Leon, A. C., Keller, M. B., et al. (1999). Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *American Journal of Psychiatry, 156,* 1000-1006.

Muller, M. J., Himmerich, H., Kienzle, B., & Szegedi, A. (2003). Differentiating moderate and severe depression using the Montgomery-Asberg depression rating scale (MADRS). *Journal of Affective Disorders, 77,* 255-260.

Mulrow, C. D., Williams, J. W. Jr., Chiquette, E., Aguilar, C., Hitchcock-Noel, P., Lee, S., et al. (2000). Efficacy of newer medications for treating depression in primary care patients. *American Journal of Medicine, 108,* 54-64.

Munoz, R. F., Beardslee, W. R., & Leykin, Y. (2012). Major depression can be prevented. *American Psychologist, 67*(4), 285-295.

Murray, E. J. (1989). Measurement issues in the evaluation of psychopharmacological therapy. In S. Fisher, & R. P. Greenberg (Eds.), *The limits of biological treatments for psychological distress* (pp 39-68). Hillsdale, NJ: Erlbaum.

Murray, C. J. L. & Lopez, A. D. (Eds.). (1996). *The global burden of disease.* Geneva: World Health Organization.

National Institute for Health and Clinical Excellence (NICE). (2009). *Depression: The treatment and management of depression in adults – National clinical practice guideline*. London: Author.

National Research Council & Institute of Medicine. (2009). *Preventing mental, emotional, and behavioural disorders among young people: Progress and possibilities.* M. E. O'Connell, T. Boat, & K. E. Warner (Eds.), Washington, DC: National Academies Press.

Nierenberg, A. A., Feighner, J. P., Rudolph, R., Cole, J. O., & Sullivan, J. (1994). Venlafaxine for treatment resistant unipolar depression. *Journal of Clinical Psychopharmacology, 14,* 419-423.

Ottevanger, E. A. (1991). The efficacy of fluvoxamine in patients with severe depression. *British Journal of Clinical Research, 2,* 125-132.

Otto, M. W., Smits, J. A. J., & Reese, H. E. (2005). Combined Psychotherapy and Pharmacotherapy for Mood and Anxiety Disorders in Adults: Review and Analysis. *Clinical Psychology: Science and Practice*, *12*(1), 72-86.

Oxman, A. D., & Guyatt, G. H. (1991).Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology, 44*(11), 1271-1278.

Pampallona, S., Bollini, P., Tibaldi, G., Kupelnick, B., & Munizza, C. (2004). Combined pharmacotherapy and psychological treatment for depression: a systematic review. *Archives of general psychiatry*, *61*(7), 714-9.

Papakostas, G. I., Thase, M. E., Fava, M., Nelson, J. C., & Shelton, R. C. (2007). Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in

treating major depressive disorder? A meta-analysis of studies of newer agents. *Biological psychiatry, 62*(11), 1217-27.

Parker, G., Roussos, J., Hadzi-Pavlovic, D., Mitchell, P., Wilhelm, K., & Austing, M. P. (1997). The development of a refined measure of dysfunctional parenting and assessment of its relevance in patients with affective disorders. *Psychological Medicine, 27,* 1193-1203.

Parker, G., Sc, D., Roy, K., & Eyers, K. (2003). Reviews and Overviews Cognitive Behavior Therapy for Depression? Choose Horses for Courses. *Psychiatry: Interpersonal and Biological Processes*, (May), 825-834.

Parloff, M. B. (1986). Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology, 54*(1), 79-87.

Patel, S. R., Bakken, S., & Rulan, C. (2008). Recent advances in shared decision-making for mental health. *Current Opinion in Psychiatry, 21,* 606-612.

Paykel, E. S. (2001). Continuation and maintenance therapy in depression. *British Medical Bulletin, 57,* 145-159.

Paykel, E. S. (2007). Cognitive therapy in relapse prevention in depression. *International Journal of Neuropsychopharmacology, 10,* 131-136.

Paykel, E. S., Scott, J., Teasdale, J. D., Johnson, A. L., Garland, A., Moore, R., et al. (1999). Prevention of relapse in residual depression by cognitive therapy: A controlled trial. *Archives of General Psychiatry, 56,* 829-835.

Petticrew, M., & Roberts, H. (2009). *Systematic reviews in the social sciences: A practical guide.* Madden, MA: Blackwell Publishing

Peveler, R., Kendrick, T., Buxton, M., Longworth, L., Baldwin, D., Moore, M., Catwin, J., Goddard, J., Thornett, A., Smith, H., Campbell, M., & Thompson, C. (2005). A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine. *Health Technology Assessment, 9,* 1-134.

Pinquart, M., Duberstein, P. R., & Lyness, J. M. (2006). Treatments for later-life depressive conditions: a meta-analytic comparison of pharmacotherapy and psychotherapy. *The American journal of psychiatry, 163*(9), 1493-501.

Pollock, K., Grime, J., Baker, E., & Mantala, K. (2004). Meeting the information needs of psychiatric inpatients: Staff and patient perspectives. *Journal of Mental Health, 13,* 389-401.

Posadzki, P., Watson., L., & Ernst, E. (2012). Herb-drug interactions: An overview of systematic reviews. *British Journal of Clinical Pharmacology, Accepted Article,* doi: 10.1111/1365-2125.2012.04350.x

Public Health Agency of Canada. (2009). *What is depression?* Retrieved Dec 15th, 2012 from http://www.phac-aspc.gc.ca/cd-mc/mi-mm/depression-eng.php

Rickels, K., Noyes Jr., R., Robinson, D. S., Schweizer, E., & Uhlenhuth, E. H. (1994).Evaluating drug treatments of generalized anxiety disorder and adjustment disorders with anxious mood. In: Prien, R. F., and Robinson, D. S. (Eds.), *Clinical evaluation of psychotropic drugs: Principles and guidelines.* New York: Raven Press Ltd.

Rose, D., Fleischmann, P., Wykes, T., Leese, M., & Bindman, J. (2003). Patients' perspectives on electroconvulsive therapy: Systematic review. *British Medical Journal, 326,* 1363.

Ross, J. (2008). Discontinuation of lithium augmentation in geriatric patients with unipolar depression: A systematic review. *Canadian Journal of Psychiatry, 53*(2), 117-120.

Royal Australian and New Zealand College of Psychiatrists. (2004). Australian and New Zealand clinical practice guidelines for the management of adult deliberate self-harm. *The Australian and New Zealand journal of psychiatry, 38*(11-12), 868-84.

Rush, A. J., Fava, M., Wisniewski, S. R., Lavori, P. W., Trivedi, M. H., Sackeim, H.A., Thase, M. E., Nierenberg, A.A., Quitkin, F. M., Kashner, T. M., Kupfer, D. J., Rosenbaum, J. F., Alpert, J., Stewart, J. W., McGrath, P. J., Biggs, M. M., Shores-Wilson, K., Lebowitz, B. D., Ritz, L., & Niederehe, G. (2004). Sequenced Treatment Alternatives to Relieve Depression (STAR*D): Rationale and design. *Controlled Clinical Trials, 25,* 119–142.

Rush, A. J., Kraemer, H. C., Sackeim, H. A., Fava, M., Trivedi, M. G., Frank, E., et al. (2006). Report by the ACNP task force on response and remission in major depressive disorder. Neuropsychopharmacology. Retrieved August 2, 2012, from http://www.nature.com/npp/journal/v31/n9/pdf/1301131a.pdf

Ryff, C. D., & Singer, B. (1996). Psychological well-being. *Psychotherapy and Psychosomatics, 65,* 14-23.

Salkovskis, P., Rimes, K., Stephenson, D., Sacks, G., & Scott, J. (2006). A randomized controlled trial of the use of self-help materials in addition to standard general practice treatment of depression compared to standard treatment alone. *Psychological Medicine, 36,* 325-333.

Santor, D. A., & Segal, Z. V. (2001). Predicting symptom return from rate of symptom reduction in cognitive-behavior therapy for depression. *Cognitive Therapy and Research, 25*(2), 117-135.

Schea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., et al. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology, 7*(10), 1-7.

Schramm, E., Zobel, I., Dykierek, P., Kech, S., Brakemeier, E. L., Kulz, A., & Berger, M. (2011). Cognitive behavioural analysis system of psychotherapy versus interpersonal psychotherapy for early-onset chronic depression: A randomized pilot study. *Journal of Affective Disorders, 129,* 109-116.

Schweizer, I., Burrows, G., Tuckwell, V., Polonowitz, A., Flynn, P., George, T., et al. (2001). Sustained response to open-label venlafaxine in drug-resistant major depression. *Journal of Clinical Psychopharmacology, 21,* 185-189.

Scott, J. (1996). Cognitive therapy of affective disorders: A review. *Journal of Affective Disorders, 37,* 1-11.

Segal, Z. V., Bieling, P., Young, T., MacQueen, G., Cooke, R., Martin, L., Bloch, R., & Levitan, R. D. (2010). Antidepressant monotherapy vs sequential pharmacotherapy and mindfulness-based cognitive therapy, or placebo, for relapse

prophylaxis in recurrent depression. *Archives of General Psychiatry 67*(12), 1256-1264.

Segal, Z., Vincent, P., & Levitt, A. (2002). Efficacy of combined, sequential and crossover psychotherapy and pharmacotherapy in improving outcomes in depression. Journal of Psychiatry and Neuroscience, 27(4), 281-290.

Shapiro, D. A., Rees, A., Barkham, M., Hardy, G., Reynolds, S., & Startup, M. (1995).Effects of treatment duration and severity of depression on the maintenance of gains after cognitive-behavioural and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology, 63*(3), 378-387.

Simon, D., Loh, A., Wills, C. E., & Härter, M. (2006). Depressed patients' perceptions of depression treatment decision-making. *Health Expectations, 10*, 62-74.

Simons, A. D., Murphy, G. E., Levine, J. L., & Wetzel, R. D. (1986). Cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry, 43,* 43-48.

Sinyor, M., Schaffer, A., & Levitt, A. (2010). The sequenced treatment alternatives to relieve depression (STAR*D) trial: A review. *Canadian Journal of Psychiatry, 55*(3), 126-135.

Spek, V., Cuijpers, P., Nyklicek, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. Psychological Medicine, 37, 319-328.

Spielmans, G. I., Berman, M. I., & Usitalo, A. N. (2011). Psychotherpy versus second-generation antidepressants in the treatment of depression: A meta-analysis. *The Journal of Nervous and Mental Disease, 199*(3), 142-149.

Stiggelbout, A. M., & Kiebert, G. M. (1997). A role for the sick role: Patient preferences regarding information and participation in clinical decision-making. *Journal of the Canadian Medical Association, 157,* 383-389.

Svartberg M, Stiles T. Comparative effects of short-term psychodynamic psychotherapy: a meta-analysis. *Journal of Consulting and Clinical Psychology 59*(5), 704–14.

Tamam, L., & Ozpoyraz, N. (2002). Selective serotonin reuptake inhibitor discontinuation syndrome: A review. *Advances in Therapy, 19*(1), 17-26.

Targum, S. D. (1984). Persistent neuroendocrine dysregulation in major depressive disorder: A marker for early relapse. *Biological Psychiatry, 19,* 305-318.

Taylor, D. (2008). Psychoanalytic and psychodynamic therapies for depression: The evidence base. *Advances in Psychiatric Treatment, 14,* 401-413.

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 68*(4), 615-623.

Trivedi, M., Rush, A., Pan, J., & Carmody, T. (2001). Which depressed patients respond to nefazadone and when? *Journal of Clinical Psychiatry, 62,* 158-163.

Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Nierenberg, A. A., Warden, D., Ritz, L., Norquist, G., Howland, R. H., Lebowitz, B., McGrath, P. J., Shores-Wilson, K., Biggs, M. M., Balasubramani, G. K., & Fava, M. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *American Journal of Psychiatry 163,* 28–40.

UK ECT Review Group. (2003). Efficacy and safety of electroconvulsive therapy in depressive disorders: A systematic review and meta-analysis. *Lancet, 361,* 799-808.

Vasiliadis, H. M., Lesage, A., Adair, C., et al. (2005). Service use for mental health reasons: Cross-provincial differences in rates, determinants, and equity of access. *Canadian Journal of Psychiatry, 50,* 614-619.

Vasiliadis, H., Lesage, A., Adair, C., Wang, P. S., & Kessler, R. C. (2007). Do Canada and the United States differ in prevalence of depression and utilization of services? *Psychiatric services (Washington, D.C.), 58*(1), 63-71.

Viguera, A. C., Baldessarini, R. J., & Friedberg, J. (1998).Discontinuing antidepressant treatment in major depression. *Harvard Review of Psychiatry, March/April 1998,* 293-306.

Vittengle, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007).Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioural activation's effects. *Journal of Consulting and Clinical Psychology, 75*(3), 475-488.

Vittengle, J. R., Clark, L. A., & Jarrett, R. B. (2009). Continuation-phase cognitive therapy's effects on remission and recovery from depression. *Journal of Consulting and Clinical Psychology, 77*(2), 367-371.

Walker, J. R., Joyce, B., Furer, P., Vincent, N., & Kjernisted, K. (2000). *Facilitating informed decision-making: A survey of patients referred to a specialty anxiety*

disorders program. Poster presented at the Anxiety Disorders Association of America Annual Convention, Washington, D.C.

Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression: Variable, substantial, and growing. *Journal of the American Medical Association, 278,* 1840-1847.

Wang, P. S., Beck, A. L., Berglund, P., et al. (2004). Effects of major depression on moment-in-time work performance. *American Journal of Psychiatry, 161,* 1885-1991.

Waraich, P., Goldner, E. M., Somers, J. M., & Hsu, L. (2004). Prevalence and incidence studies of mood disorders: a systematic review of the literature. *Canadian journal of psychiatry. Revue canadienne de psychiatrie*, *49*(2), 124-38.

Weissman, M. M., Markowitz, J. C., & Klerman, G. L. (2000). *Comprehensive Guide to Interpersonal Psychotherapy.* New York: Basic Books.

Wilkinson, P., & Izmeth, Z. (2007). Continuation and maintenance treatments for depressive disorder in older adults. *Cochrane Database of Systematic Reviews, 2007*(3), Art. No.: CD006727.

Wolf, N. J., & Hopko, D. R. (2008). Psychosocial and pharmacological interventions for depressed adults in primary care: a critical review. *Clinical psychology review*, *28*(1), 131-61.

World Health Organization. (2002). *The World Health Report 2002: Reducing risks, promoting healthy life.* Geneva, Switzerland: Author.

Young, J., & Beck, A. T. (1988).*Cognitive Therapy Scale: Rating manual.* Philidelphia: Centre for Cognitive Therapy.

Zimmerman, M., Chelminski, I., & Posternak, M. (2004). A review of studies of the Montgomery – Asberg Depression Rating Scale in controls : implications for the definition of remission in treatment studies of depression. *International Clinical Psychopharmacology*, 1-7.

Zimmerman, M., & Thongy, T. (2007). True Tachyphylaxis During Continuation Treatment Is Low. *Psychiatry: Interpersonal and Biological Processes*, (August), 1271-1276.

Zoellner, L. A., Feeny, N. C., Cochran, B., & Pruitt, L. (2003). Treatment choice for PTSD. *Behaviour Research and Therapy, 41,* 879-886.

Appendix A: Inclusion checklist

Author, date REF ID: PubMed Oct 3 2011	Yes	No	Need to check article	TO DIS-CUSS	Com-ments
A. Characteristics of the study:					
1. Adults – 18 years of age and over (not only child/adolescent or over 65) – child/adolescent or over 65 groups should be identified and labelled to discuss.					
2. Diagnosis of major depression, unipolar depression by applying diagnostic criteria.					
3. Psychosocial treatment for depression meeting inclusion criteria.					
4. Pharmacological treatment for depression meeting inclusion criteria.					
5. Published in English.					
6. Published in a journal, a dissertation, or other published format.					
B. TYPE OF STUDY					
1. Review, meta-analytic review					
2. Post-treatment follow-up of intervention and control groups (medication or psychological treatments)					
3. Maintenance study, continuation study (medication or psychological treatments)					
4. Discontinuation studies (medication or psychological treatments)					
5. Studies of the effect of some treatment (e.g., CBT, mindfulness) during the follow-up period such as booster sessions after CBT or CBT when people are going off medication					
6. Post-treatment follow-up where there is not a comparison or control group.					
7. Naturalistic studies of what happens when treatment is discontinued.					
8. Naturalistic studies of the course of depression with or without treatment.					
9. Other studies that may address the research questions.					

Title and Abstract:

Comments: