Behavioural Activation for Depression: a systematic review and controlled clinical trial

David Michael Ekers

Thesis submitted for the degree of PhD

The University of York

Health Sciences

April 2011

Abstract

Background

Depression is a common, disabling condition for which psychological treatments, in particular cognitive behavioural therapies are recommended. Promising results from randomised trials has renewed interest in behavioural therapy, which may be suitable for delivery by non-specialist therapists.

Aims

To deliver a phased research programme to examine the clinical and cost effectiveness of behavioural therapy for depressed adults, and in particular the suitability of the intervention for delivery by non specialists.

Research design

A systematic review and meta-analysis examined randomised trials of behavioural treatments for depression compared to controls or other psychotherapies. Data on symptom level, recovery/dropout rate and study level moderators (study quality, number of sessions, severity and level of training) were extracted and analysed. Based upon results of the meta-analysis a randomised controlled trial of clinical and cost effectiveness comparing BA delivered by non-specialist with treatment as usual in a primary care setting was conducted.

Results

Meta-Analysis: Seventeen randomised controlled trials including 1109 subjects were included. A meta-analysis of symptom level post-treatment showed behavioural therapies were superior to controls (SMD -0.70 CI -1.00 to -0.39, k=12, N= 459), brief psychotherapy (SMD -0.56 -1.0 to-0.12, k=3, N=166), supportive therapy (SMD -0.75 CI –1.37 to -0.14, k=2, N=45) and equal to cognitive behavioural therapy (SMD 0.08 CI -0.14 to 0.30, k=12, N=476).

Randomised controlled trial: Intention to treat analyses indicated a difference in favour of BA of −15.79 (95% CI −24.55 to −7.02) on the Beck Depression

Inventory-II, −11.12; (95% CI = −17.53 to −4.70), on the Work and Social Adjustment Scale and a 0.20 (95% CI 0.01 to 0.39) improvement in quality adjusted life year. An incremental cost effectiveness ratio of £5756 per QALY indicates with a 97% probability that BA delivered by non-specialists is more cost effective than usual primary care at a threshold value of £20,000.

Conclusion

Behavioural Activation is an effective psychological treatment for depression that appears suitable for delivery by non-specialists. Further research with larger sample sizes and longer follow up is required to expand on the findings reported in this thesis.

Table of Contents

Table of Tables

Table of figures

Acknowledgements

Firstly Professor Simon Gilbody, Professor David Richards and Professor Christine Godfrey who have provided excellent supervision, advice, support and encouragement throughout my PhD study.

Professor Joe Reilly, from Tees Esk and Wear Valleys NHS Foundation Trust/Durham University, who has been essential in supporting the delivery of this research and its dissemination.

Janet Menton from Tees Esk and Wear Valleys NHS Foundation Trust for assistance in the design and delivery of literature searches.

Professor Martin Bland from the University of York for his invaluable statistical advice and input.

Steve Parrott from the University of York for assistance with cost effectiveness analysis.

Jillian Telford, Rea Wilson, Danielle Hammond, Carloline Ogden, Adele Hayes and Ben Toner Clewes research assistants on the study.

Lesley Crawford and Terry Sharkey from Tees Esk and Wear Valleys NHS Foundation Trust for supporting me through this PhD and funding the trial therapists.

Mark Dawson and Liz Bailey the trial therapists.

My parents, their lifelong support gave me the courage to embark on this PhD.

Finally and most importantly thanks go to my wife Imelda and my children, Ciaran and Frances for putting up with me throughout this long process and being there for me when the going was hard. Their patience, encouragement and smiles kept me going.

Authors Declaration

The research outlined in this thesis has resulted in two publications in academic peer reviewed journals:

Ekers D, Richards D, Gilbody S. (2008) A Meta Analysis of Behavioural Therapy for Depression. *Psychological Medicine*; 38(5): 611-623.

Ekers D, Richards D, McMillan D, Bland M, Gilbody S. (2011) Behavioural Activation delivered by the non specialist: phase II randomised controlled trial. *British Journal of Psychiatry*; 198(1): 66-72

One additional paper is in press with the British Journal of Psychiatry:

Cost-utility of Behavioural Activation delivered by the non specialist
David Ekers, Christine Godfrey, Simon Gilbody, Steve Parrott, David Richards, Danielle Hammond, and Adele Hayes

The research reported in this thesis was led at all stages by David Ekers (DE). The research protocols for the systematic review and randomised trial were designed by DE and supported by research supervision with Professor Simon Gilbody (SG), Professor David Richards (DR) and Professor Christine Godfrey (CG). The contents of this thesis are the work of DE with the specific assistance outlined below.

Systematic Review and Meta Analysis

The systematic review and meta analysis was led by DE who adapted data extraction and quality assessment forms and conducted the primary statistical analysis. Specific assistance was given by the following people:

Janet Menton (JM) librarian at Tees Esk and Wear Valleys NHS Foundation Trust who assisted DE in the design of the search strategies and pilot searches. JM also conducted the searches for the review and produced the summaries in appendix I.

SG collaborated with DE on the design of the statistical analysis. SG conducted the regression analysis and tests for publication bias producing the figures included in the thesis of those tests.

Randomised controlled trial

The randomised controlled trial protocol was led by DE who designed the intervention, trained therapists, managed the trial and conducted the primary statistical analysis. Specific assistance was given by the following people:

Martin Bland (MB) Professor of Health Statistics at the University of York who produced randomisation computer generated randomisation codes. MB provided expert statistical consultation and conducted multiple imputation calculations included in the thesis providing the results for DE to report.

Dean McMillan (DM) Senior Lecturer in Health Services Research University of York assisted DE in the design of the clinical significance testing and in production of results and figures used in thesis.

CG Professor of Health Economics University of York advised on the design of the health economics arm of the study.

Steve Parrott (SP) Health Economist University of York who conducted bootstrapped replications and produced cost effectiveness acceptability curves and the figures included in the thesis demonstrating those results.

Adele Hayes (AH) and Danielle Hammond (DH) research assistants at Tees Esk and Wear Valleys NHS Foundation Trust collated the unit costs and used these to translate service use to costs under supervision from DE.

Introduction

This work is an evaluation of Behavioural Activation (BA) for depression in a practical setting. Having reviewed the literature to understand the theoretical background of the therapy and its existing evidence base, the next step was to explore whether this therapy was suitable to be adopted in a UK setting and specifically whether it could be delivered by non specialist therapists.
Embedded within these aims are a number of different objectives and research questions. The structure of the research and therefore the thesis follows guidance set out by the Medical Research Council in relation to complex interventions (Medical Research Council 2000, Medical Research Council 2008) (see fig 1).

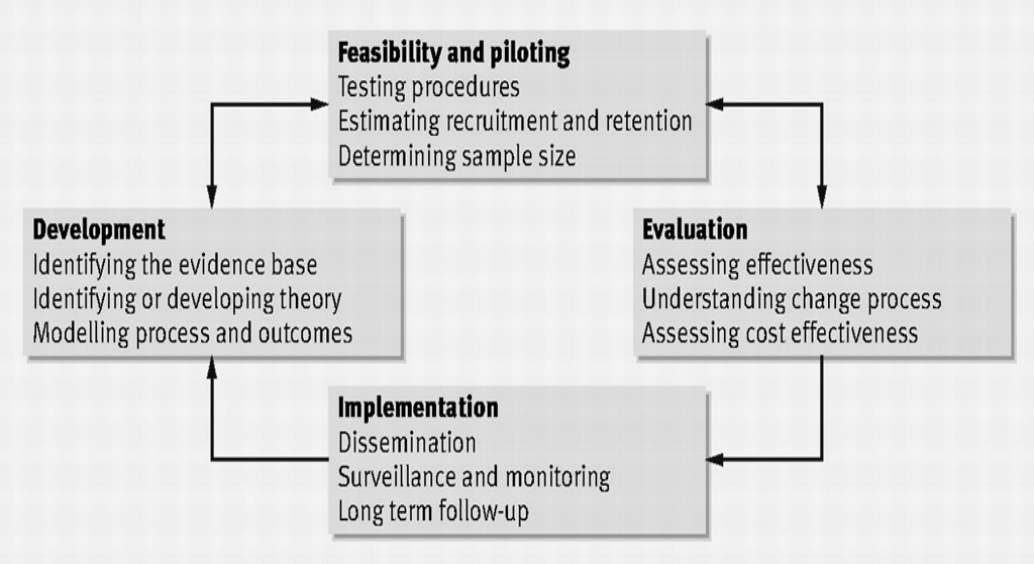


Figure 1: MRC Framework for development and evaluation of complex interventions.

The work presented in this thesis follows a logical progression through the phases of development, feasibility and piloting and into the first stages of evaluation. Results from this work have been subsequently used to inform the design of a large scale multi centre randomised controlled trial to estimate the clinical and cost effectiveness of BA delivered by non specialists with short training vs. CBT, thereby furthering the journey around the MRC framework circle.
The research for this thesis was conducted over 4 to 5 years and conducted in the context of the increasing drive to deliver effective talking therapies in the UK health

environment that offer good value for money and improve accessibility (Layard 2006). By placing research in this context and disseminating findings from the thesis as the research progressed, it was hoped that the results could help inform practice and future policy.

In this research I endeavoured to consider the durability of BA through evaluation of its delivery by non specialist qualified mental health workers. Key figures in the development of behaviour therapy have forwarded assertions that, due to its parsimony, BA was suitable for wide dissemination (Jacobson et al. 1996) however at the time of embarking on this research there was no indication such research had been conducted or was in progress (Ekers et al. 2008).

Intervention development phase

Following Figure 1, the research in this thesis can be put into the different stages required for evaluating complex interventions. The intervention development phase of this research is reported in chapters one to three of this thesis. Firstly the epidemiology of depression and its impact on individuals and society was reviewed (see Chapter One). The difficulties of access to treatment were then considered in relation to the need for the development of an effective single strand psychological therapy that lends itself to wide dissemination.

Chapter Two then examines the theoretical underpinning of behavioural activation through reference to published materials. This examination was used to reflect upon BA's adaptability across health care settings. Variants of the approach are considered and linked to the design of an intervention relevant for this research. This allowed for appropriate adaptation of BA for delivery by non-specialists in contrast to highly trained psychological therapists.

Chapter Three describes detailed analysis of the evidence base through a systematic review with meta-analysis of randomised trials of behavioural therapies for depression. This is used to establish the effectiveness of the approach from all published and unpublished randomised trial evidence and to ascertain gaps in the research literature.

This development phase is an essential component of the MRC approach (Campbell et al. 2000) allowing the development of a behavioural activation research agenda based upon need, theory and all currently available evidence.

Feasibility and Piloting Phase

Chapter four outlines the development, design and delivery of an exploratory randomised trial design based upon findings from the development phase outlined in Chapters Two and Three. Information from the meta-analysis was used to inform sample size calculation and intervention design. A new treatment manual suitable for non-specialist therapists and application in a generalised setting was written incorporating the theory explored in Chapter Two. As value for money is an important consideration in accessibility, both clinical and cost effectiveness were explored in this study. Results are outlined and discussed, relating to the potential benefit BA may offer. This process aimed to develop a new approach to the delivery of BA and present an evaluation of its feasibility in the UK. This has subsequently been used to inform the design of a large scale randomised controlled trial for the next evaluation phase in the MRC cycle of complex intervention development.

Introduction summary

This thesis contributes to the current evidence base of behavioural activation for depression by following a structured approach to exploring complex interventions. It is envisaged that findings will be disseminated widely to improve knowledge of behavioural activation and its use in the treatment of depression. The structure of the remainder of the thesis is as follows:

- Chapter 1: Epidemiology of Depression and issues in the delivery of psychological therapies.
- Chapter 2: BA and its theoretical background.
- Chapter 3: A systematic review and meta-analysis of behavioural therapies for depression.
- Chapter 4: A randomised controlled trial of clinical and cost effectiveness of behavioural activation for depression delivered by the non specialist.
- Chapter 5: Summary and conclusion.

Chapter One: Depression, epidemiology, impact and access to treatment

1.1 Depression

1.1.1 Prevalence

Mental health problems are common among the adult population. More than a quarter of all people are likely to be affected at some time in their lives by such difficulties, which contribute 43% of years lived with disability (YLD) (World Health Organisation 2001). Point prevalence rates of all mental disorders range between 10-15% (McManus et al. 2009). The most prevalent mental health conditions are made up of a range of anxious and depressive symptoms. They indicate a breakdown in 'normal' functioning, are commonly found in community settings and cause substantial distress and disability (Goldberg and Huxley 1992). Anxiety and depression diagnostic categories combined have been seen to contribute over 50% of all disability associated with mental health disorders, in contrast to less than 10% associated with schizophrenia (Henderson et al. 2001).

Depression is one of the most prevalent common mental health disorders seen globally (World Health Organisation 2001). The level of severity and disability associated with depression varies from a relatively mild fluctuating condition to major depressive disorder. There are two main classification tools used for the diagnosis of depression: in Europe the ICD-10 (World Health Organisation 1992) and in the USA the DSM-1V(American Psychiatric Association 1994). Depression can be categorised based upon severity using a symptom count approach: not depressed (fewer than four symptoms), mild depression (four symptoms), moderate (five to six symptoms) and severe (seven or more symptoms). Differences between the two classification tools (see Table 1) have been present since their development more than 50 years ago (DSM I 1952 and ICD 1948). The tools were developed on each side of the Atlantic by the American Psychiatric Association and World Health Organisation (WHO) respectively. Despite attempts to foster international consistency, slight discrepancies have remained. The development of multi-axial approaches to diagnosis in DSM III was a dramatic departure from the approaches used in ICD-9 and its predecessor DSM-II.

Table 1: Diagnostic criteria for DSM-IV and ICD10

	DSM-IV	**ICD-10**
Clinical significance	Symptoms cause clinically significant stress or impairment in social, occupational or other important areas of functioning.	Some difficulty in continuing with ordinary work and social activities, but will probably not cease to function completely in mild depressive episode; considerable difficulty in continuing with social, work or domestic activities in moderate depressive episode; considerable distress or agitation, and unlikely to continue with social, work, or domestic activities, except to a very limited extent in severe depressive episode.
Duration	Most of day, nearly every day for at least 2 weeks.	Duration of at least 2 weeks is usually required for diagnosis for depressive episodes of all three grades of severity.
Classification of severity	Five or more of following symptoms; at least one symptom is either depressed mood or loss of interest or pleasure: (1) Depressed mood (2) Loss of interest (3) significant weight loss or gain or decrease or increase in appetite (4) Insomnia or hypersomnia (5) Psychomotor agitation or retardation (6) Fatigue or loss of energy (7) Feelings of worthlessness or excessive or inappropriate guilt (8) Diminished ability to think or concentrate, or indecisiveness (9) Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or suicide attempt or a specific plan	Depressed mood, loss of interest and enjoyment, and reduced energy leading to increased fatigability and diminished activity in typical depressive episodes; other common symptoms are: (1) Reduced concentration and attention (2) Reduced self-esteem and self-confidence (3) ideas of guilt and unworthiness (even in mild type of episode) (4) Bleak and pessimistic views of the future (5) Ideas or acts of self-harm or suicide (6) Disturbed sleep (7) Diminished appetite Typical examples of "somatic" symptoms are: loss of interest or pleasure in activities that are normally enjoyable; lack of emotional reactivity to normally pleasurable surroundings and events; waking in the morning 2 h or more before the usual time; depression worse in the morning; objective evidence of psychomotor retardation or agitation; marked loss of appetite; weight loss; marked loss of libido. • Mild depressive episode, two of most typical symptoms of depression and two of the other symptoms are required. If four or more of the somatic symptoms are present, the episode is diagnosed: With somatic symptoms.

		• Moderate depressive episode, two of three of most typical symptoms of depression and at least three of the other symptoms are required. If four or more of the somatic symptoms are present, the episode is diagnosed: With somatic symptoms. • For severe depressive episode, all three of the typical symptoms noted for mild and moderate depressive episodes are present and at least four other symptoms of severe intensity are required.

Despite the variance in approach to diagnosis, relatively consistent pictures emerge regarding the epidemiology of depression and its impact on the individual and society.

Those suffering from depression are likely to experience significant distress, impaired daily functioning and an increased risk of suicide (Hirschfeld et al. 1997). The UK national psychiatric morbidity survey used a sample of close to 10,000 individuals to estimate the prevalence of mental health problems using the Revised Version of the Clinical Interview Schedule (CIS-R) (Lewis et al. 1992) to assign diagnosis. Mixed depression and anxiety was identified as the most common mental disorder with a prevalence of 8.8%, depressive disorder alone having a prevalence of 2.6% (Singleton et al. 2001). This picture did not change markedly when the survey was repeated in 2007 with 8.8% of adults found to have mixed anxiety and depression and 2.3% a depressive episode in the week prior to the survey (McManus et al. 2009) (see Figure 2). The co-occurrence of depression and anxiety disorders is long established. In the United States 58% of those with major depression have an additional anxiety disorder and two thirds of those with generalised anxiety and or panic disorder have a history of depression (Kessler et al. 1998). This pattern is a consistent finding in epidemiological studies of working age adults, older adults (Beekman et al. 2000) and those exploring sub threshold disorders (Pollack 2005).

Such point prevalence rates indicate the scale of common mental health problems and in particular depression and mixed anxiety and depression at any given time. This is expanded if we look beyond point prevalence and examine the rates of these conditions in a population over a longer time span. Using the DSM-IV in a sample of 9282 adults in the USA,12-month and lifetime prevalence of major depression was explored (Kessler et al. 2005). In this survey, rates of 6.7% for 12 month depression and 16.6% for lifetime depression were observed, with 23.3% and 37.3% of those identified experiencing severe and moderate levels of symptoms respectively. This survey identified higher rates of depression than the UK

surveys, which may be explained by the use of 12 months' and lifetime time spans.

In summary it is reasonable to assume that more than one in ten people in the community are likely to be experiencing depression or mixed anxiety and depression at any given time. These numbers are likely to increase when viewed over a longer period with approximately half of these experiencing significant symptoms that warrant intervention (National Institute for Health and Clinical Excellence 2011).

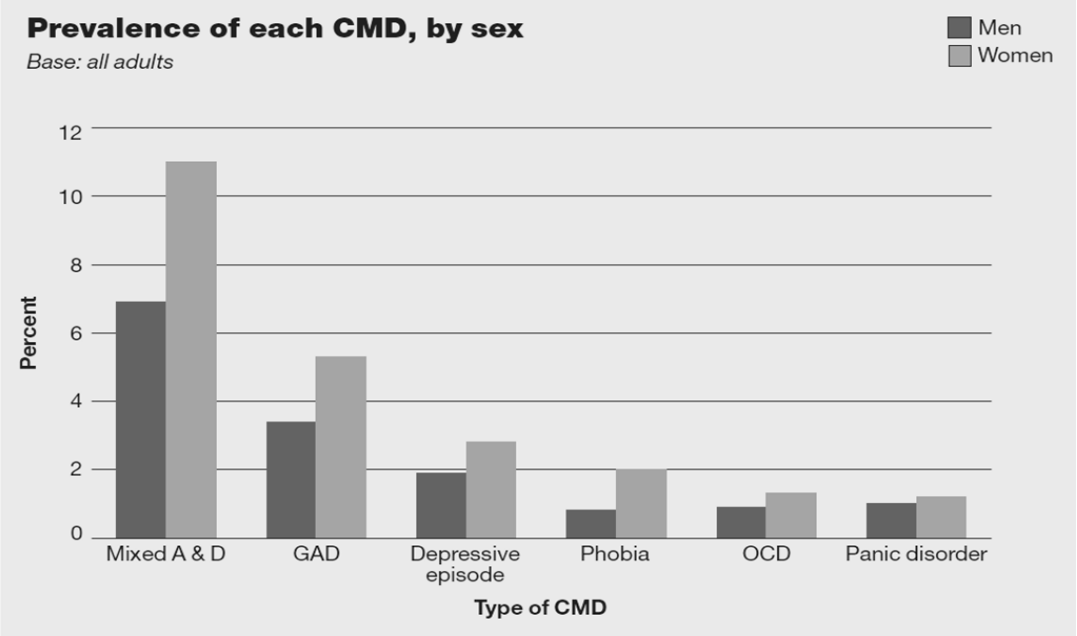


Figure 2: Prevalence of common mental health disorder in the week before assessment. Taken from psychiatric morbidity survey 2007

1.1.2 Treated depression

The majority of depression is treated within primary care (Fletcher et al. 2006) with the cost for such treatment being estimated at twice that of all inpatient psychiatric services (Pincus and Pettit 2001). Prevalence of treated depression in primary care in 1998 was 29/1000 for males and 70/1000 for females (Office for National Statistics 2000). Such data, however, is based upon a recording of diagnosis and drug treatment on

primary care systems. As such they are indicators of healthcare cost and demand, but should be viewed with caution as valid indicators of true disorder prevalence. Numbers are likely to be biased by factors such as variance in help seeking behaviour of patients and correct diagnosis, coding and prescription by practitioners.

1.1.3 Onset and Prognosis

The mean age of onset for depression is 30 years (Wilhelm et al. 2006, Kessler et al. 2005) with a strong likelihood of a recurring and chronic course. However wide variation can be seen, with a substantial proportion of people having a first episode in childhood or adolescence (Fava and Kendler 2000). At one year 60% of those treated routinely in primary care and 50% of milder depressions meet criteria for caseness (Von Korff and Goldberg 2001). It has also been observed that approximately half of people with depressive illness will experience recurrence, with further relapse being associated with poorer prognosis, following a third recurrence such risk of relapse raises to 90% (Kupfer 1991). Biological factors, social stress and life events appear related to increased duration of depressed episodes (Singleton et al. 2001). Such poor prognosis (Lloyd et al. 1996, Kennedy et al. 2004, Judd et al. 1998), reflected in older adults also (Cole et al. 1999), has led to suggestions that depression is best managed as a chronic condition rather than a series of acute episodes managed separately (Andrews 2001).

1.2 Determinates of depression

1.2.1 Socioeconomic status and depression

The association between poor socioeconomic status and ill-health has long been acknowledged and reflected in mental health conditions (Office for National Statistics 2000, Taylor et al. 1997). Low economic status is associated with increased morbidity rate of depression (Lorant et al. 2003). In common mental health conditions in general across Europe and North America, a near two-fold increase of morbidity between lowest and highest socioeconomic groups is observed (WHO International Consortium in

Psychiatric Epidemiology 2000, World Health Organisation 2001). In the United Kingdom poverty and unemployment is associated with longer duration of common mental health problems and increased financial strain (Weich and Lewis 1998, Lorant et al. 2003). The strongest predictor of the prevalence and persistence of depression in a large-scale depression management study was seen to be the measure of social deprivation of the GP practice location (Ostler et al. 2001). This was reflected in observed rates of treated depression prevalence in England and Wales. Highest rates of treated depression were found in deprived industrial areas (76.9 female, 33.5 male per 1000) (Office for National Statistics 2000) (Figure 3). It is of note that in this study those categorised from metropolitan professional areas had higher prevalence rates of treated depression than those from inner-city estates/deprived city areas. This is an interesting finding in and would appear to be at odds with other evidence linking increased rates of depression with poverty. It may be related to the use of treated depression to measure prevalence in this study. A number of potential factors may bias these results; in particular they may reflect the possibility of inverse care law factors in relation to treatment uptake, i.e. that good quality medical care is seen to vary inversely to population needs (Tudor Hart 1971). That is, those in metropolitan professional areas may be more likely to seek help and to be treated by professionals more likely to identify need and offer help. This then rewards and increases help-seeking and giving in those areas, hence increasing the treated prevalence estimate.

Changing socioeconomic status has an impact on the diagnosis and severity of depression (Lorant et al. 2007). In a sample of 11,909 individuals in Belgium, associations were noted over seven year follow-up that indicated a negative effect from worsening socioeconomic status. The impact of negative associations were far greater than positive effects from improving conditions and impacted on females and those in low income households more.

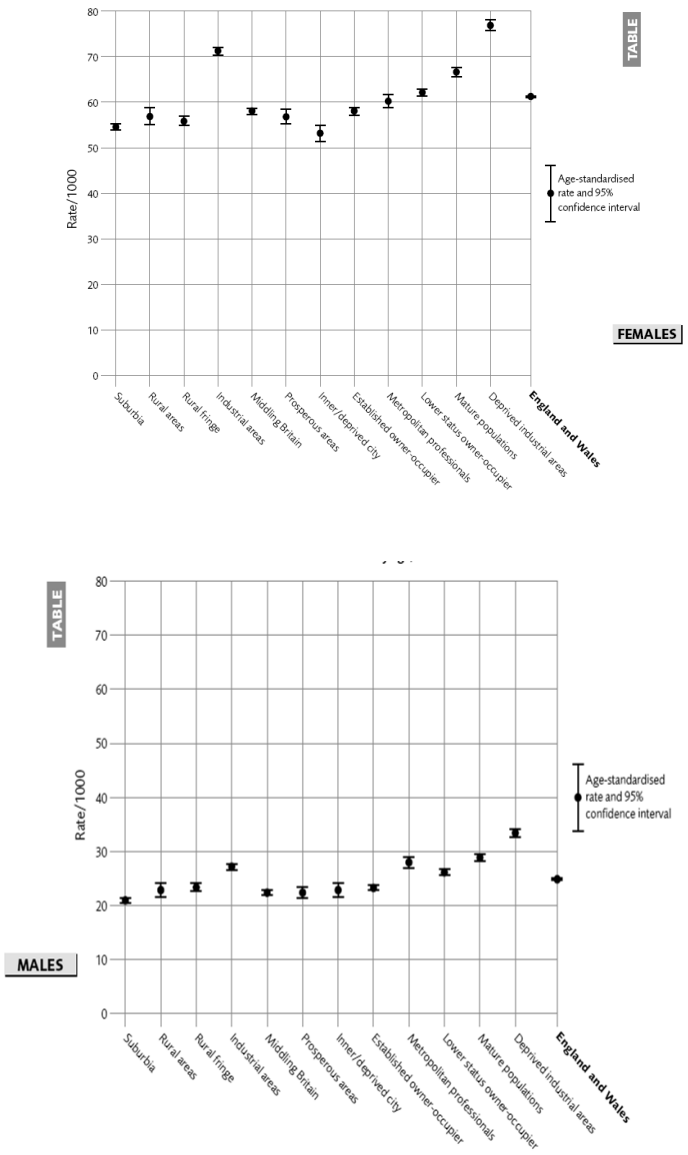


Figure 3: Prevalence of treated depression in primary care by ONS area classification, age standardised and aggregated 1994-1998 (Office for National Statistics 2000)

1.2.2 Gender and depression

As indicated in fig 3 gender has a significant effect of the prevalence rate of depression in the United Kingdom. In a large UK study (N=9792) working age women experienced 1.5-2 times more depression than is observed in males. Differences are not explained by the effect of

motherhood, employment or marital status (Bebbington et al. 2003) see fig 4. This picture has remained relatively constant across the 15 years the survey with the largest increase in prevalence across all common mental health disorders of one fifth seen in women between 45-64 (McManus et al. 2009).

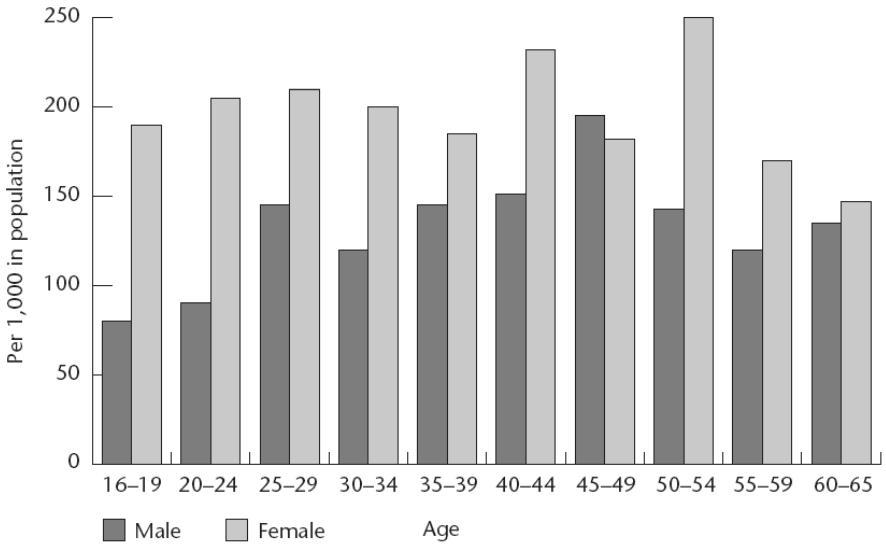


Figure 4: Common mental disorder by age and sex: Adult Population of Great Britain 2000

1.2.3 Ethnicity and depression

Higher rates of depression have been found in middle-aged Irish and Pakistani men and older Indian and Pakistani women (Weich et al. 2004). Results however are inconsistent and many confounding factors exist. Difference in cultural attitudes to mental health help-seeking behaviours and pathways to care influence findings (Bhugra and Mastrogianni 2004) alongside somatic description of depression symptoms in non-European cultures and an association between ethnic grouping and unemployment, low income, social class and support (Brugha et al. 2004). This makes it hard to define clearly any links between ethnicity and increased rates of depression compared to the general population. It is likely that depression is under-diagnosed in this group (Bhugra and Mastrogianni 2004), but any

potential indication of higher prevalence rates may be a result of other socio-economic factors commonly associated with black and minority ethnic grouping rather than ethnicity itself.

1.3 Economic factors and disease burden of depression

In the UK over 900,000 adults are claiming incapacity benefits for mental health disorders, this accounts nearly 50% of all claimants, with 38% having a mental health condition as their main disability and a further 10% as an additional disability (Layard 2004) (see Figure 5). Depression either alone of mixed with anxiety makes up a substantial proportion of this and resulted in prescribing costs of antidepressants in 2005 of £338 million (Bird 2006). This cost is only a small figure when placed against the estimates of the economic impact of common mental health disorders as a group (depression, anxiety and stress). The Confederation of British Industry assessed this cost at £25 billion, roughly 2% of the gross domestic profit, with £4 billion due to time of sick and direct loss of output, £9.4 billion due to lost output associated with economic inactivity, £4 billion time for carers and £8 billion public services cost (Layard 2006, Layard 2005). Mixed anxiety and depression have been estimated to cause on fifth of all sick days in Britain (Das-Munshi et al. 2008), with the impact of lost productivity and employment due to depression estimated at 23 times that of the direct treatment cost of these conditions (Thomas and Morris 2003).

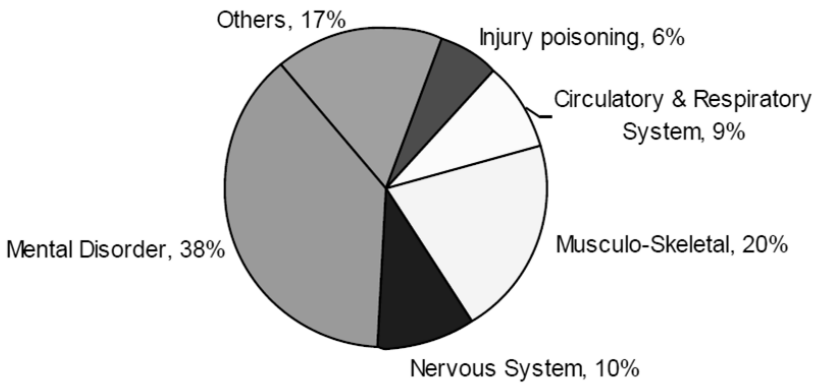


Figure 5: UK incapacity benefits recipients by medical conditions, 2004. Layard R (2005)

Depression is also associated with high levels of quality adjusted life year loss. In a large general population survey including 8028 people living in Finland depression contributed 55% of all such losses measured in the study (Saarni et al. 2007). Globally depression accounts for 4.4% of the global disease burden and 65 million disability adjusted life years annually and is set to become the second largest cause of disease burden in developed countries by 2020 (World Health Organisation 2002).

Depression is also commonly associated with chronic physical disease with a two to three fold increase in prevalence over an equivalent healthy sample (National Institute for Health and Clinical Excellence 2009). It worsens health outcomes and costs and increases the risk of death (Nicholson et al. 2006, Moussavi et al. 2007). Depression also leads to a fourfold increase in the risk of suicide compared to the general population (Bostwick and Pankratz 2000).

With signs of increasing prevalence (McManus et al. 2009) and demand for treatment (Hollinghurst et al. 2005), it has been suggested that the

estimates of the societal burden of depression have been underestimated (The Sainsbury Centre for Mental Health 2003).

1.4 Access to effective treatment for depression

It can be seen that depression is a chronic and disabling condition. It increases the risk of premature death either by suicide or worsened physical health outcome and has a high cost to society. In the decade up to 2002 a 2.8-fold increase in antidepressant prescription and a doubling of recorded GP consultations for depression were recorded (Hollinghurst et al. 2005). A number of both pharmacological and psychological treatments are effective (National Institute of Clinical Excellence 2009); however most treatment remains within primary care, and half of those with depression are likely to receive no treatment. Of those that do receive an intervention, the most common intervention is medication, with a psychological intervention being delivered in 25% of cases (McManus et al. 2009). It is of note that, while CBT is the recommended first line psychological intervention, only 7% received this treatment. Even with the roll-out of Improving Access to Psychological Therapy Services nationally in the UK aimed at increasing the availability of CBT, relatively small numbers are accessing such interventions (Glover et al. 2010). This indicates continued difficulty relating to accessibility of these treatments. Training a CBT therapist takes 1-2 years and is costly, making swift local adjustments to service provision difficult. Briefer 'single strand' interventions have for some time been seen as a possible beneficial additions to standard treatment options (Lovell and Richards 2000). These approaches may lend themselves to wider and swifter dissemination, and if effective compliment the choice of psychological interventions for depression. Behavioural activation (BA) presents itself as such an approach and for some time has been seen in randomised controlled trials to be effective and potentially suitable for wider dissemination (Jacobson et al. 1996).

This thesis will explore the background and evidence base for BA and its potential suitability for meeting some of the demands of providing widely accessible effective psychological interventions for depression raised in this first chapter.

Chapter Two: Behavioural Therapy and its application in the treatment of depression

2.1 What is behavioural activation?

Behavioural activation (BA) is a term now commonly used to describe treatments that use behavioural theory to explain depression onset and maintenance. BA views depression as a natural response to life events that interfere with a person's contact with naturally occurring positive reinforcement. This results in behavioural reductions, as behaviours are no longer positively reinforced. To cope with the feelings of sadness or anxiety which subsequently arise as a result of isolation from positive reinforcement, avoidance occurs. This further removes people from the environment in which where they have previously attained positive reinforcement, hence the context of the person-environment relationship is fundamentally changed. This is described as a contextual approach; depression is based in the context of the person's world and how they interact with it. The behavioural therapist is interested in the manipulation of these person-environment interactions to create a situation non-conducive to a depressed state. In behaviourist terms, this means reintroducing contact with regular and stable positive reinforcement from a range of sources in the person's environment. Behavioural therapists therefore are not seeking to manipulate internal states prior to change; rather they manipulate person-environment interactions, resulting in subsequent internal change. This is contrast to other approaches such as cognitive and biochemical therapies which could be described as internal deficit models, they seek to change faulty internal states (such as thoughts/serotonin levels) to then facilitate behavioural change.

The first use of the term behavioural activation in relation to psychotherapy would appear to be in 1990 as a description of the behavioural components in cognitive therapy (Hollon and Garber 1990). Prior to that time behavioural therapy was the commonly used umbrella term for this set of clinical interventions. It would appear however that behavioural activation (BA) has become the commonly adopted description over the past two decades (Martell et al. 2010). BA can be defined as a brief psychotherapeutic approach that seeks to:

- Increase activities that are associated with a sense of value to the person
- Reduce activities that result in isolation from a sense of value for the person
- Solve problems and change behaviour that limits access to a sense of value to the person.

Treatments are collaborative and focussed in the here and now. Many differing techniques are incorporated into treatment; however all use self-monitoring of a mood-environment link and scheduling of new or adaptive behaviours to meet targets (Kanter et al. 2010). BA strategies are commonly seen in other therapies (Dimidjian and Davis 2009) such as CBT and mindfulness, and were present in the early stages of cognitive therapy; they also sit well alongside the modification of interpersonal context seen in brief interpersonal therapy (Klerman et al. 1984).

2.1.1 Theoretical explanation of BA

A functional analytic perspective is used in behavioural activation to understand the development and maintenance of depression (Martell et al 2001). The function of the presenting behaviour is explored in order to understand its role in ameliorating distress and therefore how it is more likely to be repeated. The form of a behaviour (what it looks like) is of less interest in the development of a behavioural perspective and individual formulation. Skinner's writings on reinforcement (Skinner 1953) provide the basis of the behavioural understanding of depression. A reinforced behaviour is one that increases; BA seeks to understand how negatively reinforced avoidance has replaced positively reinforced 'healthy behaviours'.

Positive reinforcement occurs when an action is followed by a rewarding consequence. It is to be noted this is not a necessarily a pleasurable consequence although that may be the case. An example of positive reinforcements is the taste of a particular food. A person will see the food and take a bite which is perceived as pleasurable, hence making a second

bite and a repeat of the behaviour more likely next time the person encounters that food. This exemplifies those instances where pleasure is the reinforcing consequence. Another example is opening a door; the goal of the behaviour is to get to the other side; the behaviour is turning the handle and pushing which achieves free movement through the door, a positive outcome. While this may not be experienced as pleasure, the behaviour is likely to be repeated the next time the person is confronted by a door, since the same consequence is sought. It is of note that if one is confronted regularly with a particular 'sticky' door that requires a 'turn-lift-push' sequence of behaviour, in that particular context (by that door) the previously learnt behaviour is rapidly modified to achieve the same outcome. Soon this new behaviour is incorporated and requires no pre-thought; the new positively reinforced behaviour has replaced the previous one in this context to achieve the function of moving about the house.

Negative reinforcement occurs when an action is followed by the removal of a negative experience. An example is when a thirsty person takes a drink of water. The water removes the negative experience of thirst and therefore is likely to be repeated in the same or similar context. After time the drinking behaviour is seen to be repeated to prevent the experience of thirst. Therefore a negatively reinforced behaviour is designed to remove a direct or anticipated negative consequence.

There are two other important determinants of behaviour important to the behavioural therapist. Punishment reduces behaviour, as it is followed by a direct negative consequence. A child being shouted at as he or she put out a hand to touch a fire is seen to reduce that behaviour since shouting is perceived as unpleasant. Frustrative non-reward occurs when expected rewards do not occur. This also leads to a reduction in behaviour; an example would be stopping working when it is not followed by the expected payment, the expected positive consequence. Table 2 outlines the relationship between each of these learning conditions.

Table 2: Summary of operant conditioning

	Presented	**Omitted**
Positive	Positive reinforcement	Frustrative no reward
Negative	Punishment	Negative reinforcement

Ferster (Ferster 1973) was the first to hypothesise depressed behaviour will be maintained via operant conditioning. If a child requires a large amount of activity before positive reinforcement is achieved alongside behaviours being frequently negatively reinforced this will lead to the development of a passive approach to life. When in later life circumstances become difficult, such as after a relationship break-up or the loss of a job, this passive style then makes that individual more likely to respond by avoidance. This is due to historical learning being adopted in the current context. The person becomes isolated from positive reinforcement from their environment and increasingly avoids with the function of ameliorating distress in the short term but long term maintenance of problems.

Viewed from this perspective, behaviours that are displayed in depression appear logical, or to be more precise functional, in the context of the whole of the person's life and experience at that point in time. When formulating the maintenance of depression within a behavioural framework, five particular environmental contingencies are important to explore: (1) an increase in negative reinforcement of avoidant behaviour, (2) a reduction in positive reinforcement of non-depressed behaviour, (3) an increase in positive reinforcement of depressed behaviour, (4) a punishing/aversive environment and, (5) response cost factors.

1. An increase in negative reinforcement of avoidant behaviour

Much of the behaviour change in depressed patients functions as attempts to reduce aversive experiences. That is, either to directly ameliorate discomfort or to prevent aversive experience worsening. This process of negative reinforcement explains the increase of avoidance behaviours, since the action reduces discomfort. Withdrawal from work, social contact, family and friends can then be seen as a natural response to feeling low in mood, anxious, or physically lethargic, and become consistent responses to sources of discomfort. When, due to a significant event in a person's life, they feel the wide range of physical, emotional and cognitive symptoms that typify low mood, these avoidant responses become more than temporary coping. They become the primary method of responding to daily experiences in an attempt to cope with distress; each time the distress is felt, the avoidance reduces it. This results in increased isolation from sources of positive reinforcement hence reversal of the pattern is unlikely and depression results.

2. A reduction in positive reinforcement of non-depressed behaviour

Life events can occur to anyone. One consequence of this is that people become naturally isolated from positive reinforcement from their environment. This is clear in terms of a negative life event, such as the death of a partner. Following such a distressing event the person no longer has the direct contact with the partner, but also does not perform some of the shared activities previously enjoyed, i.e. walking. In consequence they will no longer be obtaining reinforcement from the person lost, and in addition will have discontinued the shared activities previously enjoyed. It is important to also consider how depression may also follow an apparent positive life event. A new baby, for example, may be seen to be a positive life event for a mother who was keen to conceive. The subsequent reduction in previously valued activities, such as socialising and work, however may result in her feeling somewhat isolated, a negative feeling. If in addition the care of the baby is more problematic than anticipated, there

arises a significant reduction in positive reinforcement from the mother's environment. If attempts at coping are (as Ferster suggested) passive, then negatively reinforced avoidance results and depression is more likely to occur. The behavioural activation therapist must therefore consider a range of events occurring at the time of onset or worsening of the episode of depression, and consider the impact of these on a person's interaction with environmental positive reinforcements.

As an additional consequence of their increasing avoidance the person is likely to become further removed from contact with naturally occurring positive reinforcements in their environment. This will lead to an exacerbation of their depression as the person gains less reward from their behaviours in daily life. For example, social contact or work may become more challenging and provoke anxiety as a consequence of symptoms of depression such as worsened concentration and memory. This results in a reduction in positive reinforcement from such behaviours, therefore they begin to decline or stop altogether, being replaced by negatively reinforced avoidant behaviours. This then explains the process by which the person becomes increasingly depressed.

3. An Increase in positive reinforcement of depressed behaviour

Alongside increased avoidance the depressed person may receive positive reinforcement of depressed behaviour. Behaviours like resting, missing work, and reduced general activity become common. These behaviours initially are seen as attempts to cope with distressing symptoms and can be encouraged in that person's environment. Actions may be initially positively reinforced directly by friends, relatives, work colleagues and health professionals who encourage them to 'slow down' or 'take some time out and get some rest'. This positive reinforcement for behaviour change as the person struggles to cope is also negatively reinforced by the amelioration of distress. These two factors combined can explain the how a person gets into a cycle of behaviour in depression that increases their isolation.

4. A punishing or aversive environment

It is likely that the depressed person will experience their environment as increasingly punishing. Continued efforts to manage symptoms through the methods already described may lead to environmental contingencies changing in function if not in form. Work may become increasingly problematic, and attempts to cope as outlined above worsen performance, thereby making work even more difficult. In addition, where the person's depressed behaviour was initially positively reinforced by friends and relatives, over time this behaviour stops being viewed as understandable; some people may become impatient or irritable when faced with a person's distress and avoidance, thus resulting in punishment of their social interactions, leading to a further reduction in behaviour and increased isolation.

5. Response cost factors

Once a depressed pattern of behaviour is primarily reinforced by avoidance of discomfort, attempts to reverse those behavioural patterns are more difficult than previously experienced. These attempts to change then have a greater cost than benefit and become aversive experiences in themselves. As an example one could take a person's efforts to reverse physical inactivity by re-starting a previously well-practised exercise routine. Initially this leads to a greater effort being needed than before the period of physical inactivity for the same goal (such as running one kilometre). The person feels physically worse after the exercise alongside the reduced ability to perform at previously held levels, thus the new behaviour induces despondency. This in turn results in the new behaviours (response), which were aimed at positive change, being punished (cost) and therefore their frequency is reduced.

2.1.2 Behavioural model of depression

The above factors combined lead us then to an understanding of depression based upon a behavioural theory. Life events initially result in a reduction

of naturally occurring positive reinforcement. This results in feeling bad, which leads to negatively reinforced avoidance. This then further isolates the person from positive reinforcement; new attempts to change are often punished, and so the cycle is maintained (see Figure 6).

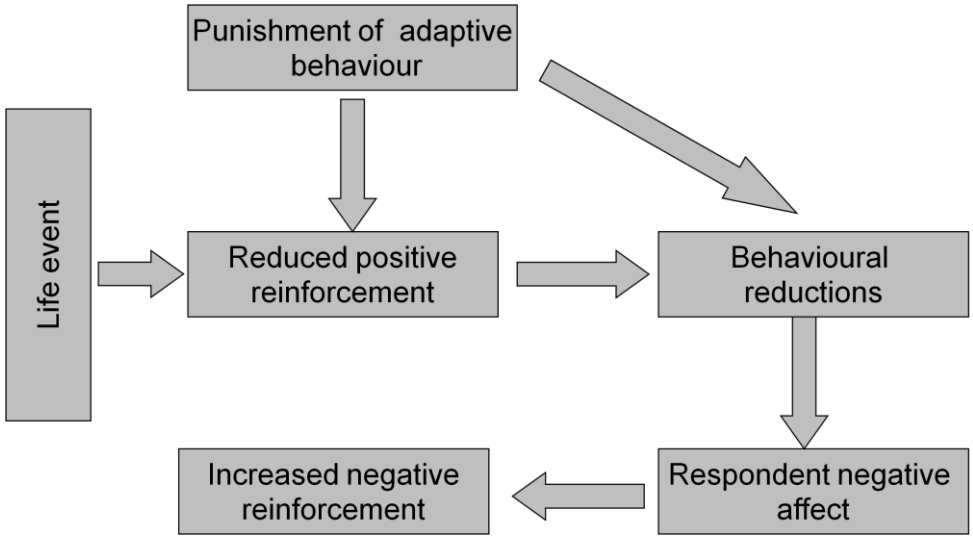


Figure 6: Behavioural Model of Depression

Early research supported the link between a reduction in positive reinforcement (Lewinsohn and Graf 1973) and higher rates of punishing environments (Rehm 1978) in depressed subjects compared to non depressed subjects. These early studies were used to inform the development of behavioural therapies for depression based upon this operant conditioning framework.

2.2 Historical development of BA as a treatment for depression

As seen above Behavioural activation (BA) is a treatment for depression which has its roots in the writings of Skinner (Skinner 1953), being focused on the role of the environment and on people's responses to it. Behavioural models of depression put forward by Ferster (Ferster 1973) and Lewinsohn (1974) took the principles of operant conditioning and used them to improve understanding of depressed behaviour (as outlined above).

Interventions based upon these principles were focused on increasing access to positive reinforcement, with the aim of interrupting the cycle seen in Figure 6. Early adopters of this model monitored the links between behaviour and mood to develop an initial understating of a person's depression. They then used activity scheduling, social skills training and problem-solving to modify a person's relationship with their environment. This aimed to increase contact with potentially 'anti-depressant' experiences, and was often focused upon 'pleasant events' (MacPhillamy and Lewinsohn 1982). Initial evaluations of these reasonably straightforward approaches showed encouraging results in early trials (Brown and Lewinsohn 1984); however, research found that the use of pleasant events did not sufficiently explain patterns seen in depression (Amenson and Lewinsohn 1981). This led Lewinsohn to review his approach, to consider a more complete understanding of reinforcement schedules and understanding of thinking as a covert behaviour that can in turn be modified. The new model identified environmental stressors as a primary trigger for depression that disrupted individuals' behavioural repertoire and resulted in negative affect and reduced positive reinforcement. It incorporated the person's efforts to cope with the stressor by avoidance and subsequent deterioration resulting in depressed 'cycles' emerging (Lewinsohn et al. 1985). This model again resulted in positive results in early studies (Lewinsohn et al. 1980), providing a therapeutic intervention for depression based within a conceptual framework of behavioural theory.

Early trials of these approaches resulted in promising findings; however purely behavioural treatments lost favour in the 1980s. The development of cognitive interventions for depression, most notably Beck's cognitive therapy (CT) (Beck et al. 1979) replaced interest in behavioural approaches. This dominance occurred despite a lack of compelling evidence demonstrating the superiority of cognitive-based treatments. For example, Zeiss, Lewinsohn and Muñoz (Zeiss et al. 1979) found no differences between cognitive and behavioural treatments on reduction in depression symptoms. Despite some exceptions (e.g. Shaw 1977), there

was very little evidence of the dominance of cognitive over behavioural approaches. The new cognitive approaches nonetheless resulted in reduced interest in behavioural models which were no longer commonly accepted as adequate explanations of human behaviour and functioning. The focus on the importance of thinking then became the main psychological approach to formulation and treatment of depression. Cognitive Behaviour Therapy (CBT) or Cognitive Therapy thus became the main focus of research agendas, and was considered the first-line psychological treatment for depression. Cognitive treatments still included self monitoring and activity scheduling in their treatment packages (Beck et al. 1979). These were used as experiments to modify assumptions relating to control of mood in contrast to re-establishing contact with positive reinforcement, and the central behavioural theory as such was lost. Throughout the 1980s and early 1990s CBT's dominance was maintained in the research literature and the early promise of a simple and effective behavioural intervention for depression was mostly forgotten.

Interest in a purely behavioural approach to treating depression returned with the trial of Jacobson et al. (Jacobson et al. 1996). Jacobson noticed when involved in trials of CT that a large amount of change in participants' symptom levels occurred early in treatment. This appeared to be at the time when the behavioural techniques (i.e. activity scheduling) were used (Jacobson and Gortner 2000). Cognitive theory asserts that modifying underlying cognitive structures were responsible for the effects of CT. Improvements however appeared prior to modification of cognitions, suggesting a possible alternative explanation: changes could be due to basic behavioural interventions which bring back contact with positive reinforcement from the person's environment. Jacobson's study used a component analysis design to compare full CT with thought challenging and behavioural techniques and with behavioural activation alone. The main finding of the study that caused controversy (Jacobson and Gortner 2000), was that there were no differences in reduction in depression symptom level between the three treatments. Results were maintained through a two- year follow up period (Gortner et al. 1998). This indicated

that BA was potentially as effective as CT in acute response, but also provided a durable intervention for consideration. This finding triggered a renewed interested in purely behavioural approaches, particularly as they might potentially offer the possibility of providing a simple effective treatment that would be more amenable to dissemination than other more complex psychological interventions (Jacobson et al. 1996). Interest in this finding has grown particularly following a replication and extension of the initial Jacobson et al. trial (Dimidjian et al. 2006) which again demonstrated long term durability (Dobson et al. 2008). Other researchers were also looking at the potential of behavioural theories to guide the development of depression treatment and were developing separate treatment packages, notably Behavioural Activation Treatment for Depression BATD (Lejuez et al. 2001).

Much of the research and development of behavioural treatments for depression have come from the USA. The UK had taken behavioural theories and developed their use in anxiety disorders over a number of years (Marks 2000); however this development did not extend to depression, where CT remained the focus of clinical and research activity. More recently, with the interest in single-strand treatments and stepped care increasing, BA has become increasingly used in the UK and is now recognised as an effective and valuable treatment of depression (National Institute of Clinical Excellence 2009).

2.3 Relating behavioural theory to intervention

As highlighted above, behavioural theory places the person's depressed condition firmly within an interrupted relationship with their environment. in contrast to other approaches that explain depression as an internal deficit due to faulty thinking, neurotransmitter problems or unconscious conflict (Martell et al. 2001). Behaviourists do not believe depression to be due to something being 'wrong' inside but that it is an understandable pattern emerging based upon events changing reinforcement contingencies; thus if these are explored, understood and manipulated, then the depression will

remit. Therefore treatment does not endeavour to resolve internal feelings or thoughts prior to behavioural change, but instead uses behavioural change to manipulate person-environment interactions, with the aim of subsequently improving those internal feelings. Simply put, behavioural therapy asks people not to 'wait to feel right to do things' but 'to do things to feel right'. When planning BA treatment, the goal is to develop sources of positive reinforcement in a person's life that are stable (the activity and reinforcing effects are repeatable) and diverse (what Kanter describes as 'not putting all of one's eggs into one basket') (Kanter et al. 2009). If this is achieved, it is more likely that the person will maintain contact with one or more sources of value/positive reinforcement, even in the face of future problematic life events.

2.3.1 The therapeutic relationship in BA

Across all cognitive behavioural therapies establishing a good therapeutic alliance is essential. The patient is being asked to give up well-practised, safe and reinforced behaviours and to replace them with activities of uncertain outcome. Results are then evaluated and reflected upon in an approach to develop new meanings or beliefs. This collaborative empiricism is also present in BA via the patient's dropping of negatively reinforced coping, replacing it with an alternative response and observing the consequences. It is crucial, then, that the therapist develops trust early in the treatment by demonstrating empathic understanding of the patient's experience. These non-specific factors are common across therapies (Rogers 1961) and reflect good clinical skills. The therapist also offers a clear understanding of the patient's problem, developing a behavioural formulation that is non-judgemental and derived collaboratively. This behavioural case formulation is then linked to a clear rationale for the treatment proposed. This approach helps the patient to realise that they can have an influence on how they feel, which can provide positive reinforcement within sessions and thus increase ongoing engagement with therapy (Kanter et al. 2009). Throughout the course of behavioural activation the therapist will observe and identify behaviours within the

session that may be examples of the patient's problem, and use their own behaviour (verbal and non-verbal) to try to decrease the frequency of these behaviours through operant methods. Similarly the therapist will attempt to identify desired changes in the patient's behaviour within sessions and to reinforce these. This is an approach used commonly in Functional Analytical Therapy (FAP) and models behavioural principles which are the basis of the treatment within each session (Kanter et al. 2010).

2.3.2 BA models of delivery

Kanter et al (Kanter et al. 2010) conducted a narrative review of behavioural therapy studies and identified several key techniques which have commonly been incorporated into treatment packages since the 1970s. The techniques used in these studies were activity monitoring, assessment of goals and values, activity scheduling, skills training, relaxation training, contingency management, procedures targeting verbal behaviour (i.e. cognition) and procedures that target avoidance. It is of note that the only components appearing consistently over all seven treatment manuals reviewed were activity monitoring and scheduling. This raises some interesting questions, which are yet to be addressed, as to the 'active' components in BA (Kanter et al. 2010). While Kanter's helpful narrative review was not published at the time of the development of the intervention study outlined later in this thesis those studies it highlights were.

There are two main approaches to the application of behavioural theory as described earlier in this chapter used currently that have a body of published data on their use in clinical settings; these are outlined below.

The first is known commonly as Behavioural Activation (BA) (Martell et al. 2001). This approach was derived from the BA intervention in the Jacobson (1996) study. The treatment focuses on the functional aspects of depressive behaviour, identifying environmental triggers to ineffective coping responses which are then linked to the maintenance of depression. Primary symptoms as such (tiredness, low mood) are not the key focus of

BA. Instead, the patient is directed to pay attention to their responses to such symptoms and to their negatively reinforced attempts to cope; these are the targets of BA techniques. Thus in this model behavioural avoidance and its function is central. Treatment is delivered in 24 sessions over 16 weeks (Dimidjian 2006).

The second approach, Behavioural Activation Treatment for Depression (BATD) was developed by Lejuez and colleagues (Lejuez et al. 2001). BATD is based on behavioural matching theory (Hernstein 1970). This model focuses on increased contact with the reinforcement of healthy (non-depressed) behaviour and reduced contact with reinforcers of depressed behaviour with the aim of decreasing depressed behaviour and increasing non-depressed behaviour. (Hopko et al. 2003). The BATD model is delivered in a 8-15 session protocol (Hopko et al. 2003).

2.3.3 Components of BA interventions

Behavioural Activation (Martell et al. 2001, Martell et al. 2010). The first sessions of BA are used to develop a therapeutic relationship with the patient, introduce the formulation and link to the treatment rationale. This aims to develop an increased awareness on the part of the patient of their attempts to cope with symptoms of depression, resulting in only short-term relief and longer-term maintenance of the condition. Initially BA uses activity and mood monitoring to build on this awareness and understanding. Patients keep a daily diary of their activities on an hourly basis with a corresponding statement and rating of mood. This information is gradually incorporated into the initial case formulation and is used to assess general activity level and range of emotion. Associations between specific activities and mood are explored through functional analysis and related to development of understanding of the maintenance of depression. This antecedent-behaviour-response (ABC) is developed and an acronym TRAP: Trigger, Response, Avoidance Pattern used to facilitate a shared language between the therapist and patient. Once patterns of avoidance are understood, the focus of treatment moves to re-establishing healthy

behaviours by developing new alternative coping mechanisms aimed at contact with valued goals. The BA goal is to create new opportunities for encountering positive reinforcement from the person's environment, again using an acronym, TRAC: Trigger, Response, Alternative Coping. BA then uses scheduling and functional analysis to reduce avoidance and increase contact with positive reinforcement throughout treatment. Identification of specific goals and alternative coping strategies are collaborative, seeking to link the patient to their particular set of values. To support this process, BA incorporates techniques such as grading activities, therapist modelling, skills training, problem-solving and mental rehearsal as methods of connecting with positive reinforcement and replacing avoidance with 'healthy behaviours' which is the aim of BA.

BATD (Lejuez et al. 2001, Hopko et al. 2003). As in BA, initial sessions of BATD are used to help establish a therapeutic relationship. It again seeks to identify the changes in reinforcement patterns, using these to describe the treatment rationale. BATD then employs "systematic activation" with the goal of increasing frequency and reinforcement of healthy behaviour. As in BA, patients are asked to monitor their activities but not mood. Activity monitoring is used to emphasise the quality and quantity of a patient's range of activities, and from this to provide possibilities for new behaviours to focus on in treatment. BATD then identifies goals in valued areas such as relationships, education, employment, hobbies and recreation, general health issues, spirituality etc. Within these areas, patients select 15 activities which are organised into a hierarchy of difficulty of achievement, and are then supported to work progressively through the hierarchy. Specific weekly goals are set in relation to frequency and duration of each activity. Additional positive reinforcement is integrated by introducing rewards for completing weekly goals.

2.3.4 Using behavioural theory to improve concordance

Often problems are encountered in activation approaches, moving people from patterns of negatively reinforced avoidance. Functional assessment

procedures can be used to explore obstacles to client progress and resolve problems (Kanter et al. 2009). By returning to basic behavioural principles it is possible to develop an understanding of the difficulties that may arise in therapy and developing approaches to resolve these.

In all approaches to behavioural treatments of depression all patients have clearly defined goals that reverse patterns of avoidance and introduce a sense of pleasure or accomplishment. Values differ from goals in that the term is used to describe a direction for life rather than an end point or target (Hayes et al. 1999). In order to reintegrate positive reinforcement, goals need to be linked to an individual's particular set of values, or they will not achieve that purpose. This means that, in order to effectively overcome potential obstacles, each set of goals used in therapy must be individually negotiated in order to correspond with each patient's particular set of values.

2.4 BA and improving access to the psychological treatment of depression

As highlighted in chapter one of this thesis depression is a highly prevalent condition with high cost to the individual and society. In the UK recent efforts have been made to improve access to psychological treatments through major investment in the IAPT programme (Improving Access to Psychological Therapies). IAPT uses treatments approved by NICE and has yielded some rewards in terms of the numbers of people offered a treatment (Clark et al. 2009, Glover et al. 2010). BA has been incorporated to some degree in this programme and would seem to lend itself to this model of delivery. Central to the delivery of IAPT is the concept of 'stepped care', that is, delivering the least intrusive and costly intervention that is likely to provide a health gain. Stepped care aims to optimise the use of the same level of resources to deliver more effective treatment aiming to make improvements in meeting demand. BA would seem to lend itself to such an approach. Some have suggested that a stepped BA delivery may be used (Kanter et al. 2009) where initial self-monitoring and scheduling are the interventions used, and only if no improvement is seen does the therapist move to delivery of the more complex functional analysis. This

would appear attractive; however, it requires further study. The other promise of BA is the relative parsimony of the approach, which has been suggested for some time (Jacobson et al. 1996). If this is shown, it may suggest that a single-strand BA intervention would lend itself to delivery by a wide range of therapists. Such therapists would be at a lower grade and require less timely training than the one year currently needed to train a CBT therapist, thereby reducing the costs of both training and treatment delivery. The approach of including single-strand therapies in stepped care has long been thought to show promise (Lovell and Richards 2000), but is rarely routinely incorporated into service design. The reason for this is unclear; however, to date it would appear that only limited evidence has supported BA's use in this way. While BA clearly is built on sound theoretical foundations and has an apparent evidence for its use, no studies appear to support the parsimony leading to effective dissemination assertion. It is an area of research that would be very beneficial for supporting effective service redesign.

2.5 Summary of chapter and research focus

Behavioural approaches to the treatment of depression are based in behavioural theory, and have over the past four decades shown promise in providing an effective psychological treatment for depression. They appear to be simple in design and delivery and as such suggest an approach that is suitable for wide dissemination. While these ideas have for some time been suggested as an option to increase availability of effective acceptable psychological treatments, clarity regarding the evidence that supports this approach is needed. Two individual trials (Dimidjian et al. 2006, Jacobson et al. 1996) offer strong support for BA, but both used experienced therapists and so have not demonstrated its durability beyond delivery by this staff group. Narrative reviews develop strong arguments supporting increased use of BA (Jacobson and Gortner 2000, Jacobson et al. 2001, Martell et al. 2001, Hopko et al. 2003) but are written by key protagonists of the approach, and are hence subject to bias.

In subsequent chapters this thesis will therefore seek to develop knowledge by exploring the evidence for BA, and in particular whether its parsimony makes it suitable for dissemination, in a structured and scientific manner. It is of note that the terms 'behavioural therapy' (BT) and 'behavioural activation' (BA) are often used interchangeably. As highlighted in Chapter Two, the term BA was first used in the description of psychological treatments of depression in 1990 (Hollon and Garber 1990); prior to this date the term BT was normally used. BA offers a more specific description of a particular intervention developed since that time. Taking into account the gradual shift over the last two decades to the use of BA as the common term, we will conduct a systematic review and meta-analysis of available evidence. This will identify gaps in knowledge and inform the design and delivery of appropriate primary research to explore such gaps. This will develop new knowledge regarding BA and its evidence base, with a particular focus on the challenges of providing effective treatment for depression as highlighted in Chapter One.

Chapter Three: Evidence base for Behavioural Activation for depression; systematic review and meta-analysis

3.1 Introduction

As we have seen in Chapter Two behavioural treatments, such as behavioural activation (BA) and behavioural activation for depression (BATD) are theory-driven collaborative treatments with randomised controlled trial evidence to support their implementation. The considerable mismatch between the prevalence of depression and the availability of evidence-based psychological treatments outlined in Chapter One creates significant problems for society (Centre for Economic Performance's Mental Health Policy Group 2006). In health care systems such as the National Health Service (NHS), behavioural therapy may offer the possibility of meeting some of this demand due to its single-strand nature (Lovell and Richards 2000) and simplicity. While the evidence cited in support of such propositions is attractive, it is far from definitive as it is in narrative reviews (Jacobson and Gortner 2000) that may only provide a selected section of the overall picture and can be misleading to the reader (Glanville and Sowden 2002).

In order to counter these problems and develop a clear picture regarding the effectiveness of an approach all available evidence should be reviewed. This process, termed systematic review, uses clearly predefined protocols to generate balanced inferences using all available evidence (Glanville and Sowden 2002). Systematic reviews, in contrast to narrative reviews, have the following characteristics (Higgins and Deeks 2008): they

- Clearly state a set of objectives with pre-defined eligibility criteria for studies;
- Use an explicit, reproducible methodology;
- Deliver a systematic search that attempts to identify all studies that would meet pre-defined eligibility criteria;
- Offer an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias;
- Produce a systematic presentation, and synthesis, of the characteristics and findings of the included studies.

Behavioural therapy has been included as a comparator in a number of systematic reviews that evaluated the effectiveness of CBT compared to other therapies (Churchill et al. 2001, Dobson 1989, Gloaguen et al. 1998). Initial searches were conducted in the Database of Reviews of Effectiveness (DARE) and the Cochrane Database of systematic Reviews. No systematic reviews were found that had examined BT/BA as the primary intervention against which other therapies were compared. The first step to conducting a comprehensive review of a clinical question is to ensure that such a body of work does not already exist that is suitable for updating (Glanville and Sowden 2002). The reviews cited above were not capable of providing a comprehensive overview of all trials using behavioural therapy, as the search terms used would have been designed to examine the subject areas addressed specifically in those reviews, such as cognitive therapy (Dobson 1989, Gloaguen et al. 1998) and brief psychological therapy (Churchill et al. 2001). This would therefore be problematic if we were expecting those reviews to provide a comprehensive picture of behavioural therapy (and/or BA). Searches would not identify studies specifically examining behavioural therapy where it was not indexed using 'brief therapy' or 'CBT' or associated terminology. Another potential problem of using those reviews looking at CBT for the purposes of this research was that they were relatively old and not conducted using the most up-to-date guidance on review design (Moher et al. 1999, Moher et al. 2009). Study selection would not have been related to the specific objectives of examination of the effectiveness of behavioural therapies of depression, so that even if a highly sensitive search strategy identified relevant studies, the review methodologies may not have included them. If such reviews were to be used as the basis of this research they would therefore potentially introduce bias at the first stage. In summary, while reviews up to the date of our work had used behavioural therapy as a comparator, relying on them would be problematic. They would potentially introduce bias and methodological limitations. Considering these limitations it was decided that a sufficiently rigorous examination of the effectiveness of behavioural therapy had not been conducted upon which appropriate primary studies could be designed. It

was therefore decided that a new systematic review of the evidence, focussing on the effectiveness of behavioural therapies in the treatment of depression, was required. The review would collate all empirical information linked to a research question via a clear pre-determined systematic approach and provide an evidence synthesis. This would minimise bias and produce a reliable estimate of the effectiveness of behavioural therapy in its own right, which could be used to inform future research based upon gaps in knowledge (Higgins and Deeks 2008).

This chapter outlines the methods used to conduct a systematic review of evidence regarding the effectiveness of behavioural therapy for depression and discusses its findings. This approach was used to contribute new knowledge regarding behavioural therapy and highlight any gaps in the evidence base. This was related to the issues summarised in Chapters One and Two of the prevalence of depression, limited access to effective psychological interventions and the potential of BA as a single-strand intervention to fill some of that gap. The review follows a transparent process to appraise and summarise all randomised controlled trials, with effectiveness explored where possible via meta-analysis. Meta-analysis provides a statistical approach to combining the results from a number of studies exploring the same construct, in this case change in level of depression symptoms. It provides a natural extension of a systematic review following the logical process of gathering all the evidence in a specific area, reviewing and appraising the quality of studies and combining overall findings. It thus provides a comprehensive picture of the true effect of an intervention compared to a particular comparator. Through its combining of study results, meta-analysis increases the precision of estimates of effectiveness of an intervention, since often individual studies may be too small and underpowered to differentiate true difference from chance findings (Moher et al. 1999). Meta-analysis increases such power, and hence the chance that observed effects of an intervention are correctly assessed. A consistency of effects from studies in differing populations or environments can be observed which is not possible through the use of an individual study or narrative reviews. This provides a more scientific

approach to exploring possible reasons for inconsistency across primary study findings, if present, as potential moderators of effect can be explored within meta-analysis. Therefore in this research a well-conducted and transparent systematic review combined with a detailed meta-analysis was used to improve statistical power, often a problem in psychotherapy trials, and produce a broader understanding of the potential benefits which behavioural therapy for depression might offer.

3.2 Development of the review question

Question development is a key first stage of the review process. It determines the review's focus, and as such is potentially most important decision which those conducting a review will take (Light and Pillemer 1984). The review protocol will be derived directly from the question asked (Counsell 1998), which guides the eligibility and search criteria, approach to data extraction and statistical methodology (Higgins and Deeks 2008). Initial literature scoping had been conducted, as outlined in Chapters One and Two, to highlight areas of uncertainty required to focus the review and identify objectives and appropriate methodology (Khan et al. 2002). Questions require clarification of several key components to provide clarity (Counsell 1997, Richardson et al. 1995): the type of population, the intervention, the comparison and the outcome. This set of markers, often referred to as PICO (Higgins and Deeks 2008) allows the development of a well formed clinical questions and subsequent objectives.

- **P**articipants. Due consideration of the participant group is needed in the definition of a clear question. Criteria set must be broad enough to allow the inclusion of sufficient studies, but specific enough to allow for meaningful collation. If such clarity is not developed *a priori*, reviewers may be left with decisions regarding the inclusion of a study that may introduce selection bias into the review. Pre-specifying criteria will limit such risks. It is important to include both the group of interest and also the method of dealing with any unforeseen issues that emerge during the review regarding inclusion

of studies related to participants. An example may be the age of participants: for example, in this review the population of interest was adults. However further clarification is required within the methods section as to what age this would mean and how to deal with studies with populations spanning such criteria (i.e. age set at 18, what approach would be used to deal with a study including participants from 16 upwards). In broad terms, when considering the population, or participants to be included in the review it is useful to clarify the condition under investigation (in this case depression) and the broad population (in this case adults in community or inpatient settings). Table 3 outlines the questions considered in defining the population grouping in this review (for detailed information regarding the handling of issues, refer to methodology section).

Table 3: Factors to considered in the development of 'types of participants' (Higgins and Deeks 2008)

- How is the disease/condition defined?
- What are the most important characteristics that describe these people (participants)?
- Are there any relevant demographic factors (e.g. age, sex, ethnicity)?
- What is the setting (e.g. hospital, community etc.)?
- Who should make the diagnosis?
- Are there other types of people who should be excluded from the review (because they are likely to react to the intervention in a different way)?
- How will studies involving only a subset of relevant participants be handled?

- **I**ntervention: In consideration of the intervention type, both the target intervention of the review and any potential comparators are clarified. It is important to consider the key intervention and how any variance will be handled. For example, through scoping we are aware that techniques central to BT, such as scheduling, are commonly present in other approaches, most notably CBT. Consideration of this aspect of the question shows us the importance of outlining how such issues will be addressed *a priori*

rather than making decisions during the review which could lead to bias.

- **C**omparators: From this position also it is clear that the definition of potential comparator groups are also key to the question development, as methodology follow directly from a well-structured question. Comparison against an inert, or control condition will allow estimation as to the effect of the intervention of interest. Comparison against other forms of active intervention will allow estimation of the performance of the intervention of interest benchmarked against other approaches. Table 4 outlines issues for consideration in development of intervention and comparison.

Table 4: Factors considered in the development 'types of intervention and comparison' (Higgins and Deeks 2008)

• What are the experimental and control (comparator) interventions of interest • Does the intervention have variations (e.g. dosage/intensity, mode of delivery, personnel who deliver it, frequency of delivery, duration of delivery, timing of delivery)? • Are all variations to be included (for example is there a critical dose below which the intervention may not be clinically appropriate)? • How will trials including only part of the intervention be handled? • How will trials including the intervention of interest combined with another intervention (co-intervention) be handled?

- **O**utcomes: Outcomes should be identified in a well-formed question in a way that guides the objectives of the review and the methodology to findings of importance to clinicians, patients and policy makers (Higgins and Deeks 2008, Sutton et al. 1998). While they will not be able to be defined within the question beyond a broad outline, such as depression symptom level, the *consideration a priori* is important to the subsequent delivery of the review. In addition to the event that will be observed (i.e. depression symptom level), the timing of observations such as post-treatment and follow-up will also have a major impact of review findings, and hence must also be considered (Gøtzsche et al. 2007). Table 5 outlines issues for consideration in deciding outcomes for a review.

Table 5: Factors considered when developing criteria for 'types of outcomes'

- Main outcomes, for inclusion in the 'Summary of findings' table, are those that are essential for decision-making, and should usually have an emphasis on patient-important outcomes.
- Primary outcomes are the two or three outcomes from among the main outcomes that the review would be likely to be able to address if sufficient studies were identified, in order to reach a conclusion about the effects (beneficial and adverse) of the intervention(s).
- Secondary outcomes include the remaining main outcomes (other than primary outcomes) plus additional outcomes useful for explaining effects.
- Ensure that outcomes cover potential as well as actual adverse effects.
- Consider outcomes relevant to all potential decision makers, including economic data.
- Consider the type and timing of outcome measurements.

In summary, the preparation of a review question requires detailed preparation, as the question is a fulcrum for the subsequent delivery of the review. Whist the above considerations may not be explicitly mentioned in a brief question, reflection on them facilitates the development of a question that is 'fit for purpose' (Counsell 1998).

3.2.1 Review question

The review question in this thesis was developed by D Ekers based upon the factors outlined above. This was then circulated to the research advisory team of Professor David Richards, Professor Simon Gilbody and Professor Christine Godfrey. Comments were received and incorporated in the design of the review question, which was subsequently agreed by email communication and face-to-face discussion between D Ekers and D Richards. The final agreed question for the review was:

> What is the relative effectiveness of individual behavioural psychotherapy on depression symptom level, recovery and dropout when compared with usual care,

inert controls and other psychosocial treatments in depressed adults?

3.2.2 Objectives

To conduct a systematic review to identify all randomised controlled trials comparing behavioural therapies with treatment as usual, control conditions and/or other brief psychotherapy (CBT, IPT, brief psychodynamic therapy and supportive counselling).

To assess the overall effectiveness, using meta-analysis where possible, of behavioural therapies for depression compared to treatment as usual, control conditions and/or other brief psychotherapy (CBT, IPT, brief psychodynamic therapy and supportive counselling).

To assess the quality of trials using a transparent process that is replicable and reliable.

To assess the internal validity, external validity and statistical power of identified trials.

To consider the effect of study level moderators on primary outcomes.

3.3 Methodology of review

3.3.1 Review team

The systematic review was led at all stages by David Ekers (DE) who produced the protocol, developed the review tools and analysed the main findings. Good reviews require a team with a range of experience to maintain quality (Green and Higgins 2008). The review team for this research consisted of DE, who had knowledge and experience of behavioural activation and the challenges of providing psychological therapies for depression; David Richards (DR), who brought extensive health services research experience alongside being a national advisor to the Improving Access to Psychological Therapies programme; and Simon

Gilbody (SG), who brought extensive health service research skills and expert knowledge of review methodology, in particular the use of meta regression as seen in a previous meta-analysis of collaborative care (Gilbody et al. 2006).

It is also recommended that an advisory board is established; however due to the financial limitations of this research this was not possible. It was agreed that the review team would act together as a steering group for the project, and that if specific advice was required this would be sought.

Additional expertise was brought into the review team for particular purposes. At the search design phase Janet Menton (JM), a librarian and information skills trainer working within the local NHS trust, was consulted to provide expert advice in the design of searches. She also conducted the searches used in the review and provided results to DE for analysis. SG brought experience of meta-regression and analysis of publication bias, and conducted these analyses included in this thesis.

3.3.2 Inclusion criteria study design.

The type of study included in a review should be guided by the question rather than the perceived values of the design itself (Sackett and Wennberg 1997). As this review focussed on the effectiveness of behavioural psychotherapy in comparison to other interventions, the appropriate study design for inclusion was randomised controlled trials (RCT). Scoping suggested that a number of such trials existed but had not been formally appraised or combined in meta-analysis. Randomised trials allow for the balancing of groups following inclusion in a study in a way that reduces the impact of confounders and bias and reduces overestimation of treatment effects (Higgins and Deeks 2008). Inclusion of non-randomised trials may be an acceptable in reviews of subject areas where there is a lack of RCTs and the endeavour is to examine the case for and design of such a trial, or where the subject under review does not lend itself to that study design (such as qualitative evaluation of a treatment approach). Also a review

including non-randomised studies may be useful to study the effects of interventions that are unlikely to be randomised (such as certain surgical procedures) or where the condition under examination is particularly rare (Higgins and Deeks 2008). In this review it was considered that BT did lend itself to randomised controlled methods, hence it was decided to include all randomised controlled trials published and unpublished in any language. The importance of inclusion of unpublished trials and all languages was emphasised as a priority due to the potential for a positive result bias if they were excluded (Khan and Kleijnen 2002). The decision to include all randomised trials identified in searches was taken after consultation with the review steering group. Study quality would be assessed (control for selection, recruitment and measurement bias) and the impact explored through sensitivity analysis. This approach would allow for the gathering of all suitable evidence which could then be assessed and graded, with findings managed accordingly. It was anticipated through scoping that several older trials would be found which were not subject to the standards that have been applied to RCTs in recent years. If, at identification, those studies were to be rejected based upon such methodological problems it was felt this review would be weakened. The objective was to identify and analyse all studies to give the most comprehensive review possible of behavioural therapies' effectiveness. One approach used to limit the potential biasing of results from the inclusion of low quality studies is to exclude such studies from the meta-analysis in a review (for an example see NICE depression update (National Institute of Clinical Excellence 2009). In this review, it was decided an adequate compromise between controlling for confounding of results by low study quality and weakening of findings due to study exclusion and subsequent low numbers would be most appropriately managed via the use of sensitivity analysis. Through this, the impact of lower quality studies on overall results could be ascertained and considered alongside overall findings.

3.3.3 Exclusion criteria: study design

Based upon the above decisions exclusion was broadly defined as non-randomised controlled trials.

3.3.4 Inclusion criteria: type of participants

Age

It was decided to focus this review upon the treatment of adults of either sex. The lower cut off age for inclusion was set at 16 years old. Studies that included participants in age categories below this would be excluded. It was considered that while BT may be a suitable intervention for this population, the purpose of the review, as defined in the question, was to consider its effectiveness in adults. Inclusion of studies with participants below the age of 16 would introduce heterogeneity which would be unhelpful in addressing the main question; while behavioural interventions may be useful for those under 16 years of age this question would require a review in its own right. Such trade-offs between reductions in wider generalisability and facilitating comparison and synthesis that deliver meaningful results are common in review design, ultimately requiring informed decision making by the researcher (Horwitz 1995). No upper limit was placed upon age. Depression and anxiety is a common condition in older adults affecting between 10-17% of this population (Chew-Graham et al. 2004). Scoping suggested trials existed in this age group, and such information would be of relevance to service providers.

Treatment settings

Studies including participants treated in community or inpatient settings were accepted into the review. The aim was to explore the efficacy of behavioural interventions rather than the settings in which they are applied, hence all treatment settings were considered relevant.

Diagnosis

Studies with participants with a primary diagnosis of depression were the main target of the review. It was anticipated, through previous narrative

reviews in this area (Martell et al. 2001), that we would identify a number of trials in this review which were published in the 1970s and 1980s. During this time the use of validated instruments to formally diagnose clinical conditions was less frequent. Based upon this expectation, studies were included if participants were identified as depressed when assessed according to standardised criteria at the time of the trial (ICD, DSM) with the use of structured diagnostic interview/validated diagnostic instruments or valid and reliable clinician/self-rated measures assessing level of depressive symptoms.

Studies including participants with a diagnosis of mixed depression and anxiety were also included. Based upon findings outlined in Chapter One, the most prevalent diagnosis seen in clinical settings is a mixed presentation of anxiety and depression. It was therefore decided to include studies that included such mixed diagnosis if the primary problem treated was depression. It was considered that, while inclusion of anxiety may introduce confounders that impact on the overall outcome within the selected study, this would be preferable to potentially losing a significant number of relevant trials. If such studies were identified their impact on overall outcome would be explored through sensitivity analysis.

3.3.5 Exclusion criteria: types of participants

Diagnosis

Related to the above inclusion criteria it was anticipated that trials would be identified that specified a mixed diagnosis of anxiety and depression when investigating a behavioural therapy approach to the treatment of anxiety. Such trials may make reference to depression symptom level as part of their overall battery of measurement. If an exposure-based rationale for the treatment of anxiety was the main intervention the study would be excluded.

In common with many other reviews, we excluded studies with participants experiencing psychosis, bipolar disorder, substance misuse problems or cognitive impairment.

3.3.6 Types of intervention

Main intervention – behaviourally-based interventions for depression

The review included studies using behavioural interventions if the underlying theoretical framework guiding the protocol related to operant conditioning approaches (Skinner 1974) as outlined in Chapter Two. Interventions would typically be based upon a formulation exploring and readjusting the participants' engagement with contextual positive reinforcement (Jacobson et al. 2001) with subsequent related behaviour change. It was anticipated that a number of trials would include a behavioural component as part of an overall intervention based upon cognitive theories (Beck 1976). Such an approach is common, and therefore required the review protocol to guide the researcher. The deciding factor in this case is the theoretical underpinning of the intervention. Behavioural theory would suggest that the focus of the intervention is the reconnection with positive reinforcement and the reduction of negatively reinforced avoidant behaviour. The use of behaviour scheduling and self-monitoring commonly found in cognitive interventions has a different goal. It is aimed to allow the patient to explore their ability to control their mood, thus is used as a behavioural experiment to modify thinking. As such, where behavioural techniques were incorporated in cognitive therapy interventions to modify beliefs, these would be classed as CBT rather than BT.

Comparators

Comparators were classed into two broad categories, those testing BT against an inactive or usual care control condition and those testing against an alternative active psychological treatment. From this approach, the review aimed to examine if behavioural therapy is effective compared to no intervention, an inert intervention or the care usually received. Also it was

intended to consider the relative effectiveness of behaviour therapy compared to other psychotherapies. Using direct comparisons in studies at post-treatment in meta-analysis is more likely to control for the influence of individual study variations (such as length of treatment, severity etc.) and hence be likely to provide a clearer picture of the superiority of one psychotherapy treatment over another (Spielmans et al. 1980).

Control conditions

Control conditions in this review included:

- Those with no contact: controls such as waiting list will include no contact with the participant for the duration of the control period.
- Placebo interventions: such interventions are designed to have no 'active component'; however they deliver an 'inert' treatment to provide control for non-specific influences such as therapist time and personal contact.
- Treatment as usual: this control condition generally offers those not randomised to the intervention arm the usual care they would receive from standard care (i.e. from their primary care practitioner). As it is assumed that those in the 'active' arm would be having such usual care with the addition of the active intervention under study, the differences post-treatment are assumed to be attributable to the active ingredient.

It was decided in incorporate these three 'control' groupings under one heading to provide a realistic estimate of the likely effect of BT in the treatment of depression.

Active comparator interventions

- Cognitive behavioural therapy (CBT): an intervention generally based upon the work of Aaron T Beck (Beck 1976) that identifies cognitive responses to situations and the emotional consequences of these as a central component of treatment. Treatment is time limited and is likely to include both behavioural and cognitive components. The treatment is psycho-educational, with therapists developing a collaborative formulation, shared with the participant, which

identifies new ways of coping. The essential difference to the behavioural model that is the main intervention is the focus towards the challenging of such thoughts and underlying belief structures. Any intervention that includes thought catching and challenging, either through thought diaries of behavioural experiments, was considered to be a CBT intervention.

- Brief Psychotherapy (BP): other brief psychotherapy approaches that utilise the interpersonal relationship with the therapist to explore the problem areas of the patient such as Brief Interpersonal Therapy (IPT) (Klerman et al. 1984) or brief psychodynamic therapy (Luborsky et al. 1995). These approaches focus on developing insight and subsequent character development. Any structured psychotherapy not included in the CBT or behavioural categories will be included in this comparison grouping.
- Supportive Counselling (SC): Generally based upon the work of Carl Rogers (Rogers 1961), supportive counselling is focused upon the therapist's use of core relationship conditions (genuineness, unconditional positive regard and empathy) to develop self-awareness by the participant leading to symptom improvement. Treatments in this category will not follow a 'structured' or protocol-led format as they are classed as non-directive however they remain brief and time limited.

Dose/mode of intervention

In the review we included studies examining 'brief therapy'; the definition of 'brief' used was up to twenty-four sessions. No clear definition has been set regarding what constitutes a brief therapy; previous reviewers have used twenty sessions (Churchill et al. 2001) however no clear rationale was provided for the choice of this number. Scoping for this review, however, had highlighted key studies described as 'brief psychotherapy', delivering a twenty-four session protocol, which were of significant relevance to this review. The researchers in this study provided a rationale regarding the number of sessions, which was aimed at giving an adequate 'dose'; this was based upon other researchers' observations regarding possible failings

of previous RCTs of behavioural therapy compared to other psychotherapies (Dimidjian et al. 2006). It was clear that such a decision could impact on review findings, therefore meta-regression was used to explore the relationship between the number of sessions and overall effectiveness of the intervention.

Scoping also indicated that studies would be likely to be identified delivering therapy to participants through both individual and group approaches. In this review the focus was specifically on individual treatments. It was considered that the inclusion of groups could bring in additional confounders (such as the role of the group support etc.) that might impact on interpretation of findings. While there is a view that inclusion of such modes of therapy would assist in the generalisation of the review findings to treatment settings, this study was specifically interested in the efficacy of individual psychotherapy. In discussions of the steering group it was felt the delivery of a separate review on the role of group delivery of psychotherapy would be the most appropriate approach for exploring this area further, rather than combining both delivery models in this review. It was also anticipated that the trials identified could include behavioural marital/couple therapy for the treatment of depression. Again it was considered that, while such approaches may be based upon a behavioural framework, the interventions which focussed upon the depressed individual, their partner and centrally their relationship would be best dealt with within a separate review.

3.3.7 Outcome measures

Psychotherapy studies often use multiple outcome measures relating to many factors such as symptoms, functioning and cognitive chance. These can be categorised into continuous or dichotomous data. As there is a paucity of economic analysis in this area (Churchill et al. 2001), trials were not required to include cost data for inclusion.

Continuous data

Studies that measured depressive episode by validated depression symptom level assessment comparing between group outcomes post-treatment and or at follow-up were included. Usual assessment consists of either self-rated measurement using scales such as the Beck Depression inventory (Beck et al. 1961) or clinician rating using a structured assessment tool such as the Hamilton Rating Scale (Hamilton 1960). Such scales commonly translate the frequency of a range of depression symptoms into a single score, with increasing scores reflecting more severe symptoms. For inclusion in the review the scales used the within studies must have been subject to validity and reliability testing. Outcomes are usually presented by means and standard deviations. As psychotherapy trials often present multiple measures, an algorithm was developed with self-report measures taking precedence over clinician-rated measures. As there is no clear standard approach to guide such a hierarchy, the review steering group had to consider available options and make a decision in consultation (Khan and Kleijnen 2002). Based upon the reviewers' experience in the field, the initial scoping of the frequency of each approach used and discussion, agreement was reached to give validated self-report measures precedence. The impact of this decision on overall results would be explored via sensitivity analysis.

Dichotomous data

Recovery is often reported in psychotherapy trials and was included as a dichotomous measure. Recovery reflects to what degree the participants in the study move from a clinical sample (based upon either symptom level measures supported by validated cut of points, or structured clinical interviews) to a non-clinical sample.

Dropout was used as a proxy for treatment acceptability. Dropout was considered an important comparison to assess the effectiveness of an intervention outside the research setting. Acceptability of treatments is an important consideration, as if there is increased attrition for a particular

intervention compared to other standard treatments this would lead to overall questions relating to overall effectiveness. Such analysis was included as it provides important information for clinicians.

3.3.8 Search strategy

The delivery of an unbiased comprehensive literature search is central to a well-designed systematic review. Results will be compromised if the search misses relevant primary studies, as it will not provide a comprehensive overview and the level of precision of the effect estimates will reduce with insufficient statistical information (Higgins and Deeks 2008). The difference between a narrative review and a systematic review is the search process involved. A systematic review delivers a comprehensive search of the literature based upon an established and piloted search strategy. The exclusion of bias through such comprehensive strategies sets the platform for the review and meta-analysis by identifying all relevant studies regardless of outcome of those studies (Easterbrook et al. 1991).

In order to deliver as comprehensive and unbiased review as possible, the search strategy used in this review was developed in partnership with an expert in the field (JM, a librarian employed within Tees, Esk and Wear Valleys NHS Foundation Trust with extensive experience in delivery of searches and of training in literature searching at a NHS trust and strategic health authority).

In designing the search strategies used in this review, the potential biases I wished to exclude were:

- Publication bias: it has long been known there is a tendency to publish studies showing a significant difference in results (Rosenthal 1979). In psychology journals it has been estimated that 95.6% of articles reported significant results (Sterling et al. 1995). Such a trend persists today, with estimates that studies with statistically significant results are up to four times as likely to be published (Hopewell et al. 2008) and those with negative results taking up to two or three years longer (Loannidis 1998). This leads

to a potential over-estimation of the effectiveness of the therapies included in a review. Therefore search strategies used in this review were designed to cover a multitude of databases aimed at identifying both published articles and 'grey literature' (i.e. material not formally or commercially published). This approach was aimed at optimising the chances of identifying studies with both positive and negative results.

- Duplicate publication bias: It is relatively common for results from an individual study to be presented in a number of publications (Higgins and Deeks 2008). This problem is particularly associated with trials with positive results, and can result in over-estimation of effectiveness in meta-analysis due to multiple inclusion of the same data set (Tramèr et al. 1997). It creates particular problems in identification, as often publications are written by different authors, or report different measures incorporated in the same study. While it was felt that this could pose a significant risk to the outcome of the meta-analysis in this research, it was decided that this problem would be better managed at the study selection phase rather than though search strategy modification. The risk would be that limiting the scope of the search could have the unintended consequence of eliminating potentially important findings.
- Citation bias: While using reference lists to identify research to be included in meta-analysis is commonly accepted (Glanville 2002), this can lead to bias in itself. Authors of book chapters and journal articles will commonly select evidence to support their view. This is particularly the case if the findings are positive (Carter et al. 2006) and hence a problem is posed for the review if such reference lists are to be used. In order to balance these findings we searched reference lists in those studies identified from our electronic searches of all databases to ensure adequate sensitivity for studies with negative findings.
- Language bias: There remains considerable debate as to the impact of language bias on review findings (Moher et al. 2003, Jüni et al.

2002). More recently there is some evidence that randomised trials are more likely to be published in English-language journals (Galandi et al. 2006). In this review however it was anticipated, based upon literature scoping, that a number of trials from the 1970s and 1980s would be identified. Therefore to exclude non-English-language articles would potentially introduce unacceptable levels of bias. Often trials are indexed and abstracted in English despite being in a non-English journal. We therefore took the approach of including non-English language articles and adopting a case by case approach to balancing the resource requirements of translation vs. the relevance and impact any non-English language paper was likely to make to the review findings.

Sensitivity vs. specificity

Sensitivity refers to the proportion of articles on a subject the search is likely to deliver. It is often expressed or measured as a percentage and measures the comprehensiveness of a search strategy. Specificity (or precision) refers to the ability of a search to exclude irrelevant articles (Glanville 2002). The balance between each of these approaches is important, especially taking into account the diminishing returns in relation to additional searching beyond a certain stage. In this strategy it was opted to employ a highly sensitive approach in designing the search strategy to ensure that a high proportion of relevant reports were identified in relation to the number in existence. It was decided that the relatively swift process of scanning results in relation to the total time employed upon the review and meta-analysis as a whole made this approach advisable. As is recommended, the inclusion of an experienced librarian in the planning and execution of the search strategy assisted in these decisions, with the search being developed over a number of meetings where pilot searches were constructed and conducted and reviewed as recommended (Glanville 2002).

Sources

It is recommended that a minimum for a comprehensive search would include the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and EMBASE (Glanville 2002). While the CENTRAL database searches MEDLINE and EMBASE for controlled trial data, there may be a lag time in that process and also some of the free text terms, particularly in EMBASE, may have been missed with older studies. To ensure that a comprehensive picture was established it was decided to search both those databases in addition to CENTRAL for this review. This approach is not necessarily recommended (Higgins and Deeks 2008) however within the discussions of the research group it was felt to be appropriate based upon the scoping of literature for this review. The search was specifically widened to include databases that include dissertations (CINAHL), specific databases of mental health research (PsycINFO), those of allied health professionals (AMED) and nursing (British Nursing Index). In addition, searches were conducted within the Database of Abstracts of Reviews of Effects (DARE) and the Cochrane Database of Systematic Reviews to identify any reviews of evidence not previously identified. Searches were conducted from inception of each database to ensure a comprehensive record was gathered. It is noted that this approach may have resulted in increased work, as MEDLINE and EMBASE results would have been already included in CENTRAL prior to 2001; however, as outlined above, it was felt that the exclusion of duplicated work would be achieved at study selection. The benefit of the delivery of a thorough search outweighed the risk of any potential unrewarded effort of duplicate searching. Appropriate filters to capture randomised controlled trials were be included in all searches other than in those in the Cochrane Library, which only includes studies using this methodology (see http://www.york.ac.uk/inst/crd/search.htm).

Hand searching

Key journals in the field of psychological interventions for depression were hand-searched for the year prior to the review to identify recently published

trials. It can take up to this time for studies to be indexed in electronic databases (Glanville 2002). Reference lists from identified trials were also searched to highlight any trials not identified by the above search strategy. This process was repeated until no new relevant reports were identified.

Grey literature

Grey literature (conference abstracts etc.) were examined via hand and internet searching for the year prior to the review, as it was anticipated those presented prior to this time would be identified through published data searching. Databases containing references to these data sources such as SIGLE (system for information on grey literature) were searched alongside the Dissertation Abstracts International database, the NHS National Research Register (NRR) (to identify current and recent relevant research), and were cross-referenced with other findings.

Key authors in the field of behavioural therapies of depression were contacted to identify any other relevant studies, which were then checked against results.

Controlled vocabulary and free text

Controlled terms such as MeSH in MEDLINE and EMTREE in EMBASE allow indexed terms to be collated in those source databases. However as these change across databases it is necessary to ensure the appropriate strategy is adopted for each database. In addition, controlled terms are indexed in subject trees and hence can be 'exploded' to widen the search scope. One difficulty of controlled vocabulary searches, however, is the variability of indexing, particularly in relation to older studies (Higgins and Deeks 2008), some of which we were expecting to identify in this review. This emphasised the requirement of the free text section of searches to be robust and that all searches were adapted to a relevant format for the search platforms used. Where relevant, search filters were employed to capture randomised trials. See Table 6 for an example search strategy for the Cochrane Database.

Table 6: Example search strategy for Cochrane database.

#1	MeSH descriptor **Depression** explode all trees in **MeSH products**
#2	MeSH descriptor **Depressive Disorder** explode all trees in **MeSH products**
#3	MeSH descriptor **Dysthymic Disorder** explode all trees in **MeSH products**
#4	(**#1** OR **#2** OR **#3**)
#5	**depress*** in **All Fields** in **all products**
#6	**dysthym*** in **All Fields** in **all products**
#7	(**#4** OR **#5** OR **#6**)
#8	MeSH descriptor **Behavior Therapy** explode all trees in **MeSH products**
#9	**behav* next therapy** in **All Fields**, in **all products**
#10	(**#8** OR **#9**)
#11	**behav* near/3 activation** in **All Fields** in **all products**
#12	**(event* or activit*) near/3 schedul*** in **All Fields** in **all products**
#13	**positive next reinforc*** in **All Fields** in **all products**
#14	**coping near/2 depression** in **All Fields** in **all products**
#15	(#10 OR **#11** OR **#12** OR **#13** OR **#14**)
#16	(#15 AND #7)

MeSH terms for depression and dysthymia have been combined with free text searches for depression and depressive disorders (combines both in free text via the use of truncation) and dysthymia to provide the diagnostic field. The MeSH term for behaviour therapy is exploded and combined with free text search of behaviour therapy (behav is again truncated to obtain all spellings, is combined using 'next' as the search is interested in the particular term). Duplication is removed to provide the intervention results. The search strategies from #11 to #14 are specific terms relevant to this intervention (combined using relevant search operators to optimise specificity); these have been combined with wider behavioural therapy terms to increase sensitivity of the search. Interventions are then combined with the diagnostic field to provide the final set of results.

The above outlines an extensive search strategy that was developed in partnership with JM. This information specialist was then contracted, based upon negotiation with the host employer, to conduct the search in

partnership with myself to deliver a mix of expertise in search delivery and behavioural theories.

Details of the main search strategies used, dates conducted and results are included in the search report in Appendix I for transparency as per recommended procedures (Moher et al. 2009).

3.3.9 Selection of Studies

Studies in this review were selected following the stages outlined below:

1. Merging the results of the numerous searches within a reference manager software programme (Endnote, (Thompson 2005) removing duplicate reports.
2. Examine the identified titles and abstracts to remove obviously irrelevant articles.
3. Obtain full text of potentially relevant reports; examine to exclude further duplication and assess relevance to inclusion criteria.
4. Review findings of searches with research team and clarify approach to studies where uncertainty existed.
5. Decide on inclusion and move studies to data extraction phase.

Merging results

Data from searches was downloaded to EndNote where possible and duplicate results searched for using the 'find duplicates' command. Each example of identified duplication was reviewed to ensure a correct decision, then one citation was deleted. This process was repeated until the master data set no longer included any duplications.

Selecting studies

Study selection is one of the most important stages in the review process, therefore it is beneficial to have a transparent process involving more than one reviewer (Glanville 2002). This helps to exclude bias as those conducting reviews may have preformed opinions that influence their decision making (Higgins and Deeks 2008). In this review, obviously

irrelevant research was excluded by title examination by one reviewer (DE). This was a pragmatic decision based upon the expectation that, due to the highly sensitive search strategies used, a high number of such reports would be included (i.e. studies identified looking at depressed thyroid gland may well be identified if RCT and the term 'depressed' were captured). Following this the reduced list was examined at abstract level by two reviewers independently (myself and DR), following which a comparison of results was conducted and a list of articles for further full text review generated.

Studies were then selected based upon their relevance to the question and inclusion criteria. Blinding of author and source details of studies may reduce bias in the selection process. Due to the resource limitations alongside the uncertain level of benefit this produces (Berlin 1997), in this study author and source details were included. Full text of selected studies were obtained and screened for relevance independently by two reviewers (myself and DR) as a measure against bias. It was outlined *a priori* that any disagreements regarding suitability were to be resolved by discussion in the first instance. If agreement was not reached, a third reviewer (SG) would be asked to independently screen the article for acceptability and relevance. A flow chart would be used to outline the numbers of studies included/excluded at any given stage with summary rationales for decisions as per recommendations (Moher et al. 1999).

Quality assessment

Two reviewers (DE & DR) rated study quality using criteria to explore bias (Khan et al 2002). Other than concealment of allocation, evidence demonstrating aspects of study quality that directly influence outcomes is unclear (Jadad et al 1996, Schultz & Grimes 2002). Many quality assessment scales have been developed however all would seem to have individual weaknesses (Moher et al. 1995). An example of problems with quality assessment tools is highlighted in a study that used 25 scales, previously adopted in published reviews, to measure the quality of 17 studies included in a meta-analysis of heparin use in the prevention post-

operative thrombosis. The association of quality and effect estimates were so varied that the authors recommended assessing individual methodological aspects of trial quality in relation to effect estimate (Jüni et al. 1999). Based upon discussions between the review steering group, it was decided to use a simple measure of quality devised for this review that rated studies against two standards relating to selection, measurement, performance and attrition bias. This would result in a possible quality score of between 0-8. Study quality assessment forms were developed and piloted by DE and finalised in collaboration with SG and DR (see Appendix 1 for quality assessment tool). A protocol regarding decision making on study quality was agreed: that disagreements were dealt with through discussion initially between myself and DR, and if required, passed to SG with issues for final scoring. In order to explore the relationship between methodological aspects of trials (study quality) and effect size, regression analysis was performed; for details see below.

3.3.10 Data Extraction and synthesis

Data were extracted from each trial at post-treatment and follow-up (six months or nearest available data set). Data extraction forms used in a previous meta-analysis of collaborative care and guided self-help for depression (Gilbody et al. 2006, Gellatly et al. 2007) were reviewed and adapted for this research. These were then tested for ease of use (see Appendix I for blank data extraction form). Forms were designed to allow identification of the unique study and the person extracting the data. Forms allowed for collection of data in relation to study populations, interventions, quality and statistical data (Higgins and Deeks 2008). Final forms were completed for each study included in the review. Ideally data should be extracted by independent reviewers, as this reduces errors (Buscemi et al. 2006). In this study, however, a balance was considered in relation to capacity and an ideal approach. There was no funding to support the systematic review, hence double data extraction, while preferred, was not possible. We approached this by initial data being extracted by DE and subsequently checked in meetings with DR. This process was conducted

with an understanding of the increased risk associated with extraction error using this approach, which was rigorously checked for in such meetings. While blinding of some study details has been suggested to remove the risk of bias from review authors, as outlined in the abstract review stage, this is not generally accepted as necessary (Berlin 1997) and, due to resource limitations was not adopted. Data were then entered into the Cochrane Collaboration Revman programme (Cochrance Collaboration 2003) for synthesis. Where missing data were identified, author contact details were sought through published papers or via the World Wide Web. Contact was attempted via email, as often authors were based in the USA. If, after two attempts no less than four weeks apart, no response was received, no further contact was attempted. Statistical data were extracted to facilitate analysis of continuous (depression symptom level) and dichotomous (recovery, dropout) variables. For continuous data the mean score of the behavioural interventions ($M_{ba)}$ and of the comparator ($M_{c),}$ standard deviation ($SD_{ba)}$ and ($SD_{c)}$ and number ($N_{ba)}$ and ($N_{c)}$ were extracted. If at this point missing standard deviation scores were identified, which was anticipated due to the number of older studies identified in scoping for this review, these were imputed from relevant studies as per accepted protocols (Furukawa et al. 2006). Such approaches appear to have minimal impact on the overall findings of meta-analysis and are considered safe, hence were deemed appropriate in this review. In the case of missing standard deviation values it was agreed that SG would identify and supply for imputation these extracted from a comparable primary study reviewed in a large meta-analysis of collaborative care for depression (Gilbody et al. 2006). For dichotomous data the count of events in each arm were extracted alongside the numbers in each arm.

3.3.11 Data pooling

Continuous data-depression symptom level

Two approaches to data pooling of continuous data are commonly used and were considered for this review. Weighted Mean Difference (WMD), more correctly described as ‘difference in means’ and standardised mean

difference (SMD). WMD measures the mean difference between the intervention under investigation compared to the comparator on a continuous scale. It is a measure that can be used when measurements to be analysed in meta-analysis are of the same scale (e.g. kilograms or points on a BDI scale). Weighted mean difference therefore combines data from each study and gives a combined mean difference in across all studies in the specific unit of measurement. This approach is possible only when the continuous data under investigation as stated is on the same scale or can be converted to the same scale (e.g. trials using kilograms and trials using ounces measuring the continuous variable of weight). In this meta-analysis this would prove problematic, as depression symptom level studies use a range of measurement scales to measure this same construct. In contrast with scales measuring weight, there is no consistent relationship between units used in depression symptom level measurement: one unit on the Beck Depression Inventory (BDI) is not constantly related to one unit on the Patient Health Questionnaire (PHQ-9) across all possible ranges on the scales. While they both asses overall depression symptoms, the BDI used a 21 item scale with each item scoring 0-3 rating feelings over the past week thus producing a possible score range of 0-63. The PHQ-9 in contrast measures depression on a nine-item scale, with each item scoring 0-3 in rating feelings over the past two weeks, producing a possible score range of 27. This results in no rational approach to convert all scales to one single 'unit of depression symptom'. While it would be possible to multiply the PHQ-9 score by 2.33 to convert it to a 63-point scale, this produces many problems, as each scale asks different questions with the BDI scoring more depression symptoms (21 vs. 9) over half the time. This would therefore clearly lead to problems in validity. For this reason the use of WMD was not possible in this meta-analysis if we wished to combine as many studies as possible rather than only those using a single scale for measurement. The alternative approach is to assess the difference in relation to the standard deviations of the means of scales used in the included studies. The standardised mean difference (SMD, sometimes described as Cohen's *d* or effect size) is commonly calculated by subtracting the mean of the control

group from the mean of the intervention group and dividing by the pooled standard deviation.

$$\text{SMD} = \frac{\text{Difference in mean outcome between groups}}{\text{Standard deviation of outcome among participants}}$$

What is being produced is a summary statistic of the giving the difference in means between groups relative to the variance (SD) in the study. This produces a measurement of the number of standard deviation units between the two means. Although the number has no quantifiable value, it allows for comparison across a number of scales as long as they are measuring the same construct (depression symptom level). In the case of the BDI and PHQ-9, the variance around the mean reflected in the standard deviation should be the same ratio even though the actual range of possible scores on the BDI scale is 2.33 times more than the PHQ-9. For example, if a moderately depressed group had a mean BDI score of 21 (SD 7) and a mean PHQ 9 score of 15 (SD 5), the ratio of the SD to mean is constant. This then allows comparison across scales using the SMD calculation outlined above. It is important to consider the use of SMD with care, ensuring that units of measurement included are measuring the same construct (such as depression symptom level) and that results viewed with reflection on the degree of heterogeneity of included participant samples. This can impact on SMD, as the pooled standard deviation may be directly related to the inclusion criteria of the study: the tighter the inclusion criteria the smaller the pooled SD. In this case the same treatment effect may result in a different SMD. Therefore a measure of heterogeneity (see later in this section) should be used to estimate its effect on results. The SMD is the only option in meta-analysis where differing scales of continuous variables are used (Deeks et al. 2009) and hence was considered appropriate in this review bearing in mind the considerations outlined above.

We assigned effect sizes values according to the standard convention where the SMD is small (0 – 0.32), medium (0.33 - 0.55) and large (0.56 and above) (Lipsey and Wilson 1993, Cohen 1988).

3.3.12 Statistical considerations

An additional issue addressed in our methodological design was related to studies that included two comparisons under the same category (i.e. CT and CBT vs. BT). Through the initial scoping it was clear this problem would be encountered. Simply entering the BA results in two separate rows would appear to be the simplest solution to this problem. This however would lead to double counting of the numbers in the BA arm of the meta-analysis, thus resulting in an over-estimation relating to the weighting of trials. That is a trial with 50 participants in a BA arm would be doubled up to 100 participants suggesting twice as many people received the approach resulting in a unit of analysis error (Deeks et al. 2009). Several options exist in relation to such problems (Higgins et al. 2008). One option commonly adopted approach is to combine groups to form a single pair-wise comparison. It was felt in this review such combinations would lose the individuality of the separate intervention arms in question. For example, in one study (Jacobson et al. 1996) both CBT and CT were compared to BA. In this case each comparator to BA was different in construct and thus merited comparison in its own right. An alternative (but slightly less favoured) approach is to enter comparisons separately but to halve numbers in the behavioural arm to avoid double counting and inaccurate weighting of trials (Higgins et al. 2008). In this review the latter approach was adopted, as it was felt more suited to the varied nature of the comparisons that may be encountered.

Where studies presented results using sub-categories (e.g. high/low depression severity), data were entered as two separate trials, provided that stratification occurred prior to randomisation. This approach was a pragmatic decision reducing the need to contact authors for combined data sets where these were not reported in publications.

Dichotomous data - recovery rate and dropout

Dichotomous (binary) data are presented when each individual in a study is in one of two states. This review analysed if patients were either

recovered/not recovered at post-treatment and follow-up and if they dropped out/ did not drop out at post-treatment. These data were presented as odds ratios, the chance of an event (improvement or dropout) in the intervention group compared to the comparison group. An odds ratio of 1 indicates no difference between an event occurring in each group, less than one a reduced chance of such an event and greater than one an increased chance. It is important to be mindful as to the preferred directions when interpreting odds ratios (Deeks et al. 2009). In this review for recovery a score of greater than one was a positive finding and for dropout less than one would be positive. Odds ratios are calculated by dividing the number of events in the intervention/control group by the number of participants in each group. This firstly establishes odds of the event happening in each arm, then to establish the odds ratio dividing the odds of the event in the control by the odds of the event in the treatment group.

$$OR = \frac{\text{odds of event in experimental group}}{\text{odds of event in control group}}$$

Effects modelling

The aim in meta-analysis is to present the combined effect of an intervention across a number of varied studies. If all studies carried an equal level of precision this would be a simple calculation: the mean of effect sizes across all included studies. Some studies, however, may be more precise in the effect estimate than others due to a number of factors, such as sampling, size and possible bias. Therefore in meta-analysis it is important to decide how the findings of each study contribute to the overall effect size. There are two approaches based upon differing assumptions that are commonly used: fixed effect and random effect modelling (Deeks et al. 2009, Hedges and Vevea 1998).

Fixed effects

Fixed effects models assume there is one true effect that underlies all studies in the meta-analysis. Therefore, if each study were infinitely large, they would all produce the same effect size. Each study is as a result

allocated a weight dependent on the degree of information that is captured. The only error assumed in this model is random error, which is reduced as sample size increases; hence weight is calculated according to study size. Therefore calculations in fixed effects models assume that all studies have a shared effect size and that all observed effects are normally distributed around this, with observed variance dependent on sample size (see Figure 7).

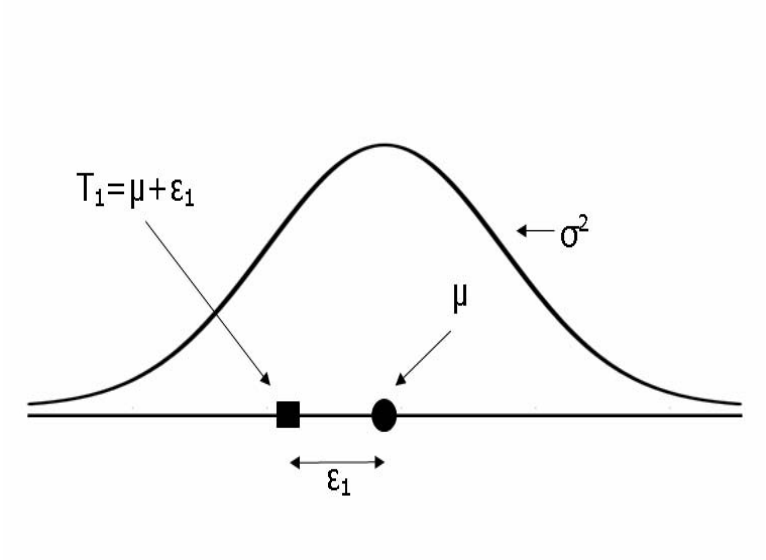


Figure 7: Fixed effect model.
The observed effects are sampled from a distribution with true effect μ, and variance σ^2. The observed effect T_1 is equal to $\mu+\varepsilon_i$ (Borenstein et al. 2007)

Random Effects

Random effects modelling, in contrast, assumes that each included study is estimating slightly different treatment effects. The model estimates the distribution assuming that effects are related and have a central value, but due to the heterogeneity of studies (i.e. age of participants, duration of treatment measurement reliability) the true effect within each study varies. The meta- analysis using this model therefore reports the mean effect estimate in this distribution. As each study is estimating a different effect size from a sample of a population, the weight of each study in a calculation is more balanced, with larger studies less dominant than they would be in a fixed effect model. This results in two levels of variance,

firstly within each study and secondly across studies, to give an estimate of the mean of the effects of studies (see Figure 8).

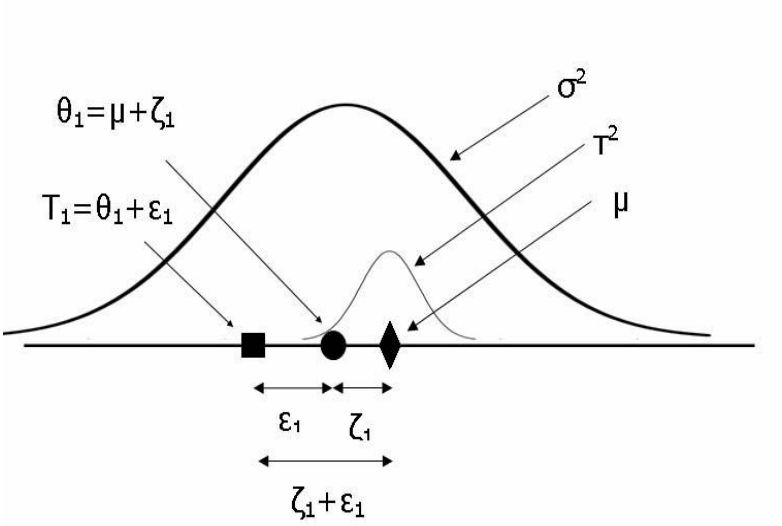


Figure 8: Random effects model.
The observed effect T_1 (box) is sampled from a distribution with true effect θ_1, and variance σ^2. This true effect θ_1, in turn, is sampled from a distribution with mean μ and variance τ^2. In turn, θ_1, is determined by the mean of all true effects, μ and the between-study error ζ_1.(Borenstein et al. 2007)

From scoping prior to this review, it was clear that studies would have heterogeneity in clinical populations (age and level of depression), and sampling methods. In addition, interventions would vary in terms of therapist experience, session number, setting and duration. As it was wished to make inferences about how results may generalise outside of the studies observed in the meta analysis, a random effects model was the most appropriate method to use (Hedges and Vevea 1998). This assumes that results would provide an estimate of the mean of a distribution of varied effects of behavioural therapy vs. comparators.

Summary of statistical issues

Based upon discussion within the review steering group of the above issues, it was decided to present pooled data of continuous outcomes by

standardised mean difference and dichotomous outcomes by odds ratios with 95% confidence intervals using a random effect model (Sutton et al. 1998), taking into account both within and between study variance.

3.3.13 Exploration of heterogeneity

Heterogeneity is a term used to describe the variability between studies in a review. Such variability can be due to clinical differences in populations or interventions (clinical heterogeneity), or the methodological approaches, such as risk of bias and study design (methodological heterogeneity). These factors alone or in combination may result in intervention effects that vary more than would be expected by chance; such variation is described as statistical heterogeneity. Scoping for this review suggested that while some clinical heterogeneity within each comparison group (differences in baseline severity and doses of intervention) was expected, this would not be sufficient to preclude meta-analysis. While slightly differing intervention results may be observed due to the impact of clinical differences in individual studies, these were assumed to be distributed around a central 'true' intervention effect (hence the use of random effects modelling). This assumption emphasised the need for estimation and reporting of heterogeneity within the review.

To maintain scientific rigour in a review, the identification of potential sources of heterogeneity should be identified *a priori* and be conservative in number (Deeks et al. 2009). Multiple *post hoc* analysis is more likely to result in spurious findings and should be avoided (Anello and Fleiss 1995). From scoping exercises three potentially important sources of clinical heterogeneity were identified:

(1) Baseline severity of depression;
(2) Training level of the therapist (graduate versus postgraduate/experienced therapist qualification);
(3) Number of treatment sessions.

Study quality was also considered a source of potential methodological heterogeneity, incorporating a cut-off point of 6 on the 8-point quality scale which had been agreed in discussions of the review steering group.

Measurement of heterogeneity

It is possible to view the variance between studies by observing the degree of overlap in confidence intervals on a forest plot. Such approaches provide a simple observation; however further formal testing is advised (Centre for Reviews and Dissemination 2009). Chi-square tests examine if the variance observed in results across studies is likely to be as a result of chance alone. Caution has been advised in the use of chi-square tests especially in meta- analysis with few studies and low numbers in some trials. This results in low power of the test, which may provide misleading results (Deeks et al. 2009). This can be adjusted for in the increase of the P value to 0.10 rather than 0.05; however the problem remains that non-significant results cannot be assumed to guarantee no heterogeneity.

An alternative approach is to assume that heterogeneity is always present in meta-analysis and the important factor to assess is the degree to which it influences results (Higgins et al. 2003). This approach uses the I^2 statistic to estimate inconsistency. The I^2 statistic uses the chi-square statistic (Q) and its degrees of freedom to describe the percentage of variability in the meta-analysis due to heterogeneity rather than chance (Higgins et al. 2003, Higgins and Thompson 2002);

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

Interpretation of the I^2 statistic is as follows:

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

Based upon scoping it was decided to adopt the I^2 statistic approach, as it was expected that a number of small studies of behavioural activation would be identified which could have resulted in difficulty with the interpretation if significance testing via chi-square approaches.

3.3.14 Sensitivity analysis

The impact of these sources of heterogeneity and study quality was explored via sensitivity analyses and meta-regression where possible (Thompson and Higgins 2002). Meta-regression examines how an outcome variable (SMD) may be predicted by pre-defined study characteristics or 'potential effect modifiers'. In comparisons with ≥10 studies, outcomes were analysed using meta-regression, specifying sources of heterogeneity as predictive co-variates. A permutation test (using 1000 Monte Carlo simulations) was used to calculate p-values, and to reduce spurious false positive findings (Higgins and Thompson 2004). The amount of heterogeneity explained by predictive co-variates was examined by reductions in the I^2 inconsistency statistic within our model. This approach was discussed between DE and SG to agree the rationale for the approach. SG had previous experience of conducting meta-regression in exploration the impact of study level moderators in a large meta-analysis of collaborative care interventions for depression (Gilbody et al. 2006). Analyses were conducted using the metan and metareg commands in Stata 8 (Stata Corporation 2003) by SG.

3.3.15 Publication and small study bias

Although the search strategies described earlier in this chapter were designed to reduce the possibility of reporting bias, they cannot eliminate the risk of it impacting on the review findings, therefore formal assessment is essential (Song et al. 2000, Rothstein et al. 2005). In this review the possibility of publication bias was assessed through a Begg funnel plot graph (Begg et al. 1994). These graphs are simple scatterplots that initially compared sample size on the vertical axis vs. intervention effect estimates of the horizontal axis. As effect estimates will be more precise as sample sizes increase, smaller studies will be more widely distributed towards the bottom of the graph. This results in the distribution representing a funnel in the absence of publication bias. If such bias is present and studies with non-significant effects are missing, the funnel will appear asymmetrical (see Figure 9). As statistical power is related to more than pure sample size

(such as standard deviation of responses in continuous outcomes), it is now recommended that standard error of intervention effect rather than sample size is plotted on a vertical axis (Sterne and Egger 2001). This places standard errors on a reversed scale, placing the most powerful studies at the top of the scale with the additional benefit of allowing the plotting of a triangular region within which 95% of studies would fall in the absence of bias or heterogeneity. Small studies, often found in psychotherapy, can still impact on the distribution of plots in a funnel plot graph. They may offer less precision and appear and introduce a positive bias in interpretation. They may also include a particularly severe or treatment resistant sample and introduce a negative bias to interpretation.

Some problems therefore exist with a reliance on funnel plots alone in estimating publication and small study bias. Firstly the subjective approach relies on the researcher's ability to identify funnel plots reflecting bias, which is unreliable (Terrin et al. 2005, Tang and Liu 2000) and SMD is naturally correlated with its standard error and can therefore produce spurious asymmetry (Deeks et al. 2009).

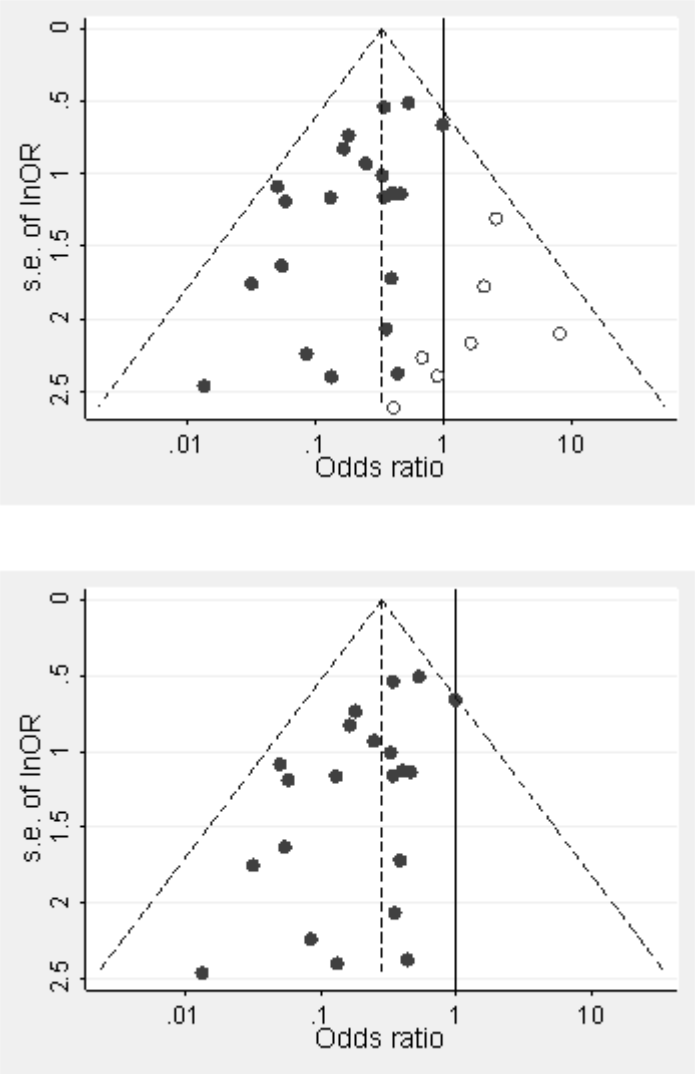


Figure 9: Examples of funnel plot graphs showing symmetrical (1) and asymmetrical (2) distribution (Deeks et al. 2009)

A sufficient number of studies are required for this approach to achieve significant power to detect real asymmetry from chance and be interpreted in light of the visual inspection (N≥10). To counter these problems, in addition visual examination of funnel plots we tested for asymmetry using Egger weighted regression test (Egger et al. 1997) where the intercept is 0 if no bias is present. Egger regression testing was conducted using Stata 8 (Stata Corporation 2003) by SG.

3.4 Results of systematic review and meta- analysis

Searches conducted between October 2005 and February 2006 identified 3353 potential studies for inclusion within the review. Following title review, 3268 reports were excluded as not relevant. Abstracts of the remaining 85 studies were reviewed, of which 52 were rejected. Of these, 9 studies included participants who did not fall within the review criteria, 30 studies reported interventions did not fall within the review criteria, two studies did not report relevant outcomes and 11 were not randomised controlled trials. A total of 33 full papers was then reviewed, of which 13 were rejected. Of these, nine studies included participants who did not fall within the review criteria, in one study the comparator did not fall within the review criteria, and three studies were not randomised controlled trials. A total of 20 relevant randomised controlled trials were finally identified as meeting the criteria for the review, which included a combined total of 1215 participants (Taylor and Marshall 1977, McLean and Hakstain 1979, Gallagher and Thompson 1982, Maldonado Lopez 1982, Wilson 1982, Wilson et al. 1983, Maldonado Lopez 1984, Skinner 1984, McNamara and Horan 1986, Thompson et al. 1987, Scogin et al. 1989, Jacobson et al. 1996, McKendree-Smith 1998b, Hopko et al. 2003 , Dimidjian et al. 2006, Cullen et al. 2006, Padfield 1976, Zeiss et al. 1979, Gardner and Oei 1981, Cole 1983). Three studies were excluded from the meta-analysis due to insufficient statistical data being available (Padfield 1976, Zeiss et al. 1979, Gardner and Oei 1981). All studies are listed in table 7. Each was allocated a number used for identification throughout the remainder of this chapter. In all, 17 randomised controlled trials, including a combined total of 1109 participants, were included in the meta-analysis. For further details of study selection processes see Figure 10. For details of included studies see Table 7. For study quality assessment scores see Table 8.

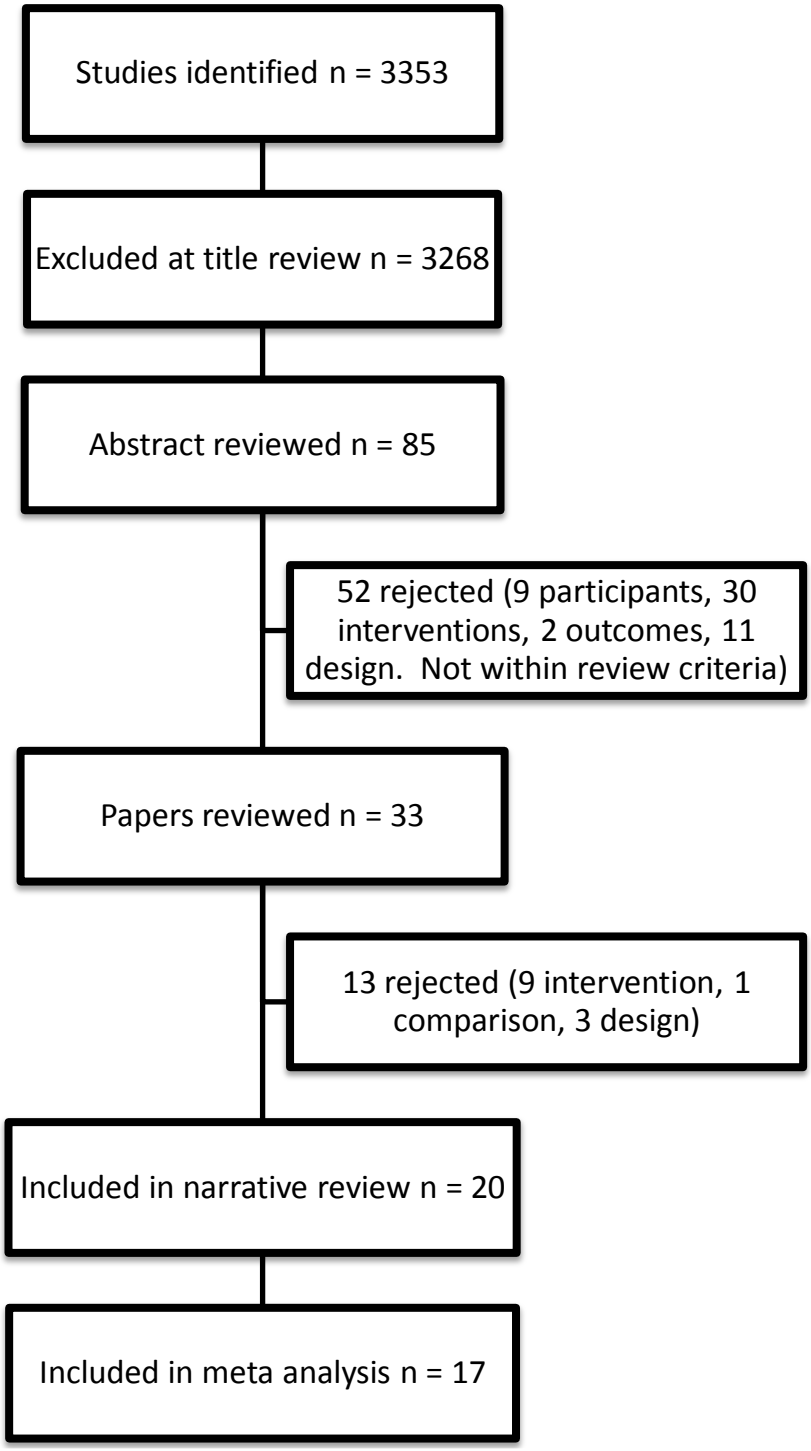


Figure 10: Flow chart of study identification

Table 7: Details of identified studies

Authors and year (study no.)	Sample/setting Mean age (SD/range) Sex (% female)	Interventions (n)	Depression level at baseline	Concurrent Pharmacology	Therapist level. Session number (duration)
Taylor and Marshall. 1977. (1)	university students 22.4(2.6) 71	behavioural (7) cognitive (7) cognitive Behavioural (7) wait List (7)	mild/moderate (21.2 BDI)	no	graduate student experience as counsellor 6 (40)
McLean and Hakstein. 1979. (2)	community outpatient39.2(10.9) 72	behavioural (42) brief Psychotherapy (44) drug Therapy (49) relaxation (43)	within or beyond moderate depression range 2 out of 3 measures used at baseline	no (other than DT arm)	licensed psychologists, physicians or psychiatrists. At least 2 years of experience as therapist 10 (1hour) not drug therapy
Gallagher and Thompson. 1982. (3)	older adult community. 67.76 (6) 76	behavioural (10) cognitive (10) psychotherapy (10)	RDC Criteria MDD	no	advanced PhD or post-doctoral therapists experience in modality 16 (90mins)

Maldonado Lopez. 1982. (4)	community outpatient. NA NA	behavioural (8) cognitive (8) drug Therapy (8)	psychiatrist diagnosis reactive depressive disorder	no (other than DT arm)	psychology dept, level of training not reported 10 (1 hour)
Wilson, P. 1982. (5)	general population media announcements. 38.8(20-55) 66	drug therapy and behavioural (12) drug therapy and relaxation (10) drug therapy and minimal contact (10) placebo and behavioural (9) placebo and relaxation (11) placebo and minimal contact(12)	BDI>19	in DT arm	graduate psychologist 7 (1hour) 2 (1hour) in min. contact arm
Wilson, P et al. 1983. (6)	general population media announcements. 39.5 (20-58) 80	behavioural (8) cognitive (8) wait list (9)	BDI >17 (moderate depression)	yes (5 subjects in trial)	not clear, university psychology clinic 8 (1hour)

Cole, M. 1983 (7)	community outpatient veterans. 56(24-71) 0	behavioural (15) treatment as usual (15)	psychiatrist diagnosis major depression. BDI>24	yes if stable	doctoral clinical psychology student 7(1hour)
Maldonado Lopez. 1984. (8)	community outpatients. NA NA	behavioural and pharmacology (8) cognitive and pharmacology (8) pharmacology (8)	psychiatrist diagnosis reactive depressive disorder	all Subjects	psychology dept, level of training not reported 10 (1hour)
Skinner, A. 1984. (9)	community volunteers. 20-61 (34) 67.5	behavioural (8) cognitive(7) control (9)	BDI>12	yes	doctoral clinical psychology Student 5(1hour)
McNamara, K and Horan, J. 1986. (10)	university students. 23(19-31) 73	behavioural (10) cognitive (10) cognitive Behavioural (10) supportive (10)	BDI>17 HRSD>20	not reported	doctoral interns in clinical psychology/masters level social worker 8 (50mins) (10 sessions in CBT arm)

Thompson, L et al. 1987. (11)	older adults community. 67.07 (5.8) 67	behavioural (25) cognitive (27) psychotherapy (24) delayed (19)	RDC major depressive disorder	if stable dose for 3 months	doctoral level psychologists plus 1 year specialised therapy training 16-20 (duration of each session not reported)
Scrogin, F et al. 1989. (12)	older adults community. 68.3(6.7) 85	behavioural bibliotherapy (23) cognitive bibliotherapy (22) delayed (22)	>9 on HRSD	if stabilised prior to trial	NA as bibliotherapy was main intervention 4 (5min) phone contacts to support exercises
Jacobson et al. 1996. (13)	community (80% HMO, 20% volunteer) 38 (not reported). 72	behavioural (56) thought challenging (43) full Cognitive (50)	major depression (DSM-IV) >19 BDI	No	experienced therapists (mean 9.5 years CT practice) 20 sessions (NA)
McKendree Smith. 1998. (14)	community volunteer. 913.17) 75	behavioural bibliotherapy (13) cognitive bibliotherapy (13) delayed (14)	mild-moderate depression	if stabilised for 3 months	NA as bibliotherapy main intervention 8(10 mins)

Hopko, D et al. 2003. (15)	inpatients. 30.5(9) 36	behavioural (10) supportive (15)	principle diagnosis of major depression	yes all patients	not clear 6 (20mins)
Dimidjian, S et al. 2006. (16)	community. 39.9 (10.97) 66	behavioural (43) cognitive (45) pharmacology/Placebo (153)	major depression (DSM-IV) >19 BDI	only in ADM arm	BA-licensed psychologists/social worker (7 years' practice. CT-Licensed Psychologists with CT training 24(50mins)
Cullen, J. 2006. (17)	community. 38.48(12.69) 32	behavioural (13) wait List (12)	MDD (Mean BDI 30.96(5.90)	yes if stable >6 weeks	previous experience in CT of depression plus 12 hours training in BA 10(50)
Studies not included in Meta-Analysis					

Padfield, M. 1976. (18)	community female rural low socioeconomic status. 21-56 100	behavioural (12) supportive (12)	moderately depressed (diagnostic tool not clear)	no	counsellor (experience not clear) 12 (NA)
Zeiss et al. 1979. (19)	community. 33.9 (19-68) NA	behavioural (22) cognitive (22) interpersonal (22)	classed as depressed using Minnesota Multiphasic Personality Inventory and Grinkler Interview Rating	not clear	graduate students in clinical psychology and counselling psychologists (master's level). At least 1 year experience. 12 (NA)
Gardner, P and Oei, T. 1981. (20)	community. 19-65 77%	behavioural (8) cognitive (8)	mild depression (BDI)	not Clear	not clear NA

3.4.1 Study quality assessment

Each study was measured against our 8 quality standards as outlined in the methods section (Table 8).

Table 8: Study quality scores

study number	selection Bias			performance bias	measurement bias		attrition bias		overall quality score
	adequate randomisation concealment	groups equal at baseline	pre-specified eligibility	adherence to treatment assessed independently	assessment of outcome independent	mean and SD reported	ITT analysis	loss to follow-up reported	
1	0	1	1	0	1	1	0	1	5
2	0	1	1	1	1	1	0	1	6
3	0	1	1	1	1	1	0	1	6
4	0	0	1	0	1	1	0	0	3
5	0	1	1	0	0	1	0	1	4
6	0	1	1	0	1	1	0	0	4
7	0	1	1	0	1	1	0	0	4
8	0	0	1	0	1	1	0	0	3
9	0	1	1	0	1	0	0	0	3

10	0	1	1	0	1	1	0	1	5
11	0	1	1	1	1	1	0	1	6
12	0	1	1	1	1	1	0	1	6
13	0	1	1	1	1	1	1	1	7
14	0	1	1	0	1	1	0	1	5
15	0	1	1	0	1	1	0	1	5
16	1	1	1	1	1	1	1	1	8
17	0	1	1	1	1	1	1	1	7
Studies included not included in meta-analysis									
18	0	0	1	1	0	0	0	0	2
19	0	0	1	0	0	0	0	1	2
20	0	0	0	0	1	0	0	0	1

3.5 Behavioural Therapy vs. control: comparison 1

3.5.1 Scope

Twelve studies with a total of 459 patients contributed data to this analysis (1, 2, 4, 5, 6, 7, 9, 11, 12, 14, 16, 17). Participants were taken from adult community sources consisting of outpatients (2,4,6,7,11,12,16,17), volunteers (5, 8,14) and students (1), two studies using older adults (11,12). Control interventions consisted of delayed treatment (1,3,9,11,12,14,16,17), treatment as usual (4,5,7) and relaxation (2,5). All comparisons were taken immediately after intervention. Interventions ranged from supported bibliotherapy (12, 14), brief therapy with six 40-minute sessions (1) to 24 50-minute sessions (16). Facilitators were advanced graduate psychology/therapy students in five studies (1, 5, 6, 7, 9), experienced psychotherapists in four studies (2, 11, 16, 17) and unclear in one study (4). Depression symptom level was assessed using either BDI self-report measure (1, 2, 4, 5, 7, 9, 17) or the HRSD assessor rating scale (12), or both (6, 11, 14, 16). Recovery was defined by clinical interview in one study (11) and by BDI score in 2 studies (2, 14).

3.5.2 Depression symptom level post-treatment

The effect size of behavioural interventions against control interventions was large, with a pooled SMD of −0.70 CI −1.00 to −0.39, demonstrating a highly significant difference in symptoms level scores favouring the behavioural group (P<0.001) (see Figure 11). There was no evidence of publication bias for this outcome, Egger's test = −1.04; 95% CI = −3.39 to 1.29 P=0.35, a funnel plot showed no evidence of asymmetry (Figure 12).

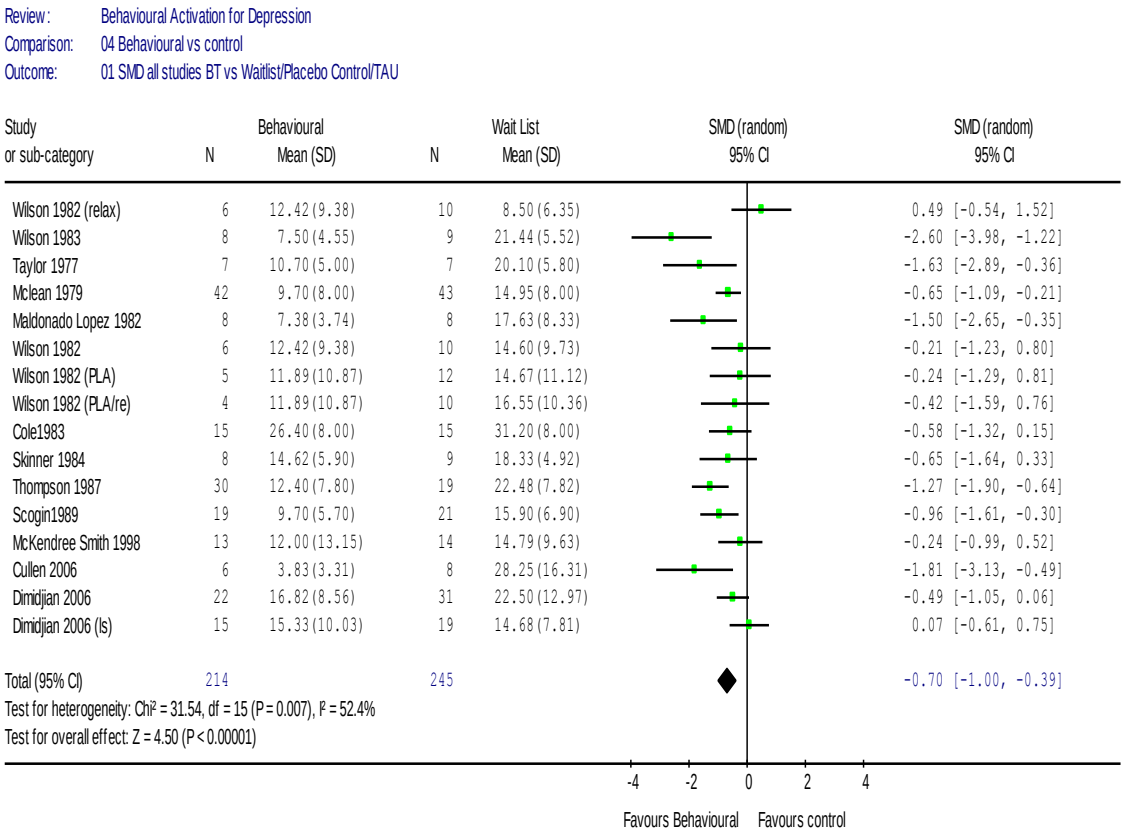


Figure 11: Forest plot BA vs. control depression symptom level post-treatment

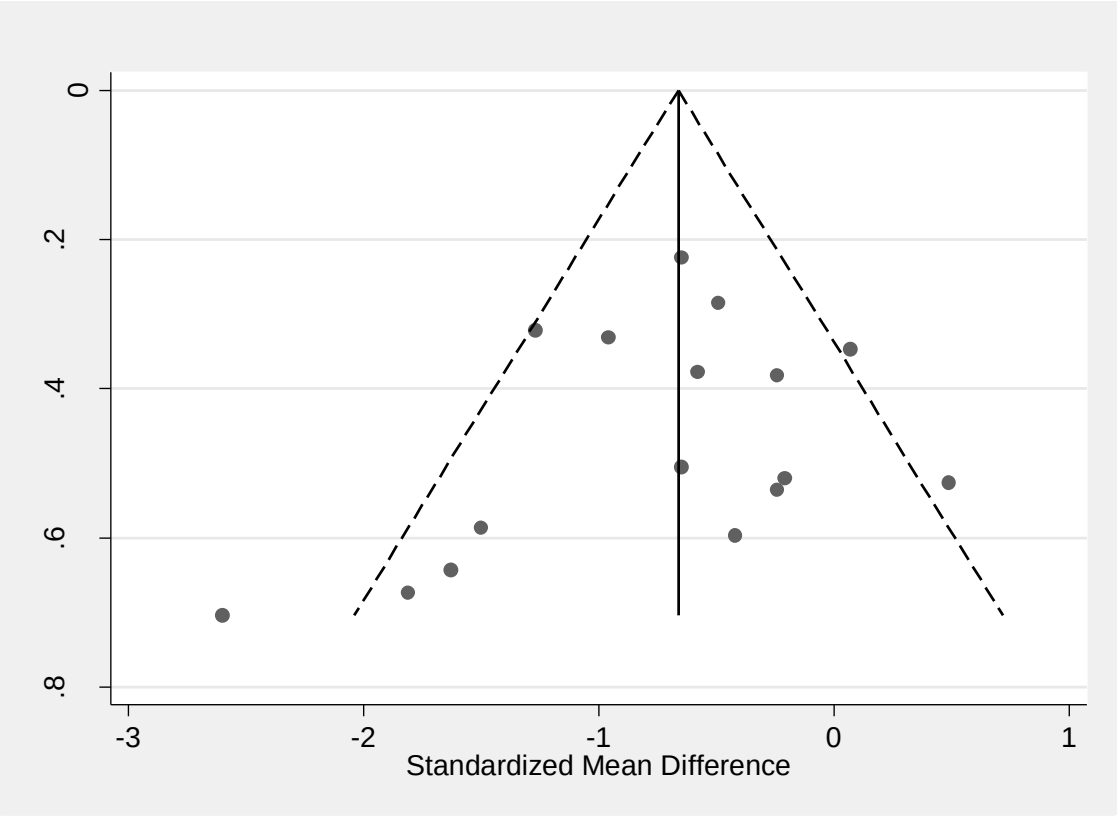


Figure 12: Begg funnel plot symptom level BA vs. control

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was moderate (55.1%). Effect size was not significantly related to the level of baseline severity, (meta-regression beta-co-efficient = 0.04, 95% CI −0.04 to 0.12; I^2 = 54%, P= 0.28); see Figure 13. Quality assessment indicated that seven studies fell below our quality threshold (1, 4, 5, 6, 7, 9, 14), and the pooled SMD was not affected by study quality (meta-regression $SMD_{low\ quality}$ = −0.67; $SMD_{higher\ quality}$ = −0.75, $P_{difference}$ = 0.77) see Figure 14. Behaviour therapists with graduate and postgraduate qualifications produced similar effect sizes (meta-regression $SMD_{graduate}$ = −0.82; $SMD_{post\ graduate}$ = −0.59, $P_{difference}$ = 0.61; I^2 = 59%) see Figure 15. There was no clear relationship between effect size and number of sessions (meta-regression beta-coefficient = 0.03, 95%CI −0.03 to 0.09; I^2 = 0.49, P=0.27) (Figure 16). Prioritising clinician-rated assessment in precedence over self-rated where possible made no significant difference to overall effect size (SMD −0.68

CI −0.98 to – 0.38).

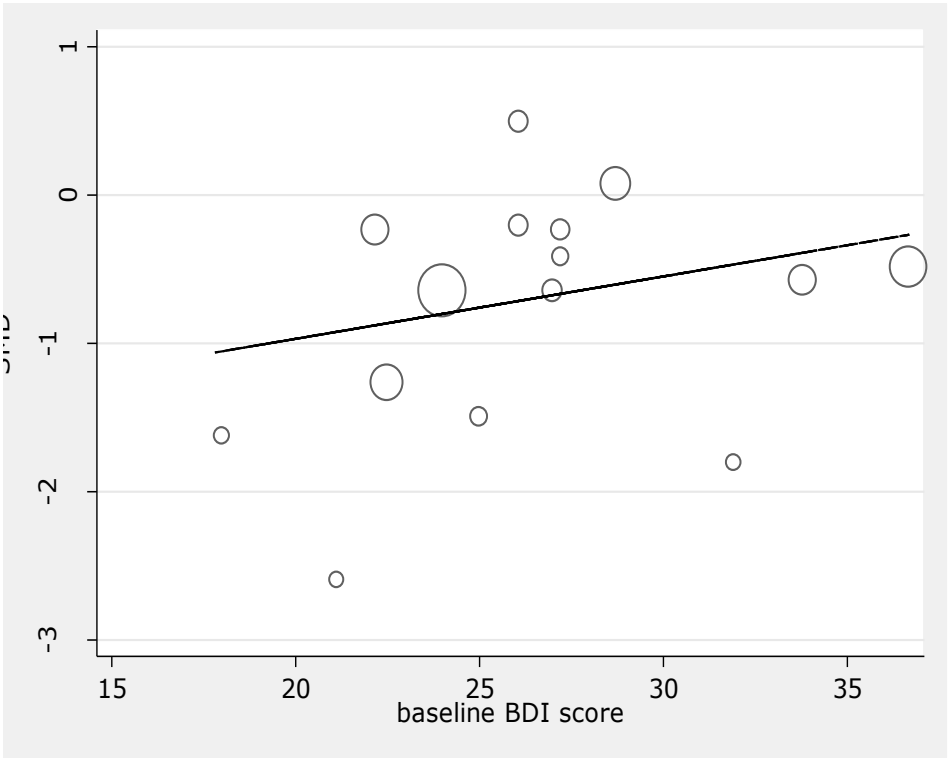


Figure 13: Meta-regression of baseline severity of depression (BDI score) versus effect size in studies comparing behaviour therapy versus control.

(Regression weighted by within-study inverse variance, represented by size of graph data points. Regression line fitted, SMD = −1.81 + 0.042*BDI score, P=0.28).

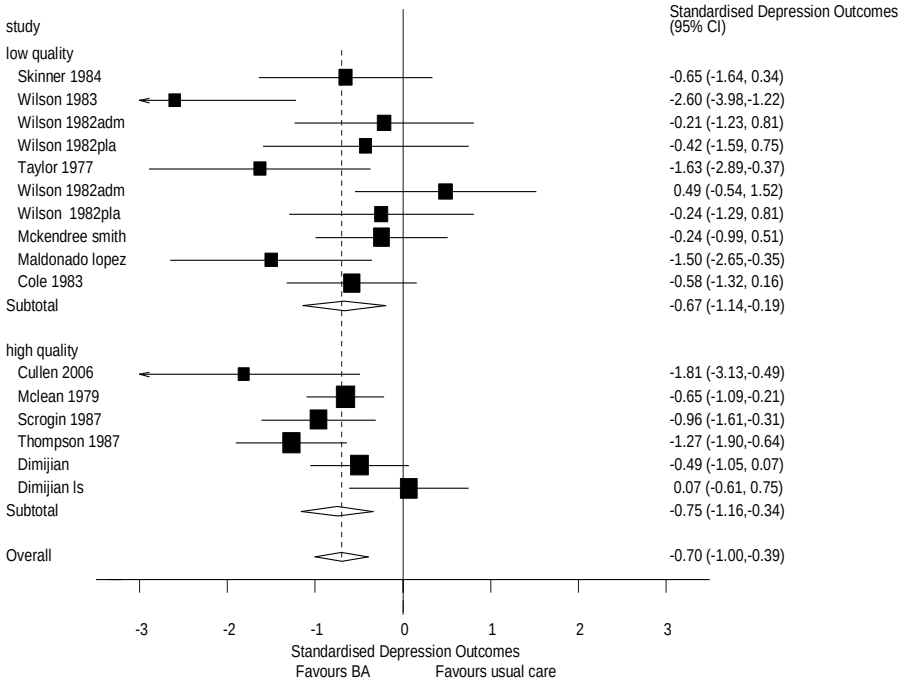


Figure 14: Study quality meta-regression

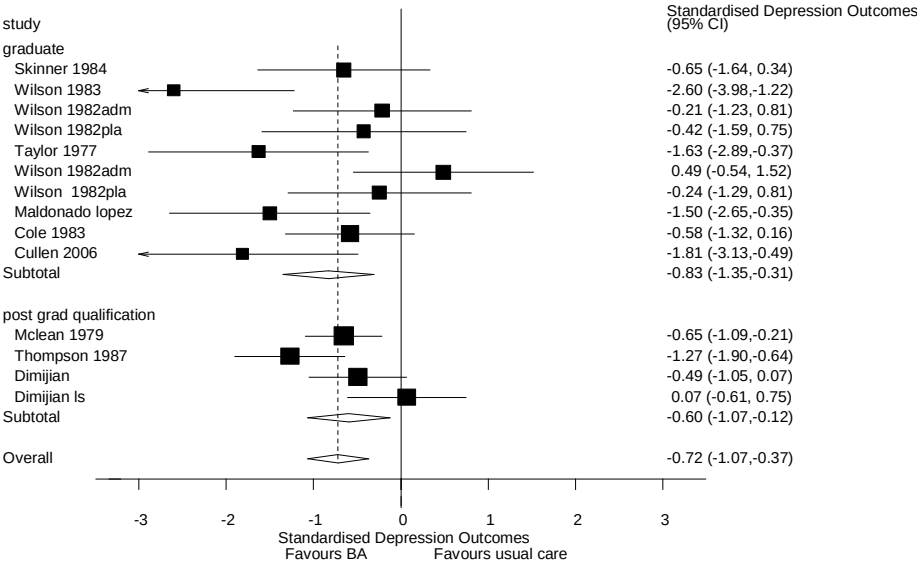


Figure 15: Therapist level meta-regression

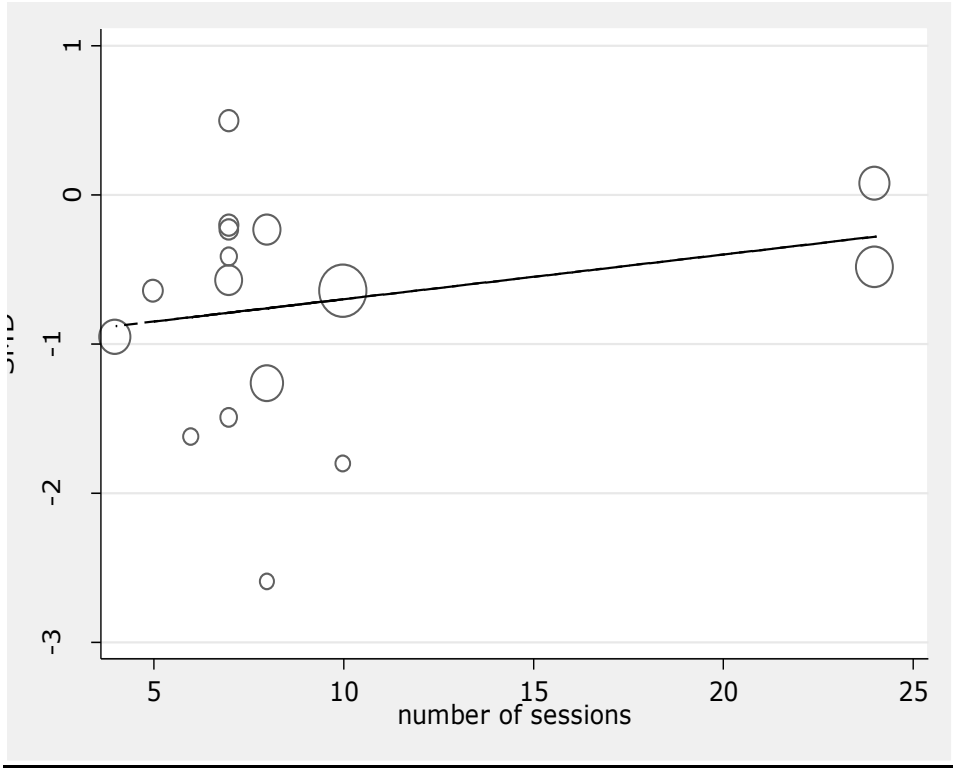


Figure 16: Meta-regression of number of sessions versus effect size in studies comparing behaviour therapy versus control.

(Regression weighted by within-study inverse variance, represented by size of graph data points. Regression line fitted, SMD = −0.99 + 0.03*sessions, P=0.27)

3.5.3 Dropout Rate

Three studies contributed data to this analysis (2, 14, and 16) on a total of 119 subjects with an average dropout rate of 19.17%. There was no difference between rates of dropout between intervention and control: odds ratio = 0.58 CI 0.28 to 1.21 (P=0.86) (Figure 17). There were insufficient studies to check for publication bias for this outcome.

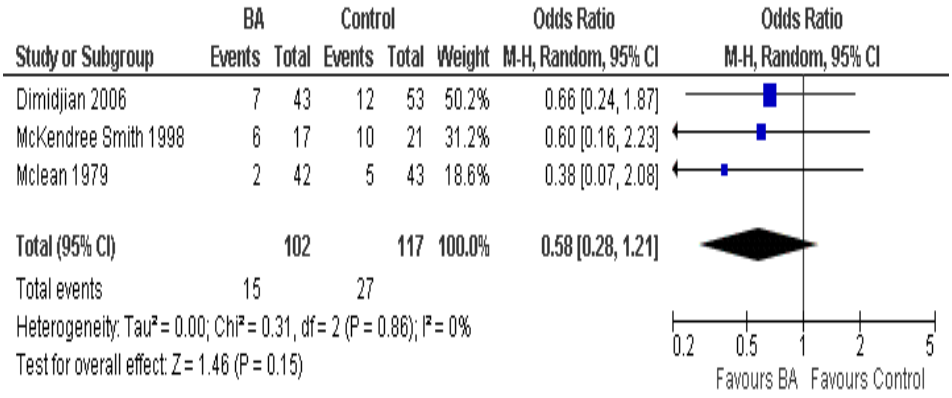


Figure 17: Forest plot of dropout data BA vs. control

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was small (0%). There were insufficient studies and negligible heterogeneity to explore the impact of our *a priori* sources of clinical heterogeneity.

3.5.4 Recovery rate

Three studies contributed data to this analysis (2,11,14) on a total of 167 subjects. There were greater rates of recovery in the behavioural intervention group (BT 52%, control 21.05%) with an odds ratio of 4.18 CI 1.14 to 15.28 (P=0.03) (Figure 16). There were insufficient studies to test for publication bias for this outcome.

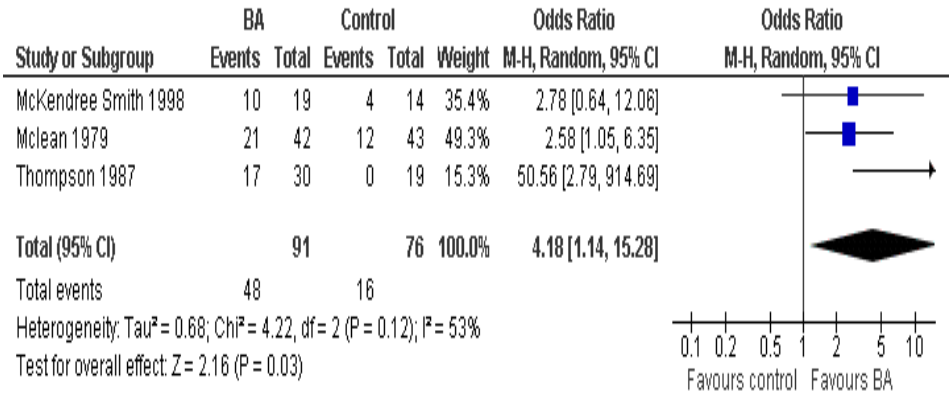


Figure 18: Forest plot of recovery rate BA vs. control

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was medium (53%). There were insufficient studies to explore the underlying causes of this heterogeneity further or to conduct meaningful sensitivity analysis.

3.6 Behavioural Therapy vs. CT/CBT: comparison 2

3.6.1 Scope

Twelve studies with a total of 476 patients contributed data to this analysis (1, 3, 4, 6, 8, 9, 10, 11, 12, 13, 14, 16). Participants were taken from adult community sources consisting of outpatients (3, 4, 8, 11, 12, 13, 16), volunteers (6,9,14) and students (1, 10), three studies using older adults (3,11,12). Interventions ranged from supported bibliotherapy (12, 14), brief therapy with six 40-minute sessions (1) to 24 50-minute sessions (16). Therapy was facilitated by advanced graduate psychology/therapy students in four studies (1, 6, 9, and 10), experienced psychotherapists in four studies (3, 11, 13, and 16); it was unclear who the facilitators were in two studies (4, 7). Depression symptom level was assessed using either BDI self-report measure (1, 4, 8, 9, 10) or the HRSD assessor rating scale (12), or both (3, 6, 11, 13, 14, 16). Recovery was defined by diagnostic interview in two studies (3, 11) and by BDI score in three studies (10, 13, and 16).

3.6.2 Depression symptom level post-treatment

No difference in effect between behavioural interventions and CBT/CT was identified, with a pooled SMD of 0.08 CI −0.14 to 0.30 (P=0.46) (see Figure 19). There was no evidence of publication bias for this outcome using Egger's test, (Intercept (0 if unbiased) = 1.07; 95% CI = −0.23 to 2.38 P=0.10), and a funnel plot showed no evidence of asymmetry (Figure 20).

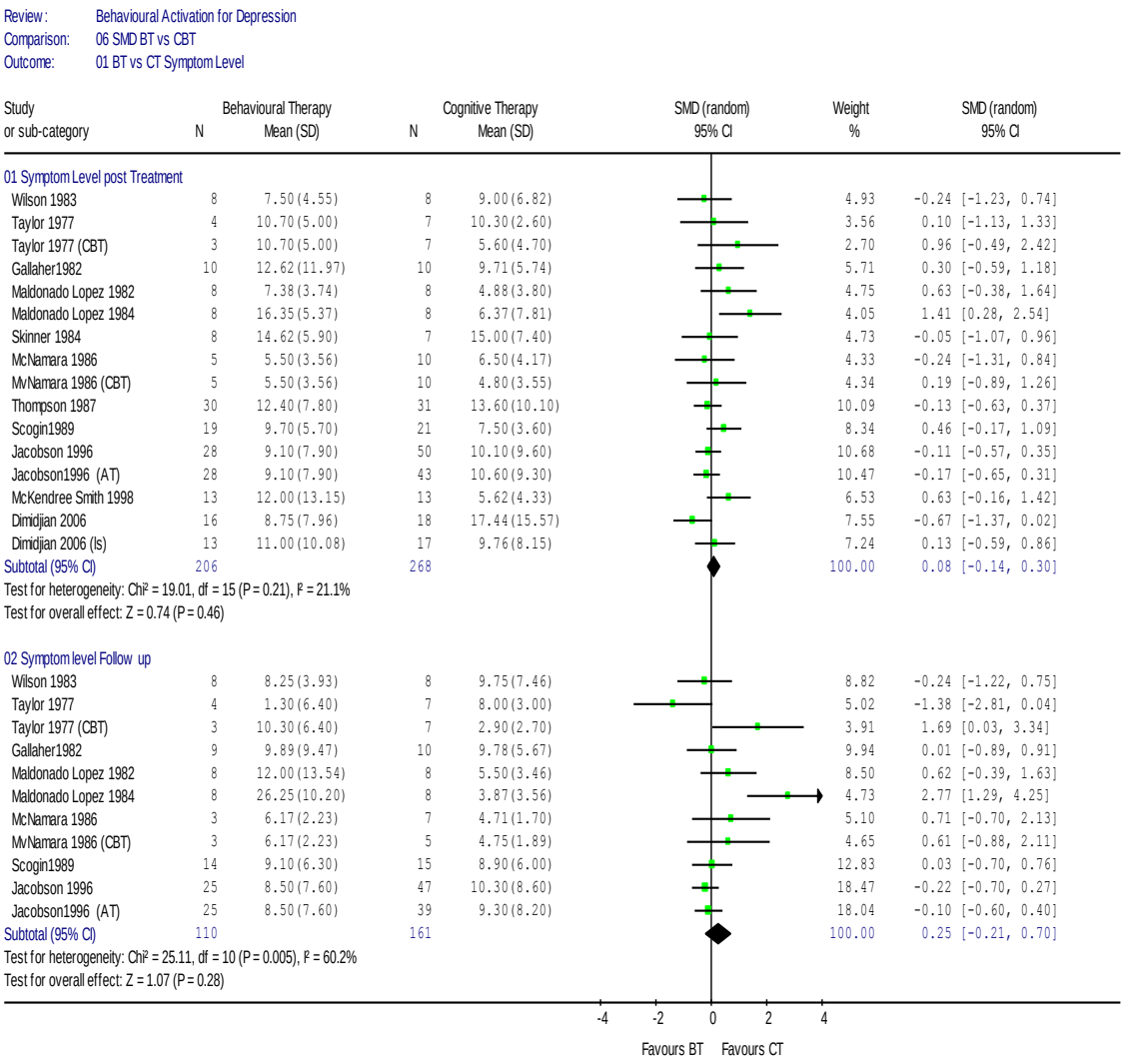


Figure 19: Forest plot BT vs. CBT/CT post treatment and follow-up

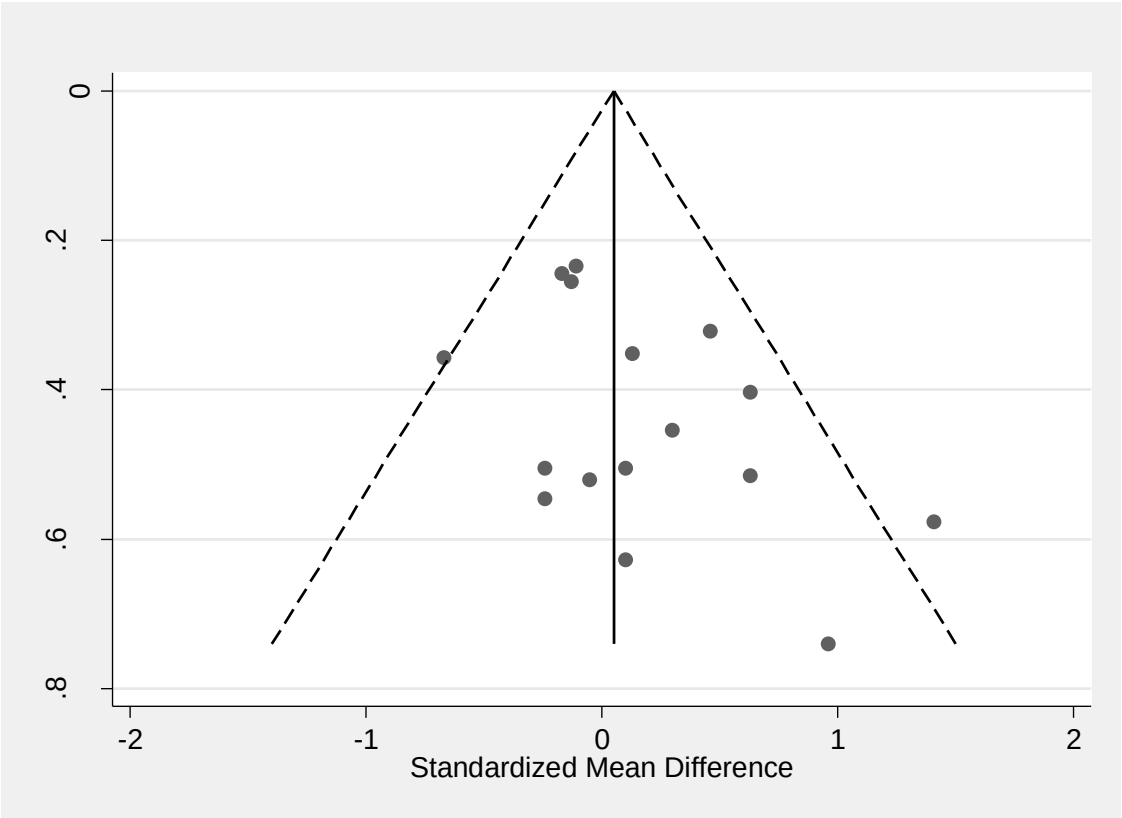


Figure 20: Begg funnel plot of BT vs. CBT

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was small (21.1%). Seven studies fell below our quality threshold (1,4,6,7,9,10,14) and pooled SMD was not significantly affected by study quality (meta-regression $SMD_{low\ quality}$ = +0.23; $SMD_{higher\ quality}$ = −0.13, $P_{difference}$ = 0.12; I^2 = 0%) (Figure 21). Comparative effectiveness of BT versus CT/CBT varied according to baseline severity of depression, behaviour therapy associated with a greater level of effectiveness at more severe levels of depression (meta-regression beta-co-efficient = −0.05, 95% CI −0.10 to −0.01; I^2 = 0%; P = 0.04) (Figure 22).

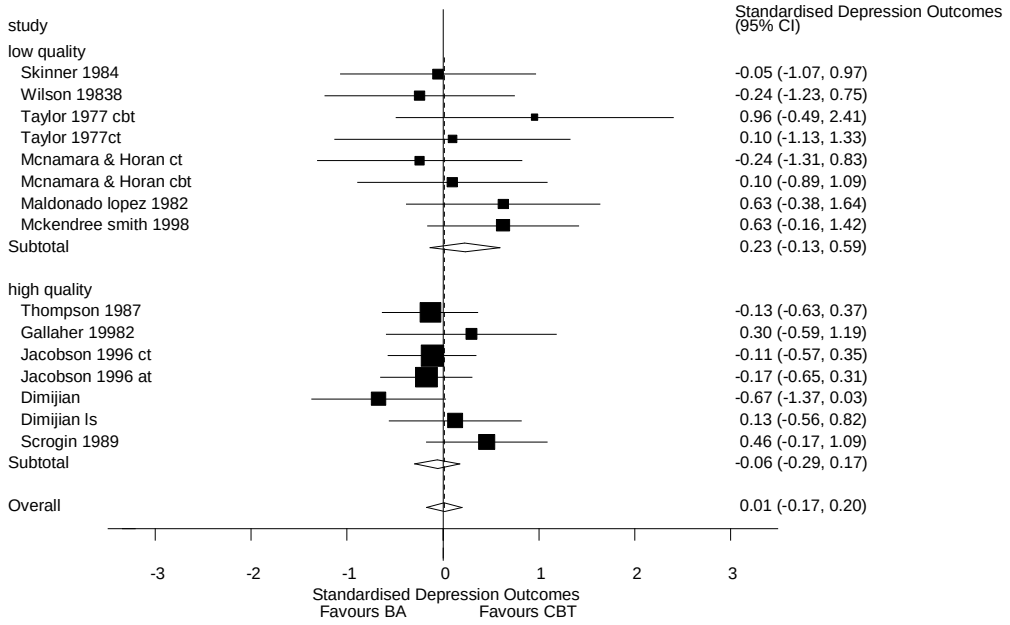


Figure 21: Study quality meta regression BT vs. CBT

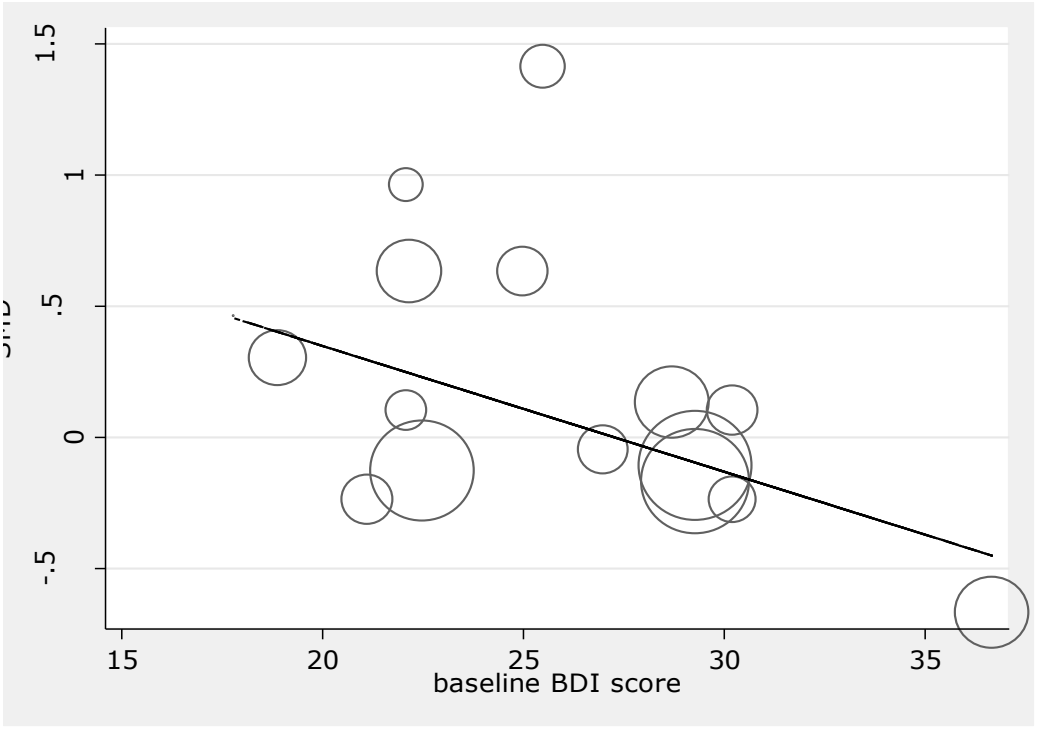


Figure 22: Meta-regression of baseline severity of depression (BDI score) versus effect size in studies comparing BA versus CBT.

(Regression weighted by within-study inverse variance, represented by size of graph data points. Regression line fitted, SMD = + BDI score, P=0.04).

Graduate level behaviour therapists produced slightly less favourable results compared to those with postgraduate qualifications in comparison to CBT, although this did not reach significance (meta-regression $SMD_{graduate}$ = 0.28; $SMD_{post\ graduate}$ =−0.135, $P_{difference}$ = 0.11; I^2 = 0%) (Figure 23). There was no clear relationship between effect size and number of sessions (meta-regression beta-coefficient = −0.025, 95% CI −0.056 to 0.006; I^2 = 0.08, P= 0.11) (Figure 24). Prioritising clinician-rated assessment over self-rated where applicable made no difference in overall effect size (SMD 0.09 CI −0.12 to 0.29).

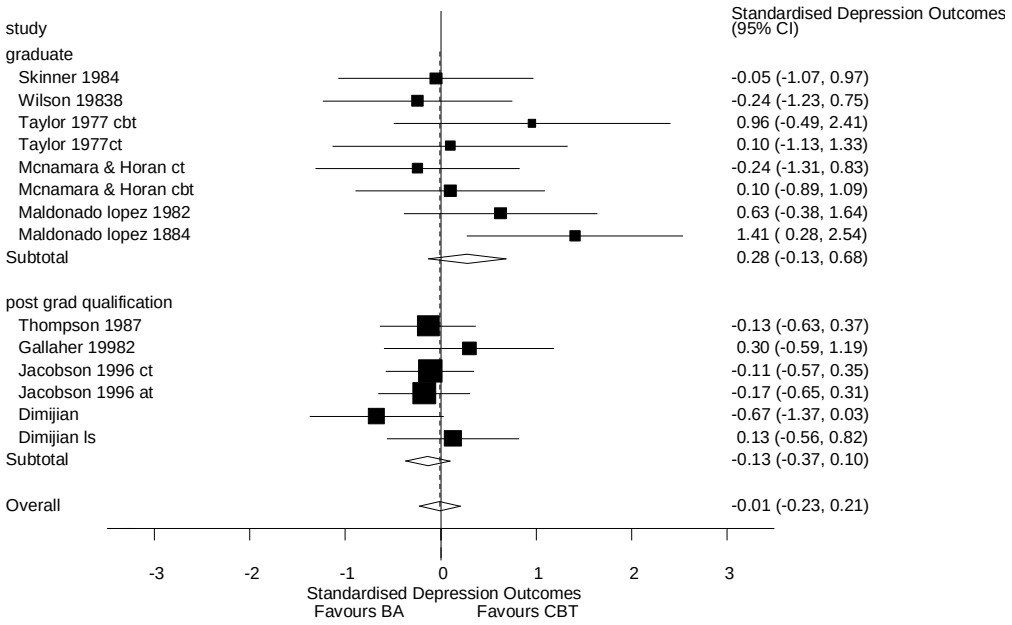


Figure 23: Therapist level meta regression BA vs. CBT

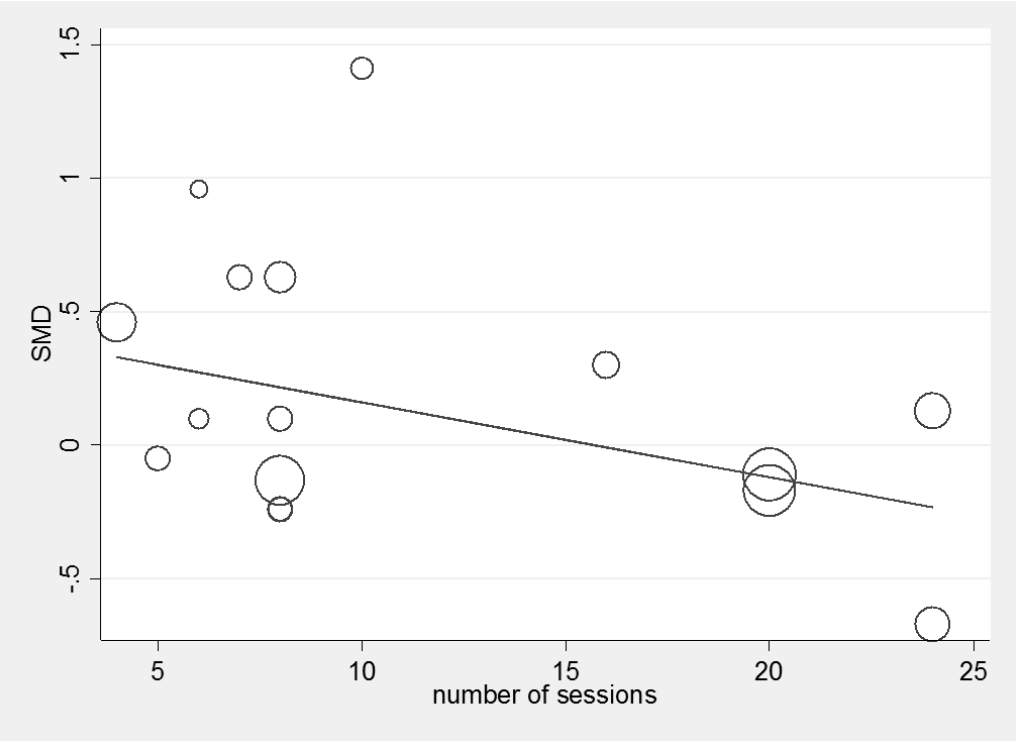


Figure 24: Meta regression number of sessions BA vs. CBT

3.6.3 Depression symptom level at follow-up

Eight studies contributed data to this analysis (1, 3, 4, 6, 8, 10, 12, and 13) on a total of 271 subjects with an average follow-up period of four months. Overall there was no difference in effect of BT compared to CBT/CT with a pooled SMD of 0.25 CI –0.21 to 0.70 (P=0.28) (Figure 19).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 60.2%. After exclusion of low-qualitystudies (1, 4, 6, 8, 10) and those with follow up of less than three months (1, 6) in a sensitivity analysis, there remained no difference in effect size between BT and CBT with a pooled SMD of −0.11 CI −0.41 to 0.19 (P=0.47). There were insufficient studies to explore the underlying causes of this heterogeneity further.

3.6.4 Dropout rate

Eight studies contributed data to this analysis (1, 3, 6, 11, 12, 13, 14, and 16) on a total of 436 subjects with an average dropout rate of 15.36%. There was no difference in rates of dropout with an odds ratio of 1.17 CI 0.57 to 2.41 (P= 0.67) (Figure 25). There were insufficient studies to check for publication bias for this outcome.

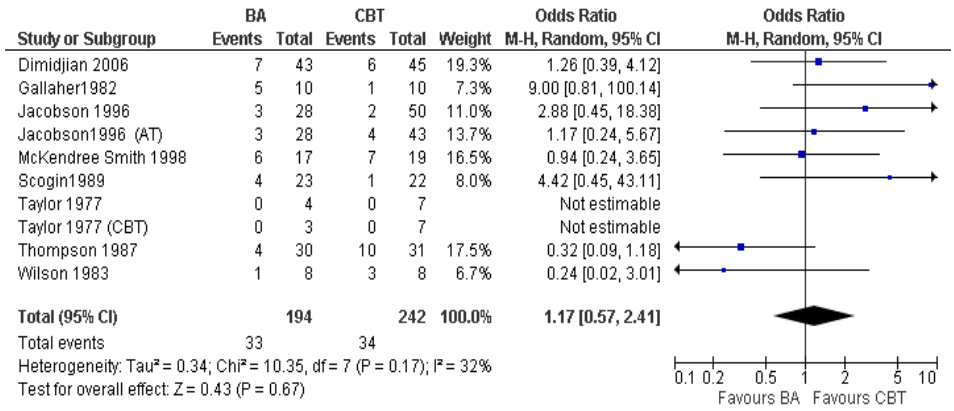


Figure 25 Forest plot of dropout rates BA vs. CBT

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was moderate 32%. Low-quality studies (1, 6, 14) were excluded in sensitivity analysis, resulting in an odds ratio of 1.47 CI 0.60 to 3.61 (P = 0.40) with an I^2 statistic of 42.9%. There were insufficient studies to explore the underlying causes of this heterogeneity further.

3.6.5 Recovery rate

Five studies contributed data to this analysis (3, 10, 11, 13, and 16) on a total of 346 subjects. There was a pooled recovery rate of 55% with no difference between the two treatment approaches: odds ratio 0.92 CI 0.59 to 1.44 (P = 0.72) (Figure 26). There were insufficient studies to check for publication bias in this outcome.

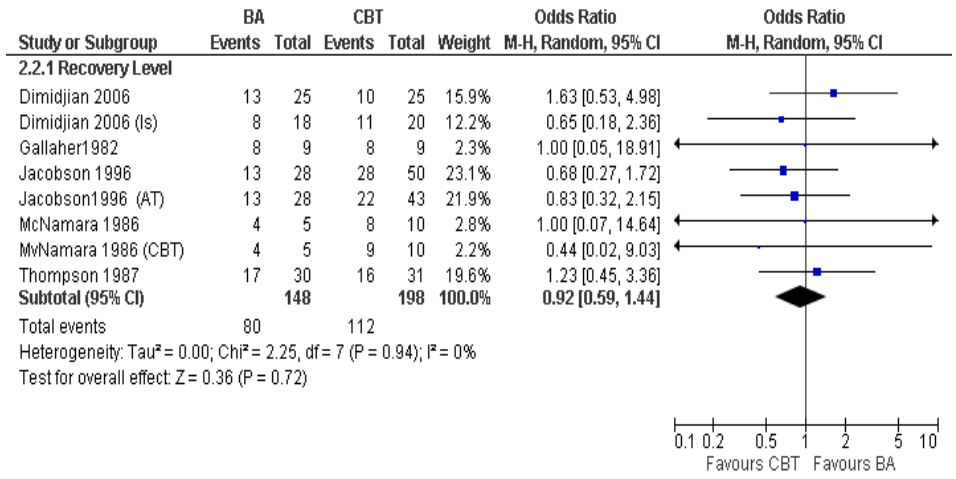


Figure 26: Forest plot BA vs. CBT recovery rates

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 0%. Low-quality studies (8) were excluded in a sensitivity analysis, resulting in an odds ratio of 0.93 CI 0.59 to 1.47 (P=0.77) with an I^2 statistic of 0%.

3.7 Behavioural Therapy vs. Brief Psychotherapy: comparison 3

3.7.1 Scope

Three studies, with a total of 166 patients, contributed data to this analysis (2, 3, and 11). Participants were from adult outpatient community sources, two studies using older adults (3, 11). Brief psychotherapy interventions were based upon a psychodynamic model in all studies. Interventions ranged from 10 to 20 sessions, and all studies used experienced therapists. Studies assessed depression symptom level using the BDI alone (2) or both BDI and HRSD (3, 11). Two studies assessed depression at intake using structured clinical interviews (3, 11), the third using cut-off points from validated self-report measures (2). Recovery was defined by clinical interview in two studies (3, 11) and by BDI score in one study (2).

3.7.2 Depression symptom level post treatment

The effect size of BT compared to brief psychotherapy was large, with a pooled standardised mean difference of −0.56 CI −1.0 to −0.12 (P=0.01) (Figure 27). There were insufficient studies to test for publication bias.

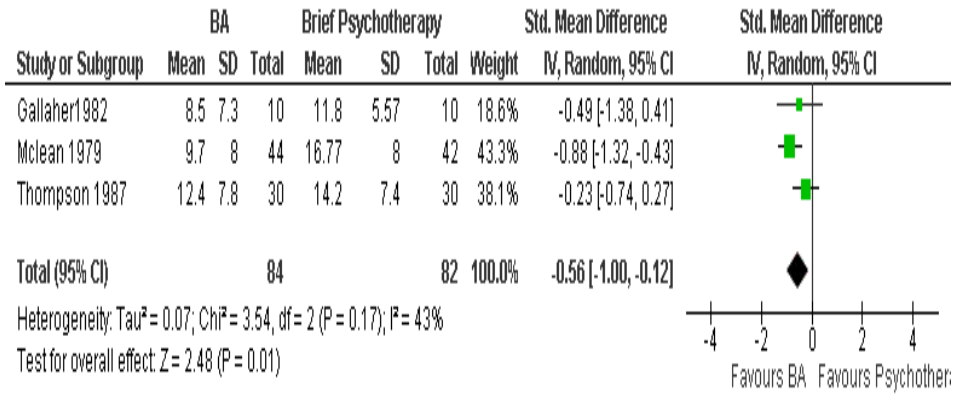


Figure 27: Forest plot BT vs. psychotherapy symptom level at post-treatment

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was moderate 43.4%. All studies were above the quality threshold, hence we performed no sensitivity analyses related to study quality. There were insufficient studies to explore the underlying causes of this heterogeneity further. Further meta-regression to explore the impact of our potential effect modifiers was not possible due to the low number of studies in this comparison. Prioritising clinician-rated assessment over self-rated assessment where applicable made no difference in overall effect size (SMD −0.52 CI −1.01 to – 0.03).

3.7.3 Depression symptom level follow up

Two studies contributed data to this analysis (2, 3) on a total of 96 subjects with an average follow-up period of 4.5 months. The effect size of behavioural interventions compared to brief psychotherapy was medium, with a standardised mean difference of −0.50 CI –0.90 to −0.09 (P=0.02).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 0%. Both studies collected follow-up beyond the three-month point and were above the quality threshold, so we performed no sensitivity analyses. No further meta-regression or testing for publication bias was undertaken due to the low number of studies in this comparison.

3.7.4 Dropout

Three studies contributed data to this analysis (2, 3 and 11) on a total of 166 subjects with an average dropout rate of 14.45% across studies. No difference in dropout was observed, with an odd ratio of 0.94 CI 0.22 to 3.96 (P = 0.11) (Figure 28). There were insufficient studies to test for publication bias.

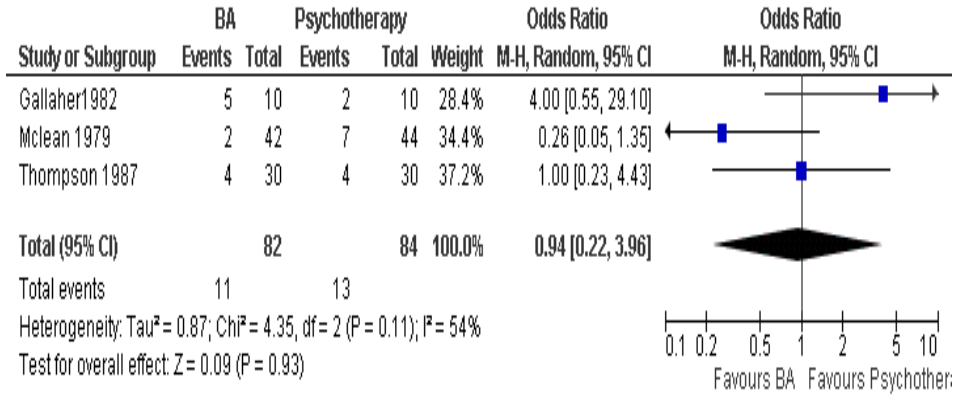


Figure 28: Forest plot BT vs. psychotherapy dropout post-treatment

Heterogeneity and sensitivity analysis

Variation in odds ratio attributable to heterogeneity (I^2) was 54.1%. All studies were above quality threshold so no sensitivity analysis was performed.

3.7.5 Recovery rate

Three trials contributed data to this analysis (2, 3 and 11) on a total of 164 subjects (note 2 subjects deceased). Greater rates of recovery were

observed in BT (56.79%) compared to brief psychotherapy (36.14%), with an odds ratio of 2.37 CI 1.23 to 4.57 (P= 0.01) (Figure 29). There were insufficient trials to test for publication bias.

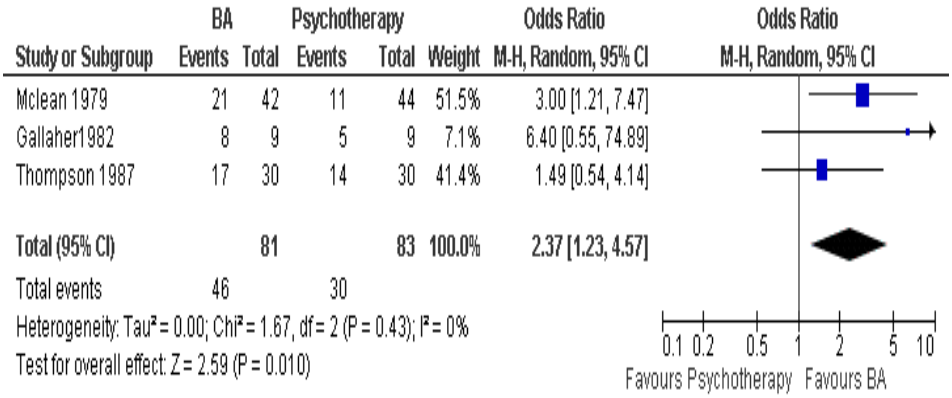


Figure 29: Forest plot BA vs. psychotherapy recovery rate

Heterogeneity and sensitivity analysis

Variation in odds ratio attributable to heterogeneity (I^2) was 0%. All studies were above quality threshold, so no sensitivity analysis was performed. No further meta-regression or testing for publication bias was undertaken due to the low number of studies in this comparison.

3.8 Behavioural Therapy vs. supportive therapy: comparison 4

3.8.1 Scope

Two studies with 45 subjects contributed data to this analysis (10, 15). Participants were university students (10) and inpatients (15). Interventions ranged from six 20-minute sessions (10) to eight, 50-minute sessions (15), delivered by doctoral clinical psychology students (10) or a clinical psychologist (15). Both studies measured depression symptom levels by self report measures (BDI), with one (10) using HRSD also.

Depression at baseline was assessed by self report measures (10) or clinical interview (15).

3.8.2 Depression symptom level post-treatment.

The positive effect in favour of BT compared to supportive therapy was large: SMD −0.75 CI –1.37 to −0.14 (P=0.02) (Figure 30). There were insufficient studies to test for publication bias.

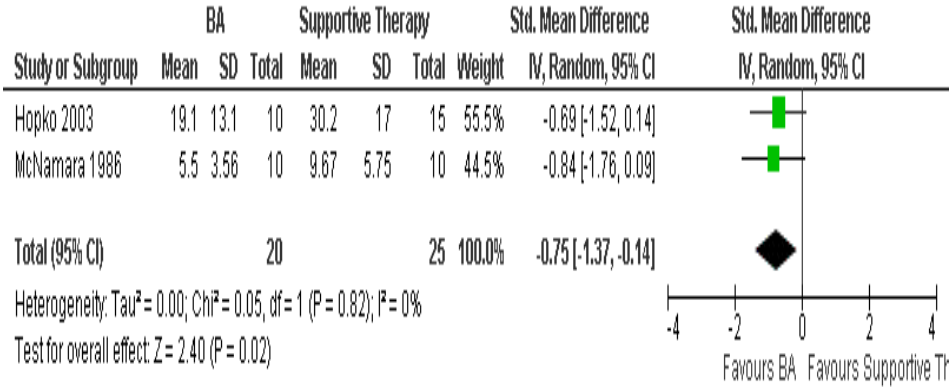


Figure 30: Forest plot BA vs. Supportive therapy post-treatment

Heterogeneity and sensitivity analysis

The variation in effect size attributable to heterogeneity (I^2) was 0%. Both studies fell below quality threshold, therefore no sensitivity analysis was performed. No further meta-regression was undertaken due to the low number of studies in this comparison.

3.9 Summary of meta-analysis findings

Behavioural therapies for depression have a large effect size difference compared to control interventions at post-treatment. When compared to CT/CBT there appears to be no difference in effect size at post-treatment or follow up. Behavioural therapies also appear to have a medium to large effect when compared to brief psychodynamic therapies and supportive

counselling; however the number of studies is small. A summary of the main statistical findings is presented in Table 9.

Table 9: Results of meta Analysis of randomised controlled trials of behavioural activation for depression ([a] indicates odds ratio)

Comparison	Number of studies	Number subjects	Standardised mean difference	95% CI	P	I^2
BT vs. Control/TAU						
Symptom level	12	459	−0.70	−1.00 to −0.39	<0.001	55.1%
Dropout[a]	3	119	0.58	0.28 to 1.20	0.86	0%
Recovery rate[a]	3	167	4.18	1.14 to 15.28	0.03	52.6%
BT vs. CT/CBT						
Symptom level Post-treatment	12	476	0.08	−0.14 to 0.30	0.46	21.1%
Symptom level Follow Up	8	271	0.25	−0.21 to 0.70	0.28	60.2%
Dropout[a]	8	436	1.17	0.57 to 2.41	0.67	32.4%
Recovery rate[a]	5	346	0.92	0.59 to 1.44	0.94	0%

BT vs. Brief Psychotherapy						
Symptom level post-treatment	3	166	−0.56	−1.0 to −0.12	0.01	43.4%
Symptom level follow-up	2	96	−0.50	−0.90 to −0.09	0.02	0%
Dropout[a]	3	166	0.94	0.22 to 3.96	0.11	54.1%
Recovery rate[a]	3	164	2.37	1.23 to 4.57	0.01	0%
BT vs. Supportive Therapy						
Symptom Level Post-treatment	2	45	−0.75	–1.37 to −0.14	0.02	0%

3.10 Studies not included in meta-analysis

It was not possible to include study numbers 17-20 in the meta-analysis due to insufficient statistical reporting. Study 18 (Padfield 1976) reported an empathy-based counselling approach compared to the same intervention with the addition of activity scheduling. A total of 24 women were recruited and randomised into two groups of 12. The study scored two in the quality assessment scale, indicating a number of methodological concerns. Results were excluded from meta-analysis, as data presented only supplied information on the mean (SD) change on the self-rated depression symptom measure. No raw scores could be extracted for baseline or post-treatment analysis. With the limited statistical and methodological approach outlined, findings were inconclusive. Study 19 (Zeiss et al. 1979) reported a comparison of a social skills-based, pleasant event-based and cognitive-based interventions. A total of 66 participants were randomised to one of the three conditions. The study scored two in the quality assessment. Results were excluded from the meta-analysis as only means were presented for the depression symptom level tool. While in other cases this had been managed by imputation methods following discussion with the review steering group it was not felt to be suitable for this study. The depression measure in use was a subscale from the Minnesota Multiphasic Personality Inventory (Hathaway and McKinley 1943). We were not confident about the imputation approaches relating to this measure, as no other studies in this review or other relevant populations could be found. Results showed a general improvement across all interventions reflective of meta-analysis results. Study 20 (Gardner and Oei 1981) reported a comparison of a basic activity scheduling approach with a basic cognitive restructuring intervention; 16 subjects were randomised. The study scored one on the quality assessment scale. Results were excluded from the meta-analysis as no raw data were presented. The Beck Depression Inventory was used as the main symptom measure, presented in a graphical format with no reference to specific mean and standard deviation. Therefore it was not possible to extract relevant

information from the document. Findings were consistent with the review that no significant difference was identified between the interventions.

It was felt after discussion within the review team that the chance of retrieving relevant information through the finding and contacting of study authors was low due to the age of the source material. As the three studies were relatively small and of low quality, and were unlikely to significantly impact results the decision was made to exclude them from the meta-analysis and include them in the narrative of the review only.

3.11 Evaluation of review against recommended standards

This meta-analysis was designed and conducted in 2006 and 2007. At that time the recommended conduct and reporting of reviews was based upon the QUOROM (quality of reporting of meta analysis) guidelines (Moher et al. 1999). This review was designed taking into account that guidance.
In recent years the science of systematic reviews and meta-analysis has advanced, resulting in updated recommendations. The preferred reporting items for systematic reviews and meta analysis (PRISMA) (Moher et al. 2009) has replaced QUOROM taking into account conceptual and practical advances. This therefore provides a standard against which this current review and meta-analysis can be measured. In Table 10 this review is considered against all 27 factors.

The review meets the majority of standards set by PRISMA. A protocol was not logged prior to commencing the study, which would be advisable as it prevents subsequent adjustments which cannot be traced by the reader. In addition a record was not kept of numbers of duplicates removed during the merging of search strategy results, which would have added to study selection transparency. In addition details of the rejected studies at both review of abstract and full report were not logged. It is not possible to return to these at the present time, but the information will be incorporated in future reviews.

The clarity of selection and analysis processes allows for a transparent approach to the review question which could be replicated. It has led to a clearer understanding of the current knowledge of the efficacy of BT with depressed adults, and has indicated areas for further research which will be followed in the subsequent chapters of this thesis.

Table 10: Review measured against PRISMA standards based upon chapter and publication (Ekers et al. 2008)

Section/topic	#	*Checklist item*	*Reported*
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Yes in title of chapter
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Yes in publication (Ekers et al. 2008)
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Yes introduction
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	Yes Introduction
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., web address), and, if available, provide registration information including registration number.	No registration of protocol
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	All included and expanded in chapter methods section
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and	Detailed description of search and dates

		date last searched.	conducted included
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Search summary used examples given in appendices (online for publication)
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Yes
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Yes reported in chapter under data extraction-limited reporting of this in publication
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Yes in methods section
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Details description of approach to study level bias in quality assessment and publication bias included
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	Yes SMD and odds ratio
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	Yes detailed description and rationale provided
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Yes for publication bias only

Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Details of sensitivity analysis and meta-regression provided with as pre-specified
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Yes, however no reporting of studies excluded by elimination of duplicates
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Yes table 6
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).	Yes in study quality assessment and where possible in publication bias
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.	yes
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	yes
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Yes publication bias
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	yes

DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).	Yes limited in publication expanded in chapter
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).	Yes
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Yes –related to other similar studies and further research based upon findings recommended
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Not specified as no funding received- would have been beneficial in publication to express with more clarity

3.12 Discussion of review findings

3.12.1 Main findings

Review question*:*

What is the relative effectiveness of individual behavioural psychotherapy on symptom reduction, recovery and dropout when compared with usual care, inert controls and other psychosocial treatments in depressed adults?

In this review a detailed systematic review of all published and unpublished research examining the efficacy of behavioural therapies for depression was conducted. In all 20 studies (including 1215 participants) were identified, 17 of which (with 1109 participants) were able to be included in meta-analysis.

Comparison 1. BA vs. controls

BT was found to be an effective treatment for depression in adults. Meta-analysis of 12 studies (including 459 participants) demonstrated superior outcomes to controls with a large effect size. BT also achieved improved recovery rates, with no difference in dropout. This indicates that, while BT is an active treatment, it would appear to be acceptable to those receiving it.

Comparison 2. BT vs. CBT

BT was found to be as effective as CBT in the treatment of depressed adults. Meta-analysis of 12 studies, including 476 participants, demonstrated that BT provided equivalent results, with no statistically significant differences in depression symptom level at post-treatment and follow-up. BT also demonstrated equivalence in recovery rate and dropout at post-treatment.

Comparison 3. BT vs. brief psychotherapy

BT was found to be superior to brief psychotherapy in the treatment of depressed adults. Three studies were identified, including 166 participants, with BT showing advantage in symptom level and recovery post-intervention. The effect size in this comparison was large and statistically significant despite low statistical power.

Comparison 4. BT vs. supportive counselling

BT was found to be superior to supportive counselling in the treatment of depressed adults. Two studies were identified, including 45 subjects, with BT showing advantage in symptom level post-intervention. The effect size was large in this comparison and statistically significant despite low statistical power.

These findings indicate that BT is an effective when delivered as a 'single strand' treatment. The addition of other components, such as cognitive restructuring, to this intervention does not provide additional benefits post-treatment or at follow-up. BT is as effective as CBT when all randomised trials are combined in a meta-analysis. It also would appear to provide improved clinical outcomes over other forms of psychotherapy and counselling, although caution is recommended regarding this finding due to the low numbers of studies.

3.12.2 Impact of study variables

The studies in this meta-analysis varied considerably in design and delivery approaches. It was possible to utilise this variability to explore factors relating to magnitude of effectiveness where sufficient studies were available (BT vs. Control and BT vs. CBT). Meta-regression analysis makes observational associations between study variables and effect. It is exploratory in nature, and as such loses the power of causal inference (Higgins and Thompson 2004). This approach provides a viable and efficient method of considering the impact of such study level variables on outcome. It increases the depth of the analysis, and highlights potential issues that may require further research. While caution is needed in such

interpretation, as associations in no way suggest definitive explanations, the alternative of planning large-scale prospective trials with many arms is costly, time-consuming and unrealistic in most cases. Therefore the meta-analysis incorporating regression provides a timely and informative association without such constraints, and at least provides quantitative data around which subsequent research questions may be built.

Meta-regression BA vs. controls.

The meta-regression found that, when compared to controls, baseline severity, length of treatment and level of qualification were not related to the overall effect size of BT. Such findings may indicate that BT is a durable approach that can be delivered in a number of ways. Therapist level is of particular interest here. The suitability of BA for wider dissemination (Jacobson et al. 1996) requires further exploration, as we found that no studies used therapists below the level of doctorate psychologists.

Meta-regression BT vs. CBT

The meta-regression found that, when BT was compared to CBT, length of treatment and therapist level were not related to the overall effect size. A statistically significant association between higher baseline severity and a larger effect size in favour of BT compared to CBT was found. This was the only significant association observed, and raises interesting clinical considerations. It may be that BT does not rely on high levels of cognitive engagement to mediate benefit. The benefit, according to the theory, will emerge when the person has re-connected with contextual relationships that provide positive reinforcement. CBT requires people to identify and challenge thinking styles and beliefs. In severe depression cognitive functioning deteriorates, hence engagement with 'thinking' may be increasingly difficult as severity increases. It may be that this that accounts for the association found in this study; however further review and research would be essential to explore this finding, with relevant approaches being

defined *a priori*. Further conclusions based upon this finding would be unwise and are beyond the scope of this review.

Summary of findings from meta-regression

Following this review it remains unclear what is the optimum delivery approach of behavioural therapy for depression. This review is the first to explore a number of pre-defined factors and their association with effect size. Further exploration is required in relation to ideal treatment duration and improving outcome in more severely depressed patients using this intervention. No association was identified between therapist level and outcomes of BT. Few studies in the review described the training and supervision therapists received in depth making detailed analysis of this factor difficult. It was possible to extract core training levels, and findings showed that all studies used relatively expensive senior grade staff to deliver a potentially simple treatment. The simple nature of BT has for some time suggested its suitability for wider training and dissemination (Jacobson et al. 1996) however no empirical support for this assertion was found in this review.

3.12.3 Strengths and limitations of this meta-analysis

This meta-analysis complements and concurs with other publications that include behavioural interventions as part of wider CBT reviews (Churchill et al. 2001, Gloaguen et al. 1998, Dobson 1989). In addition, two new reviews also focussing on BT as the primary intervention, which were published shortly before and after this work, came to similar conclusions: that BT is an effective treatment for depression (Cuijpers et al. 2007, Mazzucchelli et al. 2009). In contrast to these reviews, we chose to focus on individual rather than group interventions, and included dropout and recovery rate analyses. More studies were included in the review reported in this chapter than in previous reviews (apart from (Mazzucchelli et al. 2009), due to broader inclusion criteria and the inclusion of recent and unpublished data. Studies drew patients of varying types from a range of settings such as inpatient, psychiatric outpatient and volunteer cohorts in

adult, older adult and student settings. Interventions varied considerably across studies, from supported self-help using minimal therapist contact to full psychotherapy. The quality of included trials varied considerably, with some of low quality delivering results that deviated considerably from the overall picture (4, 8). It was attempted to account for this to by the use of sensitivity analysis, random effects modelling and meta-regression of *a priori* variables. Interpretation of the results must be made with such factors in mind. Caution must also be exercised in interpreting the comparisons of behavioural interventions with brief psychotherapy and supportive therapy, due to the low numbers of studies and/or small sample sizes that informed such findings.

The searches conducted identified a number of studies directly comparing BT with drug therapy; this was not included as an *a priori* comparator, and as such was outside the scope of this review. It might have been possible to incorporate the comparator when this omission became apparent. It was, however, felt that inclusion of relevant terms could have been missed in searches, hence it was not advised, as being prone to bias. It is clear, however, that such a comparison would be a useful addition in any future review. In the studies identified it appeared that BT produced at least equivalent results to drug therapy. A more formal protocol-based review would be needed to confirm these observations to reduce the chance of omitting relevant studies.

3.13 Implications for practice and future research

Of particular interest is the observed equivalence between BT and the more complex CBT/CT interventions recommended as the primary psychological intervention for depression (National Institute of Clinical Excellence 2009). In addition to similar levels of symptom improvement, we observed no difference in recovery or dropout. These combined findings indicate that behavioural interventions are as effective and acceptable as CBT/CT. Such findings partially endorse the BA parsimony hypothesis advanced by Jacobson and colleagues (Jacobson 1996, Jacobson

et al 2001). They question the utility of adding 'complex' cognitive techniques to simpler behavioural interventions to improve clinical outcome. One of the attractions of behavioural interventions is simplicity and dissemination, thus assisting the current scarcity of therapists and overwhelming demands outlined in Chapter One. There was no direct evidence identified in this review to support this assumption as no studies used non-psychology or psychotherapy-trained individuals to deliver BT. It may be that the effectiveness of BT is due to the therapist's skill and experience rather than the specific factors of the intervention. The impact of level of training of those who had delivered BT was examined, and no association was found between 'higher level' of qualifications and effect size. This may support the assertion that BT is be suitable for delivery by non-specialists with shorter training; however further research is clearly needed.

Another shortfall in the evidence base is the lack of economic analysis. While clinical effectiveness is a necessary consideration it is not wholly sufficient. New interventions must also be seen to provide value for money. While there is evidence that BT is effective the cost of this effectiveness has not been benchmarked against accepted threshold values, such as the £20,000 per Quality Adjusted Life Year (QALY) recommended by the National institute of Health and Clinical Excellence (National Institute for Health and Clinical Excellence 2008).

Taking into account the need to widen the availability of effective psychological interventions for depression outlined in Chapter One of this thesis, further research is required to explore these factors. Of particular note is the need to test the suitability for dissemination of BA and to begin to estimate the potential value for money that single-strand behavioural therapies might offer.

3.14 Summary of chapter 3

In summary, behavioural therapy for depression is an effective intervention that has outcomes comparable with, if not better than, those of alternative and currently recommended therapies. This review adds substantially to the literature regarding behavioural therapy for depression, as it provides a broad overview of the current evidence, reports data on recovery and dropout, and explores the effect of baseline co-variants in relation to depression symptom change. It is recommended that further research be undertaken into the clinical and cost-effectiveness of behavioural treatments of depression, in particular of Jacobson's (1996) parsimony leading to suitability to disseminate hypothesis, where the intervention is delivered by non-specialist mental health staff with specific training rather than by highly specialised psychotherapists.

Chapter Four: Randomised controlled trial of the clinical and cost effectiveness of Behavioural Activation for depression delivered by non specialist mental health workers

4.1 Introduction

4.1.1 Background to study

As outlined in Chapter One, depression will be the second largest cause of disease burden by 2020 (World Health Organisation 2001), affecting between 5-10% of the population and is the third most common reason for primary care consultation (Singleton et al. 2001). It is associated with significant distress, impairment of functioning, disturbance to interpersonal relationships and an increased risk of suicide (Hirschfeld et al. 1997). Psychological treatments, particularly cognitive behavioural therapy (CBT), are recommended to treat depression (National Institute of Clinical Excellence 2009, Hollon et al. 2002b); however less than 10% overall of those affected receive such treatment (McManus et al. 2009). Despite recent English investment in 'Improving Access to Psychological Therapies' (IAPT) services (Clark et al. 2009), the limited availability of suitably trained therapists remains a significant brake on patient access to treatment. CBT remains the standard approach to depression (National Institute of Clinical Excellence 2009), commonly based upon Beckian Cognitive Therapy (Beck 1976). This incorporates both behavioural and cognitive strategies to identify, question and modify maladaptive thought processes, life rules and core beliefs. This is in contrast to single strand behavioural interventions, as outlined in Chapter Two, that use an operant conditioning model to develop a structured daily action plan. Although CBT is recommended as a main treatment option in the treatment of depression as seen in Chapter Three of this thesis, it is unclear whether the cognitive strategies are a necessary component of the intervention. As outlined in Chapter Two the landmark study in 1996 by Jacobson et al. (Jacobson et al. 1996), comparing the full version of CBT with a reduced version including some cognitive techniques and a third intervention behavioural activation (BA), led the way to this research. There was no evidence of any differences in effectiveness between treatments at post-treatment or follow-up (Gortner et al. 1998) which led Jacobson (Jacobson et al. 1996, Jacobson and Gortner 2000) to put forward a parsimony argument in favour of BA: if CBT and behavioural interventions are

equally effective, then behavioural ones may be preferable because they are simpler and can therefore be delivered more economically by professionals with less training. Were this to be the case it may have substantial implications for the organisation and delivery of treatments (Jacobson and Gortner 2000). This is a somewhat provocative finding albeit one supported by the meta-analysis outlined in Chapter Three of this thesis. It indicates that much of what occurs in this leading psychological treatment for depression may be an unnecessary complication. A series of meta-analyses including the one outlined in this thesis (Cuijpers et al. 2007, Ekers et al. 2008, Mazzucchelli et al. 2009) are now published all reaching the same conclusion. Although a main impetus for the renewed interest in BA is the possibility of developing a simple and effective treatment for depression, it is notable that no study reviewed in the meta-analyses has explored this approach in a controlled clinical trial. As seen in Chapter Three in earlier studies the treatments were delivered by clinicians with previous experience of delivering therapy often amounting to several years. While behavioural activation may indeed by simpler to deliver, it may be the experience of the therapist that counts. If this were to be the case despite BA offering a more simple intervention it would still rely of the same set of experienced and highly trained therapists to deliver it effectively. This while possibly broadening the skills of such therapists, would be unlikely to markedly improve accessibility or value for money. The aim of this current study is to examine whether non-specialist mental health workers, without previous experience in therapy, can deliver effective behavioural interventions and if so estimate the cost utility of such an approach. This research is based upon findings of the meta-analysis reported in Chapter Three and is the first study to directly assess the parsimony leading to suitability for dissemination argument offered in favour of BA.

4.2 Research Questions

As outlined in Chapter Three the development of a clear question is essential in the research process. As in systematic reviews having a

focussed question to inform study design is essential. Again this should be based upon the PICO (participants, interventions, comparisons, outcome) format.

Our two research questions for this study are:

1. What is the impact on depression symptom level, functioning and treatment satisfaction of BA delivered by generic mental health workers to depressed adults compared to usual primary care management?

2. What is the estimate of cost utility of BA delivered by generic mental health workers to depressed adults?

These questions mapped against the PICO structure below:

- **P**articipants-depressed adults;
- **I**nterventions-BA delivered by generic mental health workers;
- **C**omparisons-BA and usual care;
- **O**utcomes-depression symptom level, functioning, service use/cost and health state.

4.3 Method

4.3.1 Choice of study design

Depression is a disorder that is likely to wax and wane over time (Kennedy et al. 2004) and it is also likely that participants would be at a significant point of severity in their depression at baseline which then led to seeking help and therefore are likely to 'regress to the mean'. Another potential confounder is that those accepting of treatment and in particular entering a research trial may be an over representation of motivated individuals who are potentially more likely to improve. Therefore any study design must control for such confounders as far as possible giving a reasonable assumption the intervention is accounting for any differences observed. The randomised control trial (RCT) is designed for this purpose; it provides

an experimental design to decide the relative effectiveness of two or more treatment groups. The use of the randomisation balances confounders between groups, reduces the potential to come up with an incorrect finding and controls for numerous types of bias (Everitt and Wessley 2008, Altman and Bland 1999b). The RCT is based upon three main components:

- A comparison of a group of participants who have been given the intervention under investigation with a group of participants who have been given a standard treatment or inert treatment.

- An approach that assigns people to each intervention under investigation in an equal manner.

- An approach to assessing effectiveness is used. Commonly a number of outcomes will be considered.

There remains some debate about when the first randomised trial was conducted in medicine but consensus tends to cite a streptomycin trial in the treatment of tuberculosis reported in 1948 (Doll 1998). The idea may however have been considered some centuries before when a Flemish physician named Jean Baptiste van Helmot advocated casting lots to decide who received the intervention of bloodletting. The outcome was to be the number of funerals in each group. No physician appeared to have accepted the challenge, which is unfortunate as it might have sowed the seeds of effective research design and prevented numerous interventions of questionable benefit being continued over many years (Chalmers 2001). It would appear that the first randomised controlled trial in psychiatry was in the mid-1950s (Healy 1997). Since then the RCT has become the accepted standard to measure clinical effectiveness and is commonly used across all areas of medicine including psychological treatment of depression (National Institute of Clinical Excellence 2009). As such the RCT was considered the suitable study design for our research question.

The study design used in this research is parallel treatment group two arm randomised controlled trial of BA facilitated by non-specialists compared to usual care for adult participants with depression.

4.3.2 Ethical issues

While randomisation provides a solution to the problems of selection bias it poses possible ethical dilemmas that require careful consideration. The trial has to consider the overall benefit of potential findings against the individual benefit for participants. This is a difficult balance in an RCT. Individual considerations relate to the need for a person to receive treatment that is likely to offer benefit (certainly to do no harm) against the need of collective consideration of evaluating competing therapies for future use. This delicate balance can also be influenced by clinicians pre-determined views. Clinicians should always act in the best interest of the patient even during clinical trials, as stated in the Declaration of Helsinki (World Medical Association 2008). The use of chance in the determination of which intervention is delivered can prove problematic for some clinicians, leading them to refuse to suggest that their patients participate in studies. Clinicians' unwavering belief in their own experience being all-important has been the subject of some debate and humour, being termed 'eminence based medicine' (Isaacs and Fitzgerald 1999), listing the following traits:

> "The more senior the colleague, the less importance he or she placed on the need for anything as mundane as evidence. Experience, it seems, is worth any amount of evidence. These colleagues have a touching faith in clinical experience, which has been defined as 'making the same mistakes with increasing confidence over an impressive number of years'. The eminent physician's white hair and balding pate are called the "halo" effect."

Clearly the trial has to counter these issues by ensuring that it is delivered in a way that is scientific and beneficial. To do so the trial must (Everitt and Wessley 2008):

- Ensure it is based in science to produce meaningful results to justify the need for the trial.
- Use a control arm that is reflects current accepted practice where available, and at worst be designed to do no harm.
- Be subject to scrutiny prior to recruiting participants.
- Provide potential participants with good information on which to base their decision to participate.
- Receive informed consent from all that take part.
- Follows the protocol and report any deviation in compliance.

These factors will be considered in more detail in relation to the current trial in the section below.

Ethical process and considerations in the current trial

The purpose of this trial was to examine if BA could be delivered by non-specialists and thus possibly provide another evidence based alternative in the psychological treatment of depression. In order to ensure a scientific robustness the trial protocol developed by DE was placed on the International Standard Randomised Controlled Trials Registry and allocated number (ISRCTN27045243). The protocol was then sent for independent scrutiny by DE. This allowed the scientific rigour to be assessed. The trial was reviewed by Professor M Bland (MB) to review the statistical methods. Dr Chris Williams a clinical expert in CBT who is a previous chair of the British Association of Behavioural and Cognitive Psychotherapists and researcher in CBT to review trial design and relevance and Professor David Richards. Their reports can be found in Appendix II. A trial steering committee was established as recommended (Medical Research Council 1998) with David Ekers (DE), Professor Simon Gilbody (SG), Professor Christine Godfrey (CG) and Professor David Richards (DR). Additional statistical rigour was provided by Professor Martin Bland (MB) who advised on statistical approaches incorporated in the protocols. The steering group approved the study protocol and received bi-monthly updates of trial progress by email.

The choice of comparator was a central ethical consideration for this study. While we had substantial evidence from our meta-analysis that BA was effective we had not seen it delivered by non-specialists. The choices considered were a waiting list control, usual care or the current main recommended psychological treatment for depression, CBT (National Institute of Clinical Excellence 2009). It was felt that a waiting list control had ethical problems. This would involve no treatment over a period of three months which was likely to result in untreated distress and disability. As there were treatment options commonly in use for depression that were likely to have benefit over no treatment, it was felt this option was unethical. The choice of face-to-face CBT would have provided a high quality active comparator. It was considered for the current trial but was felt to be beyond the resources available to this unfunded study, as availability of CBT was very limited at the time of the trial delivery. So while this comparator would have provided a robust alternative to non-specialist BA, offering CBT was not a realistic possibility or reflective of the current care most people received for depression. Usual care as offered in primary care services provided our third option. It appeared feasible and is the most commonly received treatment for depression (Office for National Statistics 2000, McManus et al. 2009). This approach could also provide cost comparisons based upon total service use in each arm with the BA costs added in the intervention arm. As this was a feasibility study regarding non-specialist delivery of BA long-term follow up was not required and we could also offer those randomised to usual care BA after three months. Based upon these factors usual GP care was adopted as the comparator.

The study developed patient information sheets and consent forms based upon good practice guidance (National Research Ethics Service 2007). I sought consent firstly to contact potential participants; this was completed by either the general practice staff or the primary care mental health team. If agreed, we then sought further consent to enter the trial prior to conducting any assessments. Potential participants were provided with

information sheets at each stage and contact details for further information. The information sheet provided study information and also details on what would happen if they consented, declined or withdrew from the study. Copies of the information sheet and consent forms can be found in Appendix II.

The study was then presented to the Northumberland local research ethics committee, the University of York research ethics committee and local NHS research governance departments. All approved this study prior to commencement. Approval letters from the above can be can be found in Appendix II.

Therefore in relation to the scientific and beneficial criteria outlined above:

- The trial was based upon a scientific design and had been reviewed by experts in therapy for depression, trial design and statistics.
- It used the standard treatment received for depression as the comparator.
- It was scrutinised by independent experts and two research ethics departments.
- It provided information based on best practice.
- It received consent prior to contact and prior to entry to trial.
- It placed the protocol on a public international database and had a steering group monitoring progress.

4.3.3 Recruitment

Potential participants aged 18 or over were recruited from either general practice directly or from primary care mental health services over a nine month period. Practices were based in a mix of rural and urban settings and following receiving information about the trial requested to participate. Recruitment often poses problems in clinical trials in primary care, even in conditions with high prevalence such as depression (Hunt et al. 2001). This may be due to a desire to protect the patient, lack of confidence in research practise by GPs and prioritising clinical administration over research

(Mason et al. 2007). Several recommended methods were used to enhance recruitment to the trial (Bower et al. 2009, Ward et al. 1999, Everitt and Wessley 2008). Meetings were arranged in each practice/primary care mental health team with DE to discuss the trial and the details of interventions, process and outcomes. Key partners were identified in each practice/team and DE maintained contact with them for the duration of the study, sending regular email updates of progress and benchmarking their contribution to recruitment against the average. A small number of practices/teams were identified initially to pilot recruitment with an additional set to add if targets were not achieved. Positive feedback was received regarding planned use of the trial as formal psychological therapy for depression at the time of the study had a wait time of well over six months. This process was felt to incorporate the major factors identified for supporting recruitment and piloting processes within the constraints of an unfunded feasibility study.

4.3.4 Participants

Participants identified by a general practitioner or primary care mental health worker were supplied with the study information sheet and asked to consider inclusion in the trial. Participants were required to have been on no antidepressant medication or a stable dose for six weeks prior to inclusion. This ensured that no crossover effect from treatment adjustment would be recorded within the trial and attributed to interventions. Following consent, eligibility was confirmed by the use of a standardised computer-based assessment tool, the Clinical Interview Schedule Revised (CSIR)(Lewis et al. 1992). Confirmation of a diagnosis and hence eligibility is an important quality measure of a trial therefore a reliable tool was required (Schulz et al. 2010). The CSIR had been extensively used for this purpose in large scale surveys in the UK population and hence was considered suitable (Singleton et al. 2001, McManus et al. 2009). Exclusion criteria were those commonly adopted in randomised trials of depression in primary care. They were aimed to limit contamination of results with factors not directly attributable to the intervention under

investigation. Exclusion criteria consisted of suicidal risk, psychotic symptoms, diagnosis of bipolar disorder, organic brain disease or the use alcohol/non-prescription drugs requiring clinical intervention. If suitable, written consent was obtained and baseline measures (outlined below) were taken prior to randomisation.

4.3.5 Allocation concealment and randomisation

While the RCT design improves objectiveness in trials there are conditions that are fundamental in the protection against bias. Allocation concealment is central to this and the degree to which it is achieved has been shown to influence trial outcome (Jadad et al. 1996). Concealment prevents any interference by researchers with the assignment of particular participants to a treatment condition and hence limits allocation bias. Clinicians will often hold a view as to which intervention is preferable and will wish for a patient to get the intervention they believe is the best (Schulz and Grimes 2002) an issue that has been discussed in the running of previous trials of BA (Jacobson and Gortner 2000). If the sequence of assignment is altered this will undermine the trial findings as random allocation can no longer be assumed and there is a risk that selected participants received the intervention under investigation. It is therefore essential that the trial design placed great emphasis of concealment of allocation.

Following assessment, if suitable for inclusion, participants in this trial were randomly allocated to two arms. A block randomisation system was used in blocks of four to enable a close balance between the numbers allocated to each group at any given time. In a small trial this is helpful due to the potential impact on resource if a simple randomisation sequence results in a high number of allocations sequentially to a single arm. It is also possible in a small trial to generate a randomisation sequence that would be weighted towards a particular arm. Block randomisation ensures a matched number within blocks that are a multiple of the number of arms in the trial and then randomly varies the sequence over consecutive blocks (Altman and Bland 1999a).

Baseline depression severity has been seen to be associated with differential effects of psychological treatments of depression (Elkin et al. 1995). With the small sample sizes in this study there was an increased risk of unequal distribution of depression severity levels across each treatment arm post randomisation. This was countered by stratifying participants by baseline depression severity prior to randomisation an acceptable approach where such imbalance could impact on interpretation of study results (Altman and Bland 1999a). Participants were allocated into two groups prior to randomisation based upon Beck Depression Inventory (BDI-II)(Beck et al. 1996) scores using the midpoint of the moderate category ($\leq$25 & $\geq$26).

Allocation bias was controlled by an allocation concealment process independent of the study team. Randomisation lists for high and low severity were generated independently of the study team by MB using the computer statistical software Clinstat (Bland 2004). These lists were supplied directly to an administration team within Tees Esk and Wear Valleys NHS Foundation Trust who were not involved in the running of or referral to the study. Lists were held in a locked location and following assessment and stratification the study team contacted and requested the next allocation code assigned. At no point did the study team have any influence on list development, communication of the list to those providing the allocation service nor any involvement in the allocation of participants to treatment arm. An ideal system of allocation is commissioning of a complete allocation service but this is often costly. As this trial had no external funding this procedure offered a suitable alternative maintaining standards of allocation concealment within study resources.

Following randomisation GPs and patients were informed of allocation automatically by letter from the study team.

4.3.6 Sample size

No previous study of BA delivered by non-specialists was available to give a precise estimate of the sample size required in this study. The present exploratory study would provide such information to future researchers however at the time of the study design such information was not available. It was decided by discussion with the trial steering group to calculate sample size based upon those studies from our meta-analysis that reflected the study design we were to use (Ekers et al. 2008). While it was accepted that this approach was not ideal, it incorporated and reflected the best available evidence at the time of planning. The studies used incorporated a range of BA approaches, and all were delivered by more senior therapists than the non-specialists being used in this study. There was a possibility that this would result in biasing sample size calculations on an over estimation of difference resulting in underestimation of numbers required if senior therapists achieve better results than non-specialists. This would potentially increase the chances of a type II error (accepting the null hypothesis when it is in fact false) in this study. In the absence of any more suitable alternative evidence calculations were based on those studies in our meta-analysis that incorporated a delayed start to psychological interventions with variable levels of concurrent usual care. It was decided this was preferable to not using a sample size calculation for the current trial.

Statistical advice on the calculation of sample size for this study was sought from MB. The current standard of 80% power at a 5% significance level was used as the basis for sample size calculations. This indicated the number of participants needed to have an 80% chance of finding a true difference between groups if with a 5% chance of a type I error. Sample size calculations were based upon standardised between-group effect size (Cohen's d) of −0.84 (CI −1.27 to −0.41) observed in a sample of nine studies including 282 subjects. Using the PS sample size calculator (Dupont and Plummer 2009) calculations based upon a continuous response variable (BDI-II) from independent control and experimental

subjects with 1 control(s) per experimental subject were used. Calculations assumed response within each subject group was normally distributed with standard deviation 1. If the true difference in the experimental and control means is 0.84 calculations indicated 23 experimental and 23 control subjects we required to be able to reject the null hypothesis with probability (power) 0.8 with a Type I error probability associated with this test of 0.05 (see Figure 31).

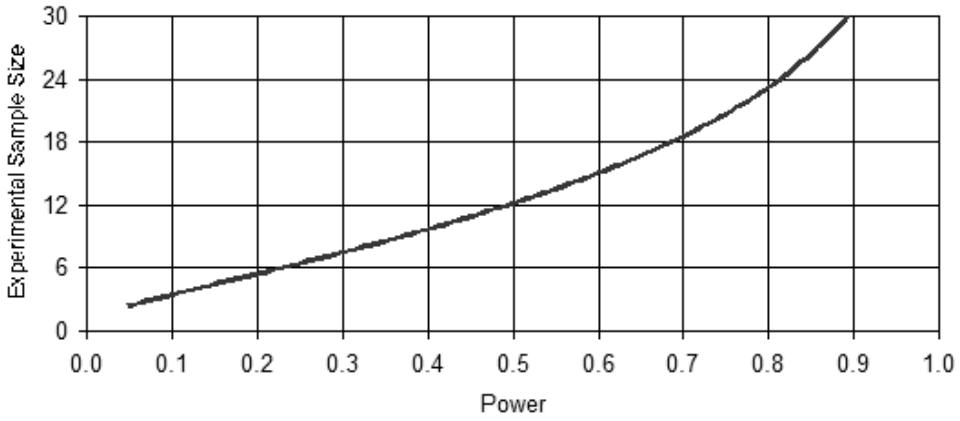


Figure 31: Sample size calculation

4.3.7 Measures

Depression symptom level

The primary clinical outcome of depression symptom level was assessed in this study using a commonly applied validated self report measure in psychotherapy studies the Beck Depression Inventory second edition (BDI-II) (Beck et al. 1996). In our meta-analysis the BDI was applied in 16 of the 17 studies of BA analysed. There is no clear correct position regarding the preference of self-rated over clinician-rated measurement tools for depression symptom level. The most frequently applied clinician rated tool found in the meta-analysis was the Hamilton Rating Scale for Depression (HRSD)(Hamilton 1960) which was used on its own in one study and combined with the BDI in 11 studies. It was decided to use the BDI-II in this study as it was the most frequently used measure identified in the meta-analysis and was practical to use with the limited resources available (no requirement for assessor training and less time needed for

assessments). The BDI-II has a score range of 0 to 63 (0-13 minimal, 14-19 mild, 20-28 moderate, 29-63 severe). It has strong concurrent validity (r= 0.71 between BDI-II and HRSD, r= 0.93 between BDI-II and BDI-I) and reliability (coefficient alpha 0.92) in an sample of 500 depressed outpatients (Beck et al. 1996).

Functioning

Functioning was measured using the Work and Social Adjustment Scale (WASA) (Munt et al. 2002). It consists of a five item scale measuring self rated impairment of functioning in relation to work, relationships, home management, social leisure and private leisure. Each item is scored on an eight point Likert scale (with anchors of 0 meaning no impairment and 8 meaning severe impairment). It has been tested for use with a depressed sample of outpatients, and has demonstrated strong reliability (Cronbach's alpha 0.80-0.94) and validity (correlation 0.76) in a sample of 380 depressed outpatients (Munt et al. 2002). The WASA also represents a commonly used measure of functioning adopted in the Improving Access to Psychological Therapies (IAPT) services across the UK. This gave an additional benefit of using this scale as results will be meaningful and familiar for the IAPT workforce.

Satisfaction

Satisfaction was measured using the Client Satisfaction Questionnaire (CSQ8) (Attkisson and Greenfield 2004). The CSQ8 is evaluates satisfaction over eight areas of service provision: The quality of the service; if it was what was wanted; if it met needs; if would recommend to others; satisfaction with amount of help; helped to deal with problems; overall satisfaction; if would use service again if needed. The CSQ8 gives a score range of 0-32. The CSQ8 was adopted for use in this study as it has a UK English version and, unusually for satisfaction questionnaires, has published validity and reliability studies related to it. Internal reliability as demonstrated across nine studies, with Cronbach's alpha ranging from 0.83-0.93 is very positive, and construct validity has been demonstrated with high correlations with other satisfaction instruments (Spearmans rank

correlation coefficient 0.6-0.8) (Attkisson and Greenfield 2004). It has been used previously in the UK in research settings (Lovell et al. 2006) and in multicentre trials of depression (see CADET trial ISRCTN 32829227).

Health State

General health state utility was assessed using the EQ-5D (EuroQol Group 1990, Brooks 1996). The EQ-5D provides a standard measure of general health status for use in clinical and economic evaluations. It is a tool that is suitable for use across a wide range of health conditions. The EQ-5D is the accepted approach to measuring health utility, and is appropriate for use in major depression in primary care settings (Sapin et al. 2004). Using the general health state measures such as the EQ-5D is increasingly recommended (Bosmans et al. 2008, Barton et al. 2009), as it allows mental health studies to be viewed in a decision making context alongside all fields of health care. The EQ-5D measures health across five dimensions of mobility, self care, usual activities, pain/discomfort and anxiety depression. Each state is divided into 3 levels of perceived problems self rated by the participant:

- Level 1 No problems
- Level 2 Some problems
- Level 3 Extreme problems.

This results in 243 possible health state values ranging from 11111 indicating no problem in each area to 33333 indicating severe problems in all areas. The EQ-5D is then converted to a single health state by applying a formula that assigns weights to the levels within each dimension resulting in a single summary score. Health state values were assigned from these scores using standard UK population values (Dolan P et al. 1995) for quality adjusted life years (QALY) between 0 (death) and 1 (full health). Specialist health economics guidance was received from Dr Steve Parrott (SP) who assigned health states using STATA (Stata Corporation 2003).

4.3.8 Costs

BA treatment costs

BA costs were calculated from therapists hourly pay rates (mid scale Agenda for Change band 5) for the year of the study plus 30% overhead costs. Overhead costs incorporate those services that serve many different departments in the health service and also management costs. In this study it was decided that overhead costs of therapist time were unlikely to make a substantial influence on the overall cost of the intervention. After consultation with health economists CG and SP and the finance department of Tees Esk and Wear Valleys NHS Foundation Trust where the trial was hosted it was decided to take the simple commonly adopted approach of assigning 30%. The effort put into overhead cost allocation should be related to the likely importance in relation to the overall analysis (Drummond et al. 2005). It was felt this was a balanced, acceptable and pragmatic way of addressing the issue of overhead costs in this study since it was unlikely there would be significant variation from the 30% value, used in the NHS hosting organisation and other cost utility studies of depression in primary care (Katon et al. 2005). In a small study it could be said variance would have a larger influence on final results. This study was considered exploratory and would include sensitivity analysis to inform potential future large definitive trials. Taking this into account, the extra resource needed to perform more detailed calculations of overhead costs was not considered necessary.

All twelve planned sessions were included regardless of attendance rates. This study was exploratory and likely to be underpowered to provide an accurate estimate of attendance. Calculating costs for only those sessions attended could therefore introduce potential bias into findings as it is unlikely observed attendance in this study would reliably translate to routine care settings. By including only those sessions attended BA costs may have been underestimated. It was therefore decided in consultation with health economists CG and SP to adopt the conservative approach of including the cost of all sessions regardless of actual attendance rate. This would provide the basis of a cost utility analysis of BA by the non specialist based upon a 12 session treatment programme and reduce the chances of underestimation.

BA training and supervision costs

Training and clinical supervision was delivered on local NHS trust premises by the DE and were considered a major factor in cost estimates. Training costs were calculated by facilitators hourly rate for the duration of the training (35 hours) plus 30% overhead cost. The hourly rate was obtained in consultation with the NHS trust's finance department. The figure obtained was then divided by the number of participants attending the training, in this case ten. BA is already an accepted intervention for depression, and as such training for such interventions as CBT, BA etc. is common within routine care. If our training costs were allocated only to the two therapists operating in the intervention arm of the trial this would have added unrealistic increased costs influencing overall economic evaluation. Based upon clinical judgement and the experience of the author a training group size of ten was considered appropriate. This group size would allow role play and feedback and be able to address any emerging issues. Those participants in the training were then asked to provide treatment to study participants finishing the usual care arm. Training costs do pose problems in economic evaluation and require balanced decisions on their management. Exploratory trials in particular, due to their small size, would carry an unrealistic cost burden of training if distribution of such costs were to be limited to the trial therapists and participants alone.

Supervision costs were based upon one hour per fortnight of supervision for each therapist and supervisor plus 30% overhead costs as outlined above.

Training and supervision costs were then adapted to reflect activity levels expected in NHS settings. As outlined above, training and supervision can burden costs in small studies in a way that limits translation into routine care. One purpose of cost-effectiveness analysis is to inform decision makers (Fenwick et al. 2001) therefore decisions are required as to how to handle training costs both in terms of per trainee and also the degree to which this influences cost per participant. Psychotherapy trials are commonly small, and if training were only assumed to have an impact on

those in the study, the distribution of such costs would be limited and hence add a substantial amount per participant. To counter this, the costs of training in this study were based upon a distribution expected outside the trial setting. This is a common approach used in the costing of trials of depression in primary care (Kendrick et al. 2006, McCrone et al. 2004) and accepted by decision makers (National Institute for Health and Clinical Excellence 2008). Through discussion between DE and experts in BA and training in UK psychological health service delivery training costs for the trial were based on an assumption that competences gained through training would be maintained over three years and be used with patients outside the trial setting. The exact number of treatments a therapist would deliver over a year was unclear. Two different estimates of therapist workload based upon local Improving Access to Psychological Therapy (IAPT) service specifications were used. This provided a routine health service approach that represented the geographical location of the study. It also provided a real-life estimate of anticipated demand in a psychological therapy service designed to increase access which is where the BA under investigation is likely to be placed. In the first estimate it was assumed each BA therapist would complete 65 treatments per year if employed in a depression-specific role (scenario A). The second scenario had a smaller estimate of the workload which was focussed specifically on depression of 33 treatments per year assuming that the therapist was operating in a common mental health service role treating both depression and anxiety (scenario B). Training and supervision costs were then distributed based upon each scenario over anticipated completed treatments for three years and added to direct BA treatment costs. This approach balanced a potential bias of overestimation of training costs in this study.

General health service resource use

As per recommendations all resource use from a health and personal social service perspective (PSS) were used for this study (National Institute for Health and Clinical Excellence 2008). As this was not intended to be a definitive study it was decided not to include cost savings that fell outside of these perspectives. Other studies have measured missed work days

(McCrone et al. 2004, Kendrick et al. 2006) allocating costs to this. There is however a methodological debate regarding this approach (McCrone et al. 2004), and results should be reported separately and not incorporated in QALY calculations in line with guidance issued on the estimation and perspectives on costs issued by NICE (National Institute for Health and Clinical Excellence 2008). Based upon these uncertainties it was decided in discussion with the trial steering group to focus on NHS and PSS costs in order to build a body of evidence usable by those designing larger scale definitive studies. Using a standard uncontroversial approach would facilitate clarity for this purpose. Direct intervention costs were calculated using total service use gathered from participants' primary care records augmented by self-report questionnaires and diaries. In the design of the cost-effectiveness protocol there was discussion regarding the method of collecting service use data. Use of self-report questionnaires and diaries is common in economic analysis; however members of the steering group who were experienced in this approach reported variable rates on reliability and completion. Based upon this it was decided to use a combination of self-report questionnaires, diaries and health service records in order to provide cross validation between methods. A questionnaire used in a previous trial of community alcohol treatment (Drummond et al. 2009) was modified for use in this study (see Appendix II). Particular alcohol related questions were adapted for depression service use and criminal justice sections removed as these were not deemed necessary for this population. The questionnaire was then taken to a local self-help group of people experiencing anxiety and depression and circulated. Verbal feedback was sought to ascertain if the questionnaire was acceptable. While it is acknowledged it would have been preferable to pilot the tool formally, this process was deemed to be an acceptable compromise and within the resources and time constraints of the study. Feedback indicated that the questionnaire would be appropriate and acceptable to use. Primary care records were used, as these were likely to include a comprehensive summary of all health care received from a variety of sources such as acute medical care and mental health care. Results from the two sources were combined and where discrepancies occurred the health service record was

used. Prioritising the health service records was agreed, as these were likely to have been recorded during or at about the time of an intervention rather than being a retrospective account based upon memory, and were therefore considered to be more reliable. The two approaches outlined were used to complement each other to supply a detailed picture of NHS and PSS service use for the six-month baseline and duration of this study.

Research assistants blinded to allocation collected baseline service use both retrospectively for six months prior to entry to the trial and for the three-month follow-up period using the above methods. They recorded all health service use including primary care, specialist services, inpatient stays and medication use, and any direct social service delivery including social worker, housing advisor, employment advisor etc. The baseline period used was beneficial to the trial due to the relatively short follow-up period. It allowed results to be balanced over a longer duration with baselines built into statistical analysis (see later under statistical methods). The total cost per patient was ascertained by multiplying each resource use by its relative cost. Costs were allocated in British pounds at 2009 rates, the year of the study. Unit costs of health and social care (Curtis 2009) and the British National Formulary (British Medical Association & Royal Pharmaceutical Society of Great Britain 2009) were used as the primary source of cost information. These sources are considered reliable and have been used in other trials of depression in UK primary care (Kendrick et al. 2006, McCrone et al. 2004). Where costs were not available from these sources we searched firstly from official listings (such as Department of Health) and then from other publicly available resources as per accepted guidance (National Institute for Health and Clinical Excellence 2008). Full details of costs adopted in this study are outlined in Table 11.

Table 11 Costs used in study

Services	Cost	Source and details
Acute medical care		
A and E visit	£126/93	Curtis (2009, p. 93) Accident and Emergency treatments (leading to admitted) £126. Accident and Emergency treatments (not leading to admitted) £93.
A and E mental health liaison services	£231 per episode	Curtis (2009, p. 93) based on national average.
Paramedic services	£240	Curtis (2009, p. 93): emergency transfers £240.
Rapid response service	£200	£200 per low-cost episode (includes assessment and travel costs.
Inpatient	£493/day	Curtis (2009, p. 93): non-elective inpatient stays (short stays) £493 national average
Outpatient	£126	Curtis (2009, p. 93) weighted average of all outpatient procedures national average £126.
Day case	£638/case	Curtis (2009, p. 93)
Primary care services		
GP	£35/ visit	Curtis (2009, p. 121): per surgery consultation lasting 11.7 minutes £35. Per telephone consultation lasting 7.1 minutes £21. Per home visit lasting 23.4 minutes (includes travel time) £117.
PN	£11/visit	Curtis (2009, p. 118): £11 per consultation; £20 per home visit.
Nurse advanced (includes lead specialist, clinical nurse specialist, senior specialist)	£16	Curtis (2009, p. 119): cost per hour of client contact £65; cost per surgery consultation £16
Counsellor	£42/visit	Curtis (2009, p. 68) £42 per hour of client contact.
Health visitor	£40	Curtis (2009, p. 115): £96 per hour of client contact; £81 per hour of clinic contact; £117 per hour spent on home visits; £40 per home visit.

Health Care Assistant	£4.49 based on 11.7 minutes contact time (as per GP)	Curtis (2009, p.117): unit costs available 2008/2009 £14 per hour; £23 per hour spent with a patient; £18 per hour in clinic contacts; £23 per hour spent on home visits; £9 per home visit.. Based on the median full-time equivalent basic salary for Agenda for Change Band 2 of the January-March 2009 NHS staff earnings estimates for unqualified nurses. Median full-time equivalent total earnings, which include basic salary plus hours-related pay, overtime, occupation payments, location payments and other payments including redundancy pay or payment of notice periods were £17,200.1 See page 178 for information on mean salaries.
Podiatrist	£11	Community chiropodist/podiatrist Curtis (2009, p. 108): £23 per hour; £21 per home visit; £11 per clinic visit.
Dietician	£34	Curtis (2009, p. 150): £34 per hour client contact; £59 per hour of home visiting.
Physiotherapist	£17	Community physiotherapist Curtis (2009, p. 105): £43 per hour of client contact; £48 per home visit; £17 per clinic visit. Hospital physiotherapist Curtis (2009, p. 147): £40 per hour of client contact; £52 per hour in home visiting.
Radiographer	£16	Curtis (2009, p. 151): £29 per hour; £48 per hour of client contact; £16 per 20-minute clinic visit.
Psychiatry services		
CPN appointment	£72	Curtis (2009, p. 136): £72 per hour of face-to-face contact
Psychiatrist appointment	£322/hour	Curtis (2009, p. 172): £322 per hour patient contact

Psychologist appointment	£75	Curtis (2009, p. 109): £75 per hour of client contact
GMHW	£32	Unit costs available 2008/2009 (costs including qualifications given in brackets): £24 (£28) per hour; £45 (£53) per hour of face-to-face contact; £32 (£37) per hour of client related work. Based on a Agenda for Change band 5.
Occupational therapist	£43	(Curtis 2009): NHS community occupational therapist £43 per hour of client contact; £47 per home visit; £17 per clinic visit. Hospital occupational therapist £44 per hour of client contact.
Drug and alcohol services	£90	Curtis (2009, p. 93): drug and alcohol services £90 national average.
Personal social services		
Social worker	£38	Curtis (2009, p. 126): £29 per hour; £38 per hour of client-related work; £140 per hour of face-to-face contact.
Debt advisor	£9.50	0.5 professional and 0.5 volunteer labour. Based on social worker Adult rate per hour of client-related work (£38 per hour) Curtis (2009, p. 126), 30 min. contact.
Benefits advisor	£19	Based on social worker dult rate per hour of client-related work (£38 per hour) Curtis (2009, p. 126): 30 min contact
Housing benefit advisor	£19	As above
Employment advisor	£19	As above
Other		
Alternative medical practitioner	£11.43/visit	Uprated to 2008/09 levels using the HCHS Pay and Prices Inflator.
NHS walk-in centre	£26	Uprated to 2008/09 levels using the HCHS Pay and Prices Inflator.

GP out of hours	£65	http://www.hsj.co.uk/news/primary-care/study-reveals-three-fold-cost-variation-in-gp-out-of-hours-services/5000264.article (Based on research from the research organisation Primary Care Foundation). Decision: worked out average.
NHS Direct	£25	http://www.telegraph.co.uk/health/3253245/Every-call-to-NHS-Direct-costs-25.html 2008
Post-Natal National Helpline	£2.35	http://www.nct.org.uk/info-centre/getting-help/helplines Accessed 12/04/2010 at 10.38am
999 call	£240	Curtis (2009, p. 93): emergency transfers £240
Phlebotomist	£9	Based on the median full-time equivalent basic salary for Agenda for Change and 2 Unit costs available 2008/2009 £14 per hour; £23 per hour spent with a patient; £18 per hour in clinic contacts; £23 per hour spent on home visits; £9 per home visit.
RELATE	£9.50	0.5 professional and 0.5 volunteer labour. Based on social worker Adult rate per hour of client-related work (£38 per hour) Curtis (2009, p. 126), 30 min contact
CAB	£9.50	As above

4.3.9 Blinding

Blinding is concerned with the minimisation of measurement bias in assessors and performance bias in participants and therapists. If achieved it ensures that investigators, patients and assessors are unaware of allocation and there can be no accusation of attempt to influence results. There are three possible levels of blinding in studies:

- Single blind-the patient is unaware of which treatment is received;
- Double blind-the patient and investigators (clinicians) are unaware of which treatment is received;
- Triple blind-the patient, investigators and assessors are unaware of which treatment is received.

Blinding poses particular problems in non-pharmacological trials such as those evaluating psychotherapy, as it would not be possible to disguise allocation to a particular intervention either from the patient or from the therapist (Boutron et al. 2007). In this study it was not realistic to suppose that we could blind participants between an allocation of usual care and BA; neither would we have been able to blind therapists. The accepted alternative commonly used in these scenarios is to use assessors who are unaware of allocation. All participants were asked not to disclose any details regarding the intervention received at the start of each appointment where measures were collected. Therefore all assessments were collected by a research worker blind to treatment allocation at each follow-up stage. Our use of self-report for clinical outcomes would also have reduced the risk of assessor bias; however it is possible that it increased the chance of performance bias by participants.

We used an additional process to reduce the possible impact of measurement bias by using two research assistants who were unaware of allocation to enter data into SPSS (SPSS for Windows 2008) to create two independent data sets, which were then checked for inconsistencies.

Initial statistical analysis was checked by MB who was blind to allocation. This was achieved by allocation of subject group codes with two randomly selected numbers. Initial analysis was then conducted on the main outcomes (BDI-II, WASA and CSQ) prior to unmasking code labels. While this was not possible for all subsequent analysis it provided a baseline which was independently analysed and blinded to allocation. This then was used as a reference point for additional analysis by investigators unblinded to allocation.

4.3.10 Interventions

Behavioural Activation

The theoretical basis for BA is outlined in detail in Chapter Two. In this study the aim was to test the dissemination of the approach when delivered in a way that was likely to be incorporated in routine NHS practice. An initial consideration was the duration of the intervention. While we required a protocol that included the key components, little evidence existed to inform us of the optimum number of sessions required. We endeavoured to address this through the use of meta-regression within our meta-analysis. The association between session number and effect size had been examined in our meta-analysis (Ekers et al. 2008) finding no association within the comparison of BA and control to guide decision making (meta-regression beta-coefficient = 0.03, 95% CI −0.03 to 0.09; I^2 = 0.49, P = 0.27). The number of sessions recommended in guidance for the approach is 16-20 sessions over 3-4 months (National Institute of Clinical Excellence 2009). In reality, however, even in Improving Access to Psychological Therapies services where therapies are designed to replicate those delivered in trials session numbers tend to fall far short of such recommendations (Clark et al. 2009, Glover et al. 2010). In these two large scale evaluations high-intensity treatments reported a wide variance in number of sessions received by service users with median numbers of therapy sessions 8.2 and 8.5 respectively. In a sample of 7,825 patients of IAPT services only 1.38% received 16 or more sessions (Glover et al.

2010). While such data had not been not published at the time of this trials protocol design, this pattern was emerging and influenced our intervention design. DR and DE had been involved in IAPT national rollout programmes and were privy to early audit data of sites. This, with additional substantial clinical experience, led to the agreement that a pragmatic approach was to have an intervention based upon session numbers falling short of those recommended in NICE guidance. A map of the components of BA was established and distributed across what was deemed realistic in one hour sessions leading to a 12-session protocol being developed (see Table 13, page 168). The protocol was based upon two behavioural approaches developed in previous research (Martell et al. 2001, Hopko et al. 2003). The BA in this study consisted of a structured programme aimed at participants increasing their contact with potentially antidepressant environmental positive reinforcement by scheduling and reducing the frequency of negatively reinforced avoidant behaviours. A shared formulation was created based upon a behavioural model in the early stages of treatment which was developed with the patient throughout the 12 sessions. Subsequent specific techniques incorporated in the 12 session protocol were self-monitoring, identifying 'depressed behaviours', developing alternative goal orientated behaviours and scheduling. In addition the role of avoidance and rumination was addressed through functional analysis and alternative responses developed. The overall goal of BA was to re-engage participants with stable and diverse sources of positive reinforcement from their environment (Kanter et al. 2009) and to enable them to understand the BA rationale, thus developing depression management strategies for future use.

Once developed, the BA treatment manual was shared with an international expert in behavioural activation (C Martell) who had been involved in two previous randomised controlled trials and was had written the most comprehensive book on the subject (Martell et al. 2001). His comments were incorporated into the manual which is presented in Appendix II.

BA therapists

This study was designed to examine the suitability for wider dissemination of BA. In order to test this realistically, it was important to consider the most appropriate staff group from which to recruit therapists. Traditionally therapy has been the domain of psychology or experienced professionals with substantial post-registration qualifications. This may be one factor in the hitherto limited availability and dissemination of effective therapies (Lovell and Richards 2000). This was reflected in our meta-analysis, where we were unable to find studies of BA delivered by non-specialists despite it being more than ten years since Jacobson had put forward the parsimony hypothesis (Jacobson et al. 1996, Ekers et al. 2008). Mental health nurses represent the largest section of the mental health workforce. They have the potential to deliver evidence-based interventions but require development to do so (Department of Health 2006). Based upon this it was decided therapists in this study would be drawn from the mental health nursing workforce. This would allow confidence in their level of baseline training and understanding of mental health. Also this workforce had been used in previous studies of psychological therapies in primary care (Kendrick et al. 2006). Two qualified mental health nurses with no previous formal psychotherapeutic training or experience were recruited to train and deliver the BA in the intervention arm of the study. This allowed a test of whether the brief training could develop sufficient skills, and would limit any potential bias of previous psychotherapy training influencing practise. Funding was supplied by Tees, Esk and Wear Valleys NHS Foundation Trust to employ the two nurses for one year. It was decided to use two full-time nurses rather than a greater number released for small periods of time to complete interventions. While there would have been potential benefit in being able to evaluate BA delivery by a wider number of therapists, the practical problems would have been considerable. Releasing staff for short time periods weekly to deliver treatments would have required considerable co-ordination of appointments and therapy rooms. Local team managers were approached; although they were interested in the study, they felt that releasing staff with minimal backfill funds would have a detrimental effect on the management of their services. It was therefore decided to recruit two nurses who would work

full-time on the trial and treat all the intervention participants. Based upon local service performance frameworks, it was thought that the numbers needed for the trial could be managed with this number of full-time therapists both for the intervention group and for those in the control group opting for treatment after the three-month study period. A potential bottleneck of participants requiring treatment midway through the study year was anticipated, as those in the control group would be entering treatment after three months alongside those entering the intervention group at the same time. It was decided to counter this by offering the training for the intervention to services on the understanding that staff released to train would treat 1-2 trial participants from the control arm. This was discussed with the trial steering group and deemed appropriate; as it offered a safety net of additional capacity if needed and also reflected normal practice (training groups for this type of intervention are not usually two members of staff).

Both trial therapists had worked in a range of services in inpatient and community settings, with three and six years' experience respectively since qualification. They were employed on Agenda for Change Band 5, the basic qualified nurse banding representing the largest section of the mental health nursing workforce. This was felt to be the most appropriate group from which to recruit, since if BA could be shown to be delivered effectively by basic grade nurses with no previous therapy experience, it would be likely that other more experienced nurses could achieve similar results.

BA training

BA training was delivered over five days in consecutive weeks and was delivered by DE. Previous experience indicated the use of 'massed practise' would be more beneficial than a day release course of longer duration (Ekers 2010). The approach also reflected the study time constraints: needing to get therapists trained and seeing patients swiftly after appointment. The course content sought to provide participants with a basic theoretical understanding of the principles and evidence base

underpinning BA. It also aimed to emphasise the clinical delivery of various BA interventions through repeated role play. The material was delivered using a range of learning methods including presentations and small and large group work, with a major component being intensive skills practice employing role-play with highly structured, constructive feedback (see Table 12). Training was structured around a typical course of treatment, initially outlining the theoretical basis and evidence base of BA. Assessment techniques were outlined which were aimed at gathering information on changes in the contingent relationships between the person and their environmental reinforcement. These were designed to facilitate the sharing of a BA rationale with the patient. This is an important skill in BA, which seeks to develop a collaborative approach, with the therapist acting as a coach to support the patient's goal attainment (Martell et al. 2010). The use of self-monitoring was covered, with participants encouraged to use it on themselves during the first two days of the course. This allowed self-reflection on the challenges presented by this that could be used within treatment to generate an empathic response when this stage of homework is or is not completed. Information gathered was used to develop a functional analysis of depressed behaviour using antecedent, behaviour and consequence (ABC). The training then explored the role of valued goal setting and scheduling in reintroducing positive reinforcement in participants' daily lives. Various procedures were employed to develop skills in collaboratively addressing negatively reinforced avoidance, such as Trigger, Response, Avoidance Pattern (TRAP) and Trigger, Response, Alternative Coping (TRAC) (Martell et al 2001). Problem-solving strategies were introduced, and then the training examined BA approaches to dealing with worrying thoughts or troubled thinking. Finally, relapse prevention planning was covered and the treatment protocol (following the same structure of the training) outlined. Participants were audio-taped for two role-play sessions (an initial assessment session and session four) to assess if a sufficient level of competency had been achieved. Training content was mapped against the 12-session protocol to ensure that all aspects in treatment were covered (see Table 13).

Table 12: BA course outline

	Day 1	Day 2	Day 3	Day 4	Day 5
Session 1	Introduction What is BA for depression? Theoretical underpinnings (reinforcement theory)	Use of self-monitoring to facilitate engagement with model Skills practice Recognising helpful and unhelpful behaviours	Using reduction in avoidance tools, skills practice continued Pulling it all together to this point Skills practice	Thinking in BA How to deal with thoughts Attention training	Whole case skills practice Competency assessment
Session 2	Evidence base for BA for depression (does it work?) A BA rationale for depression, application in practice, understanding the cycle Skills practice	Valued goal setting and its link to scheduling Activity scheduling Recap, putting it together so far and skills practice	Building in goal directed ACTION Skills practice, incorporating avoidance reduction in goal-directed scheduling	Understanding recognising and dealing with rumination Skills practice	
Lunch					

Session 3	Assessment and functional analysis. How to get ABC Skills practice	Assessing avoidance: its role and relationship to depression TRAP Skills practice using TRAP assessment	Problem solving seven stages Practice	Ending therapy and relapse prevention Building a relapse prevention plan	
Session 4	Skill practice continued Summary and discussion	Breaking avoidance, use of TRAC Summary and discussion	Graded task assignment Practice Summary and discussion	Any questions, repeated practice opportunity. Summary and discussion	The research plan Summary and discussion

Table 13: Training content mapped against 12 session protocol

Session Number	Content of session	Diary sheets etc. for session	Covered in training
1	Introduction Rational assessment	Self-monitoring	Day 1 Day 1 afternoon
2	Behaviour mood link Identifying depressed behaviours ABC formulation	Self-monitoring Goals ABC sheet	Yes functional analysis Yes day 2 a.m.
3	Further exploration of depressed behaviours and consequence Develop goal list Scheduling to break cycle	Behaviours that make me feel bad Breaking down goals sheet Scheduling sheet	Day 2 p.m. Day 2 p.m.
4	Review of scheduling Avoidance recognition	Scheduling Goals TRAP	Day 2 p.m. Day 3 a.m.
5	Scheduling Breaking avoidance What I have learnt so far	TRAP/TRAC What I have learnt sheets	Day 3 a.m.
6	Scheduling Breaking avoidance continued Assessing response to situation and making choices	Schedules Alternative behaviour worksheet Action	Day 2 p.m. and day 3 a.m.

7	Review of using schedules to reduce depressed behaviour and break avoidance How to deal with depressed thinking	Schedules Monitoring rumination	Day 2 and day 3 Day 4 a.m.
8	Continued scheduling review using TRAP/TRAC approach Dealing with rumination	Schedules RCA sheet	Day 2 and day 3 Day 4 a.m.
9	Review of treatment and any problem areas left Problem solving	Problem-solving sheet	Day 3 p.m.
10	Review scheduling and new patterns of behaviour and impact Acting 'as if' Prepare relapse prevention	Relapse flash card - what I have learnt	Day 4 p.m.
11	Review ongoing scheduling Relapse prevention	Risk signs/response to risk card	Day 4 p.m.
12	Review relapse prevention work Review progress to goals Summary and goodbye		Day 4 and all sessions

Training evaluation

Participants were provided with details of the evaluation, and written consent was obtained to report anonymous findings. Participants' experience of the training in terms of acceptability was measured on the final day via the 'Training Acceptability Rating Scale' (Davis et al. 1989). Questionnaires were self-report and completed anonymously.

Questions 1-6 asses acceptability and are scored on a six-point bipolar Likert format (1 = do not agree at all, 6 = strongly agree). Questions 7-15 refer to perceived effectiveness of the training are scored on a four-point unipolar scale (0= not at all, 1 = little, 2= quite a lot, 3= a great deal). The opportunity for free text comments was also provided at the end of the questionnaire. The TARS has good test-retest reliability (r = 0.83), internal consistency (0.99) and acceptable construct validity (Milne 2010, Milne et al. 2000). The TARS tool was chosen as it is commonly used to evaluate training in both physical and mental health settings, thereby facilitating meaningful dissemination of results. Total scoring provides a range of 6-63, with a higher score reflecting increased training endorsement. In addition scores can be separated for the acceptability and effectiveness subscales (6-36 for acceptability and 0-27 for effectiveness). Raw data are presented as range and means and then converted to a percentage score for the total and each subscale for each participant. This is achieved by dividing the achieved score by the total possible score and multiplying by 100. The mean percentage scores achieved for the group results are then presented, with scores of 70-80% being satisfactory and scores of 80% and above suggesting exceptionally good training (Milne 2010, Milne et al. 2000).

Usual Care

Usual primary care management was chosen as the comparator in this study as it reflects the most common intervention for depression (National Institute of Clinical Excellence 2009). There had been no previous evaluation of BA delivered by non-specialists, therefore the first question

would be: does it work? If so, how does it work compared to what most people receive, i.e. usual care? In usual care participants were followed up by their GP or primary care mental health worker and offered interventions deemed appropriate for their condition as per normal practice. At the three months follow-up, control participants were offered BA therapy as delivered in the intervention arm.

4.3.11 Adherence to BA and competency levels of therapists

Adherence is a complicated factor in psychotherapy trials. There are two levels of adherence that of interest to the investigator: that of the participants receiving the intervention and that of the therapists delivering the intervention.

In this study it would have been impractical to assess the fidelity to the model of participants beyond how this is managed within therapy (using and reviewing homework diaries etc.). It was assumed through randomisation equal levels of adherence would be achieved in each group, which through therapy would result in increased activation in the intervention arm alone. Therefore evaluation would be observing the results of activation in any difference between groups on post-treatment outcome measures. It was therefore important to measure the adherence of the non-specialist therapist to the protocol/model under investigation, as this would be the catalyst to increased activation.

For this purpose all treatment sessions were audio taped in the intervention arm of the study. Recordings were stratified for study phase (early, mid, late) and therapist. A sample of 20% of recordings was then randomly selected by a research assistant with no access to content. The number of sessions selected was felt by the study team to reflect a sufficient sample above the ratio used in previously reported studies of BA (Dimidjian et al. 2006) and within the studies financial constraints (independent assessors required payment). We selected two independent assessors with substantial experience both in CBT and in behavioural activation, who reviewed 38 sessions of BA therapy from the study intervention arm.

Adherence

Treatment fidelity was measured against specific behavioural factors in session/homework content. As no established tool was available for this purpose, one was devised for the study. It was felt by the study team (DE in consultation with DR and SG) that this should be simple and lend itself to simple dichotomous evaluation. Was the treatment BA? Was the homework BA? Was any other therapeutic model in evidence? The tool designed was based upon those used in manuals of therapeutic training within IAPT services (Richards and Whyte 2008) as these had been extensively used with minimal problems in clinical settings. Assessors specifically examined session content against treatment protocols, to determine whether behavioural activation was the overall modality applied and if other therapeutic models were prominent in the therapy (such as cognitive therapy); see Table 14 page 174. After reviewing each tape, assessors decided if the session could be classed as authentic BA and assigned values of 1 (yes) and 0 (no) to give a score of adherence to the model. While such rating would not be directly comparable across studies, it was felt to be an acceptable measure of performance bias in the absence of more formally validated tools.

Competence

While the evaluation of adherence measured the degree to which therapists maintained fidelity to the BA model, it did not measure the skill with which they did so. In CBT the commonly used tool for this is the Cognitive Therapy Scale Revised (CTS-R) (Blackburn et al. 2001). This tool measures competence across 14 areas. The tool is valid and reliable in the measurement of CBT, and is commonly used in clinical settings. The CTS-R was reviewed for use in this study through discussion within the steering group and contributors' (DE and DR) experience of applying it to BA. A number of the items (such as eliciting cognitions and guided discovery) were not felt to be relevant in BA. It was not felt this that tool would provide a suitably accurate measure of BA competence, especially as it had not been specifically validated for BA. An alternative was identified

through discussion with Sona Dimidjian, an experienced BA therapist and researcher in the USA. Her research group was developing a rating scale specifically for BA, adapted from the CTS-R, the Behavioral Activation Competence Scale (BACS). It was felt that, although this tool was still in development, it would provide a more accurate reflection of BA competence in this trial. It had been devised by an expert reference group and had been piloted and modified. It measures a range of BA competence across 15 categories on a six-point Likert scale. A score of 0 is poor, 1 barely adequate, 2 mediocre, 3 satisfactory, 4 good, 5 very good and 6 excellent. Guidance is given that a score of 6 would only be applied in exceptional instances of a particular skill. It was felt that this tool, despite its early stage of development and evaluation, would provide a more accurate measure of competence in this study.

Use of both the BACS and the adherence measure was piloted on two therapy tapes by David Ekers and the independent assessors. Differences in interpretation were explored through discussion to develop a consistent approach to the use of each measure.

Results from the BACS are presented as a combined score and then for each therapist individually. Any differences were compared using independent sample t-tests (Bland and Altman 1994).

Table 14: Adherence rating scale BA

Tape ID	Reviewer	Date reviewed	
	No evidence present	Some evidence present	Clear evidence present
Was there evidence that a behavioural rational underpinned interventions within sessions • Reflection on shared BA rationale in session to explain exercises • Checking understanding of BA approach with patient • Self-monitoring of mood-behaviour link • Activity scheduling • Using approaches to tackle avoidance (TRAP-TRAC, ACTION etc..) • Exploring values - goal setting • Dealing with ruminations by exploring consequence • Relapse prevention using a behavioural model	No examples of the behavioural approach were present in the session	There was a mix of behavioural approaches but these were not specific, nor were they linked to a clear shared rationale	The interventions were clearly behavioural in orientation, shared, specific and linked to a collaborative rationale

Was there evidence that homework tasks were designed primarily to re-introduce environmental positive reinforcement - reduce avoidance? • Shared understanding of homework tasks in place • Self-monitoring to draw mood-behaviour link • Developing meaningful goals linked from session to homework • Scheduling activities based upon session discussion • Exploring problems with scheduling and examples reviewed in session-homework • Use of approaches to manage avoidance as homework (TRAP-TRAC Healthy –unhealthy behaviour sheets • Monitoring rumination and/or RCA sheets explained and used • Relapse prevention tasks specific and linked to model	No examples of the behavioural approach were present in the homework-no homework discussed	There was a mix of behavioural approaches but these were not specific, nor were they linked to a clear shared rationale for the homework task	The homework interventions were clearly behavioural in orientation, shared, specific and linked to a collaborative rationale

Was there evidence that other therapeutic models (i.e. cognitive therapy) were central to session content or homework? (If so, please briefly note what was used and how.)	There was no evidence that any other therapeutic models governed session content	There appeared to be a mix of therapeutic models guiding the session, but BA was prominent	There was clear evidence that the interventions used were primarily of a therapeutic model other than BA
Overall would you say the session was authentic behavioural therapy/activation?	yes	no	

4.4 Data analysis

4.4.1 Data quality

Two research assistants independently extracted data from self-rated measures and entered into SPSS (SPSS for Windows 2008). Each data file was checked for errors and then any inconsistencies explored. Data quality was checked by producing descriptive reports of categorical and continuous variables. These were then checked to identify any outliers or values that appeared unusual; each example was examined, referring back to relevant measures as required. Once each data set appeared appropriate, descriptive statistics were produced on each variable for each data set and results compared across data sets. Any inconsistencies were examined to identify the data entry point explaining the difference between data sets. This was then referred back to the original measure and amended accordingly. Once each data set produced the same descriptive statistical results across all variables, one set was adopted as the final version and was used for statistical analysis in this thesis.

Poor data quality is a major threat to the validity of randomised controlled trials and as a result can impact negatively upon subsequent patient care. It is recommended that studies use systematic central statistical monitoring to counter this risk (Baigenta et al. 2008). In this study resources were not available to support such an approach and, as no external grant funding was received, no additional Mental Health Research Network assistance was available for this purpose. It was felt in discussion with the study steering group that the above approach provided an appropriate alternative to systematic central monitoring as a means of maintaining data quality.

4.4.2 Descriptive statistics

Baseline characteristics are presented as means and standard deviations for psychometric scales (BDI; WASS; CSQ) and counts and percentages for categorical variables (depressed/not depressed) as per current conventions (Schulz et al. 2010, Altman and Bland 1996). We analysed each group and

presented results in a table format for each group and for the overall sample.

4.4.3 Statistical approaches

Severity of depression as measured by the BDI-II, functioning as measured by the WASA and health state were compared at one-, two- and three-month follow-up between groups using analysis of covariance (ANCOVA), incorporating baseline measures from each tool respectively. The three-month analysis represented the primary analysis point in this research. Analysis of covariance increases the power of studies to detect difference while controlling for baseline imbalance (Vickers and Altman 2001). Therefore it was considered the most appropriate statistical approach in this study, as the small sample sizes to be used would increase the potential for unequal groups post-randomisation.

Satisfaction was measured using the CSQ-8 at three months and compared between groups using an independent two-sided t-test (Bland and Altman 1994).

For continuous variables between-group mean endpoint differences are presented, both in terms of scores on the instrument and as standardised effect sizes (Cohen’s d) and assigned values to effect size as per normal convention (small 0 - 0.32, medium 0.33 - 0.55 and large 0.56 and above) (Lipsey and Wilson 1993).

4.4.4 Clinical significance

While the above approaches analysed the statistical significance of findings, they did not give a clear indication of the clinical importance. The level of clinical change would give a estimation of the importance of any findings from a clinician's perspective and would complement statistical findings. The study protocol incorporated Jacobson and Truax's procedures (Jacobson and Truax 1991) for calculating reliable and clinically significance change to quantify clinical improvement in

depressive symptoms on the BDI-II. This is recommended as a standard reporting strategy for all published research involving psychological interventions (Evans et al. 1998). Reliable and clinically significant change requires a pre- to post-treatment improvement in scores that is unlikely to be due to the inherent unreliability of the measure (reliable change), accompanied by a movement from a clinical range to a non-clinical one (clinically significant change). In calculating reliable and clinically significant change criteria, the data from the BDI-II manual (Beck et al. 1996) was used for clinical means, standard deviations and the reliability estimate (Cronbach's alpha) and data from Dozois, Dobson and Ahnberg (Dozois et al. 1998) for the non-clinical mean and standard deviation. The Dozois et al. study was used for the latter due to the larger sample sizes used in that study. On the basis of these data, a participant had to improve by ten points or more from pre to post-treatment to show reliable change, and in addition had to score 17 or above pre-treatment and 16 or below post-treatment to meet criteria for clinically significant change.

As an additional measure of clinical improvement, response and remission criteria used in a previous study of BA (Dimidjian et al. 2006) was examined. Response was defined as an improvement of at least 50% or more and remission as a score of $\leq$10 on the BDI-II.

Odds ratios with 95% confidence intervals were used to compare clinically significant change in the two groups.

4.4.5 Health state

EQ-5D scores were converted to health state (QALY) values, and means of the BA and usual care groups were compared at three months using analysis of covariance with baseline health state incorporated as covariate.

4.4.6 Cost data analysis

Mean cost differences between BA and usual care were calculated for the intervention phase with 95% confidence intervals using analysis of

covariance, with six-month pre-treatment baseline cost values as covariates. This would allow for any potential baseline differences to be incorporated in calculations.

4.4.7 Cost Utility

Cost differences between BA and usual care were calculated and expressed as the ratio of cost per difference in QALY. To achieve this in a study measuring only a proportion of a year health state score differences between groups at post-treatment were divided by four to represent the three-month duration of the study. This figure would provide the proportion of a QALY benefit for the duration of this study. This was then multiplied any cost differences observed between BA and usual care over the three-month period of the study by the value required to convert the three month proportion of QALY benefit to a full QALY. This approach provided the deterministic results for the cost of a QALY found in this study.

An as additional observation the cost per point change on the BDI-II was calculated. While this is not the preferred approach to cost utility analysis, it allowed results to be considered alongside other studies using this method.

4.4.8 Sensitivity analysis of cost utility calculations

Cost effectiveness acceptability curves (CEAC) were generated to explore uncertainty around cost utility findings (Fenwick and Byford 2005). This is achieved by conducting 1000 non-parametric bootstrap replications placed upon an incremental cost-effectiveness plane, where cost is mapped against health state outcome on the *y* and *x* axis respectively. Cost-effectiveness planes are then used to map results of a sample across four quadrants: more costly and more effective, more costly and less effective, less costly and more effective and less costly and less effective (Briggs et al. 2002). Bootstrapping approaches are then used to re-sample based upon the original data set to explore uncertainty of results. This was achieved by re-

sampling a set of observations of equal size to the original data for the treatment and usual care groups. The means of cost and effect were then used to estimate the incremental cost-effectiveness of each bootstrapped sample. This procedure was repeated 1000 times and results placed upon the cost effectiveness plane as per recommended approaches (Briggs et al. 2002, Briggs 1999). This allows an estimation of the number of observations likely to fall below any particular value given to a QALY. Results were then used to generate a CEAC for scenario A and B as outlined in training costs section. The curve was created by plotting the proportion of observations indicating BA was cost-effective compared to usual care at a range of values for a QALY gain (Fenwick and Byford 2005). This approach is preferred over incremental cost effectiveness ratios with 95% confidence intervals as it allows an estimate of the probability that one intervention is preferable at a range of values and thus supports decision making (Fenwick et al. 2001). As we have an acceptable value for a QALY set at £20,000 (National Institute for Health and Clinical Excellence 2008) the CEACs used in this study would allow consideration of the probability that BA or usual care would be dominant at that value. Bootstrapping and the generation of the CEACs used in this study was conducted by SP in using data provided by DE.

This process was not repeated for the value of a point reduction on BDI-II, as no equivalent accepted tariff exists to provide appropriate context to results.

4.4.9 Missing Data

Missing data presents a common threat to the results of many trials with over half commonly missing over 10% of results (Altman 2009). Excluding cases with missing variables cannot be justified as it has been seen to bias results with the degree and direction of such bias being unpredictable (Nuesch et al. 2009) and also results in reduced statistical power (Altman 2009). In this study missing data were anticipated as dropout from active psychotherapies such as BA is commonly high (Clark et al. 2009). This

alongside the relatively small sample sizes planned in this study meant any missing data would need to be carefully handled.

Missing data can be categorised into one of three groups (Sterne et al. 2009, Altman and Bland 2007):

- Missing completely at random (MCAR) - this is where the reason for the missing observation cannot be assumed to be related to the outcome and available data. An example of this is when a sphygmomanometer machine breaks down. This explanation to the missing observations cannot be assumed to be associated with any factor other than random chance.
- Missing at random (MAR) - The missing observation is predictable and related to other observations, but not specifically the missing observation. An example might be young people having more missed blood pressure measurements. While there is a pattern to the missing data, it cannot be assumed that this depends on the actual missing value.
- Missing not at random (MNAR) – There is a relationship between the missing value and the fact it is missing. Using the above example this may be that people with high blood pressure are more likely to miss appointments due to headaches.

MNAR is the most common reason, and poses problems as the lack of such observations will introduce bias into findings (Altman and Bland 2007). It was assumed that missing data in this study was most likely to fall into either the MAR or MNAR categories; hence it was necessary to use an intention to treat (ITT) approach examining results for all participants in the group to which they were randomised. ITT is seen to be the most appropriate way to handle missing data, but there is debate with regards to the most appropriate methodology (Streiner and Geddes 2001). A commonly used approach to deal with missing data is carrying forward last observations (LOCF), as seen in previous trials of BA (Dimidjian et al. 2006). This approach, however, can significantly bias results in either direction, since a single imputation of missing values fails to reflect the

uncertainty around missed observations (Nuesch et al. 2009, Streiner and Geddes 2001). LOCF assumes that the final value or missing value is equal to the last observed value. This is problematic; if observations are MNAR there is an explanation to their absence of which we are not aware. Examples of this may be clear deterioration (treatment not acceptable or working) or improvement (treatment has worked and person needs no more) or a dislike of the therapist. Therefore LOCF takes no account of any change from outside treatment settings (natural improvement/deterioration) or the trajectory prior to dropout maintaining post dropout. Missing data are therefore likely to be related to treatment and/or prognosis in some way, making LOCF implausible (Altman 2009). To deal with such problems multiple imputation was used, where possible, for our intention to treat analysis (Sterne et al. 2009). Multiple imputation allows for uncertainty around missing data by creating many imputed data sets sampled from predictive distributions modelled on a relevant available set of observations. This does not give one replaced value but a range of replaced values reflecting the uncertainty around the missing observation based upon the variability of the observed data. The post-treatment differences of each imputed set are compared using ANCOVA and a mean of results provided. Thus results from multiple imputations do not supply individual post-treatment scores where observations are missing, nor do they provide means and standard deviations for each group. It is important where there is uncertainty as to the reason for missing data to use multiple imputation cautiously and to consider it against completer analysis, explaining any significant differences observed (Sterne et al. 2009). An intention to treat analysis was conducted for this study replacing missing data using multiple imputation by chained equations, as described by Royston using 100 imputations (Royston 2004). Baseline BDI-II, age, sex, problem duration and allocation was incorporated in modelling. Multiple imputation is a specialist statistical procedure, and becoming fully competent in its design and application is beyond the remit of this PhD. It is recommended to get expert advice and assistance in the procedure from a statistician (Sterne et al. 2009). MB advised in the design of this approach

and conducted the analysis using STATA with MVIS and MCOMBINE functions (Stata Corporation 2003).

For the clinical significance analyses it was not possible to use the multiple imputations method, as it does not supply individual participant level data (post-treatment mean). The approach of using both completer and LOCF analyses for clinical significance was adopted, taking into account the reservations regarding LOCF outlined above. In the absence of any more suitable alternative this approach was considered suitable.

An intention to treat approach (ITT) with health state data using multiple imputation, using 100 imputations incorporating baseline health state, age, sex, problem duration and allocation in modelling was used. As this approach does not replace post-treatment scores, the completer results were used for bootstrapping and production of CEACs as per recommended approaches to managing missing data in cost-effectiveness studies (Briggs et al. 2003).

4.4.10 Mapping results against previous studies

As this study was aimed at exploring the effective dissemination of BA it was felt to be important to map results against studies that had used senior therapists to deliver BA or BT. Mean and standard deviation scores for the BDI-II were entered into the meta-analysis database in review manager (Cochrance Collaboration 2003) used to produce the results in Chapter three. This was conducted for both completer analysis and ITT analysis using LOCF, as multiple imputation does not provide post-treatment mean and standard deviation scores for extraction. The results were then compared visually and by examination of their influence on SMD between groups.

4.5 Results

4.5.1 The treatment manual

The treatment manual used in the study is presented in Appendix II. It represents a major product of the research of potential use to future researchers and clinicians. Each session is laid out with an agenda, session content instructions and homework tasks. It provides the therapist with guidance on content of sessions and if followed covers all key elements of behavioural activation identified for reviewed trials. Session-by-session content is outlined in Table 13 page 168.

.

4.5.2 Training evaluation

Training was attended by 10 mental health professionals currently employed within Tees, Esk and Wear NHS Foundation Trust. Participants were mental health nurses or graduate primary care mental health workers. Brief characteristics of participants are outlined in Table 15.

Table 15: Characteristics of training participants

Characteristic	Mean (SD)
Age in years (SD)	41.7 (9.11)
Years of health service employment (SD)	13.4(9.95)
Role	Mental health Nurse-8 Graduate primary care mental health worker-2
Mean caseload number (SD)	30.5 (19.3)
Percentage of caseload experiencing depression (SD)	74.89 (21.30)

Training Acceptability Rating Scale

All participants (n=10) completed post training evaluation on the TARS questionnaire. Results are presented in Table 16.

Table 16: TARS Evaluation of five day BA training

Area (score range)	Range (min-max)	Mean
1. General acceptability (0-6)	5-6	5.7
2. Effectiveness (beneficial for staff) (0-6)	4-6	5.7
3. Negative side effects for clients (0-6)	5-6	5.7
4. Not appropriate intervention (0-6)	4-6	5.4
5. Consistent with good practice (0-6)	5-6	5.9
6. Most staff would approve of this training (0-6)	4-6	5.6
7. Did the course improve your understanding? (0-4)	2-3	2.8
8. Did the course help develop work related skills? (0-4)	3-3	3.0
9. Has the course made you more confident? (0-4)	2-3	2.3
10. Do you expect to make use of the course content in your workplace? (0-4)	2-3	2.4
11. Competency of course leaders (0-4)	2-3	2.9
12. General satisfaction (0-4)	2-3	2.9
13. Did the course meet its objectives? (0-4)	2-3	2.9
14. Did course leaders relate to group effectively ?(0-4)	3-3	3.0
15. Were the leaders motivating?	2-3	2.9

A high level of perceived acceptability and effectiveness of the training was found from the questionnaires. All training participants' scores were above the 80% 'very good' threshold of this measure (see Table 17). The mean percentage of acceptability for the training across all training participants was 94.4% (SD 6%) and for effectiveness 92.96% (SD 4.43%), with a combined rating of 93.81% (SD 4.76%).

Table 17: Percentage TARS scores for Participants in BA training

Student evaluation	Acceptability	Effectiveness	Total
1	100.00	96.30	98.41
2	100.00	88.89	95.24
3	94.44	96.30	95.24
4	94.44	100.00	96.83
5	97.22	96.30	96.83
6	94.44	92.59	93.65
7	88.89	88.89	88.89
8	100.00	92.59	96.83
9	94.44	92.59	93.65
10	80.56	85.19	82.54

Free text comments are presented in Table 18. They reflect the overall positive evaluation of the training with clear reference to the repeated skills practice being an important factor.

Table 18: Free text comments from TARS regarding five day BA training

Question	Response
Q16. What was the most helpful part of the course for you personally?	• *Role play – to practice skills in safe environment and receive constructive feedback, albeit that dreaded role play at start.* • *Skills practice.* • *Helped me understand the model much more and gave the opportunity to put model into practice which gave confidence in delivery.* • *Practising the skills in small groups to experiment with different ways of presenting the information/worksheets.* • *Observing the assessment completed by the course leader.* • *Developing skills in funnelling questions to collect information.* • *Practising techniques through role play.* • *Group discussion.* • *Working in triads.* • *Trainer with excellent knowledge base and enthusiasm for the model.* • *The assessment tool and the formulation tool made it easy to explain how it is easy for us to maintain case depression.* • *Role play.* • *Group debates.* • *Feedback from course leader during monitoring of sessions.* • *Group work i.e. role play and open discussion.* • *The role play exercises were very helpful as a learning (aid) exercise.* • *Formulation of problem solving.*

Q17. What change, if any, would you recommend (e.g. to the content or teaching)?	• *Provide handout at start.* • *None.* • *Have slides printed for staff before training to make notes as we go along.* • *No changes other than the locations were not always easy to find.* • *Handouts of slides to be optional.* • *Use of 'hand-outs'.* • *Change format of assessment form, feels all out of sync.* • *Some alternative titles for the formulation boxes.*
Q18. Please make any other comments that you would like to offer	• *Enjoyed the training very much – relaxed atmosphere – increased confidence.* • *Thank you.* • *Thoroughly enjoyed the training despite initial apprehension – feel I have learnt all appropriate techniques in a well-structured and well planned way.* • *Overall I found the training to a high standard and I'm confident that it will guide my practice in the future.* • *The course has greatly contributed to my skills in clinical work.* • *An excellent delivery of a method of treating depression and I eagerly await putting it into practice.* • *The training was very enjoyable and I feel will be very useful in my daily working practice, supervision to be organised.*

4.5.3 Data quality

Prior to analysis 40 inconsistencies on comparison of the two independent data sets were identified. These generally related to data entry errors and they were resolved with reference to source records following which appropriate amendments were made. Details of problems identified and approach taken can be found in Appendix II.

4.5.4 Baseline characteristics and study attrition

Sixty eight participants were referred to the trial, of whom 21 were excluded (17 did not meet diagnostic criteria, 2 refused randomisation, 2 had significant suicidal ideation (as measured by a score of 2 ≥ on question 9 of the BDI-II). Forty seven participants met the inclusion criteria and proceeded to randomisation. This rate (69% inclusion) is similar to other studies of BA which included 68% inclusion (Dimidjian et al. 2006) and UK primary care based problem solving (66%) (Kendrick et al. 2006). Twenty three patients were allocated to behavioural activation and 24 to control. No differences were observed in scores at baseline between the two groups on BDI-II (BA mean = 35.57 SD=9.60, usual care mean =35.08 SD=9.60), WASA (BA mean =26.39 SD 7.30, usual care mean = 25.13 SD=7.30), CSIR (BA mean = 31 SD=10.99, usual care mean=33.13 SD=8.22) or problem duration (BA mean = 186.91 weeks SD=358.49, usual care mean =195.21 weeks SD=404.64). Baseline participant characteristics are presented in Table 19 and indicate the participants in the trial represent a long term severely depressed group with substantial impairment of functioning. Data were collected from 38 participants at three month assessment, 16 in the BA arm and 22 in control. Of those opting out of the study 3 did so post randomisation (1 BA, 2 Usual care), 3 at one month (3 BA), and 3 at two months (3 BA). There were no significant differences between completers and those dropping out of treatment on baseline BDI-II depression scores (dropout mean =36.55 SD=10.77, no dropout mean =35.21 SD 9.43) or duration of problem

(dropout mean =182 weeks SD=439, no dropout mean=193 weeks SD=369). Of the 23 participants randomised to BA 11 received all 12 sessions. Of those with missed sessions 3 received 1-3 sessions, 3 received 4-6 sessions and 5 received 7-9 sessions. Study flow is presented in Figure 32.

Table 19: Characteristics of participants at baseline

Baseline Characteristic	BA (n=23)	TAU (n=24)	All (n=47)
Age in years (range)	46.43 (24-63)	43.08 (28-63)	44.72 (24-63)
Sex n (%)			
Male	8 (35)	10 (41.7)	18 (38)
Female	15 (65)	14 (58.3)	29 (62)
Employment n (%)			
Full time	13 (56.5)	8 (33.3)	21 (44.7)
Part time	1 (4.3)	7 (29.2)	8 (17)
House person	1 (4.3)	1 (4.2)	2 (4.3)
Carer	0	1 (4.2)	1 (2.1)
Retired	3 (13)	3 (12.5)	6 (12.8)
Unemployed	4 (17.4)	2 (8.3)	6 (12.8)
Incapacity benefit	1 (4.3)	2 (8.3)	3 (6.4)
Mean Duration of problem in weeks (SD)	186.91 (358.49)	195.21(404.64)	191.15 (378.61)
Mean Baseline BDI-II score (SD)	35.57 (9.60)	35.08 (9.60)	35.32 (9.50)
Mean Baseline WASA scale score (SD)	26.39 (7.30)	25.13 (7.70)	25.74(7.46)
Mean Baseline CSIR score (SD)	31 (10.99)	33.12 (8.22)	32.09 (9.63)
Prescribed anti depressants (%)	15 (65%)	17 (71%)	32 (68%)
Baseline CSIR (ICD10) diagnosis n (%)			
Mild depression	1 (4.3)	2 (8.3)	3 (6.4)
Moderate depression	13 (56.5)	9 (37.5)	22 (46.8)
Severe depression	8 (34.8)	13 (54.2)	21 (44.7)
Mixed anxiety and depression	1 (4.3)	0	1 (2.1)

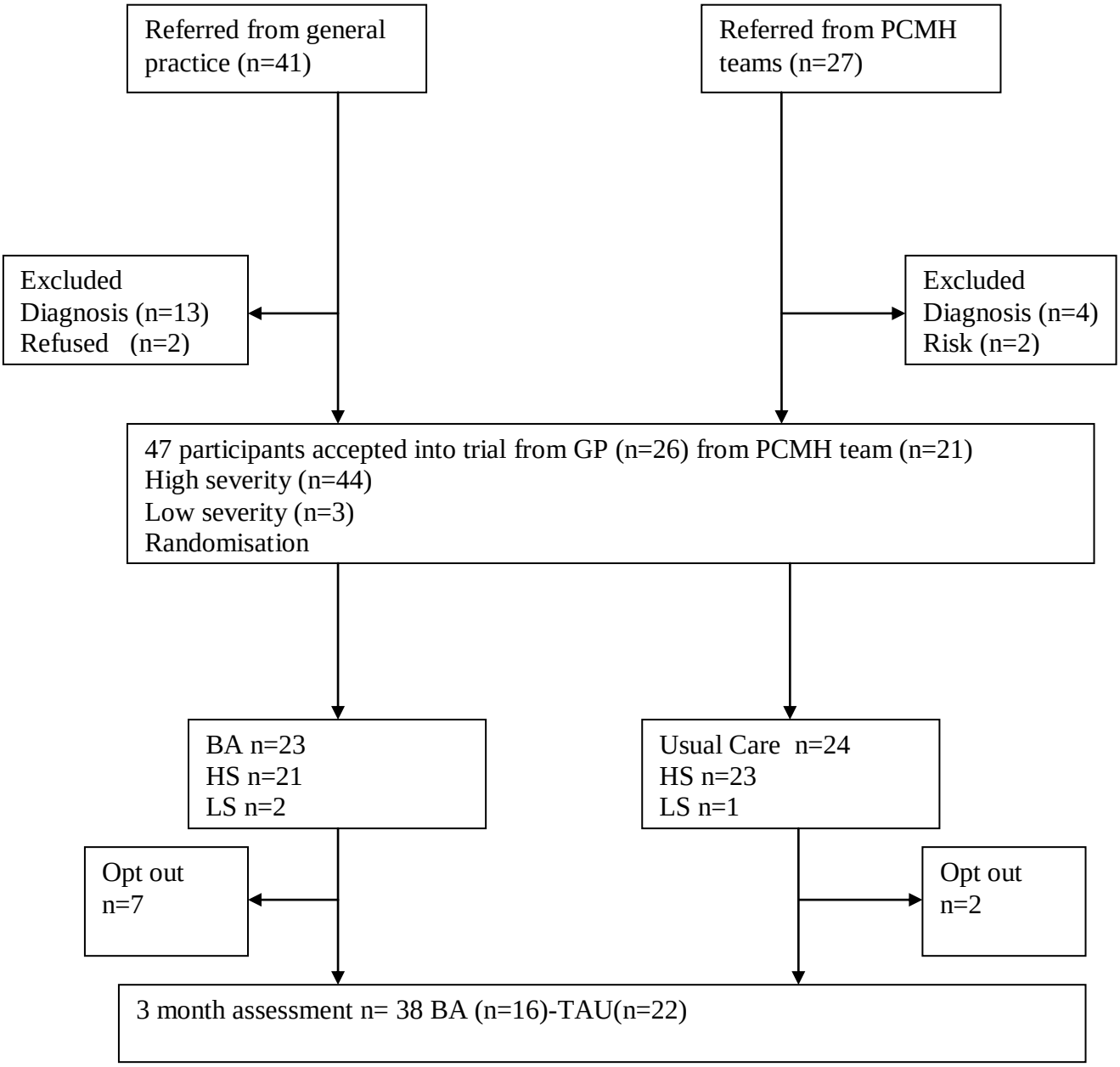


Figure 32: Study flow chart

4.5.5 Assessment of normality of baseline data

BDI-II

The 5% trimmed mean of baseline BDI-II data was 35.20, indicating minimal impact of extreme scores on distribution (overall mean 35.32 SD 9.50). Plots revealed a relatively normal distribution on a histogram, and the Kolmogorov-Smirnov statistic produced a non-significant result (P=0.13), indicating normal distribution of BDI-II scores: See Figure 33

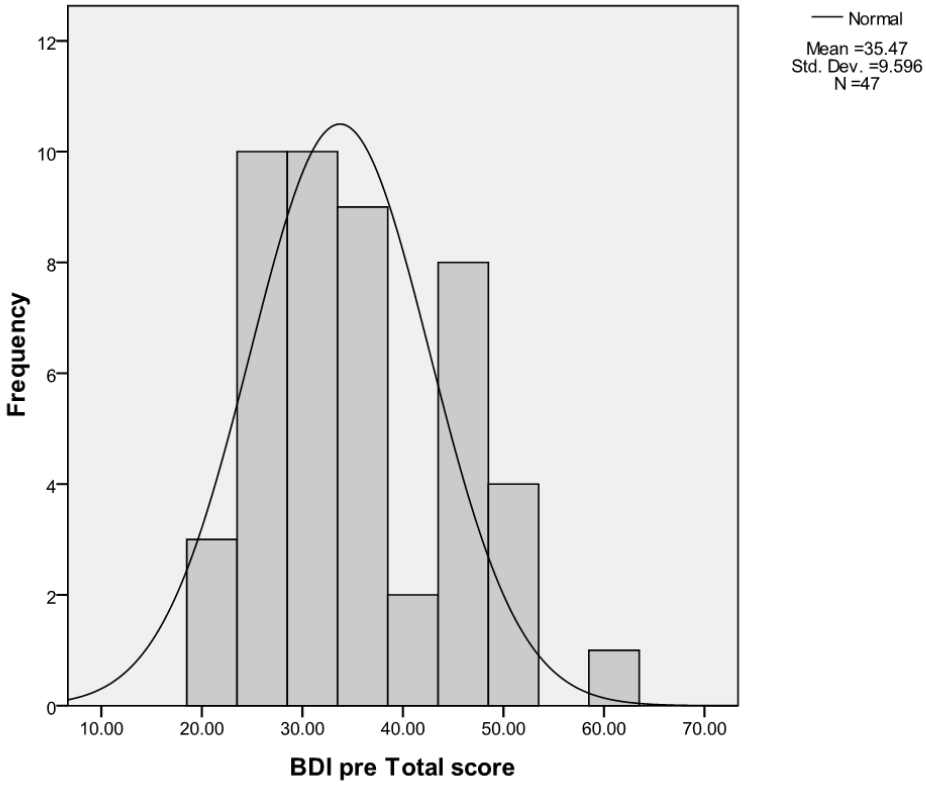


Figure 33: Histogram of baseline BDI-II pre scores combined

WASA

The 5% trimmed mean of combined baseline WASA data was 25.99, indicating minimal impact of extreme scores on distribution (overall mean 25.74 SD 7.46). Plots revealed a relatively normal distribution on a histogram, and the Kolmogorov-Smirnov statistic produced a non-

significant result (P=0.20), indicating normal distribution of BDI-II scores: See Figure 34.

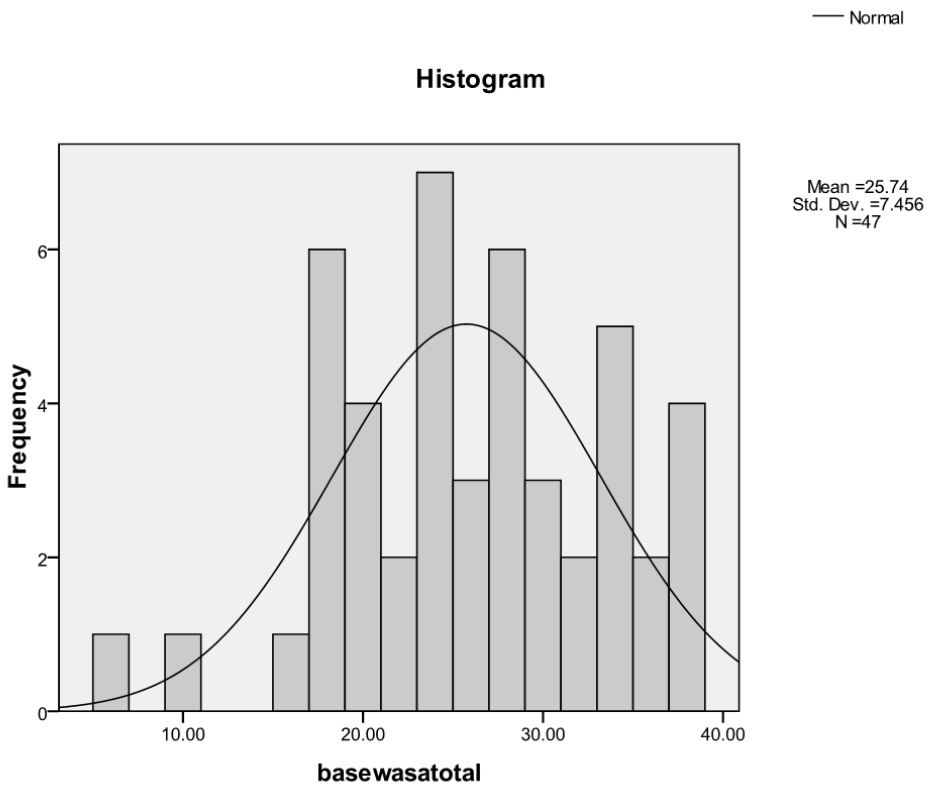


Figure 34: Histogram of Baseline WASA pre scores combined

4.5.6 Treatment adherence

Therapists closely adhered to the treatment protocol in all 38 sessions reviewed by independent specialists. A behavioural activation rationale underpinned all intervention content and homework, with all reviewed sessions scoring 1 (BA dominant) in relation to session and homework content. There was no evidence of alternative therapy models (such as cognitive therapy) being applied in any the reviewed sessions, with a score of 0 for all tapes (reflecting no other therapy modes being prominent). All sessions scored 1, being classed as an example of BA.

4.5.7 Treatment competence

Scores on the BACS scale are outlined in Table 20. They indicate a mean level of competence between satisfactory and good in all categories apart

from session structure, where the rating falls between mediocre and satisfactory. This is reflected in the overall rating of therapist competence score of 3.39 (SD 0.8239) indicating a satisfactory level BA ability across all sessions.

BACS competence ratings were then compared between the two therapists in the study. Differences between the therapists are presented in Table 21. A significant difference was found across all domains, with therapist one achieving consistently higher levels of competency across both independent assessor ratings.

Table 20: Combined BACS rating scores for therapist 1 and 2.

BACS item	N	Minimum	Maximum	Mean	Std. deviation
Structure	38	1.00	5.00	2.89	0.92
Focus BA	38	2.00	5.00	3.79	1.04
Attends to understanding	38	1.00	5.00	3.11	0.92
Validates	38	2.00	5.00	3.24	0.75
Non-judgemental	38	3.00	5.00	3.76	0.63
Warm and genuine	38	3.00	5.00	3.84	0.68
Collaborates	38	1.00	5.00	3.16	0.92
Reinforces activation	38	2.00	5.00	3.24	0.82
BA formulation	38	2.00	5.00	3.45	0.80
Reviews homework	38	2.00	5.00	3.16	0.92
Defines problems behaviourally	38	2.00	5.00	3.60	0.89
Selects appropriate targets	38	2.00	5.00	3.53	0.69
Appropriate change methods	38	2.00	5.00	3.66	0.78
Assigns homework	38	2.00	5.00	3.47	0.60
Skilful application	38	2.00	5.00	3.39	0.79
Overall rating	38	2.00	5.00	3.39	0.82

Table 21: BACs rating scale comparison between therapist 1 and 2

BACS item	therapist	N	Mean	Std. deviation	Mean difference	P value	95% CI
Structure	1	18	3.33	0.84	0.83	0.00	0.28 to 1.38
	2	20	2.50	0.82			
Focus on BA	1	18	4.50	0.61	1.35	0.00	0.82 to 1.87
	2	20	3.15	0.93			
Attends to understanding	1	18	3.83	0.51	1.38	0.00	0.98 to 1.78
	2	20	2.45	0.68			
Validates	1	18	3.78	0.54	1.03	0.00	0.67 to 1.39
	2	20	2.75	0.55			
Non-judgemental	1	18	4.22	0.43	0.87	0.00	0.57 to 1.18
	2	20	3.35	0.50			
Warm and genuine	1	18	4.33	0.49	0.93 .	0.00	0.61 to 1.26
	2	20	3.40	0.50			

Collaborates	1	18	3.79	0.55	1.18	0.00	0.71 to 1.64
	2	20	2.60	0.82			
Reinforces activation	1	18	3.83	0.51	1.13	0.00	0.74 to 1.52
	2	20	2.70	0.66			
BA formulation	1	18	3.83	0.71	0.73	0.00	0.26 to 1.20
	2	20	3.10	0.72			
Reviews homework	1	18	3.78	0.65	1.18	0.00	0.77 to 1.59
	2	20	2.60	0.60			
Defines problems behaviourally	1	18	4.28	0.67	1.28	0.00	0.87 to 1.68
	2	20	3.0	0.56			
Selects appropriate targets	1	18	3.94	0.54	0.79	0.00	0.42 to 1.17
	2	20	3.15	0.59			
Appropriate change methods	1	18	4.17	0.51	0.97	0.00	0.56 to 1.37
	2	20	3.20	0.70			

Assigns homework	1	18	4.00	0.34	1.00	0.00	0.78 to 1.22
	2	20	3.00	0.32			
Skilful application	1	18	4.06	0.42	1.26	0.00	0.94 to 1.57
	2	20	2.80	0.52			
BACS overall BA rating	1	18	4.00	.59	1.15	0.00	0.76 to 1.54
	2	20	2.85	.59			

4.5.8 Additional interventions

Antidepressant medication was prescribed at baseline to 17 (71%) participants in usual care and 15 (65%) participants in BA compared to 15/24 (62.5%) and 12/23 (52%) respectively during the intervention phase. Six participants in usual care had follow-up from a community psychiatric nurse. Two participants in BA had one session each with a psychiatrist. No other additional interventions for depression were received by participants during the study.

4.5.9 Depression symptom level

A one-way between groups analysis of covariance was conducted, with participants' scores on the BDI-II pre-treatment used as the covariate for those with completed post-treatment scores (BA n=16, control n=22). Preliminary checks were conducted to ensure there was no violation of assumptions of linearity or homogeneity of regression slopes. There was no significant interaction between BDI-II at baseline and allocation (P=0.404). After adjusting for baseline BDI-II scores, there was a significant difference in favour of BA of −15.65 (95% CI − 6.90 to −24.41) points on the BDI-II (*F (*1, 35) =13.18, P=0.001 n=38) representing a large effect size (Cohen's d = −1.15 95% CI − 0.45 to −1.85): See Table 22.

Multiple imputation analysis of missing data BDI-II

Intention to treat analyses with multiple imputation showed a mean difference on post BDI-II scores of −15.78 in favour of BA (95% CI −24.55 to −7.02, P= 0.001) with all randomised subjects (BA n=23, usual care n=24) included in analysis: See Table 22

Difference in BDI-II scores over time

One way between groups' analysis of covariance was conducted at 2 additional time points, one month into treatment and 2 months into treatment. Multiple imputation was not used due to the similar findings between imputed data sets and completers at our primary analysis time point (3 months).

At one month we observed a mean BDI-II score for BA of 27.70 (SD 3.33 n=20) and control 31.70 (SD 10.98 n= 23). After adjusting for baseline BDI-II scores, there was a non-significant difference in favour of BA of −3.90 points on the BDI-II (95% CI −10.13 to 2.31 P=0.21).

At two months, we observed a mean BDI-II score for BA of 21.31 (SD 10.42 n=16) and control 28.96 (SD 14.63 n= 23). After adjusting for baseline BDI-II scores, there was a significant difference in favour of BA of −8.47 points on the BDI-II (95% CI −16.76 to −0.17 P= 0.046 BA): see Figure 35.

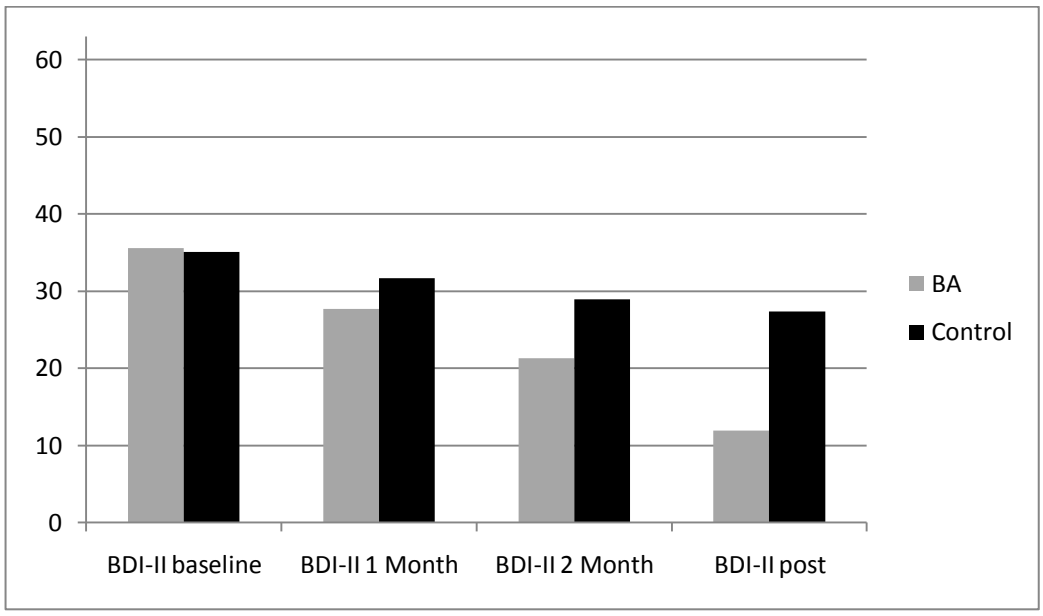


Figure 35: Change over time by group BDI-II

4.5.10 Clinical significance

Clinically significant improvement on the BDI-II for those completing the intervention

In the completer analysis (n = 38), 75% of participants in the Behavioural Activation group met the criterion for reliable improvement (improvement of 10 points or more) on the BDI-II compared to 36.4% of the control

group (OR = 5.35, 95% CI = 1.3 to 21.9). The treatment group was also more likely to meet criteria for reliable and clinically significant change (improvement by 10 points or the BDI-II and a move from ≥17 pre-treatment to ≤16 post treatment) (56.3% vs. 22.7%; OR = 4.4, 95% CI = 1.1 to 17.8). Response rates on the BDI-II were higher in the behavioural activation group (68.2% vs. 18.2%; OR = 9.9, 95% CI = 2.2 to 45.0) as were remission rates (56.3% vs. 13.6%; OR = 8.1, 95% CI = 1.7 to 39.1). Reliable and clinically significant change and response and remission for completers are presented in Figure 36 and Figure 37 respectively.

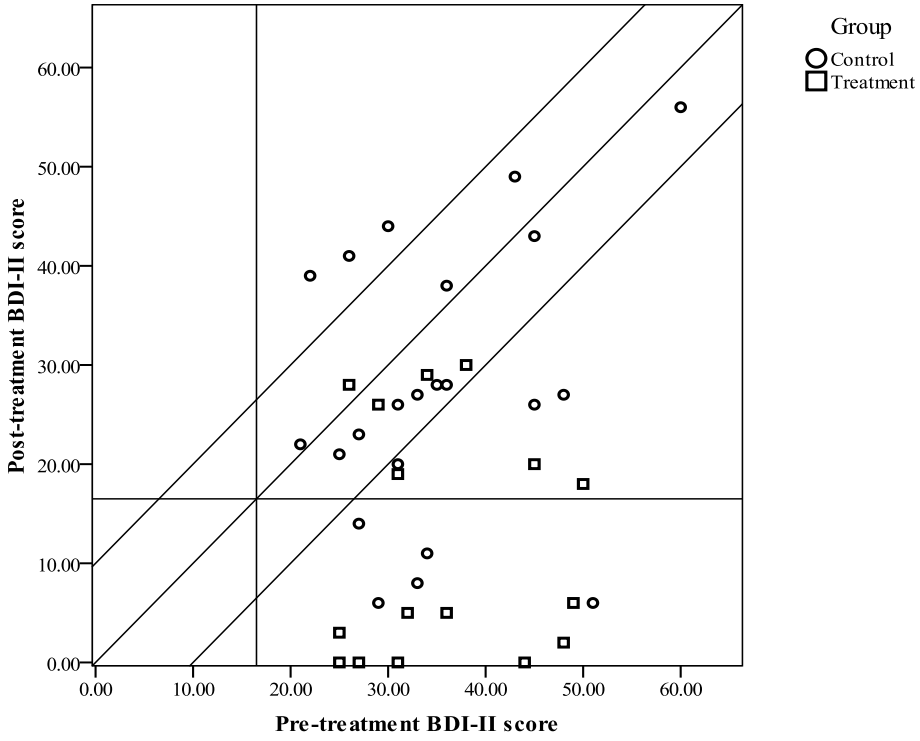


Figure 36: Reliable and clinically significant change by group for study participants completing the intervention (n=38).

Notes: reliable improvement requires a score below the lowest diagonal; scores below the lowest diagonal and below the horizontal reference line meet criteria for reliable and clinically significant change.

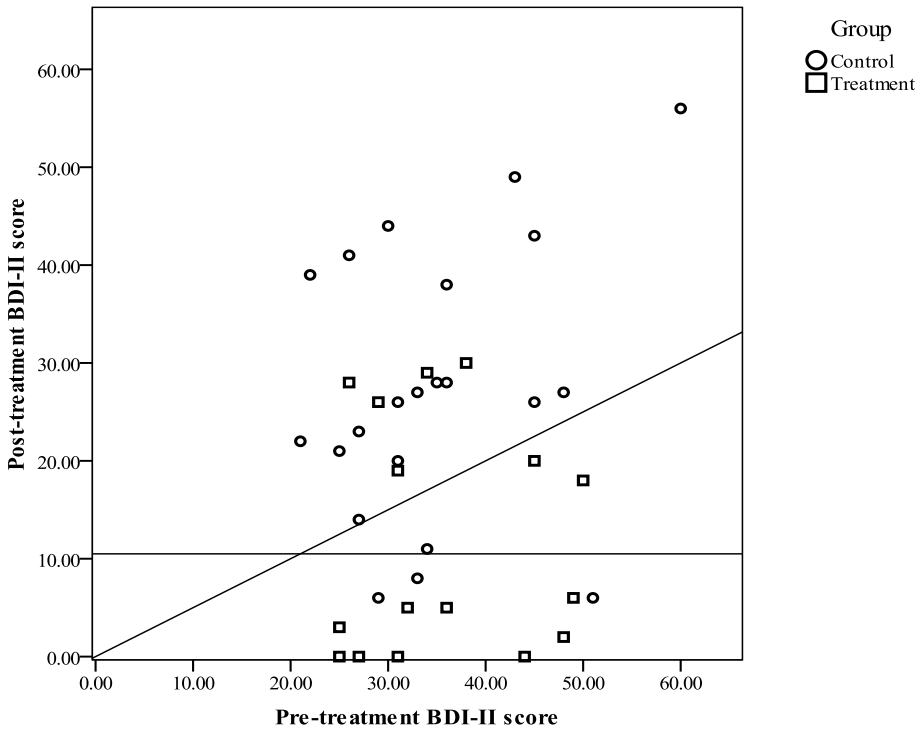


Figure 37: Participants meeting response and remission criteria by group for study participants completing the intervention (n=38).
Notes: Scores below the diagonal meet the criterion for response. Scores below the horizontal reference line meet the criterion for remission.

Clinically significant improvement on the BDI-II using LOCF

When the analysis of the BDI-II was repeated using last observation carried forward (n = 47), 65.2% of the behavioural activation group showed reliable improvement compared to 33.3% of the control group (OR = 3.8, 95% CI = 1.1 to 12.5). Although more of the treatment group (43.5%) met criteria for reliable and clinically significant change than the control group (20.8%), the confidence interval for the odds ratio included one (OR = 2.9, 95% CI = 0.8 to 10.6) hence this difference was not statistically significant. Response rates were higher in the treatment group (47.8% vs. 16.7%; OR = 4.6, 95% CI = 1.2 to 17.7) and were on the border of significance for remission (39.1% vs. 12.5%; OR = 4.5, 95% CI = 1.0 to 19.6). Four participants (16%) in the usual care arm demonstrated deterioration at three months which was not observed in BA. Figure 38 and Figure 39

summarise pre- to post-treatment change against reliable and clinically significant criteria and response and remission respectively.

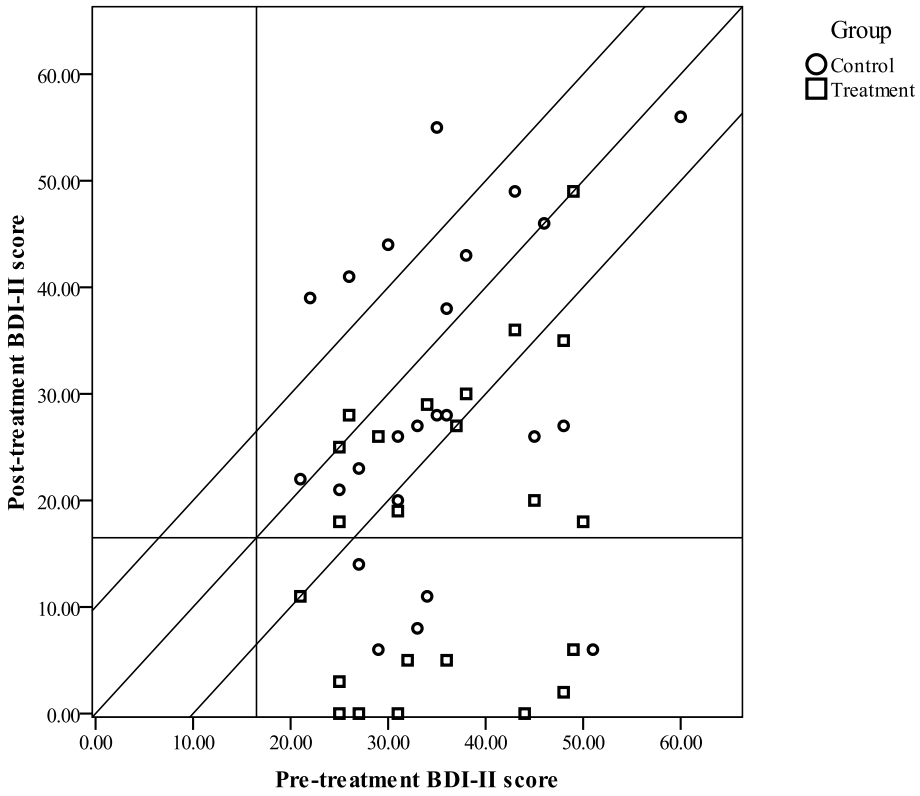


Figure 38: Participants meeting reliable and clinically significant change criteria by group (LOCF n= 47).

Notes: reliable improvement requires a score below the lowest diagonal; scores below the lowest diagonal and below the horizontal reference line meet criteria for reliable and clinically significant change.

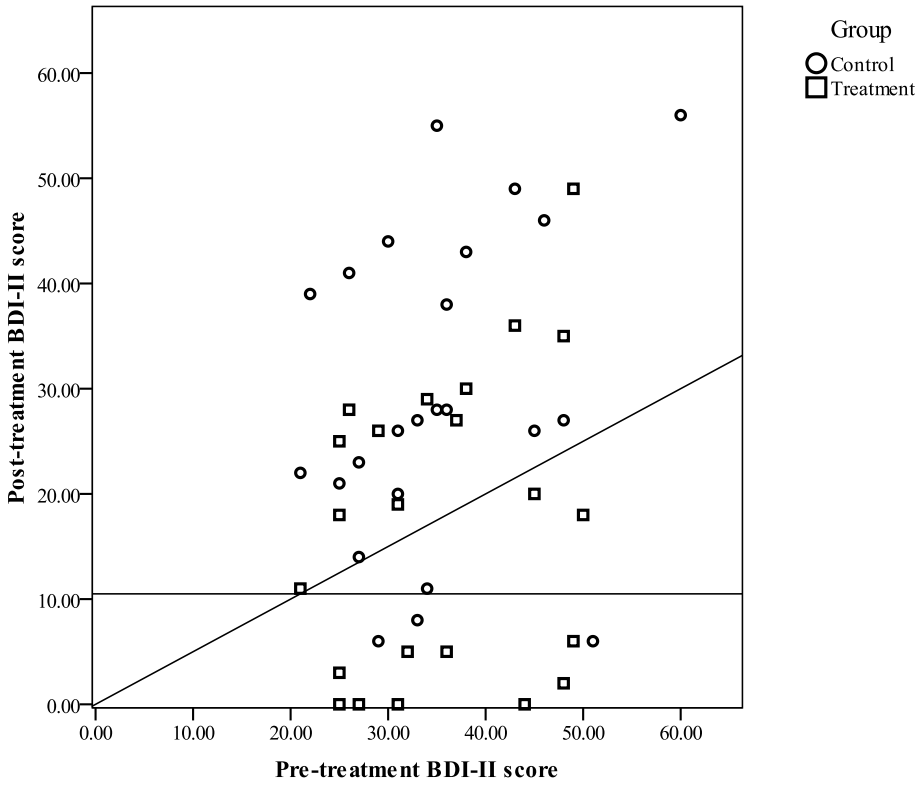


Figure 39: Participants meeting response and remission criteria by group (LOCF n=47).

Notes: scores below the diagonal meet the criterion for response; scores below the horizontal reference line meet the criterion for remission.

4.5.11 Functioning

A one-way between groups analysis of covariance was conducted with participants' scores on the WASA pre-treatment used as the covariate. Preliminary checks were conducted to ensure that there was no violation of assumptions of linearity or homogeneity of regression slopes. There was no significant interaction between WASA at baseline and allocation (P=0.97). After adjusting for baseline WASA scores, there was a significant difference in favour of BA of −11.56 (95% CI −4.79 to −18.33) points on the WASA (*F* (1,35) =12.01, P= 0.001 n=38) representing a large effect size −1.14 (−1.84 to −0.45): See Table 22.

Multiple Imputation of missing data WASA scale

Intention to treat analysis with multiple imputation showed a mean difference on post-WASA scores of −11.12 in favour of BA (95% CI −17.53 to −4.70, P= 0.001) with all randomised subjects (BA n=23, usual care n=24) included in analysis: See Table 22.

Difference in WASA scores over time

One-way between groups analysis of covariance was conducted at two additional time points, one month into treatment and two months into treatment. Multiple imputation was not conducted due to the similar findings between imputed data sets and completers at our primary analysis time point (three months).

At one month, a mean WASA score for BA of 25.55 (SD 0.01 n=20) and control 27.09 (SD 7.08 n= 23) was observed. After adjusting for baseline WASA scores there was a non-significant difference in favour of BA of −2.35 points on the BDI-II (95% CI −6.75 to 0.2.04 P=0.28).

At two months, a mean WASA score for BA of 20.44 (SD 10.52 n=16) and control 25.83 (SD 9.11 n= 23) was observed. After adjusting for baseline BDI-II scores there was a non-significant difference in favour of BA of −6.34 points on the WASA (95% CI −12.74 to 0.06 P= 0.05): see Figure 40.

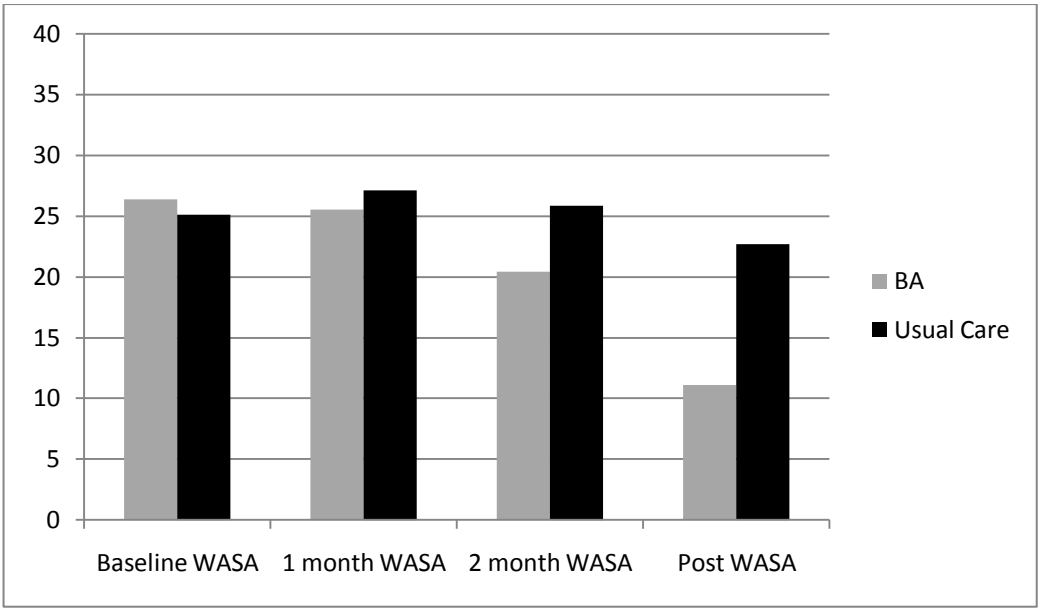


Figure 40: Change over time by group WASA

4.5.12 Satisfaction

Analyses were conducted on 38 participants (16 BA, 22 usual care) who completed CSQ questionnaires for post-treatment assessment. A mean difference in post treatment CSQ scores was found in favour of BA of 4.81 (95% CI 2.23 to7.38, P= 0.001) showing a higher level of general satisfaction in the BA group. See Table 22

4.5.13 Therapist differences

A *post hoc* analysis of results by therapist was conducted to explore any differences observed. No differences were observed at baseline in the level of depression severity in those participants treated by each therapist (BDI-II therapist 1= 38.75 SD 9.27, therapist 2= 32.09 SD 9.09) or functioning (WASA therapist 1 = 27.75 SD 6.80, therapist 2 24.91 SD 7.87). Independent samples t-tests were then conducted on post-treatment scores on the BDI-II, WASA and CSQ between therapist 1 and 2. Patients treated by therapist 1 showed a greater reduction and lower post-treatment score on the BDI-II (post-treatment BDI-II therapist 1 6.25 SD 7.09, therapist 2 17.62 SD 12.81, P = 0.05) and WASA (post-treatment WASA therapist 1 6.12 SD 4.64, therapist 2 16.12 SD 10.97, P = 0.03) using the completer analysis (N=16, N= 8 for each therapist).

When repeated using LOCF (N=23) (as multiple imputation data not available for this analysis) the differences reduced and were non-significant on the BDI-II (post-treatment BDI-II therapist 1 14.41 SD 13.92, therapist 2 19.90 SD 14.96, P = 0.366) and WASA (post-treatment WASA therapist 1 13.33 SD 11.80, therapist 2 17.45 SD 10.72, P = 0.39).

Therapist 1 achieved a higher satisfaction score post-treatment (CSQ therapist 1 31.37 SD 1.19, therapist 2 26.87 SD 4.05, P = 0.02 n=16).

Findings reflect the improved competence scores found on the BACS for therapist 1; however it is of note that this was a *post hoc* analysis with no power calculation on a non- randomised allocation (to each therapist). As such, the findings are informative but must be treated with caution.

4.5.14 Effect size findings mapped against meta-analysis

Results from this study were incorporated into meta-analysis findings reported in Chapter Three. With the completer results added, the overall effect size of BA compared to controls increased slightly with a SMD of −0.73 95% CI −1.02 to −0.44 P<0.001 observed against the reported a SMD of −0.70 CI −1.00 to −0.39 P<0.001 in the meta-analysis. With LOCF data added, the overall effect size remained unchanged, with a SMD of −0.70 (95% CI −0.98 to −0.42). These observations would indicate that BA delivered by non-specialists in this study achieved similar results to those achieved in studies using specialist therapists (see Figure 41 for visual comparison).

Table 22**:** Analysis of Outcome scores of BA vs. Usual Care at 3 month assessment

	Pre mean (SD)		Post mean (SD)		Mean difference (95% CI)	*P value*	Standardised mean difference(95%CI)
	BA	usual care	BA	usual care			
BDI-II Completers	35.57 (9.60) n=23	35.08 (9.60) n=24	11.93 (11.84) n=16	27.40 (14.01) n=22	−15.65 (−24.41 to −6.90)	0.001	−1.15 (−1.85 to −0.45)
ITT [a]			N/A n=23	N/A n=24	−15.78 (−24.55 to −7.02)	0.001	N/A
WASA Completers	26.39 (7.30) n=23	25.13 (7.70) n=24	11.12 (9.64) n=16	22.68 (10.07) n=22	−11.56 (−18.33 to−4.79)	0.001	−1.14 (−1.84 to −0.45)
ITT [a]			N/A n=23	N/A n=24	−11.12 (−17.53 to −4.70	0.001	N/A
CSQ Completers	n/a	n/a	29.13 (3.70) n=16	24.32 (3.96) n=22	4.81 (2.23 to 7.38)	0.001	N/A

[a] ITT using multiple imputation incorporates repeated imputed values of post scores for analysis hence individual post mean (SD) values are not produced.

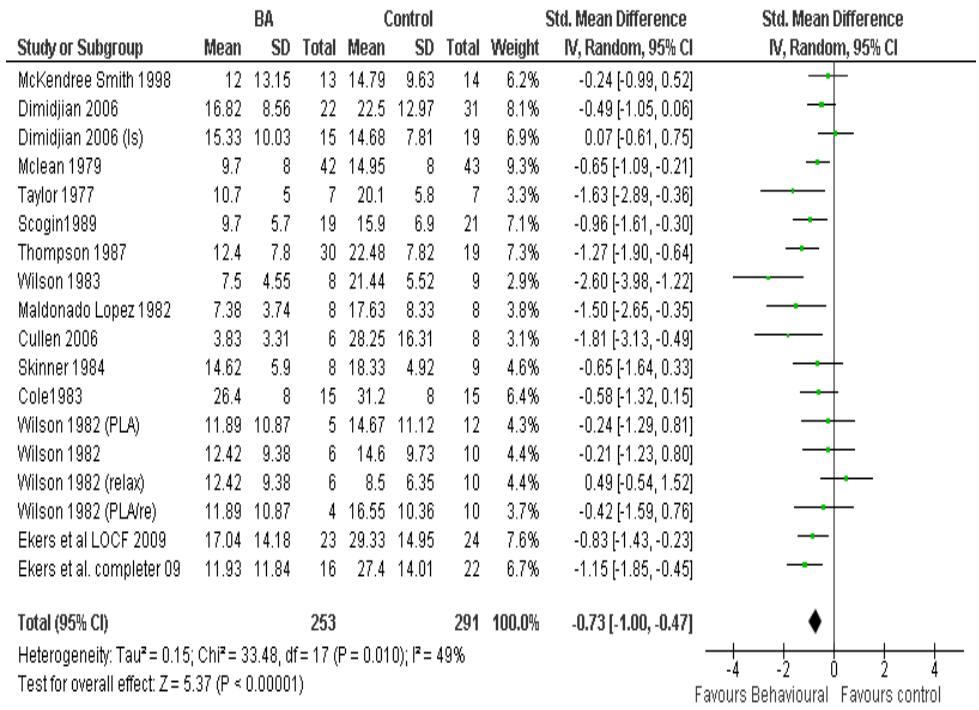


Figure 41: Forest plot of BA vs. Control with Ekers et al study included

(N.B. forest plot includes both completer and LOCF data sets. This leads to double counting. The graph is presented for visual inspection of non specialist BA compared to other studies. SMD calculations in text should be used for statistical results as these are based on analysis of each finding separately).

4.6 Economic evaluation

4.6.1 Behavioural activation by the non-specialist intervention costs

Individual therapist training costs for the five-day course were £641.55 and for delivery of the 12-session BA protocol £219.96. Training and supervisor costs for the five-day training and delivery of 40 hours of clinical supervision were £3059. Based upon scenario A, costs of delivery of a 12-session BA treatment, including therapist training and supervision, was £247 per participant, and for scenario B, £272.52 per participant. A breakdown of intervention cost calculations is presented in Table 23.

.

Table 23: BA intervention cost calculations

BA costs training	costs
Therapist mid-point £14.10 + 30%=£18.33	Training for 35 hours =£641.55
Trainer and supervisor basic costs £42.79 + 30%=£55.62	Training 35 hours at £55.62=£1947 Trained 10 people hence per therapist £194.7 Supervision per therapist 1 hour each 2 weeks for 40 weeks (one year- leave and sickness etc.)=20 hours of supervision per therapist £1112-supervsior, £366-therapist
Trainer costs divided by 195 treatments over three years assuming depression treatment alone	£1.00 per patient (rounded up from 0.99)
Trainer costs divided by 98 treatments over three years assuming half treated cases depression	£1.99 per case
Supervision costs one hour per fortnight divided by 65 cases per year (supervisor + therapist)	£17.11 + £5.64=£22.75 per case
Supervision costs one hour per fortnight divided by 33 cases per year	£33.70 + £11.11= £44.01 per case
BA costs for training-based upon mid-point Band 5 for 35 hours	£641.55

BA training cost estimate per patient based upon 65 completed treatments per year (HI trained estimates from local IAPT specification) over three years 195	£3.29
BA training cost estimate per patient based upon 65 completed treatments per year (HI trained estimates from local IAPT specification) with 50% over 3 years depression related 98	£6.56
Total add on costs to BA Treatment per patient (based upon195 treatments over three years total number treated per training)	£1.00+£22.75+£3.29= £27.04 per case
Add-on cost to BA for training and supervision (based upon 98 treatments over three years treated per training)	£1.99+£44.01+£6.56= £52.56 per case
Cost per patient therapist time for treatment (assuming received all 12 sessions)	£219.96
Cost of BA treatment plus training plus supervision (based upon 195 treatments over three years assumptions)	**£219.96 + £27.04= £247**
Cost of BA treatment plus training plus supervision (based upon 98 treatments over 3 years assumptions)	**£219.96 + £52.56= £272.52**

4.6.2 Health care costs

Complete resource use costs, excluding intervention costs, were available for all 47 recruited participants for both the six-month baseline and the three-month intervention phase (see Table 24). There were no differences in costs observed at baseline between the BA and usual care (BA mean £1050.12 SD = £1907.75, usual care mean £899.31 SD=£1131.33) or during the intervention phase (BA mean £336.01 SD = £434.48, usual care mean £412.81 SD=£380.29).

4.6.3 Cost comparison between study arms

With BA delivery costs included a one-way between groups analysis of covariance was conducted on health service costs during the intervention

phase, with baseline cost used as the covariate. In scenario A, after adjusting for baseline costs BA was £149.24 more costly than usual care (95% CI −£354.82 to £56.34 P= 0.15). In scenario B, after adjusting for baseline costs BA was £174.74 more costly than usual care (95% CI -£380.34 to £30.82 P= 0.09) (see Table 25).

Table 24: Service use and cost BA vs. Usual Care

Cost Category	Treatment - Pre (n=23)		Control - Pre (n=24)		Treatment - Post (n=23)		Control - Post (n=24)	
	Volume per patient Mean (s.d.)	Cost per patient Mean (s.d.)	Volume per patient Mean (s.d.)	Cost per patient Mean (s.d.)	Volume per patient Mean (s.d.)	Cost per patient Mean (s.d.)	Volume per patient Mean (s.d.)	Cost per patient Mean (s.d.)
Primary Care NHS (General)								
GP	6.13 (4.911)	£214.57 (171.887)	4.92 (2.812)	£172.08 (98.410)	248 (1.648)	£86.74 (57.675)	2.67 (2.869)	£93.33 (100.419)
GP telephone	0.83 (2.498)	£17.35 (52.463)	0.46 (1.285)	£9.63 (26.978)	0.17 (.388)	£3.65 (8.139)	0.29 (0.624)	£6.13 (13.106)
GP out of hours	0	£0	0.13 (.338)	£8.13 (21.959)	0	£0	0.08 (.282)	£5.42 (18.351)
GP home visit	0.09 (.288)	£10.17 (33.708)	0.04 (.204)	£4.88 (23.883)	0	£0	0	£0

Nurse	1.74 (2.684)	£19.13 (29.519)	1.83 (2.297)	£20.17 (25.265)	0.74 (1.322)	£8.13 (14.539)	0.96 (1.429)	£10.54 (15.718)
Nurse telephone	0	£0	0.17 (.381)	£1.11 (2.543)	0	£0	0.04 (.204)	£0.28 (1.364)
Nurse home visit	0.04 (0.209)	£0.87 (4.170)	0	£0	0	£0	0.08 (0.282)	£1.67 (5.647)
Health Care Assistant	0.35 (0.647)	£1.56 (2.906)	0.17 (0.482)	£.75 (2.162)	0.13 (0.458)	£0.59 (2.055)	0.25 (0.676)	£1.12 (3.034)
Health Visitor (HV)	0	£0	0.21 (1.021)	£8.33 (40.825)	0	£0	0	£0
HV In Surgery	0	£0	0.08 (0.408)	£6.75 (33.068)	0	£0	0.08 (0.408)	£6.75 (33.068)
HV Telephone	0	£0	0.08 (0.408)	£0.95 (4.638)	0	£0	0.13 (0.612)	£1.42 (6.957)

HV Home	0	£0	0	£0	0	£0	0.08 (0.408)	£3.33 (16.330)
NHS Direct	0	£0 0	0.17 (0.482)	£4.17 (12.039)	0	£0	0.08 (0.408)	£2.08 (10.206)
Walk-in Centre	0.09 (0.417)	£2.26 (10.843)	0.04 (0.204)	£1.08 (5.307)	0	£0	0	£0
Secondary Care NHS (General)								
Dietician	0.04 (.209)	£1.48 (7.089)	0	£0	0	£0	0.13 (0.612)	£4.25 (20.821)
Ambulance	0	£0	0.04 (.204)	£10.00 (48.990)	0	£0	04 (.204)	£10.00 (48.990)
Rapid Response	0	£0	0	£0	.04 (.209)	£8.70 (41.703)	0	£0

Ambulance	0	£0	.04 (.204)	£10.00 (48.990)	0	£0	04 (0.204)	£10.00 (48.990)
A&E	0.30 (0.559)	£28.30 (51.968)	0.25 (0.676)	£23.25 (62.837)	0	£0	0.08 (0.282)	£7.75 (26.257)
In-Patient	0.13 (0.458)	£63.00 (221.067)	0.29 (1.429)	£140.88 (690.144)	0	£0	0	£0
Out-Patient	1.35 (2.516)	£169.83 (316.961)	1.21 (2.431)	£152.25 (306.344)	0.61 (1.234)	£76.70 (155.432)	0.71 (1.367)	£89.25 (172.199)
Day-Patient	0.43 (1.376)	£277.39 (877.866)	0.13 (0.448)	£79.75 (286.097)	0.09 (0.288)	£55.48 (183.810)	0.08 (0.282)	£53.17 (180.126)
Alternative Medical Practitioner	0	£0	0.25 (1.225)	£2.86 (13.999)	0	£0	0.04 (0.204)	£0.48 (2.333)
Radiographer	0.65 (3.128)	£10.43 (50.043)	0.04 (0.204)	£0.67 (3.266)	0	£0	0	£0

Phlebotomist	0	£0	0.21 (0.658)	£0.94 (2.961)	0.04 (0.209)	£0.20 (0.938)	0.13 (0.612)	£0.56 (2.756)
Physiotherapist	0.30 (0.876)	£12.17 (35.027)	0	£0	0.04 (0.209)	£1.87 (8.966)	0.04 (0.204)	£1.79 (8.777)
Occupational Therapist	0	£0	0.13 (0.612)	£5.38 (26.332)	0.04 (0.209)	£1.87 (8.966)	0.04 (0.204)	£1.79 (8.777)
Total physical Medication prescribed (a)	5.70 (8.657)	£44.70 (94.428)	9.29 (12.791)	£53.36 (120.437)	2.91 (5.567)	£28.52 (83.066)	2.96 (5.344)	£29.57 (70.266)
Sub-total Primary Care		£254.91 (239.130		£238.02 (140.318)		£99.11 (64.610)		£132.07 125.673
Sub-total Secondary Care		£562.61 (1431.361)		£415.96 1023.698		£144.80 (278.560)		£169.04 (277.616)
Primary Care NHS: Mental Health Services								

Link Worker	0.87 (1.392)	£62.61 (100.201)	0.88 (1.424)	£63.00 (102.513)	0	£0	0.63 (1.715)	£45.00 (123.459)
Graduate Mental Health Worker	0	£0	0.42 (2.041)	£9.38 (45.928)	0	£0	0	£0
Counselling	0.22 (0.736)	£9.13 (30.906)	0.17 (0.816)	£7.00 (34.293)	0.04 (0.209)	£1.83 (8.758)	0	£0
Post Natal National Helpline	0	£0	0.04 (0.204)	£0.10 (.480)	0	£0	0	£0
Psychologist	0.04 (0.209)	£3.26 (15.639)	0	£0	0	£0	0	£0
Psychiatrist	13 (0.626)	£42.00 (201.425)	0	£0	0.13 (0.458)	£42.00 (147.378)	0	£0
Psychiatrist telephone consultation	0	£0	0	£0	.04 (.209)	£1.66 (7.944)	0	£0

Liaison Psychiatry	0	£0	0.08 (0.408)	£19.25 (94.305)	0	£0	0	£0
Sub-total (2)		£117 (225.73)		£98.73 (152.19)		£45.49 (154.20)		£45 (123.46)
Secondary Care NHS Services:								
Nurse specialist	0.04 (0.209)	£.70 (3.336)	0	£0	0	£0	0	£0
Community Psychiatric Nurse	0.17 (0.491)	£12.17 (34.372)	0.46 (1.444)	£32.08 (101.080)	0	£0	0.21 (0.658)	£15.00 (47.376)
Mental Health Medication prescribed (b)	3.09 (4.833)	£17.62 (44.983)	2.50 (2.207)	£10.91 (15.644)	1.43 (1.950)	£10.66 (27.887)	1.71 (1.876)	£5.90 (8.459)
Total Medication (a) + (b)		£62.32 (19.15)		£64.27 (30.08)		£39.18 (12.62)		£35.47 (16.74)
Sub-total (3)		£12.87 (34.28)		£32.08 (101.08)		£0 (.000)		£15 (47.38)

Community/Social Services:								
Social Work	.04 (.209)	£6.09 (29.192)	.17 (.816)	£23.33 (114.310)	0	£0	0	£0
Community Alcohol Service	0.04 (0.209)	£3.91 (18.766)	.00 (.000)	£0 (.000)	0	£0	0	£0
Employment Advisor	0.83 (2.933)	£15.70 (55.734)	0.13 (0.338)	£2.38 (6.419)	0.26 (0.752)	£4.96 (14.284)	0.29 (1.042)	£5.54 (19.792)
Citizens Advice Bureau	0.13 (0.458)	£1.24 (4.348)	0.12 (0.338)	£1.19 (3.209)	0	£0	0.08 (0.282)	£0.79 (2.682)
Relate	0	£0	0.50 (2.449)	£4.75 (23.270)	0	£0	0.17 (0.816)	£1.58 (7.757)
Legal/Debt Advisor	0.09 (0.417)	£0.83 (3.962)	0.13 (0.448)	£1.19 (4.260)	0	£0	0.37 (1.173)	£3.56 (11.140)
Benefit/Housing Advisor	0.04 (0.209)	£0.83 (3.962)	0.37 (1.279)	£7.13 (24.301)	0.09 (0.288)	£1.65 (5.474)	0.17 (0.482)	£3.17 (9.149)
Job Centre	0.04 (0.209)	£0.83 (3.962)	0.54 (2.449)	£10.29 (46.533)	0.04 (0.209)	£0.83 (3.962)	0.08 (0.408)	£1.58 (7.757)

Sub-total (4)		£29.43 (68.22)		£50.26 (121.62)		£7.44 (17.88)		£16.22 (39.47)
Total NHS (1+2+3)		£958.38 (1836.62)		£784.80 (1065.85)		£289.39 (400.29)		£361.11 (356)
Total Mental Health (2+3)		£129.87 (226.74)		£130.80 (175.25)		£45.48 (154.20)		£60 (155.54)
Total Cost of Care (1+2+3+4)		£987.80 (1866.43)		£835.03 (1113.67)		£296.82 (400.75)		£377.34 (374.71)
Total Cost of Care (1+2+3+4) and Medication Cost (a+b)		1050.12 (1907.75)		899.31 (1131.33)		336.01 (434.48)		412.81 (380.29)

4.6.4 Health State comparison between study arms

Preliminary checks were conducted to ensure there was no violation of assumptions of linearity or homogeneity of regression slopes. There was no significant interaction between Health state at baseline and allocation (P=0.85). A one-way between groups analysis of covariance with participants' health state pre-treatment used as the covariate showed a significant difference at post-treatment in favour of BA (n=16) over usual care (n=22) of 0.24 (95% CI 0.052 to 0.437 P=0.01).

Intention to treat analyses with multiple imputation showed a mean difference on post health state scores of 0.20 in favour of BA (95% CI 0.01 to 0.39, P= 0.04) with all randomised subjects (BA n=23, usual care n=24) included in analysis (see Table 25).

4.6.5 Economic deterministic results

In scenario A, with BA £149.24 more costly than usual care and a health state difference of 0.20 generates a cost per QALY earned through BA valued at £2985 (0.20/4 = 0.05 QALY gain per year, £149.24 × 20 = £2984.8 per full QALY).

In scenario B with BA £174.74 more costly than usual care, the resulting ratio is £3495per QALY (0.20/4 = 0.05 QALY gain per full year, 174.74 × 20 = £3494.80 per full QALY).

A cost per point reduction on the BDI-II earned through BA is calculated at £9.45 and £11.04 for scenarios A and B respectively.

Table 25: comparison of cost and health state at 3 month follow up

	Baseline Mean (SD)		Post Mean (SD)		Mean difference post (95% CI)	*P value*
	BA	Usual Care	BA	Usual Care		
Costs	£1050.12 (1907.75) n=23	£899.31 (1131.33) n=24	£608.53(434.48)[a] n=23	£412.81 (380.29)[a] n=24	£174.74 (-£380.34 to £30.82)[a]	0.09
Health State						
Completer	0.40 (0.24) n=23	0.46 (0.35) n=24	0.79 (0.24) n=16	0.58 (0.39) n=22	0.24 (0.052 to 0.437)	0.01
ITT			N/A[b] n=23	N/A[b] n=24	0.20 (0.01 to 0.39)	0.04

[a] Using scenario B costs as these represent the most conservative analysis

[b] Multiple imputation does not report individual mean (SD) scores

4.6.6 Exploration of uncertainty of economic evaluation

One 1000 bootstrap replications were conducted for scenarios A and B which were mapped on a cost effectiveness plane see (Figure 42 and Figure 43 respectively). To examine the uncertainty of the results these findings were placed upon a cost effectives and acceptability curve. At a threshold value of £20,000/QALY there was a 97.7% probability when adopting scenario A that BA delivered by generic mental health workers is more cost effective than usual care. Adopting the same threshold value there was also a 97% probability when adopting scenario B that BA delivered by generic mental health workers is more cost effective than usual care. At a value of £30,000/QALY the probabilities converged at 98.9% for both scenarios (see Figure 44). Results suggest an incremental cost effectiveness ratio of £5006 per QALY for scenario A and £5756 per QALY for scenario B. Both scenarios indicate that the additional cost of BA over usual care per QALY gained is less than the current UK accepted value of £20,000 (National Institute for Health and Clinical Excellence 2008).

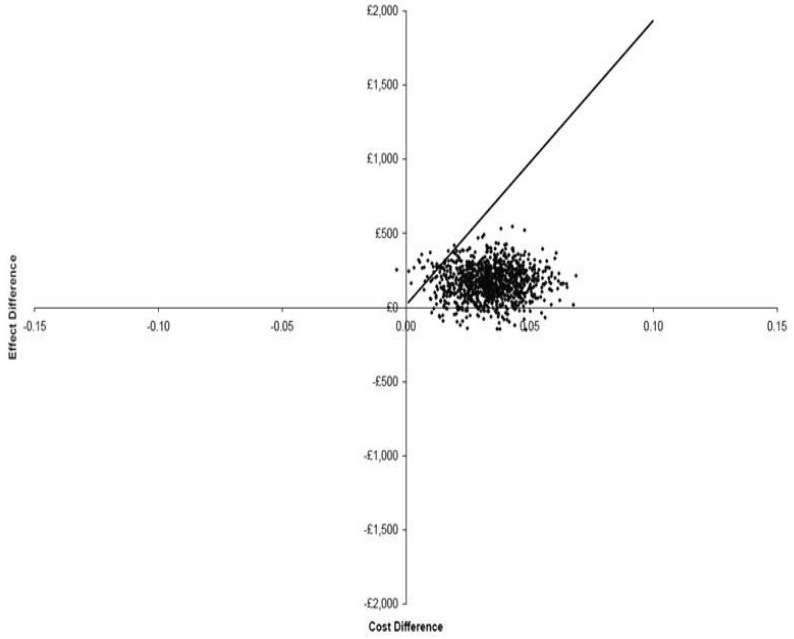


Figure 42: Scenario A Cost effectiveness plane for BA vs. Usual Care (reference line indicates threshold value of £20000 per QALY)

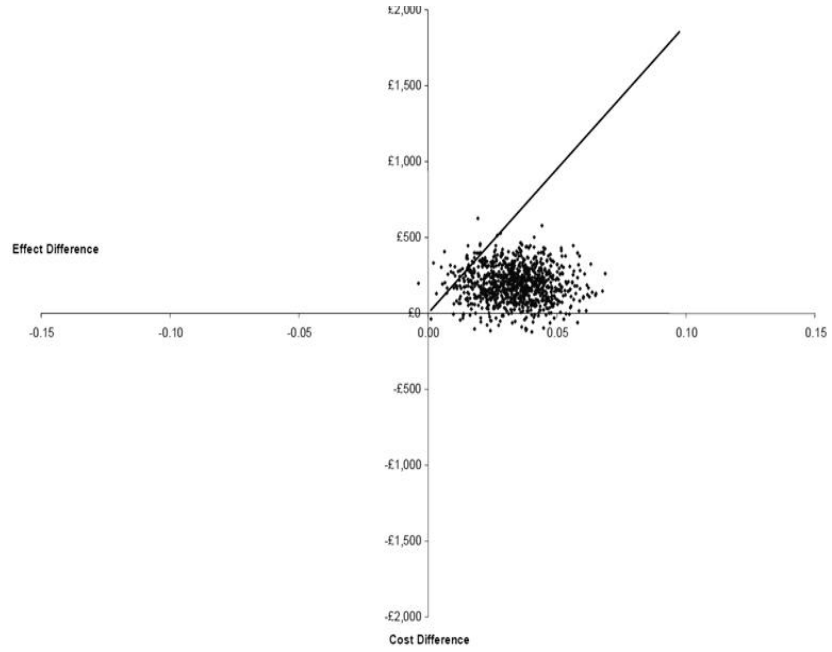


Figure 43: Scenario B Cost effectiveness plane for BA vs. Usual Care (reference line indicates threshold value of £20000 per QALY)

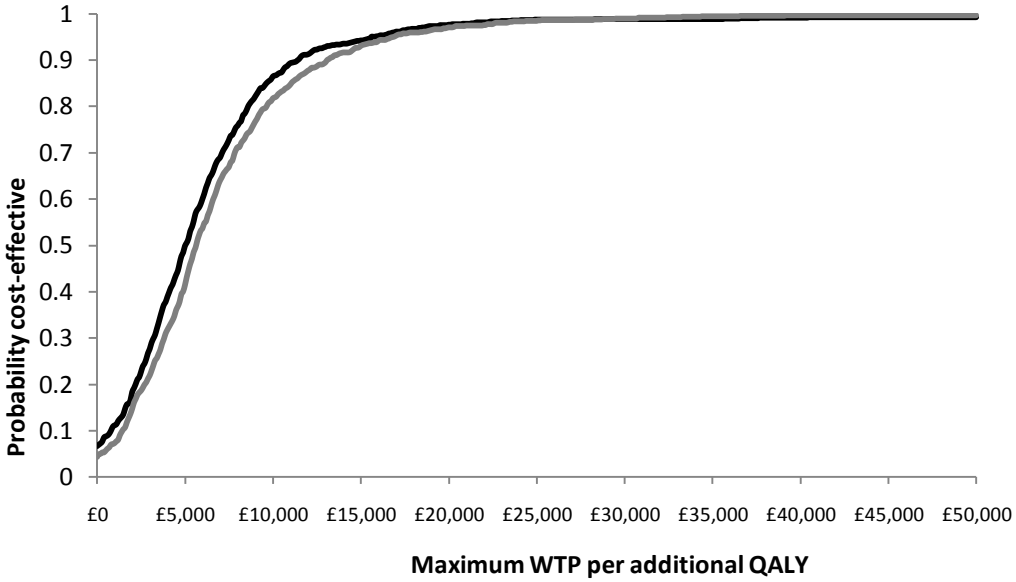


Figure 44: Cost Effectivenss and acceptability curve of BA vs. Usual care scenario A and B

4.7 Discussion

4.7.1 Main findings

In this study behavioural activation was found to be an effective individual therapy for depression compared to usual care when delivered by non-specialist mental health staff. This approach had significantly greater benefits over primary care management in terms of our primary outcome of depression symptom and our secondary outcomes of functioning and satisfaction. Detailed analysis of reliable and clinically significant change and response and recovery criteria also demonstrated significant levels of difference in favour of the BA group. Results appeared to reflect effect sizes found in our previous meta-analysis in which all studies had used therapists trained to the level of clinical psychologist or psychotherapist. Such an *ad hoc* approach to comparing specialist against non-specialist delivery of BA is not conclusive; this would require a randomised controlled trial. It does however provide a first observation based on data observed in practice of the potential benefit that the wider dissemination of BA may offer.

This study represents an important addition to the BA evidence base. As discussed throughout this thesis repeated publications have made reference to the potential BA offers due to its relative simplicity (Jacobson et al. 1996, Jacobson and Gortner 2000, Martell et al. 2001, Dimidjian et al. 2006, Kanter et al. 2010). This potential in was explored in the meta-analysis finding, reported in Chapter 3. No previously published randomised study that had scientifically examined this assertion. On this basis the findings reported here will be of interest to clinicians and researchers as they present new knowledge in this important area. While the systematic review illustrated that BA offers clinical benefit (Ekers et al. 2008) we were not clear if this was the result of the senior therapists therapeutic skill and experience or the BA approach (Chapter 3). If it were the therapeutic experience and skill of the therapist that explained outcomes, this would indicate limited benefit of increasing the capacity of effective psychological therapies for depression. Therefore while small this

study would appear to be the first randomised controlled trial to test directly the feasibility of the parsimony-dissemination assertion put forward by Jacobson and colleagues well over a decade ago (Jacobson et al. 1996). Results would also appear timely related to evidence emerging from the Improving Access to Psychological Therapies (IAPT) initiative in the UK where much attention is given to increasing the numbers of people treated (Clark et al. 2009). Results also suggest that BA by the non-specialist may offer an economically viable and rapid alternative to CBT which is currently the mainstay intervention of IAPT services for depression. These results, which represent the first economic evaluation of BA appear very promising and are discussed in more detail later in this section. With depression being cited as a primary problem in up to 95% of those referred to IAPT (Clark et al. 2009) or 30% by those diagnosed within the service (Glover et al. 2010) a 'semi high intensity' treatment that lends itself to rapid training and is effective would appear attractive. Commonly such services experience blocks in the stepped care system at access to step 3 (Glover et al. 2010) which is precisely where this BA intervention may sit. Training additional CBT therapists takes one year and produces therapists on senior clinical bandings. In this context results from this study suggesting BA by non specialists may offer both clinical and cost effectiveness should be of interest.

Mental health nurses were used in this study as they represent the largest mental health workforce that require improved training in psychological interventions (Department of Health 2006). If this workforce can be trained then the possibilities for dissemination would therefore be substantial. Those trained in BA for this study appeared to find the training both acceptable and an effective approach to developing their skills. All trainees gave an exceptionally positive score on the TARS well above the 70-80% normally achieved (Milne et al. 2000). Further dissemination outside of mental health nurses specifically may offer additional benefit if results can be maintained. Large, adequately powered trials are required to examine this area further and to replicate these findings.

4.7.2 Clinical outcomes

A large effect size of −1.15 on depression symptom level post-treatment favouring BA against usual care using intention to treat analysis was found. This compares favourably to an overall effect size of −0.70 of twelve studies (459 participants) comparing BA to controls using experienced therapists in our previously reported meta-analysis (Ekers et al. 2008). The difference between BA and usual care at post-treatment of between 15 and 16 points on a BDI-II measure is clearly clinically important. It reflects the difference between a minimal depression level score of 13 and below and a severe depression level score of 29 and above. The findings also compare favourably with a review of three studies with 208 participants where CBT has been compared to usual primary care where minimal difference was identified (National Institute of Clinical Excellence 2009). This research may add to the evidence available to NICE when their depression guidance is further reviewed. Currently NICE identified no cost effectiveness evaluation for BA which they reported as a weakness of its evidence base. They recommended BA for mild to moderate depression with 16 to 20 sessions over three to four months. This research adds to knowledge as if now provides the first cost utility analysis of BA with less sessions than previously recommended in a severely depressed sample. The research recommendations of the NICE depression guideline also indicate BA should be directly compared to CBT. This research would support the design of such a study if researchers were to look at non-specialists delivering the BA intervention.

The various clinical significance criteria used in this study further strengthen the results. Differences between the groups were significant or close to significance for all comparisons using reliable and clinically significant change criteria (Jacobson and Truax 1991). These stringent approaches to data analysis provide further insight into the clinical meaning of the results. They use normative data from scales to inform the degree to which participants move from a clinical to a non-clinical group and the number whose changes during therapy are considered clinically meaningful. The results found in this study compare favourably with other

major studies of BA that achieve remission rates of around 50% (Jacobson et al. 1996) although they are below the exceptional and unusually high findings of one important trial (Dimidjian et al. 2006).

Overall clinical findings in this study appear very positive and lend weight to the promise that BA may offer as a simple and effective psychological treatment. There are a number of limitations to the trial outlined later in this discussion section but as an initial exploration of the feasibility of non specialist delivery of BA this study has provided encouraging clinical results.

4.7.3 Functioning

Similar findings in functioning were observed to those found in depression symptom level; whilst both groups improved, in BA improvement was substantial whilst in usual care it was marginal. BA participants on average achieved a 15 point drop on the WASA scale compared to a 2 point drop for usual care. Both groups scored above 20 on this measure pre treatment. This score range reflects severe symptoms and functional impairment (Munt et al. 2002). While usual care remained in that category the BA group reduced to a mean WASA score of 11.12. This is slightly above the subclinical cut off point of 10 identified for this measure (Munt et al. 2002). It would appear from these findings that BA delivered by non-specialists also improves functioning alongside depression severity.
The study findings would support the view that BA may lend itself to improvement of functioning as this is a direct intervention target aimed at the reintroduction of positively rewarding environmental contingencies (Kanter et al. 2010). BA focuses specifically on the 'outside in' approach of activating and accessing positive reinforcement which as a result increases functional behaviour. Those participants undertaking BA would be encouraged to reintroduce, or to develop new, behaviours that are personally viewed as meaningful and valued despite internal feelings. New behaviours are then paired with positive reinforcement and hence occur with increased frequency. In contrast, other standard treatments seek to

change internal deficits or to change internal feelings prior to functional change. This approach of seeking internal improvement prior to activity reflects the pattern of 'I will do it when I feel better' is often seen in people with depression (Martell et al. 2001). It may be that as BA is directly targeting the problems measured on the WASA as a component of the intervention this to some degree accounts for the large improvement in functioning in the BA group compared to usual care. It is of note that functioning improvement is consistent with symptom level findings, it would appear BA is not simply getting people more active whilst feeling as depressed. This supports the theoretical assumption that depression is unlikely to be paired with a high level of contextual positive reinforcement (Kanter et al. 2010). It was beyond the scope of this small study to test further the interpretation of these findings. Future researchers may be able to consider this rationale for BA's effectiveness in more detail with larger sample sizes. Following the MRC complex intervention guidance the feasibility and piloting in this research has identified this as a potential area of interest. With the introduction of relevant process measurements and qualitative analysis larger studies could explore this area further.

4.7.4 Satisfaction

Satisfaction with BA appeared extremely good with a mean of score of 29 out of a possible 32. This is significantly better than usual care which received a reasonably positive evaluation score of 24. Across six United States based studies with 7838 participants using the CSQ8 the mean satisfaction level was 27.17 (Attkisson and Greenfield 2004). In the UK the CSQ 8 has been used in trials of both face-to-face and telephone based CBT with scores of 29 reported as indicative of very high satisfaction (Lovell et al. 2006). Patients are able to distinguish between clinical benefit and satisfaction (Larsen et al. 1979) so it would appear from our results that BA may offer both a satisfactory and effective intervention. It is of note however that only those completing post-treatment evaluation completed the CSQ8 forms in this study. We were unable to use any method of imputation as no previous scores were collected to inform

calculations. No measure of satisfaction was made at baseline or at one and two month data collection points. This weakens our results in this area and future research should consider taking satisfaction measures throughout treatment at key stages to allow for ITT analysis in the event of dropout.

4.7.5 Cost utility analysis

Clinical effectiveness is necessary but not sufficient for policy makers who need to be sure that innovative treatments represent good value for money. To my knowledge this is the first study to examine cost effectiveness of BA and in particular this innovative approach to its delivery. Mental health economic analysis should be conducted to allow consideration within a general decision making context (Bosmans et al. 2008, Barton et al. 2009). NICE suggests that a QALY gain valued below £20,000 provides good value for money across the health care spectrum (National Institute for Health and Clinical Excellence 2008). It was in this context the economic evaluation of BA delivered by non-specialists was conducted using generic health outcome measures as recommended (Barton et al. 2009, Sapin et al. 2004, Bosmans et al. 2008). BA delivered QALY gains significantly below accepted threshold values using the EQ5D generic health state measurement tool. Changes in health state were consistent with clinical, functioning and satisfaction benefits. The results indicate BA may offer lower cost per QALY or point reduction on the BDI-II than brief problem solving (Kendrick et al. 2006) or online CBT (Hollinghurst et al. 2010, McCrone et al. 2004), interventions also aimed at increasing accessibility. Both the baseline severity of the participants in this study and the lower grade staff used for treatment may have contributed to the strong indication of cost effectiveness we found when benchmarked against other trials.

These gains were calculated using a short follow-up period of three months. At this point we identified a QALY difference of 0.20. In order not to overestimate the benefit of the intervention this difference was divided by four to provide a quarter of the QALY gain, a value of 0.05 and based our calculations on this figure. This approach was sufficient to demonstrate

QALY gains at a cost which falls below the threshold values set out by NICE. It makes no assumptions about what might happen to the health status of patients beyond the three month follow-up period. Previous studies have found BA to provide lasting benefit (Dobson et al. 2008, Gortner et al. 1998) with between 60% to 70% of those who show improvement maintaining gains a year following treatment. A similar pattern would be expected to emerge in this study however the analysis was restricted to observed results at 3 months building in no assumptions regarding any difference between groups following this point. If BA when delivered by non-specialists does provide any lasting benefit beyond the completion of therapy this would increase the QALY gain we used beyond 0.05 and subsequently reduce the cost per QALY found.

The economic findings were also modelled using a senior grade of therapist, Agenda For Change band 8c, to train and supervise workers and assumed all sessions were attended. This again reflects a conservative baseline from which further studies can design their economic models. Considerably less expensive staff are commonly used in the NHS to deliver training and supervise psychological treatments and not all patients needed all treatment sessions. If built into calculations the cost of a course of BA would have dropped, again reducing the cost assigned to a QALY gain. Costs were not based on actual number of sessions attended to prevent an underestimation of costs given the small sample size. The small numbers was likely to result in the study being underpowered to estimate attendance rates. Future studies with more participants may supply more clarity on the number of required sessions for optimal benefit and may provide greater precision on the true cost effectiveness of BA compared to usual care.

4.7.6 Recruitment issues

Recruitment rates to the study were well below those anticipated in planning. Despite developing communication links and regularly updating practices and primary care mental health teams on recruitment progress the rate of referral to the study was low. Monthly updates were sent to specific contacts within practices with their individual referral rate benchmarked

against the overall rate. This approach was unsuccessful in increasing recruitment rates even when supported by personalised email communication. The study was rolled out to additional general practice sites to compensate for this poor response. While the number of participants required as per our power calculations was achieved (23 in each arm) it would have been beneficial to have recruited to the target set in initial planning of 70 participants (35 in each arm). This would have allowed analysis of completers to have been above the 23 required in each arm.

Explanations for these difficulties in recruitment are unclear. General practice clearly has many patients experiencing depression who would have been suitable for this study. The study offered a structured psychotherapy with a waiting time, even in the control arm, well below the local service average. It is possible that the experience of referring participants who did not meet the study inclusion criteria was an aversive experience for their general practitioner. As this was the most common cause for potential participants being excluded, 13 out of 41 referrals (31%), this may have reduced their desire to pass subsequent suitable participants to the trial. While this is difficult to confirm it may have added to the existing problems encountered in encouraging general practitioners to refer to research (Mason et al. 2007). The exclusion rates were not particularly high when compared with the 40-45% exclusion found in other studies of primary care depression in the UK (Richards et al. 2008, McCrone et al. 2004) or the 32% found in previous BA for depression studies (Dimidjian et al. 2006). Unlike those studies however the study was reliant solely on the behaviours of others such as GPs and mental health staff to refer to the trial. As this study did not have external funding it could not access the resource of clinical trial officers from research networks to support recruitment process. Therefore recruitment in this study was solely reliant on the behaviour of referrers which is known to be unreliable (Hunt et al. 2001, Mason et al. 2007).

It is clear that future studies may require of a modified recruitment procedure in order to be successful. The use of patient electronic records

can be used to identify suitable participants through repeated in practice searches based upon depression coding. Potential participants can then be contacted by practice staff via letter and asked if they are willing to be contacted by the research team. This would allow study information sheets and consent forms to be forwarded directly to the participant and not be considered overly intrusive. Those that do not respond could be reminded by telephone as this has been seen to improve recruitment rates (Watson and Torgerson 2006). This approach has the benefit of removing variance in referral practice and hence would minimise the risk of selection bias and improve external validity. Such approaches have been used in trials and have improved recruitment rates (Richards et al. 2006); they would minimise the impact of the behaviour of those referring on study recruitment or validity. Externally funded trials would also have access to clinical study officers to monitor the process closely and to perform searches, which was outside the scope of this small unfunded trial.

4.7.7 Participants

Participants in this study had high baseline depression and functional impairment, scoring 35 on the BDI-II and 25 on the WASA. This score represents severe depression with high levels of disability equal to or greater than those studies included in the meta-analysis outlined in Chapter Three. The main diagnosis in allocated to those in the study by the CSIR was moderate or severe depression and these problems were of long duration, on average 3.5 years. The study therefore offered treatment to a cohort of participants with a higher proportion of severe problems and with longer mean duration of symptoms than those commonly seen in primary care therapy settings (Clark et al. 2009). One of the criticisms of randomised controlled trials is that the populations treated often do not reflect clinical reality. This study would appear to go some way towards that in the group treated which contributes to the external validity of findings.

4.7.8 Randomisation and allocation concealment

As outlined in the methods section of this chapter randomisation and in particular allocation concealment are fundamental aspects of trial quality (Altman and Bland 1999b). Owing to financial constraints, this study was unable to use a completely independent clinical trials unit to perform this task. Such services are ideal as they hold randomisation lists and automatically notify therapists GPs and trial coordinators of allocation. I was not able to use such services due to the costs associated and the unfunded nature of the study. It would have posed a problem of internal validity in the study, with the potential for selection bias, if the randomisation had been conducted or able to be influenced by those involved with the study. The approach of having a computer-generated stratified allocation sequence generated independently, as per recommended methods (Altman and Bland 1999a), and held by a medical secretary independent of the study team was a balanced compromise that maintained quality and allocation concealment. At no point in the study was the list able to be seen by those involved as it was held in an envelope in a locked location and viewed only by the secretary when asked for a participant allocation.

The randomisation would appear to have been successful as the groups appeared equal in terms of age, severity, duration of symptoms and antidepressant use.

4.7.9 Therapist effects

Therapists used in this study were representative of the vast pool of generic mental health workers in that they had no previous therapy training, were relatively recently qualified and were employed at the base level of registered psychiatric nurses. Two therapists were used and a consistent approach to BA was ensured by developing a structured 12-session protocol, assessing competence post training and providing fortnightly supervision. It is of interest that while overall measures on the BACS indicated a reasonable level of competence in the delivery of BA there was clear variation between therapists. These observations are reflected in

individual outcome scores on the BDI-II observed at post-treatment. So while it would appear based upon independent evaluation with very little training it is possible overall to equip staff with the skills to deliver BA effectively, there may be some variation to how this is delivered in practise between each therapist. This is not a new finding: in an early study comparing BA with CT there was clear variation between therapists scores on an established CT rating scale (Jacobson and Gortner 2000). Agreement on therapist competence level is hard to establish even when reviewed by those experienced in using the tool. In this study a competency scale was used that is new and unfamiliar and not currently supported by peer reviewed validation publications.

There are several major risks in interpretation of these findings. This examination is a *post-hoc* analysis. The study was not powered to detect difference on the BACS nor any between-therapist variations. It is very likely that this would require a much larger trial and more background information on which to base any such calculations. Participants were not randomised between therapists, and in such a small trial we cannot with any confidence assume that confounding variables are equally distributed between the groups treated by different therapists. It is possible that therapist 2 received by chance a more challenging and treatment-resistant group of participants who were less amenable to such a structured BA protocol. This may have resulted in the inferior clinical outcome and ability to apply BA techniques in structured manner found with therapist 2. Therefore while it is important to observe the variation in therapist competence and outcome mindfully it cannot be considered definitive. The debate between therapist and intervention effect in RCTs is longstanding (Wampold et al. 1997) and beyond the scope of this study to resolve.

In general I have found BA to be amenable to dissemination in this study albeit with signs that there may be variation in competency and outcome between therapists requiring further examination in a larger study.

4.7.10 Study limitations

This was an exploratory study and there are a number of limitations of note.

Firstly the relatively small numbers of participants and therapists recruited limits the generalisability of our results. Sample numbers were based on power calculations from the meta-analysis, see Chapter 3, suggesting a sample size of 23 in each arm was sufficient to detect previously observed effect sizes. Despite this, it would have been beneficial to have recruited more participants to allow for our completers analysis to reflect these numbers rather than our intention to treat analysis. The small sample sizes may account for the wide confidence intervals found on post-treatment measures. Due to the lack of previous economic analyses of BA, no power calculation was conducted for this cost effectiveness analysis. Power calculations for economic analysis within clinical trials are rare and usually based upon costs and effects separately. If results are directed at cost effectiveness then power calculations should be focussed on this directly (Briggs et al. 2002). The use of sensitivity analysis exploring uncertainty using CEAC goes some way to moderating this shortfall. BA was found to have a 97% probability of offering a cost effective intervention at currently accepted thresholds based upon 1000 bootstrap replications. This is a strong finding using conservative assumptions from our data (no assumption of benefit maintenance post three months and use of expensive staff to train and supervise). Improving recruitment and increasing sample sizes would have allowed more confidence in findings. I have now established baseline data from which future studies can conduct cost effectiveness power calculations based upon recommended approaches (Briggs et al. 2002).

More people dropped out in BA than in usual care. This in itself is not surprising, as BA is an active intervention relying on the person receiving treatment to complete homework on a regular basis. Of those discontinuing treatment, three did so in the first month and three in the second. The

dropout rate in this study (of approximately 30%) is similar to those seen in large data sets of CBT provision in primary care (Clark et al. 2009), so viewed in this context our dropout rate for BA is not unusual. Usual care in contrast does not involve as much investment from the participant and is "nested" within a person's overall health care thereby reducing the likelihood of dropout. Larger sample sizes would have allowed a more precise estimate of this finding, as it is likely this limited sample was underpowered to accurately estimate dropout rates from BA delivered by this workforce. It is possible that BA was considered an aversive treatment by those that dropped out of therapy. This would be in marked contrast to the strong satisfaction found in those with post-treatment measures. Future large scale studies would be better placed to explore this finding further and allow more precise estimates of likely dropout from BA. This is an important consideration as high dropout rate markedly influences a treatment's overall clinical effectiveness.

The lack of follow-up is also a limiting factor in the interpretation of the study. This was due to the limitations of the study in terms of funding and available time scales. Previous studies of BA have demonstrated its durability to be equal to that of other therapies such as CBT (Dobson et al. 2008, Gortner et al. 1998); however such studies have been delivered by experienced therapists. The aim of this study was to explore the feasibility of dissemination of BA; while we found results supporting this I was unable to conduct follow up assessments due to my financial and time constraints. The results showed comparable gains at post-treatment to those of previous BA studies. As those studies demonstrated prolonged improvement comparable to that of CBT, it is reasonable to assume that benefits in this study are likely to follow a similar maintenance pattern. This however has not been monitored and as such requires future testing. It is possible that improvement was due to the 'Hawthorne Effect' , showing response due to being the subject of a study rather than the intervention being studied (McCarney et al. 2008). The Hawthorne effect first identified when methods of increasing productivity were examined in the Western Electrical Company's Hawthorne depot in Chicago in the 1920s and 1930s.

They observed that no matter what method of increasing productivity was used results increased, indicating that it was being the focus of such methods rather than the methods themselves that caused change (Mayo 1993). Whilst the increased focus on participants in the current study was equal in terms of researcher attention the Hawthorne Effect should have been equally distributed in both groups. It may have resulted in an over-inflation of the effect size, both groups showing more improvement than would be anticipated in usual care. While it is difficult to control for this in trials the short follow-up may not have provided sufficient time for these effects to settle, i.e. subjects to become habituated to the stimulus of being studied. The impact of the Hawthorne effect may have been moderated in longer term follow-up; this requires consideration in the interpretation of our results and consideration in the design of further study. It is of note, however, that this study showed comparable differences to other studies (see Figure 41) and a number of those studies used longer follow up demonstrating continued improvement where one would have expected the of treatment to be reduced if the Hawthorne effect had been operating. It is not possible to be sure if and how much the Hawthorne Effect may have influenced our results however the potential impact should be considered.

The lack of follow-up also limits the long term assessment of QALY gains and associated costs. While it is clear that treatment costs in BA would be front loaded in the first three months it is difficult to assess the differential effects on costs and health state of relapse from our methods. In the cost utility analysis the use of a six month retrospective cost measurement at baseline used as a covariate in our calculation of service use costs over the study duration helped redress this issue. Also only the proportion of QALY gained during the three months was used within our calculations, with no modelling of maintenance. An ideal follow-up period of 18 months to 2 years would allow for this limitation to be accounted for in future research. This would then pose ethical considerations for a study adopting usual care as the comparator. If BA offers clear clinical benefit at three months over usual care, following up as randomised would exclude those in the usual care arm from such benefit. While the durability of BA delivered by non-

specialist is an important consideration this may require exploration in a study with an alternative therapeutic intervention.

A self-report measure was used in the analysis of depression symptom level post intervention. This is a source of potential information bias as in psychological therapy trials participants will be aware of their treatment allocation and hence not blind to allocation. This may lead to an over reporting of benefit in those participants receiving the intervention under investigation. It is again difficult to assess the degree of impact, if any, that this limitation places on our results. The BDI-II (Beck et al. 1996) was used as the primary outcome for several reasons. Firstly it is a well validated measure and fell within the financial and manpower limitations of the trial. It was felt that as it had been used very regularly in previous trials it was an acceptable solution to not having grant funding to support more detailed and costly assessor-rated depression outcomes. I also noted that in previous studies identified in the meta-analysis (Ekers et al. 2008) the BDI was the most commonly used measure, used in 16 of the 17 studies analysed. The alternative assessor-rated measure was the Hamilton Rating Scale for Depression (Hamilton 1960). This was used on its own in only one study. Both measures were used in 11 studies and showed minimal difference in outcome between the tools. Sensitivity analysis was used to explore potential impact using our meta-analysis data which suggested that using an assessor rated measure would have minimal impact on the results (BA vs. Control prioritising BDI SMD −0.70 CI −1.00 to −0.39 vs. prioritising HRSD SMD −0.68 CI −0.98 to – 0.38). It is unclear the degree to which this limitation may have influenced findings, but on balance the approach adopted appears justified, as a commonly applied scientifically tested set of measures was used. The inclusion of the CSIR at end-point may have gone some way to moderating this possible limitation also; however this tool is generally used as a diagnostic tool to confirm eligibility, and as such serves its purpose at assessment. Repeating the CSIR at end-point may have been useful, but it is debatable whether it would have provided any additional scientific rigor over a well validated self rated measure of depression such as the BDI-II. While minimising the

number of measures used is advised, as it decreases the chance of a Type 1 error of finding a difference when it is not there the use of both self-rated clinician and blind assessor measures should be given due consideration in future research.

Another issue was the choice of the QALY approach for cost utility analysis. Some previous studies in depression have used the cost of a 'depression free day' based upon depression symptom scores (Katon et al. 2005, McCrone et al. 2004). However, no accepted cost threshold value for 'depression free day' is available, whereas we have an accepted tariff for a QALY (National Institute for Health and Clinical Excellence 2008). While the cost per QALY approach may limit direct comparison with such studies I chose, in discussion with the trial steering group, the more conventional approach that allows BA to be contrasted with other interventions across the health services spectrum, as this is increasingly recommended for mental health economic research (Barton et al. 2009, Bosmans et al. 2008). A basic analysis of cost per point change on the BDI-II was included to allow comparison alongside other research that had used that approach.

4.8 Review of study design and reporting

This study examined the impact of BA by non-specialists compared to usual primary care. The accepted method for such comparisons is the randomised controlled trial; however this methodology is subject to several forms of bias, as outlined in the methods section of this chapter, if not conducted and reported appropriately. The accepted benchmarks to review this methodology are included in the CONSORT statement (Schulz et al. 2010). The trial will be reviewed against these standards below.

CONSORT item 1: title and abstract

Items 1a and 1b indicate the title should clearly report the study as an RCT with a structured abstract. The study is clearly identified as a randomised controlled trial with the second item being included as an abstract of the whole research process at the start of the thesis.

CONSORT item 2: background and objectives

Items 2a and 2b indicate a scientific background and rationale for the study should be included and clear objectives should be set. These criteria have been met as the study was based upon a meta-analysis and the need for investigation has been highlighted. Objectives for the study have been outlined and reviewed in the conclusion section.

CONSORT item : reporting of methods (CONSORT item number included in brackets)

- **Trial design**: the trial is described as a parallel design in methods section (3a). It is also recommended that any changes in methods after trial commencement are reported (3b). This is not relevant to this study, as no changes were made.
- **Participants:**- eligibility criteria are clearly reported and justified in the methods section (4a), and the method of data collection from participants described (4b).
- **Interventions:** interventions are clearly described in methods section with details of staff group used, training approach and

manual provided. This would allow for adequate replication as recommended (5).

- **Outcomes:** outcome measures used are clearly outlined with justification for use in the methods section (6a). There were no changes to these so no requirement was needed to report variation (6b).
- **Sample size:** sample size was justified in the methods section (7a) with details of calculations based upon our meta-analysis results. No interim analysis or stopping guidelines were used, so were not reported (7b).
- **Randomisation:** the methods used to create the randomisation sequence are clearly reported (8a) and the rationale and approach to stratification and use of blocked randomisation reported (8b). The method of allocation concealment is explained in detail, with the computer program referenced (9) and the implementation of this outlined in terms of who generated the sequence, who enrolled and who allocated (10).
- **Blinding:** the approach to blinding is clearly reported in the methods section alongside the challenges this poses in psychotherapy trials (11a). The interventions were clearly different, and descriptions of both have been reported in methods. There was no attempt to blind participants to interventions as in pill placebo trials as this is clearly not feasible, hence reporting similarity of interventions was not relevant (11b).
- **Statistical analyses:** these are clearly outlined in methods with approaches to missing data outlined *a priori* (12a and 12b).

CONSORT item: reporting of results (CONSORT item number included in brackets)

- **Participant flow:** the numbers of participants assessed, excluded, included, randomised, received treatment and were assessed or lost to follow-up at post-treatment was clearly reported including a diagram as recommended (13a and 13b). Recruitment study dates

are reported, as is trial duration (14a); the trial ran its duration, so reporting reasons for stopping (14b) was not relevant.

- **Baseline data:** a table is included as recommended (15).
- **Numbers analysed:** is clearly reported in text and tables in results section as recommended (16).
- **Outcomes and estimation:** each outcome is reported with both SD and 95% confidence intervals as recommended (17a). Binary outcomes of recovery and clinical significant improvements are reported as odds ratios with 95% CI as recommended (17b).
- **Ancillary analysis:** adjusted analysis (such as ITT) is clearly indicated at all stages, as is reporting of subgroup and exploratory analysis in the text when these were not pre-specified (such as results by therapist) (18).
- **Harms:** no harm or unexpected events in the study were identified and as such these are not reported (19).

CONSORT item: discussion of findings (CONSORT item number included in brackets)

- **Limitations:** a section is included in the discussion outlining potential limitations of the study in terms of participants, follow-up, measures making reference to impact of potential precision (wide confidence intervals due possible to low numbers) and bias (issues of blinding and use of self-report measurement). This is as recommended (20).
- **Generalisability:** issues and limitations to generalisability are considered in the discussion under participants and therapist sections and in terms of approach to cost utility analysis. Issues are also addressed in relation to number of participants in study limitation section (21).
- **Interpretation:** interpretation of the results are balanced against limitations with findings benchmarked against other studies exploring BA to give context (22).

CONSORT item other information: (CONSORT item number included in brackets)

- **Registration and protocol:** the trial was registered with the International Standard Randomised Controlled Trials Registry, number ISRCTN27045243 where the protocol is available (23, 24).
- **Funding:**- funding sources and role of funders are important to report, as this study received no external funding this was not a focus. The support of Tees, Esk and Wear Valleys NHS Foundation Trust in funding the therapists was outlined (25).

Summary of trial against CONSORT standards

It can be seen from the above that while this was a small unfunded study it has been conducted and reported in line with the current recommended standards for randomised controlled investigations.

4.9 Conclusion, clinical implications and future research

The questions for this study were:

1. What is the impact on depression symptom level, functioning and treatment satisfaction of BA delivered by non-specialist mental health workers to depressed adults compared to usual primary care management?

2. What is the estimate of cost utility of BA delivered by generic mental health workers to depressed adults?

This study is the first randomised controlled clinical trial to test feasibility of dissemination of BA to a wider mental health workforce than psychologists and psychotherapists and as such represents a major step forward in our understanding of the intervention. In answer to the first question I have demonstrated that with limited training, non-specialist mental health nurses can be trained to deliver clinically effective BA to people with severe long standing depression. In answer to the second

question results demonstrated an incremental cost ratio of between £5-6,000 per QALY with 97% probability that this would fall below the currently recommended thresholds.

If these findings can be replicated and translated into routine health care, then clinical and cost implications are substantial. Depression is a disabling condition of high prevalence (World Health Organization 2004) and finding and testing potentially effective new treatments should be considered a priority (Hollon et al. 2002a). Now such feasibility has been shown, future research with a larger sample and multiple therapists should investigate the longer-term durability of this approach. This would build on the limitations identified in this study. Comparison of this BA delivery mode against an active psychological treatment such as CBT would provide an ideal comparator. This would allow longer term follow-up and provide meaningful clinical and cost effectiveness results against the current gold standard psychological intervention for depression. Such findings could inform future decision making regarding service design for depression.

Chapter 5: Summary, dissemination and link to future research

Summary of thesis

This research has followed a process recommended for the examination of complex interventions outlined in the introduction. The research has followed the development phase by identifying an area of need and scoping the evidence base. The demands for accessible psychological therapies for depression were highlighted in Chapter One. The theoretical basis of BA was examined in Chapter Two and its potential parsimony highlighted it as a single strand approach that possibly made it a suitable to explore in relation to needs outlined in Chapter One. In Chapter Three the available evidence base was scoped via a systematic review of all randomised trials in which BA has been used to treat depression. From this it became clear that while BA appeared effective and as effective as the current gold standard psychological treatment for depression, no research had adequately explored the dissemination of the approach. This was clearly an area of need highlighted in preparatory work outlined in Chapters One and Two of the thesis.

As the development phase found no relevant studies that had adapted accepted BA approaches for use by non specialists, the research moved to the feasibility and piloting phase in the MRC cycle. In this phase the theory of BA was matched with information from the meta-analysis to create a short-term 12-session protocol that had potential to be delivered by non-specialists. This approach had no previous trial evidence, so required feasibility testing in order to develop the intervention, to estimate effect size and ascertain potential recruitment and delivery issues. The trial outlined in Chapter Four was conducted over one year and measured clinical and cost outcomes, providing information that can now be used to design larger definitive trials beyond the scope of this thesis.

By following this process the research outlined in this thesis has contributed both to a body of evidence and created a platform for future studies. It has explored the idea that the parsimony of BA makes it suitable for delivery by non-specialists. This was a surprising gap in the research

literature, as the first presumption of such an idea in relation to BA had been written a decade before we embarked on our meta-analysis. The study had limitations due to the sample size and duration. It was however able to provide a clear indication that BA may be an intervention that is relatively simple and suitable for dissemination. As highlighted in the trial results and discussion sections, there was some indication that variation in therapist effect requires further examination. Larger studies should include more therapists and the detailed examination of therapist effect should be clearly defined in future studies methods. As dissemination is a central factor in the advantage BA may offer variation between therapists is clearly an important factor. The post hoc methods viewed in this study require cautious interpretation but they do shed light on the need for further research. The BA used in this study reflected a comprehensive model and there has been some discussion in literature as to the effective elements in BA which may indicate a simpler approach may be as effective (Kanter et al. 2010). The findings at one and two months in both the BDI-II and WASA demonstrated gradual improvement throughout the duration of the intervention. It was therefore unclear if any particular aspect of the intervention provides the benefit. Simplified models do exist (Lejuez et al. 2011) that include a less complex BA model. Future research may benefit from exploring to what degree simplicity provides improved consistency in dissemination whist maintaining treatment benefit. The inclusion of a qualitative analysis would also benefit future research. Qualitative interviews may offer insight into how BA is perceived by those receiving the intervention and themes relating to key aspects of therapist and intervention effect. No well conducted qualitative examination of BA was found in the development or analysis of this thesis. This represents it is a clear gap in the evidence and therefore knowledge of BA that should be addressed in future research design. The study is also the first to estimate the potential cost utility of the BA approach. Again a surprising finding in the review was that the cost effectiveness had not been investigated for BA. Initial results are promising and future study should build on these with larger sample sizes and longer term follow-up. It would also be of benefit to examine the cost effectiveness of BA delivered by non-specialists

against CBT, the current gold standard psychological treatment for depression. This would provide meaningful data for the design of clinical services aimed at improving access to effective depression treatment.

While this study is limited by lack of follow-up and size, which is understandable of any unfunded research programme embarked on as a PhD, it has however made a meaningful contribution to the evidence providing new knowledge. It has resulted in wide dissemination both by conference presentation and publications in peer-reviewed journals which have been widely cited. The research has formed the basis of a large multi-centre trial funding application to the National Institute of Health Research (NIHR) under the Health Technology Appraisal (HTA) funding stream. The study Cost and Outcome of Behavioural Activation (COBRA) is currently in the second phase of the review process. This research formed the basis of the grant application informing the design of the research (see appendix III for relevant section of that protocol).

This research suggests that BA may be an intervention that is important for clinicians and policy makers alike. The research highlighted in this thesis indicates BA may provide a clinically and cost effective psychological intervention for depression that can be widely disseminated. If larger trials are able to build on the results outlined here it is anticipated BA will become a standard treatment for depression used commonly across health care settings. Research outlined in this thesis will have made a significant contribution to such developments and hence have provided a meaningful and important addition to the evidence base for psychological treatment of depression.

Appendices

Appendix I: Systematic review and meta-analysis of behavioural therapies for depression: additional materials

Search Reports

14 February 2006

To : David Ekers

Search Report

Topic
Comparison of Behavioral Treatments for depression with treatment as usual, placebo or other brief psychotherapy

Date requested
5 October 2005

Date completed
- 8 February 2006

Databases searched (all via Dialog DataStar except Cochrane)
- Medline 1951 to date
- Embase 1974 to date
- Cinahl 1982 to date
- PsycInfo 1806 to date
- BNI 1994 to date
- AMED 1985 to date
- Cochrane Library Issue 2006-1

Limits
Adult or adolescent populations from age 16 upwards
Randomised Controlled Trials

Results
- Results in ASCII tagged format (for import to EndNote) were emailed to david.ekers@blueyonder.co.uk and to your Trust email.

  You requested both *Short* format results for initial scan, and *Medium* format results to provide abstracts.
- Search histories were provided for each database. For completeness, I include them in this report. The histories show numbers of results provided for each database.
- Duplicates across databases were not removed.

Issues
- Search methodology was geared to a highly sensitive search. Guidance from the CRD was used in order to retrieve all RCTs.

- Full use was made of descriptors and other special options, as well as free text searching.

Conclusion

Comprehensive searches have been conducted in 7 databases.
Please let me know if you discover any missed references or gaps in the search strategies.

Janet Menton

CDDPS Library Services Training Coordinator

0191 333 3465

Allied & Complementary Medicine - 1985 to date (AMED) (9 Nov 05)

Search history – Bold indicates results taken from this set

No.	Search term	Results
1	depression.ti. or depressive$1.ti.	1105
2	depressive-disorders.de.	739
3	dysthym$	18
4	1 or 2 or 3	1365
5	behavio$ adj therapy	1044
6	behavior-therapy.de.	853
7	treatment$1 or intervention$1 or therap$7 or activat$5 or psychotherap$9 or psycholog$7 or psychosocial$3	88365
8	behavi$7 with 7	2383
9	positive adj reinforce$ or (event$1 or activ$6) near schedul$3	54
10	9 same 7	31
11	10 or 8 or 6 or 5	2402
12	11 and 4	77
13	(children or child or adolescen$3) not (adult$ or elderly or geriat$ or old adj age$1)	11480
14	12 not 13	69
15	rct or rcts	180
16	randomized-controlled-trials.de.	1308
17	random$ near (clinical or trial$ or control$ or allocat$)	4357
18	pt=randomized-controlled-trial	542
19	15 or 16 or 17 or 18	4365
20	pt=controlled-clinical-trial	20
21	pt=clinical-trial$	679
22	clinical-trials#.de.	2803
23	(clinic$4 adj trial$1).ti,ab.	1752
24	(single adj blind adj method).de. or double-blind-method.de.	352
25	((singl$1 or doubl$1 or tripl$1 or trebl$1) adj (blind$2 or mask$2)).ti,ab.	1239
26	placebo$1.ti,ab.	1597
27	placebos.w..de.	480
28	19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27	7199
29	**14 and 28**	**12**

BRITISH NURSING INDEX- 1994 to date (6 Feb 06)

Search history (6 Feb 2006)

Bold indicates results taken from these sets

Comments

- Search results have not been compared with results from other databases, i.e. there may be duplicates with results from Medline, Embase, Cinahl.
- **21** results from **set 18** have been emailed in both short and medium formats.

No.	Search term	Results
1	RCT	22
2	RANDOM$ NEAR (CLINICAL OR TRIAL$ OR CONTROL$ OR ALLOCAT$ OR STUDY OR DESIGN$2 OR VOLUNTEER$ OR PROSPECTIVE)	1095
3	(CLINIC$4 ADJ TRIAL$1).TI,AB.	227
4	((SINGL$1 OR DOUBL$1 OR TRIPL$1 OR TREBL$1) ADJ (BLIND$2 OR MASK$2)).TI,AB.	58
5	PLACEBO$1.TI,AB.	80
6	RESEARCH-AND-DEVELOPMENT.DE. OR RESEARCH	30116
7	1 OR 2 OR 3 OR 4 OR 5 OR 6	30373
8	DEPRESSION OR DEPRESSIVE OR DEPRESSED	1677
9	Depression.W..DE.	1035
10	DYSTHYM$ OR DYSPHOR$	14
11	8 OR 9 OR 10	1686
12	11 NOT (MATERN$3 OR INFANT OR BABY OR BABIES OR MIDWIF$4)	1482
13	BEHAVIO$7 ADJ THERAPY	311
14	TREATMENT$1 OR INTERVENTION$1 OR THERAP$7 OR TECHNIQ$3 OR ACTIVAT$5 OR MODIF$8 OR PSYCHOTHERAP$9 OR PSYCHOLOG$7 OR PSYCHOSOCIAL$3	20236

15	BEHAVI$7 WITH 14	642
16	POSITIVE ADJ REINFORCE$ OR (EVENT$1 OR ACTIV$6) NEAR SCHEDUL$3	2
17	13 OR 15 OR 16	643
18	**17 AND 12 AND 7**	**21**

Search history – CINAHL 1982 TO DATE (4 Jan 2006)

Bold indicates results taken from these sets

Comments

- Search results have not been compared with results from other databases, i.e. there will be duplicates with previous results from Medline and Embase.
- *Medline has been reloaded in the intervening time and it gives different numbers of results compared with the original Medline search, so an attempt to remove duplicates within Dialog would be inappropriate.*
- 275 results from set 26 have been emailed in both short and medium formats.

No.	Search term	Results
1	depression.ti. or depressive.ti. or depressed.ti.	7129
2	(depressive adj disorder$ or major adj depression or organic adj depression or reactive adj depression).ab.	1064
3	(dysthym$ or dysphor$).ti,ab,de.	409
4	depression.w..de. or depression-reactive.de. or dysthymic-disorder.de.	13204
5	1 or 2 or 3 or 4	14506
6	behav$ adj therap$6	5399
7	behavior-therapy.de.	2063
8	behav$ near activat$3	167
9	behavi$ with (treatment or intervention or therap$7 or activat$5 or modif$8 or psychotherap$9 or psycholog$7 or psychosocial$3)	24259
10	positive adj reinforce$ or (event$1 or activ$6) near schedul$3	520
11	10 or 9 or 8 or 7 or 6	24621
12	rct	3202
13	random$ near (clinical or trial$ or control$ or allocat$ or assign$ or study or design$2)	63682
14	random$8.ti.	8248
15	12 or 13 or 14	64378
16	pt=clinical-trial$	16820
17	clinical-trials#.de.	44186
18	((clinic$4 or control$) adj trial$1).ti,ab.	25182
19	((singl$1 or doubl$1 or tripl$1 or trebl$1) adj (blind$2 or	5736

	mask$2)).ti,ab.	
20	placebo.ti,ab.	10795
21	placebos.w..de.	3260
22	15 or 16 or 17 or 18 or 19 or 20 or 21	84197
23	5 and 11	1444
24	23 and 22	470
25	adolescence.de. or adult.de. or middle-age or aged.de..w. or aged-80-and-over	273576
26	**24 and 25**	**275**

Search history for Cochrane Library – performed on 8 Feb 2006

Cochrane Library Issue 2006-1

Comments

- Search results have not been compared with results from other databases.
- I have not limited results to any keywords for RCT
- **602 citations** from **set 14** have been emailed, divided up by database (see list beneath search history)

  When exporting results, **abstracts** are only available for items from **CDSR** and **Central** databases, so **581** items **citations with abstracts** have been emailed.

ID	Search	Hits
#1	MeSH descriptor **Depression** explode all trees in MeSH products	2599
#2	MeSH descriptor **Depressive Disorder, Major**, this term only in MeSH products	575
#3	MeSH descriptor **Depressive Disorder**, this term only in MeSH products	3305
#4	dysthymic disorder in All Fields in all products	208
#5	dysthym* or dysphor* in Title, Abstract or Keywords in all products	815
#6	depression or depressive or depressed in Record Title in all products	7730
#7	MeSH descriptor **Behavior Therapy**, this term only in MeSH products	2071
#8	(event* or activ*) near schedul* in All Fields in all products	272
#9	behavio* near (therap* or treat* or techniq* or modif* or interven* or activat*) in Title, Abstract or Keywords in all products	7649
#10	positive next reinforc* in All Fields in all products	100

#11	coping near/3 depressi* in All Fields in all products	65
#12	(#1 OR #2 OR #3 OR #4 OR #5 OR #6)	10944
#13	(#7 OR #8 OR #9 OR #10 OR #11)	8013
#14	**(#12 AND #13)**	**602**

Results by database:

Cochrane Reviews [10] | DARE [11] | CENTRAL [571] | HTA [6] | NHS EED [4] |

- » There are **10** results out of **4200 records** for: "(**#12** AND **#13**) in **The Cochrane Database of Systematic Reviews**"
- » There are **11** results out of **5859 records** for: "(**#12** AND **#13**) in **Database of Abstracts of Reviews of Effects**"
- » There are **571** results out of **470139 records** for: "(**#12** AND **#13**) in **The Cochrane Central Register of Controlled Trials**"
- » There are **6** results out of **5378 records** for: "(**#12** AND **#13**) in **Health Technology Assessment Database**"
- » There are **4** results out of **17015 records** for: "(**#12** AND **#13**) in **NHS Economic Evaluation Database**"

Search history – EMBASE 1974 TO DATE (13 Dec 05)

Bold indicates results taken from these sets

no.	database	search term	results
46	embase	depression.ti. or depressive.ti. or depressed.ti.	47158
47	embase	depressive adj disorder	8412
48	embase	major-depression.de.	4363
49	embase	dysthym$.ti,ab,de.	3065
50	embase	depression.w..de.	102512
51	embase	endogenous-depression.de.	791
52	embase	46 or 47 or 48 or 49 or 50 or 51	117527
53	embase	behav$ adj therapy	19129
54	embase	behavior-therapy	17099
55	embase	behavior-therapy.mj.	6716
56	embase	behav$ near activat$3	2564
57	embase	behavi$ with (treatment or intervention or therap$7 or activat$5 or modif$8 or psychotherap$9 or psycholog$7 or psychosocial$3)	55494
58	embase	positive adj reinforce$ or (event$1 or activ$6) near schedul$3	2621
59	embase	58 or 57 or 56 or 55 or 54 or 53	58146
60	embase	59 and 52 and human=yes	5434
61	embase	rct	2351
62	embase	randomized-controlled-trial$.de.	101719

63	embase	random$ near (clinical or trial$ or control$ or allocat$ or study or design$2)	193305
64	embase	randomization.de.	17198
65	embase	61 or 62 or 63 or 64	203604
66	embase	clinical-trial#.de.	388595
67	embase	(clinic$4 adj trial$1).ti,ab.	89647
68	embase	single-blind-procedure	5667
69	embase	double-blind-procedure	60529
70	embase	crossover-procedure	17018
71	embase	((singl$1 or doubl$1 or tripl$1 or trebl$1) adj (blind$2 or mask$2)).ti,ab.	83851
72	embase	placebo$1.ti,ab.	93718
73	embase	placebo.w..de.	93940
74	embase	prospective-study	51906
75	embase	65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74	616257
76	embase	75 and 60	1726
77	embase	76 and animal=yes	7
78	embase	77 not (76 and human=yes and animal=yes)	0
79	embase	76 and adult=yes	624
80	embase	adolescent.de.	351050
81	embase	(76 and 80) not 79	126
82	*embase*	*79 or 81*	*750*
96	medline	40 or 42 *these were the results from Medline*	1685
97	embase medline	combined sets 82, 96 *combined results from Embase and Medline*	2435
99	embase medline	unique records from 97 *(dropped 250 duplicates from the databases)*	2185
101	**embase**	**split set 99** (these are the unique records from EMBASE)	**522**

Search history – EMBASE 1974 TO DATE PART 2 (19-12-05)

Bold indicates results taken from these sets

Objectives

1. Include additional descriptors for types of depression (see set 16)
2. Include additional study types (see set 33)
3. Obtain results from these which are additional to results despatched on 13 Dec 05 (set 82 before being compared with Medline for duplicates).

About the Extra Results

These have **not** been compared against Medline search results, i.e. may contain duplicates.

- Set 35 equivalent to set 76 of initial search) contains an extra 4 results due to database update – all emailed (medium format only)
- Set 41 contains an extra 16 results due to additional descriptors for types of depression (from set 16)
- Set 51 contains an extra 11 results which mention these **study types**: (controlled ADJ study OR major ADJ clinical ADJ study OR clinical ADJ article) combined with **random$** and also combined with only

selected major terms/descriptors for depression and behaviour therapy

- Set 53 contains an extra 53 results (compared with set 51) which mention those **study types** and **random**$ without limit to selected terms for depression and behaviour therapy.

No.	Search term	Results
1	DEPRESSION.TI. OR DEPRESSIVE.TI. OR DEPRESSED.TI.	47212
2	DEPRESSIVE ADJ DISORDER	8424
3	MAJOR-DEPRESSION.DE.	4391
4	DYSTHYM$.TI,AB,DE.	3069
5	DEPRESSION.W..DE.	102652
6	ENDOGENOUS-DEPRESSION.DE.	792
7	1 OR 2 OR 3 OR 4 OR 5 OR 6	117704
8	BEHAV$ ADJ THERAPY	19150
9	BEHAVIOR-THERAPY	17120
10	BEHAVIOR-THERAPY.MJ.	6721
11	BEHAV$ NEAR ACTIVAT$3	2569
12	BEHAVI$ WITH (TREATMENT OR INTERVENTION OR THERAP$7 OR ACTIVAT$5 OR MODIF$8 OR PSYCHOTHERAP$9 OR PSYCHOLOG$7 OR PSYCHOSOCIAL$3)	55573
13	POSITIVE ADJ REINFORCE$ OR (EVENT$1 OR ACTIV$6) NEAR SCHEDUL$3	2626
14	13 OR 12 OR 11 OR 10 OR 9 OR 8	58230
15	14 AND 7 AND HUMAN=YES	5448
16	(dysphoria OR melancholia OR major ADJ depression).TI,DE.	10559
17	14 AND (7 OR 16) AND HUMAN=YES	5555
18	RCT	2361
19	RANDOMIZED-CONTROLLED-TRIAL$.DE.	101876
20	RANDOM$ NEAR (CLINICAL OR TRIAL$ OR CONTROL$ OR ALLOCAT$ OR STUDY OR DESIGN$2)	193608
21	RANDOMIZATION.DE.	17256
22	18 OR 19 OR 20 OR 21	203937
23	CLINICAL-TRIAL#.DE.	389428
24	(CLINIC$4 ADJ TRIAL$1).TI,AB.	89823
25	SINGLE-BLIND-PROCEDURE	5681
26	DOUBLE-BLIND-PROCEDURE	60584
27	CROSSOVER-PROCEDURE	17036
28	((SINGL$1 OR DOUBL$1 OR TRIPL$1 OR TREBL$1) ADJ (BLIND$2 OR MASK$2)).TI,AB.	83915
29	PLACEBO$1.TI,AB.	93806
30	PLACEBO.W..DE.	94167
31	PROSPECTIVE-STUDY	52053
32	22 OR 23 OR 24 OR 25 OR 26 OR 27 OR 28 OR 29 OR 30 OR 31	617381
33	controlled ADJ study OR major ADJ clinical ADJ study OR clinical ADJ article	3717141
34	32 OR 33	3920524
35	32 AND 15	1730
36	34 AND 17	3457
37	32 AND 17	1774
39	37 NOT 35	44

40	ADULT# OR AGED.DE. OR ADOLESCENT.DE.	2054682
41	**39 AND 40**	**16**
49	(36 AND 40 AND (10 OR 11 OR 13) AND (1 OR depression.MJ.)) NOT 37	181
50	49 AND random$	11
51	**49 AND (random$ OR 25 OR 26 OR 28 OR 30)**	**11**
52	(36 AND 40 AND random$) NOT 37	64
53	**52 NOT 51**	**53**

MEDLINE 1951 TO DATE Search history (30 Nov 05)
Bold indicates results taken from these sets

No.	Search term	Results
1	depression.ti. or depressive.ti. or depressed.ti.	54179
2	(depressive adj disorder$).ab.	7365
3	depressive-disorder#.de.	47822
4	dysthym$.ti,ab,de.	1937
5	depression.w..de.	41033
6	1 or 2 or 3 or 4 or 5	102398
7	behav$ adj therapy	20048
8	behavior-therapy.de.	17604
9	behavior-therapy.mj.	10631
10	behav$ near activat$3	2841
11	behavi$ with (treatment or intervention or therap$7 or activat$5 or modif$8 or psychotherap$9 or psycholog$7 or psychosocial$3)	83625
12	positive adj reinforce$ or (event$1 or activ$6) near schedul$3	2819
13	12 or 11 or 10 or 9 or 8 or 7	86486
14	13 and 6 and human=yes	4270
15	rct or rcts	3397
16	randomized-controlled-trials.de.	40249
17	random$ near (clinical or trial$ or control$ or allocat$ or study or design$2)	229268
18	pt=randomized-controlled-trial	209366
19	random$8.ti.	49476
20	15 or 16 or 17 or 18 or 19	336906
21	pt=controlled-clinical-trial	69809
22	pt=clinical-trial$	420643
23	clinical-trials#.de.	172371
24	(clinic$4 adj trial$1).ti,ab.	97767
25	(single adj blind adj method).de. or double-blind-method.de.	93219
26	((singl$1 or doubl$1 or tripl$1 or trebl$1) adj (blind$2 or mask$2)).ti,ab.	83013
27	placebo$1.ti,ab.	95342
28	placebos.w..de.	25241
29	research-design.de.	42243
30	20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29	687321
31	comparative-study.de.	1247963
32	evaluation-studies#.de.	540159
33	follow-up-studies.de.	309308
34	prospective-studies.de.	195627
35	(control$3 or prospectiv$3 or volunteer$).ti,ab.	1640684
36	31 or 32 or 33 or 34 or 35	3233487
37	animals.de. not (human.de. and animals.de.)	2977974

38	(30 or 36) not 37	2678930
39	38 and 14	2083
40	**39 and adult#**	**1377**
41	adolescent.de.	1090548
42	**(39 and 41) not 40**	**308**

PSYCINFO Search history – (20 Jan 2006)

Bold indicates results taken from these sets

Comments

- In set 11, classification category 3312 covers *Behaviour Therapy and Behaviour Modification*; it falls within the hierarchy *Health and Mental Health Treatment and Prevention*.
- Search results have not been compared with results from other databases, i.e. there will be duplicates with previous results from Medline, Embase, Cinahl.
- 592 results from set 33 have been emailed in both short and medium formats.

no.	search term	results
1	depression.ti. or depressive.ti. or depressed.ti.	45219
2	(depressive adj disorder).ab.	8250
3	dysthymic-disorder.de.	1119
4	(dysthym$ or dysphor$).ti,ab.	5182
5	major-depression.de.	44673
6	1 or 2 or 3 or 4 or 5	65690
7	behav$8 adj therapy	48875
8	behavior-therapy#.de.	12773
9	behav$8 near activat$6	4889
10	psychological-interventions.kw.	276
11	'3312'.cc.	13087
12	behav$8.ti,ab,de,kw. and (10 or 11)	10563
13	positive adj reinforce$ or (event$1 or activ$6) near schedul$3	1888
14	13 or 12 or 9 or 8 or 7	56194
15	14 and 6	4174
16	rct or rcts	586
17	random$ with (clinical or trial$ or control$ or allocat$ or study or design$2 or volunteer or prospective)	4919
18	16 or 17	5485
19	treatment-outcome-clinical-trial.at.	9992

20	clinical-trials#.de.	609
21	(clinic$4 adj trial).ti,ab.	6105
22	((single or double or triple or treble) adj (blind$2 or mask$2)).ti,ab.	10365
23	placebo.ti,ab.	17423
24	placebos.w..de.	1478
25	18 or 19 or 20 or 21 or 22 or 23 or 24	36521
26	prospective-studies.de.	274
27	at=followup-study or at=prospective-study or at=meta-analysis or at=treatment$	39867
28	25 or 26 or 27	65549
29	animals.de. not (human.de. and animals.de.)	78090
30	28 not 29	65074
31	15 and 30	753
32	age=adolescence-13-17-yrs or age=adulthood-18-yrs-and-older	796544
33	**31 and 32**	**592**

Study Quality Assessment tool

Study quality checklist BA systematic review
Describe briefly and allocate: yes, no, not clear

Study details (title author year)

Adequate randomisation procedure used (computer generated/random number table, independent, describe)

Assessment of outcome blinded to intervention arm (describe)

Pre-specified eligibility criteria for study

Groups equal at baseline

Was adherence to treatment procedures monitored in each arm

Mean and Standard Deviation recorded for primary outcome measure

Power calculation

loss to follow up reported

Intention to treat analysis

Study Data Extraction Form

Behavioural Activation SYSTEMATIC REVIEW DATA EXTRACTION FORM

INSTRUCTIONS

- **Please enter data in each BLANK shaded data entry field**
- **Use the TAB key to move DOWN from one data entry field to the next one**
- **Use SHIFT+TAB together to move UP from one data entry field to the previous one**
- **Boxes will expand to fit additional data**
- **If data is not available, please enter 'Not clear'**
- **When the data extraction is completed, please save the file with the following filename:**
    - **[First author] [Date] [Reviewer initials].doc**
    - **e.g. Scogin 2005 PB.doc**
    - **If two studies share the same identifier, please use 2001a, 2001b etc**

General information

Study ID (First author + Year)

Title

Source

Diagnosis

Population

Country where data collected

Recruitment context (e.g. primary care, specialist setting, community, other)

Recruitment method (description of actual recruitment procedure)

Study population

Target population (broad description e.g. depressed students)

Inclusion criteria

Exclusion criteria

Baseline sex

Baseline age

Baseline social economic status

Baseline education

Baseline ethnicity

Baseline other

Trial design

Number of groups, conditions or arms

Unit of randomisation (i.e. who was randomised? e.g. patient, group, class?)

Method of randomisation (i.e. how was it conducted? e.g. coin toss, central telephone)

Power calculation (i.e. was a power or sample size calculation reported?)

Was a main aim identified a priori?

Did they conduct an intention to treat analysis? (e.g. analysed all patients irrespective of adherence to treatment)

What assessment of depression outcomes were included? (e.g. BDI, CES-D)

Secondary outcomes measured

How many patients were eligible for the trial? (i.e. met the criteria for inclusion)

How many patients actually took part in the trial? (i.e. randomised)

How many patients were lost to follow up? (e.g. failed to return outcome assessments, or left the trial)

Lost patients reasons recorded

Blinding method adopted

Behavioural Intervention

Description of the intervention (please enter a one sentence)

Level of training of therapist (general)

Level of specific training of therapist for BA intervention

Number of sessions

Duration of sessions

Frequency of sessions

Concurrent pharmacology

Did the person providing the intervention receive supervision?

Therapist/researcher independence?

How was adherence to intervention assessed?

Incentives for participants?

Additional information not covered above

Control Groups

Comparison group 1

Level of training for comparison group intervention

Balanced for non specific factors (describe eg time/contact)

Comparison group 2

Level of training for comparison group intervention

Balanced for non specific factors (describe eg time/contact)

Comparison group 3

Level of training for comparison group intervention

Balanced for non specific factors (describe eg time/contact)

Results

Were groups equal at baseline (comment)

Study quality rating

Post treatment measurement at what time point (post intervention, specific time following first assessment)

Number of follow up assessments

Time period of follow up assessments

Additional comment

Results Table (n means sd)

Intervention	Pre	Post	Fu1	Fu2	Fu3	Fu4
Measure 1						
Measure 2						
Measure 3						

Appendix II: A randomised controlled trial of clinical and cost effectiveness of behavioural activation for depression delivered by the non specialist: additional materials

Independent peer reviews of study protocol for ethics

Dr Chris Williams MBChB BSc MMedSc MD FRCPsych
Senior Lecturer and Honorary Consultant Psychiatrist

Direct Line: 0141 211 3912
Secretary: 0141 211 0685
Fax: 0141 357 4899
E-mail: Chris.Williams@clinmed.gla.ac.uk
http://www.gla.ac.uk/departments/psychologicalmedicine/index.html

Re: CW/AG/Research

22 February 2008

Dear David

Thank you for asking me to review your protocol entitled 'Trial Protocol for Behaviour Therapy for Depression delivered by specifically trained generic mental health staff'.

This study is timely and will pave the way for a possible larger funded randomised controlled study. There is a need to discover which sorts of staff best can deliver these sorts of interventions and it may well be the case that junior mental health staff with appropriate training and supervision are able to offer this very effective intervention. I am less convinced that 12 sessions of 30-45 minutes really saves so much time though compared to traditional CBT so maybe the issue is more about widening the scope of practitioners who may be able to offer this effective intervention.

Please find my specific comments below:

Methodology:
It is very sensible at this stage to do a small randomised controlled study in order to establish the feasibility of recruiting, training, delivering and retaining both patients and staff in this approach. The primary care sample is a logical group to focus on and the confirmation of diagnosis using the Clinical Interview Schedule Revised is very appropriate (though it is not stated who will do this). It would be helpful to define exactly what you mean by currently actively suicidal (for example a score of 2 or more on the Beck Depression Inventory suicide item might be used for consistency).

The main design aspect I would question is the control arm. In practice it might be better to compare BT + TAU with TAU + monitoring rather than monitoring alone. In practice many GPs will have prescribed/referred already but ethically both arms should be treated by G.P.s as they would usually do. If there is going to be advice/expectation that GPs don't offer other treatment for a time this may fit with watchful waiting initiatives but only for 3-4 weeks and I am concerned that TAU should be available to both arms as needed.

It is sensible to recruit from General Practice and to use two forms of recruitment. A major feature that needs to be addressed is that of contamination as it is a small sample and a proportion will be already referred to other practitioners/may have new

Sample size: this seems appropriate based on the previous research.

Ethics of Recruitment:
antidepressants etc. This needs to be recorded in detail.

Recruitment: For practice case-notes search it is not stated who will do this. Will this be one of the G.P. practice staff or a member of the research team? If it is a member of the research team it may not be appropriate for records or contact details to be taken away from the G.P. surgery (and if they are they need to be encrypted) and in the first instance it is probably sensible for the letter sending the patient information sheet to the person to come on behalf of the practice saying that the practice is interested in supporting this research and including the PIS. I assume that a written consent form will then be returned as described in your protocol allowing the research team to provide further contact. This seems very appropriate. There is no mention of whether an advert will also be used.

The area that is not described in detail is the direct referral from the health care provider. Much depends on the severity of depression but there is a need to avoid any hints of coercion. A model that can work well is for the GP to provide the potential participant with the Participant Information sheet (PIS) and an opt-in card they can sign and either post off and/or drop in a box in reception to be collected by the research team who can then contact the person and proceed with recruitment or not.

Overall I think this is a well-planned, detailed protocol which has a high chance of success and will add significantly to what is known.

Yours sincerely

Dr Chris Williams
Senior Lecturer in Psychiatry & Honorary Consultant Psychiatrist

THE UNIVERSITY of York

DEPARTMENT OF HEALTH SCIENCES

Room 222, 2nd Floor (Area 4)
Seebohm Rowntree Building
University of York
Heslington
York YO10 5DD

Telephone (01904) 321908
Fax (01904) 321382
Email dr17@york.ac.uk

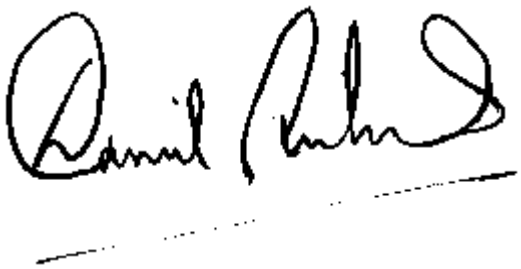
www.york.ac.uk/healthsciences

Sunday, 11 November 2007

Re: Research Proposal from David Ekers

I have been involved in considerable peer review of this proposal, which David Ekers has produced for his PhD studies. I am happy that all aspects of his background, methods and analysis plan are of high quality, having been rigorously scrutinised by myself, Professor Simon Gilbody and others in our department, not least Professors Godfrey (for the health economics) and Bland (for the statistical analysis plan). The proposal has been through several drafts and iterations and has been considerably edited as a consequence of this process. It is based on a published systematic review and meta-analysis of the literature which Mr Ekers has conducted and recently published to considerable acclaim from the scientific and psychology communities[1]. He has also taken advice on ethics and research governance. I have no doubt, therefore, that it is a suitably rigorous proposal.

Yours sincerely,



[1] Ekers, D, Richards, D and Gilbody S. (2007). A Meta Analysis of Randomised Trials of Behavioural Treatment of Depression. *Psychological Medicine* doi: 10.1017/S0033291707001614, published online 01 October 2007.

David Richards
Professor of Mental Health
Sample size

The difference to be detected is what is indicated by the meta-analysis as most likely, i.e. the point estimate of the difference. The chosen power, 80%, is accepted by funders such as MRC and HTA as acceptable for trials. The power may be increased by the use of baseline depression as a covariate. We have allowed for 30% dropout in this trial. In the earlier meta-analysis, the dropout was 19% and in the collaborative care for depression trial is was 16%. Hence 30% is, if anything, generous.

Power calculations are most appropriate when the study is intended to be the final, definitive study. This is an exploratory study, in that no other trials of this intervention have been done. It will be used to help design a further, larger and definitive trial if the treatment appears promising.

Martin Bland
17 January 2008

Study information sheet and consent forms

Behavioural Therapy for Depression Information sheet version 2 (17.03.08)

Treatment of Depression by Behavioural Therapy Participant Information Sheet

You are being invited to take part in a research study. Before you decide whether you want to take part or not it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

Depression causes misery to many patients and is a major health problem in the UK. Effective talking treatments are available however access to these is often difficult. The recommended treatment in the UK is Cognitive Behavioural Therapy. However Behavioural Therapy alone has been seen to be as effective as full Cognitve Behavioural Therapy, although these treatments have mainly been developed in the United States. We do not know who is best placed to deliver this treatment; therefore, this study will investigate this question further by training mental health workers to deliver behavioural therapy to approximately 35 people with depression. This treatment will then be compared with approximately 35 people being treated by their GP as per usual practice (with monthly brief phone calls from the research team) over 12 week period (who will then be offered the behavioural therapy).

Why have I been chosen?

Your health worker has identified you as suffering from depression. This letter asks you to consider taking part in the research study.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

What will happen to me if I take part?

This study is a randomised controlled trial. Sometimes, because we do not know which way of treating patients is best, we need to make comparisons. People are put into groups and then compared. The groups are selected by a computer which has no information about the individual – i.e. by chance. Patients in each group then have a different wait time before starting behavioural therapy and are compared before treatment and at 12 weeks. Firstly, we would like to interview you to make sure you are eligible for the trial and to ask you to fill in some questionnaires. We would then like to interview you again after three months to repeat the questionnaires. We may also need to collect some information from your medical records. The research study will last for approximately one year. You will be invited to commence behavioural therapy either immediately or at most 12 weeks after agreeing to take part in the study. We would also like to record the therapy sessions you attend, however if you do not wish us to do so your inclusion in the study will not be affected. A proportion (20% picked at random) of these recordings will be listened to by an expert in this field (bound by standard NHS confidentiality rules) outside of the core research group. This will allow independent verification of the content of the intervention. All recordings will treated as confidential information as per NHS standards and will be destroyed at the end of the study. No personal information (such as names & address) will be associated with the recordings.

What do I have to do?

There are no restrictions in your lifestyle from taking part in this research. You should continue to follow the advice of your GP as they will remain responsible for your overall medical care throughout your involvement in this study. In addition to the treatment offered from your GP you will be given access to 12 40-50 minute sessions of behaviour therapy for depression. This will be at a local health centre either straight away or following 12 weeks.

What is the drug or procedure that is being tested?

The treatment is called Behavioural Therapy for Depression. You will be allocated a therapist (a mental health worker) trained in this method of helping people with depression. Behavioural therapy helps you consider how what you do affects the way you feel. It considers any changes in your life and in partnership with your therapist you explore your response to these, trying out changes in agreement between you. In this way it is aimed to address those things that contribute to your depression and based on this consider changes to make to overcome this problem.

Are there any side effects, disadvantages and risks of taking part?

We are not aware of any side effects, disadvantages or risks to you of taking part in this research.

What are the possible benefits of taking part?

You will have access to 12 weeks of a talking therapy aimed at improving your depression. We hope that this treatment will teach you techniques that can make you feel better. The information we get from this study may help us to treat future patients with depression more effectively.

What will happen if I don't want to carry on with the study

You can withdraw from the study at any time and your GP will continue to treat you depression. We would like to keep in contact with you to check your progress however this will be with your consent.

What happens when the research study stops?

Throughout the study and afterwards, your GP will continue to treat your depression as s/he feels is best for you and with your agreement.

What if something goes wrong?

We do not anticipate any harm coming to you from involvement in this study. However, if you wish to complain, or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the normal National Health Service complaints mechanisms are available to you. In the event that something does go wrong and you are harmed during the research and this is due to someone's negligence then you may have grounds for a legal action for compensation against (name of Sponsor Organisation, NHS Trust, Private Clinic) but you may have to pay your legal costs.

Will my taking part in this study be kept confidential?

All information collected about you during the course of the research will be kept strictly confidential. Any information about you will have your name and address removed so that you cannot be recognised from it. As your GP/health care worker has suggested you might like to take part in this study and is involved in your treatment s/he will be informed of your progress as part of the research study with your permission.

What will happen to the results of the research study?

We will publish the results of this research study widely. As well as producing a research report and writing articles for health professionals to read you will be given a summary of the findings. We will ensure service organisations such as Depression Alliance are informed of the results of the trial. You will not be personally identified in any publications from this trial.

Who is organising and funding the research?

The research is being supported by your local NHS mental health trust as part of is normal running costs, no additional funding has been provided. Your GP is not being paid any extra money for being involved in the study.

Who has reviewed the study?

All research in the NHS is reviewed by an independent group of people called a research ethics committee to protect your rights, safety, wellbeing and dignity. The Northumbria Research Ethics Committee of the NHS has reviewed the study and approved it.

Next Steps

If you would like to take part in this research study please fill in the attached consent form and send it off in the enclosed stamped envelope. If you need further information to help you decide, please contact David Ekers at the address below. Once the research team receives your consent form, a researcher will contact you and arrange to see you in the next few days.

Thank you for reading this and for considering taking part in this study.

You are welcome to keep this Information Sheet and if you agree to take part in the research study you will also be given a copy of a signed consent form to keep.

Contact for Further Information

If you need further information about this study please contact the research lead:
David Ekers,
Nurse Consultant Primary Care Mental Health
Health Centre,
Newcastle Road,
Chester Le Street,
Co Durham,
DH3 3UR

Telephone: 01913336038.

Email: david.ekers@cddps.nhs.uk

For independent advice re participation in the study please contact:

Ms Jacqui Lovell
R & D Manager

TEWV NHS Tust
TAD Centre
Ormsby Rd
Berwick Hills
Middlesbourgh
TS3 7SF

Telephone 01642 516981

Permission to contact consent version 2 17.03.08

GP Practice Letter Headed Paper

Patient Identification Number for this trial:

Permission for researcher to contact

practice name

Treatment of Depression by Behavioural Therapy

I confirm that I have read and understand the information sheetfor the above study and am happy for a researcher to contact me.

I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.

Name
_____

Address
_____

Signature
_____

Telephone contact details
day_____
Evening_____
Mobile
Email address

Return in enclosed pre-paid envelope to:
David Ekers
Treatment of Depression by Behavioural Therapy Trial
Health Centre,
Newcastle Road,
Chester Le Street,
Co Durham,
DH3 3UR

Patient Identification Number for this trial:

CONSENT FORM

Treatment of Depression by Behavioural Therapy
David Ekers,
Nurse Consultant Primary Care Mental Health
Health Centre,
Newcastle Road,
Chester Le Street,
Co Durham,
DH3 3UR

Telephone: 01913336038.

Email: david.ekers@cddps.nhs.uk

Please initial box

1. I confirm that I have read and understand the information sheet dated 06/03/2008 (version 2) for the above study and have had the opportunity to ask questions ☐.

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐

I understand that sections of any of my medical notes may be looked at by research staff from Tees Esk & Wear Valleys NHS Trust. I give permission for these individuals to have access to my records. ☐

4. I agree to take part in the above study ☐

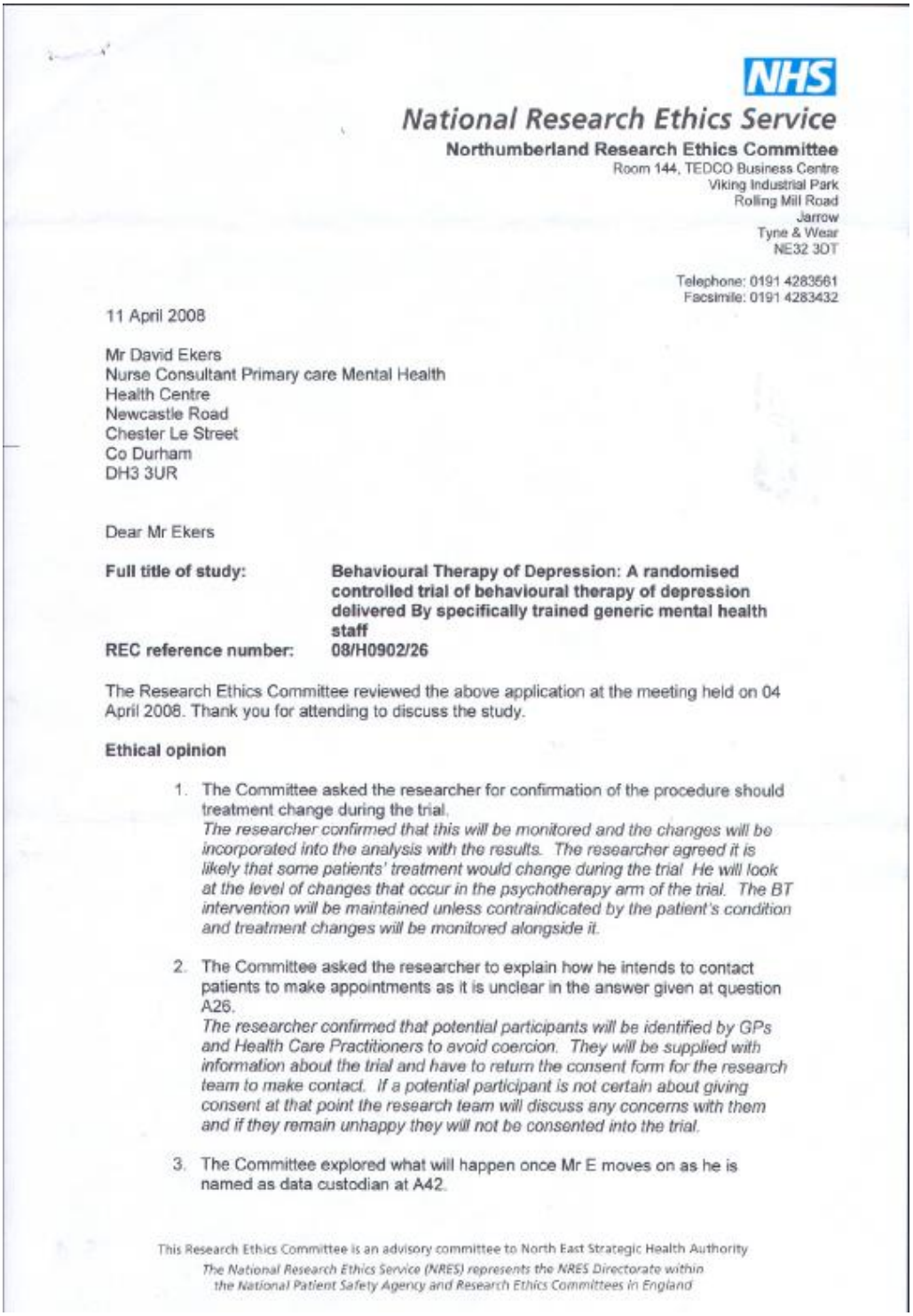
5. I agree to my GP being informed of my participation in this study ☐

6. I agree to sessions being recorded and listened to by an independent expert to verify content of the intervention ☐

_____

Name of Patient Date Signature

Copies of research ethics and governance approvals



NHS

National Research Ethics Service

Northumberland Research Ethics Committee

Room 144, TEDCO Business Centre
Viking Industrial Park
Rolling Mill Road
Jarrow
Tyne & Wear
NE32 3DT

Telephone: 0191 4283561
Facsimile: 0191 4283432

11 April 2008

Mr David Ekers
Nurse Consultant Primary care Mental Health
Health Centre
Newcastle Road
Chester Le Street
Co Durham
DH3 3UR

Dear Mr Ekers

Full title of study: **Behavioural Therapy of Depression: A randomised controlled trial of behavioural therapy of depression delivered By specifically trained generic mental health staff**

REC reference number: **08/H0902/26**

The Research Ethics Committee reviewed the above application at the meeting held on 04 April 2008. Thank you for attending to discuss the study.

Ethical opinion

1. The Committee asked the researcher for confirmation of the procedure should treatment change during the trial.
   The researcher confirmed that this will be monitored and the changes will be incorporated into the analysis with the results. The researcher agreed it is likely that some patients' treatment would change during the trial. He will look at the level of changes that occur in the psychotherapy arm of the trial. The BT intervention will be maintained unless contraindicated by the patient's condition and treatment changes will be monitored alongside it.

2. The Committee asked the researcher to explain how he intends to contact patients to make appointments as it is unclear in the answer given at question A26.
   The researcher confirmed that potential participants will be identified by GPs and Health Care Practitioners to avoid coercion. They will be supplied with information about the trial and have to return the consent form for the research team to make contact. If a potential participant is not certain about giving consent at that point the research team will discuss any concerns with them and if they remain unhappy they will not be consented into the trial.

3. The Committee explored what will happen once Mr E moves on as he is named as data custodian at A42.

This Research Ethics Committee is an advisory committee to North East Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

The researcher confirmed that data is held as per Trust policy, if Mr E moves on other arrangements will be made by the Trust in accordance with this policy.

4. The Committee asked whether the researcher had considered that the service user questionnaire could uncover benefit fraud.
   The researcher confirmed that the questionnaire is anonymous.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation.

Ethical review of research sites

The favourable opinion applies to the research sites listed on the attached form.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Application	2	20 March 2008
Investigator CV		01 December 2007
Protocol	2	18 March 2008
Covering Letter		17 March 2008
Summary/Synopsis	2	18 March 2008
Letter from Sponsor		19 November 2007
Peer Review		22 February 2008
Statistician Comments		25 January 2008
Questionnaire: Service Use	1	07 December 2007
Questionnaire: Validated: Patient Health		
Questionnaire: Validated: EQ-5D		
Questionnaire: Validated: Work & Social Adjustment		
Questionnaire: Validated: Beck Depression Inventory		
Questionnaire: Validated: Client Satisfaction		
Letter of invitation to participant	1	16 November 2007
GP/Consultant Information Sheets	2	06 March 2008
Participant Information Sheet	2	17 March 2008
Participant Consent Form	2	17 March 2008
Permission for researcher to contact slip	2	17 March 2008
Supervisor CV		01 December 2007

R&D approval

The study should not commence at any NHS site until the local Principal Investigator has obtained final approval from the R&D office for the relevant NHS care organisation.

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Website > After Review

Here you will find links to the following

a) Providing feedback. You are invited to give your view of the service that you have received from the National Research Ethics Service on the application procedure. If you wish to make your views known please use the feedback form available on the website.
b) Progress Reports. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
c) Safety Reports. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
d) Amendments. Please refer to the attached Standard conditions of approval by Research Ethics Committees.
e) End of Study/Project. Please refer to the attached Standard conditions of approval by Research Ethics Committees.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nationalres.org.uk .

08/H0902/26	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

Eggleston

pp **Dr Saul Miller**
Chair

Email: verity.eggleston@suntpct.nhs.uk

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*

Standard approval conditions

Site approval form (SF1)

Copy to: Professor Joe Reilly, R&D Office, Tees Esk & Wear Valley NHS Trust

THE UNIVERSITY of York

DEPARTMENT OF
HEALTH SCIENCES

30th April 2008

Re: Behavioural Therapy Of Depression

Dear David Ekers,

Thank you for sending your application for external ethical approval of the above project to me as Chair of the Health Sciences' Research Governance committee (HSRGC). I am writing on behalf of the HSRGC to confirm that approved the research can proceed.

Good luck with your project.

Dr Stephen Holland
Chair, HSRGC

Tees, Esk and Wear Valleys NHS
NHS Foundation Trust

Research & Development Dept
The TAD Centre
Ormesby Rd
Berwick Hills
Middlesbrough
Cleveland
TS3 7SF
01642 516984
jenny.brock@tewv.nhs.uk

10.07.2008 REC: 08/H0902/26

Mr David Ekers
Nurse Consultant Primary Care Mental Health
Health Centre
Newcastle Road
Chester Le Street
Co Durham
DH3 3UR

Dear David Ekers,

I am pleased to inform you that you have successfully gained research governance approval from the TEWV NHS Trust. This approval is specifically in relation to your research entitled: ***Behavioural Therapy of Depression: A randomised controlled trial of behavioural therapy of depression delivered by specifically trained generic mental health staff.***

We look forward to receiving a short report upon completion and would like to take this opportunity to wish you every success in your research.

If there is any way that we can assist you in the future please contact us.

Yours sincerely

Jennifer M Brock
Research Governance and Accreditation Co-ordinator

NHS

County Durham Primary Care Trust
Darlington Primary Care Trust

Research Management & Governance Unit
Henson Close
South Church Enterprise Park
Bishop Auckland
County Durham
DL14 6WA
OSR 574 621/177
Tel: 01388 452297
Fax: 01388 452290
Safehaven Fax: 01388 452297
www.countydurhampct.nhs.uk

Our ref: RE-MM324
Your ref:

29 May 2008

Direct line: 01388 452299
Switchboard: 01388 458835
Fax: 01388 452290
Email: richard.errington@cdpct.nhs.uk nhs.net

Mr David Ekers
Nurse Consultant Primary Care Mental Health
Health Centre
Newcastle Road
Chester-le-Street
County Durham
DH3 3UR

Dear David

Behavioural Therapy of Depression: A randomised controlled trial of behavioural therapy of depression delivered by specifically trained generic mental health staff

The Research Management & Governance Unit of County Durham & Tees Valley Primary Care Trusts gives **approval** for this project to begin on behalf of **County Durham Primary Care Trust** and **Darlington Primary Care Trust** subject to the following conditions:

- Approval from the Research Ethics Committee with site-specific approval where appropriate.
- Honorary Contracts have been issued where relevant.
- Any Accidents and Complaints related to the research are reported to the PCT(s) and RM&G Unit through the usual systems.
- Serious Adverse Events affecting local patients are reported to the PCT(s) and RM&G Unit promptly.
- The RM&G Unit is provided with copies of any updated documentation after NRES approval and before it is implemented
- The Researchers will provide assistance with any Monitoring or Audit requests from the RM&G Unit or the PCT(s).
- The research will not require any financial support from the PCT(s), unless there is a written agreement to the contrary.
- The PCT(s) and RM&G Unit are informed when the project ends.

Best wishes in your research.

Yours sincerely

R. Errington

Richard Errington
RM&G Unit Lead

Service use questionarres baseline and 3 months

SERVICE USE QUESTIONNAIRE

These questions refer to the previous **6 months**.

USE OF HEALTH SERVICES (excluding specific treatment for depression problem)

1. Have you visited a hospital as a patient during the last six months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1e.

1a. Have you visited a hospital A & E dept in the last six months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1b.

How many times? | | | |

How many nights did you stay in total? | | | |

How many times did you use an emergency (999 call) ambulance? | | | |

1b. Have you visited hospital as an inpatient for a medical problem in the past 6 months? (i.e. stayed overnight, but excluding A&E) Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1c.

How many nights did you stay in total? | | | |

1c. Have you visited hospital as an outpatient in the last 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1d.

How many times? | | | |

1d. Have you visited hospital as a day patient in the last 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1e.

How many days? | | | |

1e. Have you seen your GP or a practice nurse at the surgery in the last 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1f.

How many times have you visited the GP?			
How many times have you visited the practice nurse?			

1f. Have you been visited at home in the last 6 months by a GP, practice nurse? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1g

How many times? GP			
Practice nurse			

1g. Have you contacted NHS Direct over the past 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1h

How many times?			

1h. Have you visited a walk in centre over the past 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 2

How many times?			

MEDICATION

2. Have you received any prescriptions in the last 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 3.

How many?			

DEPRESSION SERVICES CONTACTS

3. Have you received treatment for a depression over the last six months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 4.

3a. Have you been prescribed medication? Yes ☐ No ☐

If so what was the name of the medication
If "NO", please skip to 3b.

What dose in mg per day? | | | | ☐

If current dose is unknown
tick this box.

3b. Were you offered any other services for a problem with depression (such as seeing a counsellor, CPN, Psychiatrist etc)? Yes ☐ No ☐

If "NO", please skip to 3c.

Service	When and how long were the appointments	How many times

3c. Were you prescribed any other medication for psychological problems (such as sleeping, anxiety etc)? Yes ☐ No ☐

If "NO", please skip to 3d.

Drug	Dose	No of prescriptions

This section records other contacts with depression services over the last six months.

3d. Have you been treated for a depression problem in the last 6 months where you stayed in a hospital? Yes ☐ No ☐ Don't know ☐

If "NO", please skip to 3e.

How many times were you admitted? | | | |

How many nights did you stay in total? | | | |

3e. Have you received specific counselling or advice in a specialist depression clinic? Yes ☐ No ☐ Not answered ☐

If "NO", please skip to 3f.

What type of counsellor did you see?	Number of occasions seen		
Psychologist			
Keyworker/general counselling			
Social worker			
Psychiatrist			
Other (please specify):			

……

OTHER SERVICE RECEIPT

4. Have you used any of the following services during the last 6 months?

		Number of contacts		
(i)	Advisor regarding state benefits or housing issues			
(ii)	Social Worker (at home)			
(iii)	Occupational Therapist (at home)			
(iv)	Citizens Advice			
(v)	RELATE			
(vi)	Alternative medical practitioner			
(vii)	Advisor on legal or debt issues			
(viii)	A homeless persons agency			

(ix) An employment advisor

(x) Other
(please specify……..……..……)

EMPLOYMENT

5a. Have you had a job in the past 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 5i.

5b. How many weeks have you been working? Not Answered ☐

5c. Have you been unemployed in the last 6 months? Yes ☐ No ☐ Not Answered ☐

5d. Do you *currently* have a job where you are on a contract and pay income tax? Yes ☐ No ☐ Not Answered ☐

If "NO" or "Not Answered", please skip to 5e.

If "Yes", is it full time or part time? Full time ☐ Part time ☐ Not answered ☐

What is your weekly wage before tax?

5e. Do you earn money from 'cash in hand' work? Yes ☐ No ☐ Not Answered ☐

If "NO" or "Not Answered", please skip to 5f.

If "Yes", is it full time or part time? Full time ☐ Part time ☐ Not answered ☐

How much, approximately, do you earn per week from cash in hand work?

5f. How many days have you been absent from work in the last 6 months?

5g. Do you think your performance at work has been affected by your depression in the past 6 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 5h.

On how many days in the past 6 months has your productivity been affected?

Was your productivity affected . . . . . Slightly ☐ Moderately ☐ Considerably ☐ Extremely ☐

5h. Have you had an accident at work in the last 6 months? Yes ☐ No ☐ Not Answered ☐

If Yes, how many times?

5i. If you have claimed benefits, which of the following have you claimed and for how long? Please tick as many as apply.

	Yes	No	Not Answered	If Yes, no of weeks claimed in last 6 months
Income support	☐	☐	☐	
Invalidity benefit	☐	☐	☐	
Unemployment benefit	☐	☐	☐	
Sickness benefit	☐	☐	☐	
Housing benefit	☐	☐	☐	
Severe disablement allowance (DLA)	☐	☐	☐	
Mobility allowance	☐	☐	☐	
Family credit	☐	☐	☐	
Child benefit	☐	☐	☐	
Lone parent benefit	☐	☐	☐	
Attendance allowance	☐	☐	☐	
Other (specify below)	☐	☐	☐	

_____

EDUCATION

6a. Please indicate which of the following qualifications you have achieved.

Degree or equivalent ☐
Higher education below degree level ☐

GCE 'A' Level or equivalent

One ☐
Two ☐
Three or more ☐

HND / HNC/ etc.	☐
5 or more GCSE Grades A-C or equivalent	☐
GCSE Grades D-G or equivalent/commercial/apprenticeship	☐
Foreign or other qualifications	☐
No qualifications	☐

LIVING ARRANGEMENTS and ACCOMMODATION

7a. Which of the following best describes your current accommodation?

Owner occupied - owner outright/with mortgage	☐
Rented from council/housing association	☐
Privately rented flat or house	☐
Rented flat or house shared with other people other than family	☐
Temporary accommodation (B&B, hostel, etc)	☐
Temporary accommodation (NHS or other treatment facility)	☐

7b. Are you:

- married or cohabiting with a partner? ☐
- married but not living with partner? ☐
- single and in a current relationship? ☐
- single and not in a current relationship? ☐

7c. Do you / your partner have children whom you live with? Yes ☐ No ☐

- How many children under 16? ☐1 ☐2 ☐3 ☐4 ☐5 ☐6
- What is your relationship to the children?

Parent	☐
Legal Guardian	☐
Other	☐

7d. If you are 'single' and with children are you?

Widowed	☐
Divorced	☐
Separated	☐

SERVICE USE QUESTIONNAIRE

These questions refer to the previous 3 **months**.

<u>USE OF HEALTH SERVICES</u> (excluding specific treatment for depression problem)

1. Have you visited a hospital as a patient during the last three months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1e.

1a. Have you visited a hospital A & E dept in the last 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1b.

How many times? | | | |

How many nights did you stay in total? | | | |

How many times did you use an emergency (999 call) ambulance? | | | |

1b. Have you visited hospital as an inpatient for a medical problem in the past 3 months? (i.e. stayed overnight, but excluding A&E) Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1c.

How many nights did you stay in total? | | | |

1c. Have you visited hospital as an outpatient in the last 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1d.

How many times? | | | |

1d. Have you visited hospital as a day patient in the last 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1e.

How many days? | | | |

1e. Have you seen your GP or a practice nurse at the surgery in the last 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1f.

How many times have you visited the GP? | | | |

How many times have you visited the practice nurse? | | | |

1f. Have you been visited at home in the last 3 months by a GP, practice nurse? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1g

How many times? GP | | | |

Practice nurse | | | |

1g. Have you contacted NHS Direct over the past 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 1h

How many times? | | | |

1h. Have you visited a walk in centre over the past 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 2

How many times? | | | |

MEDICATION

2. Have you received any prescriptions in the last 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 3.

How many? | | | |

DEPRESSION SERVICES CONTACTS

3. Have you received treatment for a depression over the last 3 months other than seeing the trial therapist? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 4.

3a. Have you been prescribed medication in the past 3 months? Yes ☐ No ☐

If so what was the name of the medication
If "NO", please skip to 3b.

What dose in mg per day? | | | | ☐ **If current dose is unknown** *tick this box.*

3b. Were you offered any other services for a problem with depression (such as seeing a counsellor, CPN, Psychiatrist etc)? Yes ☐ No ☐

If "NO", please skip to 3c.

Service	When and how long were the appointments	How many times

3c. Were you prescribed any other medication for psychological problems (such as sleeping, anxiety etc)? Yes ☐ No ☐

If "NO", please skip to 3d.

Drug	Dose	No of prescriptions

This section records other contacts with depression services over the last 3 months.

3d. Have you been treated for a depression problem in the last 3 months where you stayed in a hospital? Yes ☐ No ☐ Don't know ☐

If "NO", please skip to 3e.

How many times were you admitted? | | | |

How many nights did you stay in total? | | | |

3e. Have you received specific counselling or advice in a specialist depression clinic? Yes ☐ No ☐ Not answered ☐

If "NO", please skip to 3f.

What type of counsellor did you see?	Number of occasions seen		
Psychologist			
Keyworker/general counselling			
Social worker			
Psychiatrist			
Other (please specify):			

……

OTHER SERVICE RECEIPT

4. Have you used any of the following services during the last 3 months?

		Number of contacts		
(i)	Advisor regarding state benefits or housing issues			
(ii)	Social Worker (at home)			
(iii)	Occupational Therapist (at home)			
(iv)	Citizens Advice			
(v)	RELATE			
(vi)	Alternative medical practitioner			
(vii)	Advisor on legal or debt issues			
(viii)	A homeless persons agency			
(ix)	An employment advisor			
(x)	Other			

(please specify……..……..……) | | | |

EMPLOYMENT

5a. Have you had a job in the past 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 5i.

5b. How many weeks have you been working? | | | Not Answered ☐

5c. Have you been unemployed in the last 3 months? Yes ☐ No ☐ Not Answered ☐

5d. Do you *currently* have a job where you are on a contract and pay income tax? Yes ☐ No ☐ Not Answered ☐

If "NO" or "Not Answered", please skip to 5e.

If "Yes", is it full time or part time? Full time ☐ Part time ☐ Not answered ☐

What is your weekly wage before tax? | | | | |

5e. Do you earn money from 'cash in hand' work? Yes ☐ No ☐ Not Answered ☐

If "NO" or "Not Answered", please skip to 5f.

If "Yes", is it full time or part time? Full time ☐ Part time ☐ Not answered ☐

How much, approximately, do you earn per week from cash in hand work? | | | | |

5f. How many days have you been absent from work in the last 3 months? | | | |

5g. Do you think your performance at work has been affected by your depression in the past 3 months? Yes ☐ No ☐ Not Answered ☐

If "NO", please skip to 5h.

On how many days in the past 3 months has your productivity been affected? | | | | |

Was your productivity affected . . . . . Slightly ☐ Moderately ☐ Considerably ☐ Extremely ☐

Manual for the 12 session treatment protocol used in study

Behavioural Activation for Depression Facilitated by Mental Health Practitioners

Treatment Protocol

This is a protocol for the delivery of Behavioural Therapy for Depression over 12 sessions. It draws heavily of work by C Martell, S Dimidjian, M Addis and N Jacobson
(see Martell, C. R., Addis, M. E. and Jacobson, N. S. (2003) *Depression in context: Strategies for guided action*, W. W. Norton, New York)

There is also material adapted from the work of D Hopko and colleagues (see Hopko, D., Lejuez, C., Ruggiaro, K. and Eifert, G. (2003) Contemporary behavioural activation treatments for depression: procedures, principles and progress. *Clinical Psychology Review*, **23**, 699-717).

The content of each session is specified to ensure all aspects related to this treatment are covered. With many sessions there are associated diaries and exercise sheets that can be located in the appendix. The process should be followed. If there is any reason to deviate from the protocol this should be discussed in clinical supervision, deviation is acceptable if maintaining the protocol would be to the detriment of the patient.

Session 1

Agenda for session 1

Introductions

Measures

Rational for BA for depression

Assessment information relevant to intervention

Setting homework

Introductions

Introduce self and role (part of a team exploring the use of this evidence based intervention), that your role is as a therapist and this means your first priority is to work in partnership with patients to understand their experiences of depression/low mood and identify ways that should help to make it better over time. Emphasise that it is by working together that we will get the best out of the sessions.

Outline the process, that you will meet for 12 weekly sessions that will be will be structured (using an agenda like this). That you hope the person finds the sessions helpful and are able to discuss any problems that arise in relation to how they are run, getting to them etc, so this can be worked out. Emphasise that the main person who can make changes is the patient and that there will be work to continue between the sessions that builds on what is covered within the sessions. This can be termed as homework (although some do not like that term), and it is the regular use of this homework that will make the most difference.

Also point out that the sessions are focused upon this approach, and that as a therapist your job is to ensure they get the most out of each session by keeping focussed on the problems at hand.

Make sure the person has contact details in case they need to make changes to appointments

Rationale for BA treatment

The treatment that we will be providing for you is based on the belief that the best way to reduce depressive symptoms and to make long-term changes in your life is by looking at how you spend your time, understanding what is important in your life (your values), relating this to the way you feel and then making appropriate changes to your daily activity. In other words, we are going to teach you a different way of structuring your days that will make it more likely that you will experience more positive situations and gradually reduce some of life's problems. It is much more difficult to feel depressed and have low self-esteem if you are regularly engaging in activities that bring you a sense of pleasure and/or accomplishment.

Keep in mind that it is possible for an individual to be quite active yet still depressed. This might occur when activities occur often but are not overly fulfilling or gratifying to an individual. In other words, the activities may be inconsistent with their life goals and values. For example, although a person may be active at work and complete chores at home, sacrificing time and activities with his/her child or spouse (an important life goal) may result in depressive emotions. Thus, it is not how often you engage in different activities, but rather how rewarding or gratifying these activities are for you and how you engage with them.

Depression trap

From what we have discussed, we can see there is a trap in depression; the lower you feel, the less energy you have and the less you do. This leads to a build up of problems, which in turn, make you feel worse and like doing less things. Eventually you end up avoiding many things in life you used to find pleasurable and rewarding. This again results in reduced mood.

This treatment is behavioural in nature, which means that we will work toward changing your behaviour (how you spend your time) as a method for improving your thoughts, mood, and overall quality of life. We will be working on the here and now and making changes designed to improve the future. Some people seek to find a reason/explanation for depression in the expectation that that will make things better. Often these things are hard to identify, or even impossible, even if you do manage you still have to deal with the current situation you are in and make changes in order to move forward. Many individuals with depression often feel lethargic and lack the motivation to engage in various activities. In this treatment, behaviour is changed first, despite feelings of depression or anxiety. Once behaviour is changed, you'll then see fairly dramatic changes in your energy level, motivation, positive thinking, and better moods (less depressed and anxious). In focusing on behaviour change, we do not ignore thoughts and feelings. Instead, we suggest negative thoughts and feelings will change only after you change your behaviour and are experiencing positive events and consequences more frequently. We call this working from **'the outside in'** rather than the '**inside out'**.

Check out rationale with patient – "How does this sound to you?"

Assessment of current depression and problems

Start by asking about current problems in life. For many people, the beginning of depression is clearly related to stressful life events (e.g., loss of a loved one, financial difficulty, job loss, maybe being diagnosed with an illness). For others, the specific causes of depression may be unclear and it may occur from "out of the blue." ***Relate to patient's prior discussion of his/her depression.*** Whether it lasts a couple of weeks or as long as several years, depression may produce:

Significant problems in living (e.g., unable to work, cook, or take care of children). Individuals with depression also may isolate themselves from others, which may result in reduced social support from friends and family, divorce, decreased job satisfaction or unemployment, and educational failure.

Get a picture of this for this person by using semi structured interview, this is major work for this session, take about 30 minutes.

Then ask about depression symptoms. Ask how they have been feeling, use this to introduce the PHQ 9 and (other measures) and explain that you will be using this every session to review such symptoms in a standardised way.

If there is any indication of suicide risk, explore this and check for any plans, intent, methods or contingencies (refer to risk protocol). If this is positive, you will need to inform/discuss this with your supervisor. If it presents in your view a high risk, follow Trust protocols for urgent assessment.

Recap content of session so far. Check out understanding and any problems with the patient. Focus on the 'outside in' nature of the approach. Recap some examples of this.

Then introduce formulation template to begin collaboration and to lead into homework design. Note this is to share the ideas that have been generated via semi structured interview. This will be built on throughout treatment.

Setting Homework

First homework is to self-monitor activities and related mood over the coming week. Emphasise that this does not have to be the whole week, but the more the better. Outline it is to start exploring the relationship between their day and their mood, and that this will help build a treatment plan relevant for them. Ask the patient to have a look at this before next session to see if they observe any patterns.

Introduce Diary 1 (self monitoring) and in session firstly complete **'two hours from 48 hours ago'**. Then use to do '**last 2-3 hours'**. Discuss the difference in ability to complete with accuracy. Discuss the longer you leave it the more how you felt will be influenced by what has followed.

Check out any questions re self-monitoring

Use diary 1 in appendix

Summarise session and confirm next appt

Tools for use in this session

- Semi Structured interview sheet
- Formulation Template
- Daily/Weekly activity schedule (DAS/WAS)
- PHQ 9 (all standard measures for IAPTus)

Session 2

Agenda for Session 2

Measures

Review homework: diaries, discuss any problems

Behaviour-mood link and relate to formulation

Explore the concept of what is depressed behaviour and what are healthy behaviours and the reward that explains why behaviours continue (relief from feelings of pressure) *(note positive/negative reinforcement)*

Treatment goals

Homework

Review diaries

Look at this together; ask what patterns if any they see (what may be healthy or depressed behaviours, check out what this reveals to them, how others might feel if there range of activities were the same). How might that help us consider what changes may help?

Introduce behaviour mood link

Discuss that

Often behaviour is out of awareness and is automatic

Much is done due to habit

To change you have to recognise the habit and its consequence so you know what and when to change.

What types of behaviours were linked to what types of moods in diaries? Is this a common picture?

Through discussion, distinguish between healthy and depressed behaviours in others or prior to depression. What are 'depressed behaviours' (e.g., passivity, avoidance of situations, staying in bed) and what are 'healthy behaviours' (dealing with jobs, meeting with friends, exercising etc)?

Then look over diary and consider some examples of depressed and healthy behaviours from the week. What were the consequences of acting in each way?

Complete a contextual functional analysis (Antecedent/Behaviour/Consequence ABC) in session to look at examples of behaviour and emphasise consequences (and the reinforcing nature of short term relief). Then go back to behavioural activation formulation sheet used in last session and consider how relates.

Use example such as staying in bed instead of getting up and having a shower

Then discuss the function of behaviour. All behaviour serves 1 of 2 purposes, to obtain something positive or to escape something negative. Does it serve to reduce discomfort (avoidance) i.e. we eat dinner because it "feels good" to fill our stomachs and also to obtain essential vitamins and minerals and we can avoid the unpleasant feeling of being hungry, or even worse, starving.

In the session, start to look at examples of depressed and healthy behaviours and what type of rewards were maintaining them (note that depressed behaviours can have a positive reward, such as obtaining sympathy, or increased contact, this can explain why they are hard to break).
How might this help us work to break depression loops? What types of behaviours should be aimed for? Discuss this in session.

Treatment goals

To build on this, you need to consider what the goals are the person has for treatment (that relate to their values). Not 'feeling less depressed' but what behaviours would they like to be doing in the medium term that would indicate things had improved. Use Goals sheet, explain the SMART (specific, measurable, achievable, relevant, timed) nature of goals. Create one in session and then give as homework.

PHQ9

Recap session and check understanding. Elicit any questions.

Homework

Self-monitoring, make notes of behaviour types (healthy/depressed) and the consequence of these.

Goals sheet.

Check out next session planning

Tools for use in session

- ABC formulation sheet
- Behavioural activation formulation sheet
- Goal sheet
- WAS/DAS

Session 3

Agenda for Session 3

Measures

Review Home Work

Link findings to our understanding of depression

Goals and how to move towards them

Activity scheduling

Home Work review

Check out how it went, what observations were made? What different behaviours did they find in their time? Look at diaries together and discuss.

Link findings to our understanding of depression

How does this link to the depression cycle and the need to work from the 'outside in'? Link back to the rationale that depression is strongly linked to the way that we relate to our environment. If we are generally avoiding (due to escape from difficult feelings) and then having subsequent problems with an absence of positive or healthy experiences, this may help us to understand why such feelings are maintained.

Therefore the patient needs to make some changes and start to schedule activities that will re-introduce such healthy behaviours in place of avoidant behaviours.

Seven steps to change:

Identify situations/behaviours that depress you

Develop ideas of alternative behaviours

Schedule alternative behaviours – small steps not giant leaps. Experiment with alternative behaviours

Develop experimental attitude to depression, challenge negativity and try none the less

Try out behaviours – attend to the action and not thoughts/evaluations. Keep trying

Evaluate results

Continue to experiment – no quick fixes. Examine results – place in schedule 2 – 3 weeks

Goals

We are now in steps 2-3.
Look at homework goal sheet.

Now look at alternative behaviours. Introduce activity sheets (routine, pleasurable, necessary) and consider in relation to goals set for treatment. Start to generate a list of SMART behaviours in session that move the person towards goals. Do about 10, across all areas and then list in order of difficulty.

Then in session you can use 'behaviours that make me feel bad' exercise sheet and explore some options for behaviour change. Use self-monitoring sheets and discuss 2-3 situations and what may have been useful 'alternative behaviours' using previous worksheet (remember in all these sessions you must be using collaborative discussion, getting the patient to generate ideas, not therapist's actions, your job is as a coach).

Activity Scheduling

Then introduce scheduling, ask patient to consider what would be the best time to schedule in some of the changes discussed in today's session. Then do for 2 of those days together considering alternative behaviours that are linked to the goals. Remind the patient of the 'outside in' approach and that it is important not to wait to feel ready to do these things, but to do them to help feel better over time. Discuss this as an experiment, so on the days scheduled, follow plan and continue to monitor mood to see what happens.

Finally recap what you have covered in treatment so far, how it links up to help to have a practical plan to overcome depression. Use this recap to highlight the structured step-by-step approach we are adopting, rather than the reactive to internal state approach (inside out).

Before end of session complete PHQ9.

Homework

Follow scheduling and monitor. Continue to try to recognise what is depressed and what is healthy behaviour. Introduce alternatives if you can.

Tools for use in session

- Goal sheet 2, routine, necessary & pleasure
- Behaviours that make me feel bad
- Use self monitoring forms to schedule
- PHQ9

Session 4

Agenda for Session 4

Measures

Review homework

Discuss changes made and the consequence of them

TRAP and TRAC

Home Work review

What did they observe in relation to scheduling, what are the healthy behaviours that they endeavoured to use to replace depressed behaviours*? **Look through dairies/behaviours that make me feel bad worksheets together in discussion***. Congratulate any success discuss any problems. Use results to consider in relation to treatment rationale. What may be the use of scheduling in relation to goals? Acknowledge how difficult it is to change and how it will have a mix of success and failure. Speak of using the approach to understand how best they can make changes in life to give "***best chance of feeling better"***. Look at where scheduling didn't work so well and consider what may be the explanation (are they working from the 'inside out', waiting to feel ready, was the goal set too high, was it not specific enough, is it focussed towards a goal that is useful for them).

Introduce the role of avoidance in depression maintenance, relate to the cycle of depression you have been using throughout sessions. **Explain TRAP/TRAC:**

Trigger	**T**rigger
Response	**R**esponse
Avoidance	**A**lternative
Pattern	**C**oping

This framework can be used to look at where depressed behaviours exist and help explain why these become so entrenched; that is, the reward for a trap is initial relief from discomfort. What however will be the long-term consequence on mood (**discuss in session**)? Helps us identify what are the things being avoided and why we are avoiding them. In session use 2 TRAP/TRAC sheets based upon previous weeks experiences as examples. Then give as homework to complete when engaging in depressed (avoidant) behaviour.

Homework

To schedule incorporating goal orientated activities.
Use TRAP/TRAC worksheet when not engaging in scheduled activities.

Tools for use in session

- Monitoring scheduling sheets
- TRAP/TRAC forms

Session 5

Agenda for Session 5

Measures

Review homework

Getting on TRAC (Review again TRAP/TRAC)

What have I learnt so far re my mood?

Homework planning

Homework review

Look at scheduling and mood, discuss how using to move towards goals and the activities placed in line with this. Discuss how they have found doing this and what it may mean for managing depression in the future.

Next look at TRAP/TRAC sheets and discuss how they help identify avoidances that are 'blocking' access to new more adaptive and useful experiences. Review the role of negative reinforcement in increasing such blocks and how as you become aware of them you have some power in relation to change.

Discuss how TRAC helps to overcome blockages and decide to do something different that increases interaction with important 'events' in the environment hence increasing positive reinforcement and moving towards goals. Go over 2 examples related to TRAP sheets for past couple of weeks together.

Treatment review

Next in session spend some time looking at outlining what you have covered in treatment and discuss what it means for the person. Introduce 'what I have learnt about my depression so far worksheet' and ask for it to be reviewed during the following week. The aim is to spend some time to review the different approach to managing depression so the person can become their own therapist.

Complete PHQ9

Homework

Continue to schedule. Review how they are progressing in relation to task lists and goals.

Using TRAP/TRAC sheets when noticing **'depressed behaviour'** to help finding alternative behaviours.

Spend time on **'what I have learnt about my depression'** worksheets

Tools for use in session

- Self monitoring/scheduling
- TRAP/TRAC sheets
- What I have learnt about my depression
- PHQ9

Session 6

Agenda for Session 6

Measures

Review homework

What TRAP/TRAC helped me do

Using ACTION to help summarise

Review homework as per previous sessions

Look at 'what I have learnt about my depression worksheet' and discuss findings. Anything that is not clear discuss openly. Discuss the role of changing avoided behaviours to improve how you feel and to deal with life problems. Consider any problems openly and discuss, recognise the powerful drive of responding to internal feelings, and how it explains reduced/avoidant behaviours. Remember to consider the pros & cons of choices made and how it is best to focus on own behaviours rather than others. Review examples of TRAP/TRAC sheets and discuss how these have been used. Then introduce the alternative behaviours worksheet to reduce need for writing. This however is to be used using the TRAP/TRAC process.

Next introduce **ACTION** as a way of analysing response to situations:

Remind that:

Preparing to change:

Open mind
Experiment
Small steps
Expect ups and downs in process
Not 'Just do it' as would have already.

Assessing situations:

Understanding patterns
Recognising when happening

Choosing alternatives:

Recognise TRAPS and work out how to get on TRAC.
Make choice based on knowledge.

Try:

Pick out time and day/s
Put on schedule
Expect to be hard
Don't give up

Integrate:

Try several times before making decision
Expect some unexpected results

Observe results:

Get used to checking feeling – just before, during and after

Now Evaluate/**N**ever Give Up:

Did behaviours help mood
Did you manage
Was it helpful within life
Continue?
What learned

ACTION – reminder

A - Assess mood and behaviour
C - Choose alternatives
T - Try out
I - Integrate into life
O - Observe results
N - Now evaluate/never give up

Go over steps and give sheet to person to use. Discuss how scheduling is at the core of behavioural treatment, and that TRAP/TRAC & ACTION are designed to help identify what to do in schedules to make things better. Revisit Goal sheet and re rate to assess progress towards them. You can use turning reactive-proactive behaviour sheet.

Home Work

Continue to schedule mix of activities. Use worksheets and learning to identify behaviours to incorporate and how to break down to small steps towards goals.

Tools for use in session

- Alternative behaviours worksheet
- ACTION /reactive-proactive sheet
- WAS
- PHQ9

Session 7

Agenda for Session 7

Measures

Review homework

How to deal with depressed thinking/worry

Homework

Homework review. How do you schedule now, how do you choose what to incorporate. Reflect on tools used so far

<u>Review homework as per previous session</u>

Pay particular attention to progress towards goals linked to schedules. Discuss how using ACTION sheets and alternative behaviour sheets to consider how to plan each day. Reinforce valued direction and reflect on benefits of this.

Worry and thoughts – How to respond

Exercise in session can we control thinking:

Use elephant exercise!

If they can't stop such an inert thought what chance is there of doing so with emotional thought

Content of thinking)
> Distinguish
Function of thinking)

Commonly we focus on content but in this treatment we are most interested in function.

Common examples of function:

Try to figure why depressed, trying to solve it.
Think about difficulties and feelings whilst in bed (rather than get up).
Worry about possible bad events in the future, so as not to think about the present.
Hold on to bad thoughts as to let go would be to give up / let others off the hook.

Thinking in these situations can be seen as a behaviour you do, hence you could choose to do something else.

<u>Golden rule:</u>

What is the thinking doing for you at a given time?
Are you aware of it?
What else is there to do?

> *A metaphor for thoughts and urges is traffic on a road. Engaging with thoughts is akin to standing in the road and trying to divert the cars (and getting run over) or trying to get one and find a parking space for it. However, even if one manages to divert or park one car there are always more to be dealt with. The goal is to acknowledge the thoughts but not to attempt to stop or control or answer back at them. The aim is to accept fully aversive thoughts and to 'walk along the side of the road', engaging with life despite the traffic, which one can quietly ignore. (Taken from Veale 2008)*

Rumination

Outline rumination – dredging up and turning over such as worry, brooding, turning problems over and over, often about:

Relationship problems
Past hurts
Future worries
Money problems
Bad decisions
Feelings
What others think

Exercise in session:

What do you ruminate on?

Induce rumination on this subject for 2 minutes within session as a collaborative exercise.

Then:

How do I feel when I do this?

Then note rumination stops problem solving and leads to internal focus. This can lead to avoidance of actions that change the situation.

Relate to noisy road example!

Homework

Continue scheduling towards goals

Monitoring rumination using worksheet

Noticing rumination is the start – use examples of this

Consider this when:

You notice thinking over and over the same thing
It doesn't give you any solutions

After two minutes ask:

Has it helped solve a problem?
Do I have any new understanding?
Am I less self critical/depressed?

Tools for use in this session

- WAS
- Monitoring rumination sheets

Session 8

Agenda for Session 8

Measures

Homework review

Consider rumination, use diaries

Introduce RCA

Homework review as per previous sessions

Go over scheduling and discuss any issues, review progress towards goals and consider blocks or developing further goals.

Review rumination diaries

Review rumination diaries and discuss the consequences of rumination on mood and activity. Where was attention focussed at this time? Even if active can be focussed on internal rather than external cues. Go back to discussion regarding the function/consequence of rumination and discuss. Review the noisy road metaphor. To lead on to RCA (Martell et al 2004).

Introduce RCA

Session theme RCA:

Rumination
Cues
Action

Rumination can make you feel worse, plus you have tried to get better at recognising so now practise changing. Use RCA.

Remember the importance of attending to your experience.

In session exercise:

Fill out RCA form for three ruminating situations:

	Situation	Rumination	Cues	Action
1			→	1. 2.
2				
3				

Use at home for others and practise.

Negative thoughts

Discuss the following:

When they come:

How do you feel?
Do you try to debate, stop them?
Does it work?

Try to notice, label and move on.

Remember focus attention on what is happening what you are doing, self-soothe through your five senses. Focus intensely on sights, sounds, smells, tastes, and touch. Appreciate and understand the world around you. Go through a practice exercise with the patient where you have them sit back and relax. Have them take a couple deep breaths. Then, have them describe everything they see, hear, smell, and feel (in great detail). At the end, ask the patient how this experience related to negative thoughts. Did they have any (very unlikely if they were doing the exercise correctly)? Ask the patient how this exercise might be incorporated into their daily life. When are the times (days, hours, or situations) when the patient is most likely to ruminate? Can you schedule activities at these times to minimize the likelihood of ruminative experiences?

Note how this is different to ignoring thoughts, or dismissing them. It is about recognizing them, how they make us feel and deciding that rather then ruminate on them getting on with goal orientated activity is more likely to be productive, and help feel better.

Home Work

Practise RCA alongside Scheduling.

Tools for use in this session

- Scheduling and goal sheets
- Monitoring rumination
- RCA sheets
- Paying attention sheets

Session 9

Agenda for Session 9

Home Work review (RCA)

Recap topics covered in treatment

Problem solving

Treatment review

Recap what has been used in treatment up to this point (e.g. Cycle of depression, Reduced activity and avoidance, Mood and action, Working from outside in, Scheduling, Avoidance = TRAP/TRAC, ACTION, Thoughts and ruminations and depression, RCA)

Here on in it is practise of building into daily life and accepting varied responses. Focus on small steps, one at a time. Do you need to wait for motivation?

Establish if any area remains troublesome. Review goal list and progress towards this. What areas remain problematic? Return to and use tools relevant to this.

How to solve a problem

Review 7 stages of problem solving

Identify problem

Establish SMART goals

Generate list of possible steps to meet goals

Review pros & cons of each

Pick one Outline detailed steps

Do

Evaluate

Finally return to what induces my behaviours, and discuss with the patient the role of reactive/proactive behaviours.

Use PST worksheet.

Home Work

Schedule, use worksheets as needed to help have balance and goal orientated approach.

Use problem solving sheets for difficulties in week

Tools for use in this session

- Scheduling/goals
- Any identified in treatment review
- Problem solving sheet
- PHQ9 measures

Session 10

Agenda for Session 10

Measures

Review Home Work tasks

Acting 'as if'

What I have learnt about depression

Home Work review

What did they use, what problems? Remember action and working from 'outside in'. Review any use of problem solving.

Discuss the role of how we act in relation to how we feel

Exercise in session:

Role Play social situation. ***1st five minutes act as if low (poor eye contact, slow speech, etc), 2nd five minutes act as if o.k.***
Compare and contrast. How will affect mood and goal attainment. What does this mean?

What you have learnt about depression

There are now only two more sessions. Emphasise that treatment is aimed at making them their own behaviour expert. Discuss continued rehab after treatment:

Discuss treatment as road to recovery
Plan what ongoing goals may be
What have they learnt from sessions?
How may this help with 'life ahead'
Learning to solve problems
Seven stages of problem solving

Introduce concept of flash card to manage mood in future. First step is to review what they have learnt about depression. Start worksheet in session and complete as homework.

Home Work

Problem solving and link to schedules

Use acting as if and monitor outcomes

Complete ‘what I have learnt about my depression’

Tools for use in session

- What I have learnt about my depression
- WAS
- Measures

Session 11

Agenda for Session 11

Measures

Review Home Work

Progress towards goals/Scheduling

What I have learnt about depression

Relapse prevention Flash cards

Home Work review

Any problems, solving/acting as if review outcome.

Review 'what I have learnt about depression' sheets and discuss how this differs from before treatment. Then link to the **'Relapse Prevention flashcard**'.

Give **'Relapse Prevention flashcard'** to keep with what I have learnt flashcard. Discuss working on these over next week, do 1-2 examples in session.

Home Work schedule

Do the relapse prevention keycard

Tools for use in session

- Relapse prevention flashcard
- WAS

Session 12

Agenda for Session 12

Measures

Homework review

Keycards and how to use them

Home Work review

Recap of treatment and what learning has taken place. Have full discussion on the process adopted throughout treatment. Explore goals set by person and how these have been progressed. At all times re emphasise that proactive, non-avoidant behaviour and working from the outside in has helped get things done, feel better although it is often hard and there is desire to withdraw.

Note that all the tools used in treatment are there for future use, and keeping these up will help maintain a healthy mood in the future. Review keycards and how to use them.

Inform the patient that they will be passed back to GP (or worker who referred them) and a brief letter will be sent, that will be copied to them.

Take some time to say goodbye as this may be difficult for the person

Data error checklist

Error	Solution/ Action taken
Different Freq employment scores between data sets	Found inconsistencies (Subject No's 3, 18, 25, 35 – was PT when referred, 45) and checked original file for correct data
CSQ data inconsistencies	Found errors (Subject No's 16 and 44), corrected according to original file.
Basecsirdiag different Freq scores	Found errors (Subject No's 3 and 18), corrected according to original file.
BDIcatpre different Freq scores	Found error (subject no 31), check BDI total to get correct categories
BDIpostsev different Freq scores	Found error (Subject no 35), checked BDI total to get correct category
BDIpostcat different Freq scores	Found error (subject no's 8, 15, 18, 38, 43) checked BDI total to get correct category
CSQ1 and CSQ2 different Freq scores	Could not find difference when reading data manually but could have been changed inadvertently when editing earlier data.
CSQ6 different Freq scores	Corrected subject 4's error by referring to original file
CSQ8 Freq error	Could not find difference when reading data manually but could have been changed inadvertently when editing earlier data.
Referrer difference freq scores	Could not find difference when reading data manually
BDI1Month 3	Compared files, found errors (4, 24), and referred to original data.
BDI1month17	Compared files, found errors (7), and referred to original data.
onemonthwasa	Compared files, found errors (3, 4, 9), and referred to original data.
BDIpost all data	Subject 9 all data entered incorrectly
BDI9post	Compared files, found error (28), and referred to original data.
CSQ1, CSQ2, CSQ6 and CSQ8 Scores	Could not find reason for differences in mean scores
BDI Pre scores	Compared files, found error (3, 4, 18), and referred to original data.
basewasafull	Compared files, found error (3, 4, 7, 9, 31, 45, 47), and referred to original data.
Basewasahome	Compared files, found error (3, 4, and 31) and referred to original data.
All 2month data	Compared files, found errors, and referred to original data.

Appendix III: dissemination, research publications and subsequent research protocols

This research has been disseminated through conference presentations both nationally and internationally. It has also led to publication in high impact peer reviewed journals with an international audience.

List of Conference presentations

Meta Analysis

2007

British Association Of Behavioural and Cognitive Psychotherapies- July 2007 Brighton UK. Oral presentation open paper session. BA for depression a systematic review.

Network for Psychiatric Nursing Research-September 2007 Cambridge UK. Oral presentation symposium new approaches. A systematic review of BA treatment of depression and the potential relevance for psychiatric nurses.

2008

Australian College of Mental health Nurses-September 2008, Melbourne, Australia. Open oral presentation. Behavioural Treatments of Depression: Do they Work And Can Mental Health Nurses Deliver Them Effectively.

Meta Analysis and Trial

2009

British Association Of Behavioural and Cognitive Psychotherapies- July 2009, Exeter UK. Open oral paper presentation,. Behavioural Activation for depression summary of the evidence base and development in a randomised controlled trial by non specialists.

2010

European psychiatric Nursing Congress. HORATIO. Aril 2010, Prague Czech Republic. Oral presentation. Behavioural Activation for depression delivered by mental health nurses. A systematic review of the evidence and controlled clinical trial.

Royal College of Nursing Research Society. International Nursing Research Conference. May 2010, Gateshead UK. Open oral paper. Behavioural Activation for depression delivered by mental health nurses. A systematic review of the evidence and controlled clinical trial.

Health Care Events National Depression Conference June 2010 London UK. Oral presentation. Behavioural Activation for Depression.

British Association Of Behavioural and Cognitive Psychotherapies- July 2010, Exeter UK. Open oral paper presentation. Behavioural Activation for depression summary of the evidence base and results of a randomised controlled trial by non specialists.

Network for Psychiatric Nursing Research-September 2010 Cambridge UK. Oral presentation. Behavioural Activation for depression summary of the evidence base and results of a randomised controlled trial by non specialists.

York Primary Care Conference-March 2011 York UK. Oral Presentation. Behavioural Activation for depression summary of the evidence base and results of a randomised controlled trial by non specialists.

3rd International Nursing and Midwifery Conference- April 2011 Galway Ireland. Oral Presentation. Behavioural Activation for Depression by Mental Health Nurses. A randomised Controlled Trial of Clinical and Cost Effectiveness.

Publications from research

Copies of papers are included below

Ekers D, Richards D, Gilbody S. (2008) A Meta Analysis of Behavioural Therapy for Depression. *Psychological Medicine*; **38**(5): 611-623.

Ekers D, Richards D, McMillan D, Bland M, Gilbody S. (2011) Behavioural Activation delivered by the non specialist: phase II randomised controlled trial. *British journal of psychiatry*; **198**(1): 66-72

Contribution to new research development

CASPER

The Meta analysis was used in the design of the HTA funded multi centre study Collaborative care for screen positive elders: the CASPER Trial. D Ekers is a co applicant on this study due to experience and knowledge gained in BA as part of this thesis.

COBRA

The Meta Analysis and RCT have been developed into a large scale multi centre study design COBRA (Cost and Outcome of BehaviouRal Activation): a Randomised Controlled Trial of Behavioural Activation versus Cognitive Therapy for Depression. The COBRA study is designed to expand on the findings reported in this thesis as a definitive trial. The results from this research form the basis of this study design and findings have informed sample size, recruitment intervention design and therapist issues. The relevant section of the COBRA protocol is included in appendix III.

D Ekers is a co applicant of the COBRA study which has been submitted for HTA funding and has progressed to second stage review.

D Ekers is also on the advisory board of 2 further studies exploring dissemination of BA. Firstly to examine BAs effectiveness in Muslim communities and secondly adapting BA for delivery to children and adolescents. Both studies are in early phases of development and have been influenced by the research findings reported in this thesis.

Copy of Ekers et al 2008 A Meta Analysis of randomised trials of behavioural treatments of depression

Psychological Medicine (2008), 38, 611–623. © 2007 Cambridge University Press
doi:10.1017/S0033291707001614 Printed in the United Kingdom

REVIEW ARTICLE

A meta-analysis of randomized trials of behavioural treatment of depression

D. Ekers[1]*, D. Richards[2] and S. Gilbody[2]

[1] *Tees Esk & Wear Valleys NHS Trust/University of York Department of Health Sciences, The Health Centre, Newcastle Road, Chester le Street, Co. Durham, UK*
[2] *Department of Health Sciences, Seebohm Rowntree Building, University of York, York, UK*

Background. Depression is a common, disabling condition for which psychological treatments, in particular cognitive behavioural therapies are recommended. Promising results in recent randomized trials have renewed interest in behavioural therapy. This systematic review sought to identify all randomized trials of behavioural therapy for depression, determine the effect of such interventions and examine any moderators of such effect.

Method. Randomized trials of behavioural treatments of depression *versus* controls or other psychotherapies were identified using electronic database searches, previous reviews and reference lists. Data on symptom-level, recovery/dropout rate and study-level moderators (study quality, number of sessions, severity and level of training) were extracted and analysed using meta-analysis and meta-regression respectively.

Results. Seventeen randomized controlled trials including 1109 subjects were included in this meta-analysis. A random-effects meta-analysis of symptom-level post-treatment showed behavioural therapies were superior to controls [standardized mean difference (SMD) −0.70, 95% CI −1.00 to −0.39, $k=12$, $n=439$], brief psychotherapy (SMD −0.56, 95% CI −1.0 to −0.12, $k=3$, $n=166$), supportive therapy (SMD −0.75, 95% CI −1.37 to −0.14, $k=2$, $n=45$) and equal to cognitive behavioural therapy (SMD 0.08, 95% CI −0.14 to 0.30, $k=12$, $n=476$).

Conclusions. The results in this study indicate behavioural therapy is an effective treatment for depression with outcomes equal to that of the current recommended psychological intervention. Future research needs to address issues of parsimony of such interventions.

Received 30 March 2007; Revised 25 July 2007; Accepted 31 July 2007; First published online 1 October 2007

Key words: Behavioural therapy, cognitive behavioural therapy, depression, meta-analysis, psychotherapy.

Introduction

Depression causes substantial disability, is set to become the second largest cause of disease burden by 2020 (WHO, 2001), affects between 5% and 10% of the population and is the third most common reason for primary-care consultation (Singleton *et al.* 2001). It is associated with significant distress, impairment of functioning, disturbance to interpersonal relationships and an increased risk of suicide (Hirschfeld *et al.* 1997). Psychological treatments, particularly cognitive behavioural therapy (CBT) are recommended to treat depression (Hollon *et al.* 2002; NICE, 2004), however, less than 10% of those affected receive such treatment (Singleton *et al.* 2001).

* Address for correspondence: D. Ekers, Tees Esk & Wear Valleys NHS Trust/Department of Health Sciences, The Health Centre, Newcastle Road, Chester le Street, Co. Durham DH3 3UR, UK.
(Email: David.Ekers@cddps.nhs.uk)

CBT combines both behavioural and cognitive techniques in each treatment programme. The standard approach is Beck's cognitive therapy (Beck, 1976) using both behavioural and cognitive techniques to identify, question and modify maladaptive thought processes, life rules and core beliefs.

However, recent research has suggested that pure behavioural models utilizing an operant conditioning formulation to develop a structured daily action plan may be as effective as full cognitive therapy (CT) (Jacobson & Gortner, 2000, Jacobson *et al.* 2001). Ferster (1973) pioneered the early incorporation of learning theory to the treatment of depression in the 1970s followed by the establishment of the 'coping with depression' intervention (Lewinsohn & Graf, 1973). With the development of cognitive models, behavioural interventions lost popularity, until recent renewed interest led to research reminding us of their potential (Jacobson & Gortner, 2000). The optimum combination of behavioural and cognitive techniques within CBT is unknown (Jacobson *et al.* 1996).

Behavioural therapy (BT) may provide a more parsimonious treatment option as it may be simpler to deliver (Jacobson *et al.* 1996). If similar health outcomes could be achieved with such a lesser 'dose' of psychotherapy, service and training procedures could be radically overhauled. Narrative reviews conducted by advocates for behavioural approaches suggest positive outcomes (Martell *et al.* 2001, Hopko *et al.* 2003), however, such reviews are prone to bias. Alternatively, systematic reviews of psychotherapy for depression have looked at behavioural interventions in the context of considering the effect of other psychological approaches (Dobson, 1989; Gloaguen *et al.* 1998; Churchill *et al.* 2001).

Therefore, we conducted a systematic review of randomized controlled trials of behavioural interventions for depression compared to other psychological approaches and controls. We explored effectiveness in terms of depressive symptoms, dropout and recovery rates.

Method

Identification of suitable studies

We searched a range of databases from inception to January 2006 (Medline, EMBASE, PsycINFO, Cochrane Library DARE, CINAHL, AMED and the British Nursing Index), incorporating randomized controlled trial filters. We reviewed reference lists of identified studies to find additional trials. Two authors (D.E. and D.R.) considered abstracts and screened the full text of selected studies for relevance.

Inclusion criteria

We included all available randomized controlled trials in any language to reduce the potential for publication bias (Khan & Kleijnen, 2002). Studies included participants who were adults (aged $\geq$16 years), treated in community or in-patient settings with a primary diagnosis of depression. We excluded studies including participants with psychosis or bipolar disorder, substance misuse problems, cognitive impairment or without depression as primary diagnosis. We included trials of individual time-limited behaviourally orientated psychotherapeutic approaches to the treatment of depression with an alternative psychotherapy, control with confirmation of randomized allocation.

BT

We included trials in the behavioural intervention group if the treatment was based upon the rescheduling of activities to reintroduce positive reinforcement and reduce avoidance. Such interventions manipulate the behavioural consequence of a trigger (environmental or cognitive) rather than directly interpret or restructure cognitions.

Comparators

Treatment as usual/control. A range of standard treatments or non-treatment options (waiting list, usual general practitioner treatment, inert control conditions) delivered to the patient in the absence of any 'active' psychotherapy.

CBT/CT. Interventions that directly identified, questioned and modified cognitive responses to situations and their emotional consequences. We included any intervention conceptualized as an intervention to directly challenge thinking including 'thought catching' and 'challenging' through diary-keeping or behavioural experiments.

Brief psychotherapy. Approaches that focused on developing insight and subsequent character development through interpersonal relationships with the therapist, including brief interpersonal therapy (IPT; Klerman *et al.* 1984) or brief psychodynamic therapy (Luborsky *et al.* 1995).

Supportive counseling. We included any approach which focused upon the therapist's use of core relationship conditions (Rogers, 1961) to develop self-awareness by the patient.

We excluded marital, couple or group therapy as the change in therapist contact coupled with other group-member interaction would introduce substantial clinical heterogeneity and was outside the aims of this review.

Outcome measures

Our primary outcome measure was depression symptom-level self-rated [e.g. Beck Depression Inventory (BDI); Beck *et al.* 1961] or clinician-rated [e.g. Hamilton Depression Rating Scale (HAMD); Hamilton 1960], presented by means and standard deviations (continuous data) or clinical improvement/non-clinical improvement (dichotomous data). As psychotherapy trials often present multiple symptom measures we adopted an algorithm so that validated self-report measures took precedence over clinician-rated measures and performed sensitivity

analysis to explore the impact of this approach. We entered recovery and dropout rates as dichotomous data, dropout being viewed as a proxy for acceptability.

Quality assessment

Two authors (D.E. and D.R.) rated study quality using criteria to explore bias (Khan *et al.* 2002). Other than concealment of allocation, clear guidance on aspects of study quality that directly influence outcomes is unclear (Jadad *et al.* 1996; Schultz & Grimes, 2002). We assessed studies against two standards each relating to selection, measurement, performance and attrition bias resulting in an overall score of between 1 and 8. Disagreements regarding study quality were dealt with through discussion.

Data extraction and synthesis

We extracted data from each trial at post-treatment and follow-up (6 months or nearest available dataset). We synthesized data using the Cochrane collaboration RevMan program (Cochrane Collaboration, 2003). We sought missing data from study authors by email. We imputed missing standard deviation (S.D.) scores from other relevant studies where these data were not available following the above procedure (Furukawa *et al.* 2006).

Data pooling

We combined continuous data to estimate the standardized mean difference (SMD) across trials to facilitate analysis of the same outcome (depression symptom level) using different scales as a standardized unit (SMD). Where studies included two comparisons under the same category (i.e. CT and CBT) we entered these comparisons separately but halved numbers in the behavioural arm to avoid double counting and inaccurate weighting of trials. Where studies presented results using subcategories (e.g. high/low depression severity), we entered data as two separate trials, provided that stratification occurred prior to randomization. We assigned effect sizes according to the standard convention where the SMD is small (0–0.32), medium (0.33–0.55) and large (≥0.56) (Lipsey & Wilson, 1993). We present dichotomous data for dropout and recovery rate as odds ratios (OR), which demonstrates the chance of an event (improvement or dropout) in the intervention group compared to the comparison group. We present pooled data with 95% confidence intervals (CI) using a random-effects model (Sutton *et al.* 1998) taking into account both within- and between-study variance. We consider such a model as appropriate based upon anticipated heterogeneity for this review (number of sessions, therapy approaches and setting, etc.).

Exploration of heterogeneity

We measured statistical heterogeneity using the I^2 statistic for statistical variation across studies (Higgins *et al.* 2003); values of 25% are low, 50% moderate and 75% high.

Three sources of clinical and statistical heterogeneity were identified *a priori*: (1) baseline severity of depression; (2) training level of the therapist (graduate *versus* postgraduate/experienced therapist qualification); (3) number of sessions. We considered study quality as a source of potential heterogeneity, by assessing the impact of lower quality studies on overall outcomes; using a cut-point of 6 on the 8-point quality scale.

We explored the impact of these sources of heterogeneity using sensitivity analyses and meta-regression (Thompson & Higgins, 2002). We analysed outcomes using meta-regression, specifying sources of heterogeneity as predictive covariates. We used a permutation test (using 1000 Monte-Carlo simulations) to calculate p values, and to reduce spurious false-positive findings (Higgins & Thompson, 2004). The amount of heterogeneity explained by predictive covariates was examined by reductions in the I^2 inconsistency statistic within our model. Analyses were conducted using the MIXAN and METAREG commands in Stata 8 (Stata Corporation, 2003).

The possibility of publication bias was assessed through a Begg funnel plot graph (Begg, 1994) and testing for asymmetry using the Egger weighted regression test (Egger *et al.* 1997) where the intercept is 0 if no bias is present.

Results

Searches conducted between December 2005 and February 2006 identified 3353 studies (see Fig. A1 for study flow chart; available in online Appendix). We identified 20 randomized controlled trials (Table 1), three of which were excluded from the meta-analysis [18–20] due to insufficient reported data. (*Note*: throughout the following sections numbers within square brackets refer to the Study numbers listed in Table 1.) We meta-analysed the remaining studies which included 1109 subjects (Table 2).

Table 1. *List of studies included in review*

Study no. (first-named author and year)	Sample/setting Mean age [s.n. (range)] Sex (% female)	Interventions (*n* in cell)	Depression level at baseline	Concurrent pharmacology	Therapist level Session number (duration)
[1] Taylor (1977)	University students 22.4 (2.6) 71	Behavioural (7) Cognitive (7) Cognitive behavioural (7) Wait list (7)	Mild/moderate (21.2 BDI)	No	Graduate student experience as counsellor 6 (40 min)
[2] McLean (1979)	Community out-patient 39.2 (10.9) 72	Behavioural (42) Brief psychotherapy (44) Drug therapy (49) Relaxation (43)	Within or beyond moderate depression range 2 out of 3 measures used at baseline	No (other than DT arm)	Licensed psychologists, physicians or psychiatrists. At least 2 years of experience as therapist 10 (1 h) not drug therapy
[3] Gallagher (1982)	Older adult community 67.76 (6) 76	Behavioural therapy (10) Cognitive therapy (10) Psychotherapy (10)	RDC Criteria MDD	No	Advanced Ph.D. or post-doctoral therapists experience in modality 16 (90 min)
[4] Maldonado Lopez (1982)	Community out-patient N.A. N.A.	Behavioural (8) Cognitive (8) Drug therapy (8)	Psychiatrist Diagnosis reactive depressive disorder	No (other than DT arm)	Psychology Dept, level of training not reported 10 (1 h)
[5] Wilson (1982)	General population media announcements 38.8 (20–55) 66	Drug therapy & Behavioural (12) Drug therapy & Relaxation (10) Drug therapy & Minimal contact (10) Placebo & Behavioural (9) Placebo & Relaxation (11) Placebo & Minimal contact (12)	BDI > 19	In DT arm	Graduate psychologist 7 (1 h) 2 (1 h) in min contact arm
[6] Wilson (1983)	General population media announcements 39.5 (20–58) 80	Behavioural (8) Cognitive (8) Wait list (9)	BDI >17 (moderate depression)	Yes (5 subjects in trial)	Not clear, University Psychology Clinic 8 (1 h)
[7] Cole (1983)	Community out-patient veterans 56 (24-71)	Behavioural (15) Treatment as usual (15)	Psychiatrist Diagnosis major depression. BDI > 24	Yes if stable	Doctoral clinical psychology student 7 (1 h)
[8] Maldonado Lopez (1984)	Community out-patients N.A. N.A.	Behavioural & Pharmacology (8) Cognitive & Pharmacology (8) Pharmacology (8)	Psychiatrist Diagnosis reactive depressive disorder	All subjects	Psychology Dept, level of training not reported 10 (1 h)
[9] Skinner (1984)	Community volunteers 20–61 (34) 67.5	Behavioural (8) Cognitive(7) Control (9)	BDI > 12	Yes	Doctoral clinical psychology student 5 (1 h)
[10] McNamara (1986)	University students 23 (19-31) 73	Behavioural (10) Cognitive (10) Cognitive behavioural (10) Supportive (10)	BDI > 17 HAMD > 20	Not reported	Doctoral interns in clinical psychology / masters-level social worker 8 (50 min) (10 sessions in CBT arm)

[11] Thompson (1987)	Older adults community 67.07 (5.8) 67	Behavioural (25) Cognitive (27) Psychotherapy (24) Delayed (19)	RDC MDD	If stable dose for 3 months	Doctoral-level psychologists plus 1 year specialized therapy training 16–20 (duration of each session not reported)
[12] Scogin (1989)	Older adults community 68.3 (6.7) 85	Behavioural bibliotherapy (23) Cognitive bibliotherapy (22) Delayed (22)	>9 on HAMD	If stabilized prior to trial	N.A.-as bibliotherapy was main intervention 4 (5 min) phone contacts to support exercises
[13] Jacobson (1996)	Community (80% HMO, 20% volunteer) 38 (not reported) 72	Behavioural (56) Thought challenging (43) Full cognitive (50)	Major depression (DSM-IV) > 19 BDI	No	Experienced therapists (mean 9.5 years CT practice) 20 sessions (N.A.)
[14] McKendree Smith (1998)	Community volunteer 44.88 (13.17) 75	Behavioural bibliotherapy (13) Cognitive bibliotherapy (13) Delayed control (14)	Mild-moderate depression	If stabilized for 3 months	N.A.-as Bibliotherapy main intervention 8 (10 min)
[15] Hopko (2003a)	In-patients 30.5 (9) 36	Behavioural (10) Supportive (15)	Principle diagnosis of major depression	Yes all patients	Not clear 6 (20 min)
[16] Dimidjian (2006)	Community 39.9 (10.97) 66	Behavioural (43) Cognitive (45) Pharmacology/Placebo (153)	Major depression (DSM-IV) > 19 BDI	Only in ADM arm	BA-licensed psychologists/social worker (7 years practice. CT-licensed psychologists with CT training 24 (50 min)
[17] Cullen (2006)	Community 38.48 (12.69) 32	Behavioural (13) Wait list (12)	MDD (Mean BDI) 30.96 (5.90)	Yes if stable > 6 weeks	Previous exp. in CT of depression plus 12 h training in BA 10 (50)
Studies not included in meta-analysis					
[18] Padfield (1976)	Community female rural low socio-economic status 21–56 100	Behavioural (12) Supportive (12)	Moderately depressed (diagnostic tool not clear)	No	Counsellor (experience not clear) 12 (N.A.)
[19] Zeiss (1979)	Community 33.9 (19–68) N.A.	Behavioural (22) Cognitive (22) Interpersonal (22)	Classed as depressed using Minnesota Multiphasic Personality Inventory & Grinkler Interview Rating	Not clear	Graduate students in clinical psychology & counselling psychologists (masters level). At least 1 year experience 12 (N.A.)
[20] Gardner (1981)	Community 19–65 77%	Behavioural (8) Cognitive (8)	Mild depression (BDI)	Not clear	Not clear N.A.

ADM, Antidepressant medication; BT, behavioural therapy; BDI, Beck Depression Inventory; CBT, cognitive behavioural therapy; CT, cognitive therapy; DT, drug therapy; HAMD, Hamilton Depression Rating Scale; HMO, health maintenance organization; MDD, major depressive disorder; n.a., not available; RDC, research diagnostic criteria.

Table 2. *Meta-analyses of studies examining the effects of behavioural therapy*

Comparison	No. of studies	No. of subjects	SMD	95% CI	*p*	I^2
BT *versus* Control/TAU						
Symptom level	12	459	−0.70	−1.00 to −0.39	<0.001	55.1%
Dropout[a]	3	119	0.58	0.28 to 1.20	0.86	0%
Recovery rate[a]	3	167	4.18	1.14 to 15.28	0.03	52.6%
BT *versus* CT/CBT						
Symptom-level post-treatment	12	476	0.08	−0.14 to 0.30	0.46	21.1%
Symptom-level follow-up	8	271	0.25	−0.21 to 0.70	0.28	60.2%
Dropout[a]	8	436	1.17	0.57 to 2.41	1.17	32.4%
Recovery rate[a]	5	346	0.92	0.59 to 1.44	0.92	0%
BT *versus* Brief psychotherapy						
Symptom-level post-treatment	3	166	−0.56	−1.0 to −0.12	0.01	43.4%
Symptom-level follow-up	2	96	−0.50	−0.90 to −0.09	0.02	0%
Dropout[a]	3	166	0.94	0.22 to 3.96	0.11	54.1%
Recovery rate[a]	3	164	2.37	1.23 to 4.57	0.01	0%
BT *versus* Supportive therapy						
Symptom-level post-treatment	2	45	−0.75	−1.37 to −0.14	0.02	0%

BT, Behavioural therapy; CBT, cognitive behavioural therapy; CT, cognitive therapy; CI, confidence interval; SMD, Standardized mean difference; TAU, treatment as usual.

[a]Indicates odds ratio.

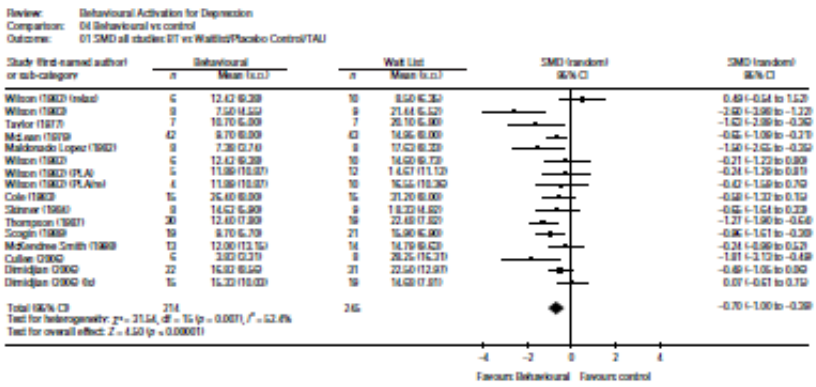


Fig. 1. Behavioural therapy (BT) *versus* wait list/control/placebo symptom-level post-treatment.

Comparison 1: Behavioural interventions versus *waiting list/placebo control*

Scope

Twelve studies with a total of 459 patients contributed data to this analysis [1, 2, 4–7, 9, 11, 12, 14, 16, 17]. Participants were taken from adult community sources consisting of out-patients [2, 4, 6, 7, 11, 12, 16, 17], volunteers [5, 8, 14] and students [1], two studies used older adults [11, 12]. Control interventions consisted of delayed treatment [1, 3, 9, 11, 12, 14, 16, 17], treatment as usual [4, 5, 7] and relaxation [2, 5]. All comparisons were taken immediately after intervention. Interventions ranged from supported bibliotherapy [12, 14], brief therapy with six 40-min sessions [1] to 24 50-min sessions [16]. Facilitators were advanced graduate psychology/therapy students in five studies [1, 5, 6, 7, 9], experienced psychotherapists in four studies [2, 11, 16, 17] and unclear in one study [4]. Depression symptom level was assessed using

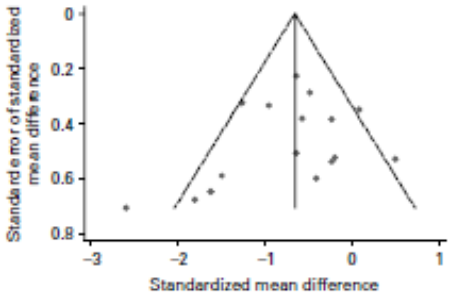


Fig. 2. Begg funnel plot symptom level: behavioural therapy *versus* control interventions/usual care.

either BDI self-report measure [1, 2, 4, 5, 7, 9, 17] or the HAMD assessor rating scale [12], or both [6, 11, 14, 16]. Recovery was defined by clinical interview in one study [11] and by BDI score in two studies [2, 14].

Outcome 1: Depression symptom level post-treatment

The effect of behavioural interventions against control interventions was large with a pooled SMD of −0.70 (95% CI −1.00 to −0.39), demonstrating a highly significant difference in symptom-level scores favouring the behavioural group ($p < 0.001$) (Fig. 1). There was no evidence of publication bias for this outcome (Eggers test −1.04; 95% CI −3.39 to 1.29, $p = 0.35$), a funnel plot showed no evidence of asymmetry (Fig. 2).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 55.1%. Effect size was not significantly related to the level of baseline severity (meta-regression β-coefficient 0.04, 95% CI −0.04 to 0.12; $I^2 = 54\%$, $p = 0.28$) (Fig. A2 online). Quality assessment indicated seven studies fell below our quality threshold [1, 4–7, 9, 14], and the pooled SMD was not affected by study quality (meta-regression $SMD_{low\ quality} = -0.67$; $SMD_{higher\ quality} = -0.75$, $p_{difference} = 0.77$). Behavioural therapists with graduate and postgraduate qualifications produced similar effect sizes (meta-regression $SMD_{graduate} = -0.82$; $SMD_{postgraduate} = -0.59$, $p_{difference} = 0.61$; $I^2 = 59\%$). There was no clear relationship between effect size and number of sessions (meta-regression β-coefficient 0.03; 95%CI −0.03 to 0.09; $I^2 = 0.49$, $p = 0.27$) (Fig. A3 online). Prioritizing clinician-rated assessment in precedence over self-rated where possible made no significant difference to overall effect size (SMD −0.68, 95% CI −0.98 to −0.38).

Outcome 2: Dropout rate

Three studies contributed data to this analysis [2, 14, 16] on a total of 119 subjects with an average dropout rate of 19.17%. We found no difference between rates of dropout between intervention and control (OR 0.58, 95% CI 0.28–1.20, $p = 0.86$).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 0%. There were insufficient studies and negligible heterogeneity to explore the impact of our *a priori* sources of clinical heterogeneity.

Outcome 3: Recovery rate

Three studies contributed data to this analysis [2, 11, 14] on a total of 167 subjects. There were greater rates of recovery in the behavioural intervention group (BT 52%, control 21.05%) with an odds ratio of 4.18 (95% CI 1.14–15.28, $p = 0.03$). There were insufficient studies to test for publication bias for this outcome.

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 52.6%. Low-quality studies [14] were excluded in a sensitivity analysis resulting in an odds ratio of 8.56 (95% CI 0.40–182.63, $p = 0.04$) with an I^2 statistic of 76.4%. There were insufficient studies to explore the underlying causes of this heterogeneity further.

Comparison 2: BT versus CT/CBT

Scope

Twelve studies with a total of 476 patients contributed data to this analysis [1, 3, 4, 6, 8–14, 16]. Participants were taken from adult community sources consisting of out-patients [3, 4, 8, 11–13, 16], volunteers [6, 9, 14] and students [1, 10], with three studies using older adults [3, 11, 12]. Interventions ranged from supported bibliotherapy [12, 14], brief therapy with six 40-min sessions [1] to 24 50-min sessions [16]. Therapy was facilitated by advanced graduate psychology/therapy students in four studies [1, 6, 9, 10], experienced psychotherapists in four studies [3, 11, 13, 16] and was unclear in two studies [4, 7]. Depression symptom level was assessed using either the BDI self-report measure [1, 4, 8–10] or the HAMD assessor rating scale [12], or both [3, 6, 11, 13, 14, 16]. Recovery was defined by diagnostic interview in two studies [3, 11] and by BDI score in three studies [10, 13, 16].

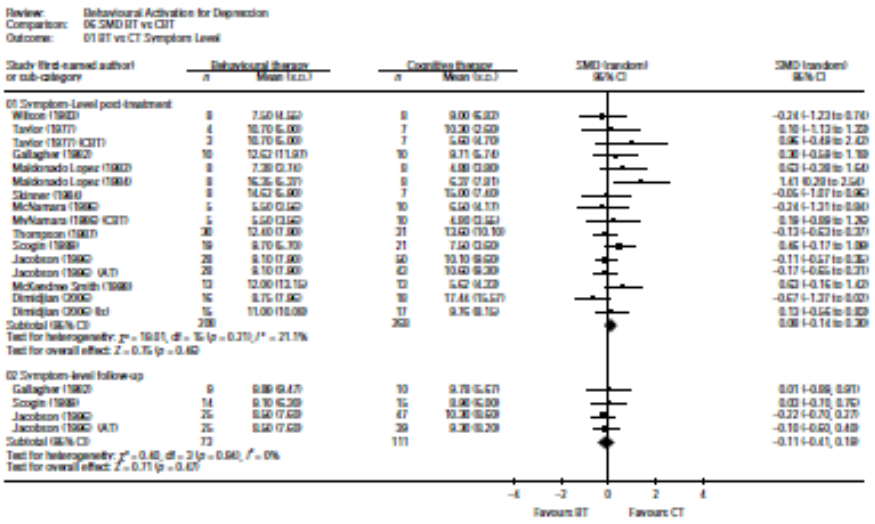


Fig. 3. Behavioural therapy (BT) *versus* cognitive behavioural therapy (CBT) symptom-level post-treatment and follow-up.

Outcome 1: Depression symptom level post-treatment

No difference in effect between behavioural interventions and CBT/CT was identified with a pooled SMD of 0.08 (95% CI −0.14 to 0.30, $p = 0.46$) (see Fig. 3). There was no evidence of publication bias for this outcome using Egger's test [intercept (0 if unbiased) = 1.07; 95% CI −0.23 to 2.38, $p = 0.10$], and a funnel plot showed no evidence of asymmetry.

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 21.1%. Seven studies fell below our quality threshold [1, 4, 6, 7, 9, 10, 14] and the pooled SMD was not significantly affected by study quality (meta-regression $SMD_{low\ quality} = +0.23$; $SMD_{higher\ quality} = -0.13$, $p_{difference} = 0.12$; $I^2 = 0\%$). Comparative effectiveness of BT *versus* CT/CBT varied according to baseline severity of depression, BT demonstrating a greater level of effectiveness at more severe levels of depression (meta-regression β-coefficient −0.05, 95%CI −0.10 to −0.01; $I^2 = 0\%$, $p = 0.04$) (Fig. A4 online).

Graduate-level behavioural therapists produced slightly worse results compared to those with post-graduate qualifications in comparison to CBT, although this did not reach significance (meta-regression $SMD_{graduate} = 0.28$; $SMD_{post\ graduate} = -0.135$, $p_{difference} = 0.11$; $I^2 = 0\%$). There was no clear relationship between effect size and number of sessions (meta-regression β-coefficient −0.025, 95% CI −0.056 to 0.006; $I^2 = 0.08$, $p = 0.11$). Prioritizing clinician-rated assessment in precedence over self-rated where possible made no significant difference in overall effect size (SMD 0.09, 95% CI −0.12 to 0.29).

Outcome 2: Depression symptom level at follow-up

Eight studies contributed data to this analysis [1, 3, 4, 6, 8, 10, 12, 13] on a total of 271 subjects with an average follow-up period of 4 months. Overall there was no difference in effect of BT compared to CBT/CT with a pooled SMD of 0.25 (95% CI −0.21 to 0.70, $p = 0.28$) (Fig. 3).

Heterogeneity and sensitivity analysis

Variation in effect size (I^2) attributable to heterogeneity was 60.2%. After exclusion of low- quality studies [1, 4, 6, 8, 10] and those with follow-up of <3 months [1, 6] in a sensitivity analysis the pooled SMD was −0.11 (95% CI −0.41 to 0.19, $p = 0.47$). There were insufficient studies to explore the underlying causes of this heterogeneity further.

Outcome 3: Dropout rate

Eight studies contributed data to this analysis [1, 3, 6, 11–14, 16] on a total of 436 subjects with an average dropout rate of 15.36%. We found no difference in rates of dropout with an odds ratio of 1.17 (95% CI 0.57–2.41, $p = 0.67$).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 32.4%. Low-quality studies [1, 6, 14] were excluded in a sensitivity analysis resulting in an odds ratio of 1.47 (95% CI 0.60–3.61, $p=0.40$) with an I^2 statistic of 42.9%. There were insufficient studies to explore the underlying causes of this heterogeneity further.

Outcome 4: Recovery rate

Five studies contributed data to this analysis [3, 10, 11, 13, 16] on a total of 346 subjects. We found a pooled recovery rate of 55% with no difference between the two treatment approaches (OR 0.92, 95% CI 0.59–1.44, $p=0.72$).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 0%. Low-quality studies [8] were excluded in a sensitivity analysis resulting in an odds ratio of 0.93 (95% CI 0.59–1.47, $p=0.77$) with an I^2 statistic of 0%.

Comparison 3: Behavioural interventions versus brief psychotherapy

Scope

Three studies with a total of 166 patients contributed data to this analysis [2, 3, 11]. Participants were from adult out-patient community sources, two studies using older adults [3, 11]. Brief psychotherapy interventions were based upon a psychodynamic model in all studies. Interventions ranged from 10 to 20 sessions, all studies used experienced therapists. Studies assessed depression symptom level using the BDI alone [2] or both BDI and HAMD [3, 11]. Two studies assessed depression at intake using structured clinical interviews [3, 11], the third using cut-off points from validated self-report measures [2]. Recovery was defined by clinical interview in two studies [3, 11] and by BDI score in one study [2].

Outcome 1: Depression symptom post-treatment

The positive effect of BT against brief psychotherapy was large with a pooled SMD of −0.56 (95% CI −1.0 to −0.12, $p=0.01$). There were insufficient studies to test for publication bias.

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 43.4%. All studies were above the quality threshold, hence we performed no sensitivity analyses. There were insufficient studies to explore the underlying causes of this heterogeneity further. Prioritizing clinician- rated assessment in precedence to self-rated assessment where possible made no difference in overall effect size (SMD −0.52, 95% CI −1.01 to −0.03).

Outcome 2: Depression symptom level follow-up

Two studies contributed data to this analysis [2, 3] on a total of 96 subjects with an average follow-up period of 4.5 months. The positive effect of behavioural interventions against brief psychotherapy was medium with a SMD of −0.50 (95% CI −0.90 to −0.09, $p=0.02$).

Heterogeneity and sensitivity analysis

Variation in effect size attributable to heterogeneity (I^2) was 0%. Both studies collected follow-up beyond the 3-month point and were above the quality threshold so we performed no sensitivity analyses.

Outcome 3: Dropout

Three studies contributed data to this analysis [2, 3, 11] on a total of 166 subjects with an average dropout rate of 14.45% across studies. No difference in dropout was observed with an odd ratio of 0.94 (95% CI 0.22–3.96, $p=0.11$). There were insufficient studies to test for publication bias.

Heterogeneity and sensitivity analysis

Variation in odds ratio attributable to heterogeneity (I^2) was 54.1%. All studies were above the quality threshold so no sensitivity analysis was performed.

Outcome 4: Recovery rate

Three trials contributed data to this analysis [2, 3, 11] on a total of 164 subjects (note two subjects deceased). Greater rates of recovery were observed in BT (56.79%) compared to brief psychotherapy (36.14%) with an odds ratio of 2.37 (95% CI 1.23–4.57, $p=0.01$). There were insufficient trials to test for publication bias.

Heterogeneity and sensitivity analysis

Variation in odds ratio attributable to heterogeneity (I^2) was 0%. All studies were above the quality threshold so no sensitivity analysis was performed.

Comparison 4: Behavioural interventions versus *supportive therapy*

Scope

Two studies with 45 subjects contributed data to this analysis [10, 15]. Participants were university students [10] and in-patients [15].

Interventions ranged from six 20-min sessions [10] to eight, 50-min sessions [15], delivered by doctoral clinical psychology students [10] or a clinical psychologist [15]. Both studies measured depression symptom levels by self-report measures (BDI), with one [10] using HAMD also. Depression at baseline was assessed by self-report measures [10] or clinical interview [15].

Outcome 1: Depression symptom level post-treatment

The positive effect of BT against supportive therapy was large (SMD −0.75, 95% CI −1.37 to −0.14, $p=0.02$). There were insufficient studies to test for publication bias.

Heterogeneity and sensitivity analysis

The variation in effect size attributable to heterogeneity (I^2) was 0%. Both studies fell below the quality threshold therefore no sensitivity analysis was performed. Insufficient data were available in this comparison for further analysis.

Discussion

We found clear evidence that BT is an effective treatment for depression. It provides superior outcomes to control, supportive counselling and brief psychotherapy. BT and CBT provided equivalent results with no statistically significant differences in post-treatment and follow-up symptom levels, in recovery rate or dropouts.

The BT trials were variable in design and delivery. To some degree, we have been able to utilize this variability to explore factors relating to magnitude of effectiveness. Such meta-regression analysis makes observational associations and is exploratory in nature and as such loses the power of causal inference (Higgins & Thompson, 2004). We considered such an approach viable and efficient in this review as the alternative of planning large-scale prospective trials with many arms is costly and time consuming. Sufficient data for this analysis were available only where BT was compared to controls or CT/CBT post-treatment on symptom level. Our meta-regression found that compared to controls, baseline severity, length of treatment and level of qualification were not related to BT effect although there is a positive relationship between greater baseline severity and BT efficacy compared to CT/CBT. Such findings provide direction in the development of BT for future research. They indicated that further exploration is needed into length of treatment and skill level required for optimum BT delivery. Our review identified a number of trials directly comparing BT with drug therapy; this was not included as an *a priori* comparator. Such a comparison would be a useful addition in any future review as BT would appear equivalent, if not superior, to pharmacology in the included studies.

Our meta-analysis complements and concurs with other publications that include behavioural interventions as part of wider CBT reviews (Dobson, 1989; Gloaguen *et al.* 1998; Churchill *et al.* 2001), or focus on activation alone (Cuijpers *et al.* 2007). In contrast to these previous reviews we chose to focus on individual rather than group interventions, and included dropout and recovery rate analyses. Our review includes more studies than previous reviews due to our broader inclusion criteria and the inclusion of recent and unpublished data. The studies drew patients from a range of settings such as in-patient, psychiatric out-patient and volunteer cohorts in adult, older adult and student settings. Interventions varied considerably across studies from supported self-help using minimal therapist contact to full psychotherapy. The quality of included trials varied considerably, with some of low quality delivering results that deviated considerably from the overall picture [4, 8]. We attempted to account for this by the use of sensitivity analysis, random-effects modelling and meta-regression of *a priori* variables. Interpretation of our results must be made with such factors in mind. Caution must also be exercised in interpreting the comparisons of behavioural interventions with brief psychotherapy and supportive therapy due to the low numbers of studies and small sample sizes.

Of particular interest is the observed equivalence between behavioural interventions and the CBT/CT strongly recommended in guidelines (e.g. NICE, 2004). In addition to similar levels of mean symptom improvement, we observed no difference in recovery or dropout. These combined findings indicate that behavioural interventions are as effective and acceptable as CBT/CT. Such findings partially endorse the BT parsimony hypothesis advanced by Jacobson and colleagues (Jacobson *et al.* 1996, 2001). They question the utility of adding 'complex' cognitive techniques to simpler behavioural interventions to improve clinical outcome. One of the attractions of behavioural interventions is that they may lend themselves to shorter training of less-qualified individuals, thus assisting

the current scarcity of therapist availability and overwhelming demand (Centre for Economic Performance's Mental Health Policy Group, 2006). We found no direct evidence in this review to support such an assumption, as we found no studies using non-psychology- or non-psychotherapy-trained individuals delivering BT. However, when we examined the impact on level of training of those who had delivered BT in meta-regression, we did not find that superior outcomes were associated with 'higher level' qualifications. Such findings may support the assertion that BT may be suitable for shorter training and hence improve access by increasing available therapists within limited resources. We recommend further research of this question based upon our findings.

In summary, BT for depression is an effective intervention that has equal, if not better, outcomes than alternative and currently recommended therapies. Our review adds to the literature in the area as it provides a broad overview of the current evidence, reports data on recovery, dropout and explores the effect of baseline covariants in relation to depression symptom change. We recommend further research into the efficacy of behavioural treatments of depression, in particular Jacobson *et al.*'s (1996) parsimony hypothesis where the intervention is delivered by 'technicians' rather than therapists.

Declaration of Interest

None.

Note

Supplementary information accompanies this paper on the Journal's website (http://journals.cambridge.org).

References

Beck A (1976). *Cognitive Therapy and the Emotional Disorders*. International Universities Press: New York.

Beck A, Ward C, Mendelson M, Mock J, Erdbaugh J (1961). An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571.

Begg C (1994). Publication bias. In: *The Handbook of Research Synthesis* (ed. H. Cooper and L. V. Hedges), pp. 399–409. Russell Sage Foundation: New York, NY.

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.

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 of depression. *Health Technology Assessment* **5** (35).

Cochrane Collaboration (2003). Review Manager (RevMan) [computer program]. Version 4.2 for Windows. The Nordic Cochrane Centre: Copenhagen.

Cole M (1983). Behavioural treatment of chronic depression. *Dissertation Abstracts International*, vol. 45, no. 3. Univ Microfilms International, US.

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

Cullen J, Spates R, Pagoto S, Doran N (2006). Behavioral activation treatment for major depressive disorder: a pilot investigation. *Behavior Analyst Today* **7**, 151–166.

Dimidjian S, Hollon S, Dobson K, Schmaling K, Kohlenberg R, Addis M, Gallop R, McGlinchey J, Markley D, Gollan J, Atkins D, Dunner D, Jacobson N (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 (1989). A meta analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology* **57**, 414–419.

Egger M, Davey-Smith G, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.

Ferster C (1973). A functional analysis of depression. *American Psychologist* **28**, 857–870.

Furukawa T, Barbui C, Cipriani A, Brambilla P, Watanabe N (2006). Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* **59**, 7–10.

Gallagher D, Thompson L (1982). Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy* **19**, 482–490.

Gardner P, Oei T (1981). Depression and self-esteem: an investigation that used behavioural and cognitive approaches to the treatment of clinically depressed clients. *Journal of Clinical Psychology* **37**, 129–135.

Gloaguen V, Cottraux J, Cucherat M, Blackburn I (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* **49**, 59–72.

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

Higgins J, Thompson S (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* **23**, 1663–1682.

Higgins J, Thompson S, Deeks J, Altman D (2003). Measuring inconsistency in meta-analyses. *British Medical Journal* **327**, 557–560.

Hirschfeld R, Keller M, Panico S, Arons B, Barlow D, Davidoff F, Endicott J, Froom J, Goldstein M, Gorman J, Guthrie D, Marek R, Maurer T (1997). The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of

depression. *Journal of the American Medical Association* **277**, 333–340.

Hollon S, Thase M, Markowitz J (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest* **3**, 1–39.

Hopko D, Lejuez C, Lepage J, Hopko S, McNeil D (2003*a*). A brief behavioral activation treatment for depression; a randomized pilot trial within an inpatient psychiatric hospital. *Behavior Modification* **27**, 458–469.

Hopko D, Lejuez C, Ruggiaro K, Eifert G (2003*b*). Contemporary behavioural activation treatments for depression: procedures, principles and progress. *Clinical Psychology Review* **23**, 699–717.

Jacobson N, Dobson K, Traux P, Addis M, Koerner K, Gollan J, Gortner E, Prince S (1996). A component analysis of cognitive-behavioural treatment of depression. *Journal of Consulting and Clinical Psychology* **64**, 295–304.

Jacobson N, Gortner E (2000). Can depression be de-medicalised in the 21st century: scientific revolutions, counter revolutions and the magnetic field of normal science. *Behaviour Research and Therapy* **38**, 103–117.

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

Jadad A, Moore R, Carroll D, Jenkinson C, Reynolds D, Cavaghan D, McQuay H (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials* **17**, 1–12.

Khan K, Kleijnen J (2002). Phase 4: Selection of Studies. Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. In Report **4**, 2nd edn (ed. K. Khan, G. ter Riet, J. Galanville, A. Sowden and J. Kleijnen). Centre for Reviews and Dissemination, University of York, UK.

Khan K, ter Riet G, Papay J, Nixon J, Kleijnen J (2002). Phase 5: Study Quality Assesment. Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. In Report **4**, 2nd edn (ed. K. Khan, G. ter Riet, J. Galanville, A. Sowden and J. Kleijnen). Centre for Reviews and Dissemination, University of York, UK.

Klerman G, Weissman M, Rounsaville B, Chevron E (1984). *Interpersonal Psychotherapy of Depression*. Basic Books: New York, NY.

Lewinsohn P, Graf M (1973). Pleasant activities and depression. *Journal of Consulting and Clinical Psychology* **41**, 261–295.

Lipsey M, Wilson D (1993). The efficacy of psychological, educational and behavioural treatment. *American Psychologist* **48**, 1181–1209.

Luborsky L, Mark D, Hole A (1995). Manual for supportive expressive dynamic therapy for depression: adaptation of the general SE manual. In *Psychodynamic Psychotherapy for Axis I Disorders* (ed. J. Barber and P. Crits-Christoph). Basic Books: New York.

Maldonado Lopez A (1982). Behavioural therapy and depression. *Revista dePsicologia General y Aplicada* **37**, 31–56.

Maldonado Lopez A (1984). Behavioural therapy and depression: an experimental analysis of the interaction between cognitive and behavioural therapies and pharmacological therapy in depressed people. *Revista de Psicologia General y Aplicada* **39**, 517–535.

Martell C, Addis M, Jacobson N (2001). *Depression in Context. Strategies for Guided Action*. Norton: New York..

McKendree-Smith N (1998). Cognitive and behavioural bibliotherapy for depression: an examination of efficacy and mediators and moderators of change (unpublished Ph.D. thesis). University of Alabama.

McLean P, Hakstain R (1979). Clinical depression: comparative efficacy of outpatient treatment. *Journal of Consulting and Clinical Psychology* **47**, 818–836.

McNamara K, Horan J (1986). Experimental construct validity in the evaluation of cognitive and behavioral treatments for depression. *Journal of Counselling Psychology* **33**, 23–30.

NICE (2004). *Management of Depression in Primary and Secondary Care*. National Institute for Clinical Excellence: London.

Padfield M (1976). The comparative effects of two counseling approaches on the intensity of depression among rural women of low socioeconomic status. *Journal of Counselling Psychology* **23**, 209–214.

Rogers C (1961). *On Becoming a Person*. Houghton Mifflin: Boston, MA.

Schulz K, Grimes D (2002). Allocation concealment in randomised trials: defending against deciphering. *Lancet* **359**, 614–618.

Scogin F, Jamison C, Gochneaur K (1989). Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting and Clinical Psychology* **57**, 403–407.

Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer Y (2001). *Office of National Statistics: Psychiatric Morbidity Among Adults Living in Private Households, 2000*. HMSO: London.

Skinner D (1984). Self-control of depression: A comparison of behaviour therapy and cognitive behaviour therapy. *Dissertation Abstracts International*, vol. 45, no. 3. Univ Microfilms International, US.

Stata Corporation (2003). Version 8 [program]. Stata Corporation: Texas.

Sutton A, Abrams K, Jones D (1998). Systematic reviews of trials and other studies. *Health Technology Assessment* **2** (19).

Taylor F, Marshall W (1977). Experimental analysis of a cognitive-behavioral therapy for depression. *Cognitive Therapy and Research* **1**, 59–72.

Thompson L, Gallagher D, Breckenridge J (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology* **55**, 385–390.

Thompson S, Higgins J (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* **21**, 1559–1573.

WHO (2001). *The World Health Report 2001: Mental Health: New Understanding, New Hope*. World Health Organisation: Geneva.

Wilson P (1982). Combined pharmacological and behavioural treatment of depression. *Behavior Research and Therapy* **20**, 173–184.

Wilson P, Goldin J, Charbonneau P (1983). Comparative efficacy of behavioral and cognitive treatments of depression. *Cognitive Therapy and Research* **7**, 111–124.

Zeiss A, Lewinsohn P, Munoz R (1979). Nonspecific improvement effects in depression using interpersonal skills training, pleasant activity schedules or cognitive training. *Journal of Consulting and Clinical Psychology* **47**, 427–439.

Evidence based mental health review of Meta Analysis

Review: Individual behavioural therapy reduces symptoms of depression

QUESTION

Question: How effective are behavioural therapies for depression?

Outcomes: Depressive symptoms (either self-rated—eg, Beck Depression Inventory, or clinician rated—eg, Hamilton Depression Rating Scale).

METHODS

Design: Systematic review with meta-analysis.

Data sources: MEDLINE, EMBASE, PsycINFO, Cochrane Library, DARE, CINAHL, AMED and the British Nursing Index were searched from inception to January 2006 for randomised controlled trials (RCTs). Reference lists of identified studies were hand searched.

Study selection and analysis: Two reviewers appraised studies and selected RCTs of behavioural therapy in adults (aged ≥16 years) with a primary diagnosis of depression. Treatments could be delivered in a community or inpatient setting. Behavioural therapy had to be individual, time limited, and based on changing the behavioural consequences of environmental or cognitive triggers. Comparator treatments could include cognitive behavioural therapy (CBT) or cognitive therapy, brief psychotherapies, supportive counselling, treatment as usual, waiting list or other inactive control. Two reviewers rated study quality and extracted data. Authors were contacted to obtain missing data, but if this was not successful, standard deviations were imputed from other related studies. Random effect meta-analyses were conducted using RevMan software. Separate meta-analyses were carried out for different comparator groups. Heterogeneity was investigated using the I^2 statistic.

MAIN RESULTS

Twenty RCTs were identified which met inclusion criteria, but 3 of the RCTs did not have enough data to be included in the meta-analyses. Behavioural therapy reduced depressive symptoms compared to control at the end of treatment (12 RCTs, 459 participants; controls included waiting lists, treatment as usual and relaxation). There was no significant difference between behavioural therapy and cognitive therapy or CBT in depressive symptoms post-treatment or at an average of 4 months' follow-up (post-treatment: 12 RCTs, 476 participants; follow up: 8 RCTs, 271 participants). Behavioural therapy reduced depressive symptoms post-treatment and at an average of 4.5 months' follow-up compared to brief psychotherapy (post-treatment: 3 RCTs, 166 participants; follow up: 2 RCTs, 96 participants). Compared to supportive counselling, behavioural therapy reduced symptoms of depression post-treatment (2 RCTs, 45 participants). See online table.

CONCLUSIONS

Individual behavioural therapy reduces symptoms of depression more than control, supportive counselling or brief psychotherapy. It provides similar reductions in depressive symptoms to cognitive therapy and CBT.

ABSTRACTED FROM

Ekers D, Richards D, Gilbody S. A meta-analysis of randomized trials of behavioural treatment of depression. *Psychol Med* 2008;**38**:611–23.

Correspondence to: D Ekers, Tees Esk & Wear Valleys NHS Trust/University of York Department of Health Sciences, The Health Centre, Newcastle Road, Chester le Street, County Durham, UK; David.Ekers@cddps.nhs.uk

Source of funding: Unclear.

► Additional notes and a table are published online only at http://ebmh.bmj.com/content/vol11/issue4

COMMENTARY

Cognitive behaviour therapy (CBT) is the most well-researched and established psychotherapeutic approach for treating depression.[1] However, as with antidepressant medications,[2] many unanswered questions remain about the mechanisms through which CBT for depression produces its therapeutic effects. Is it the distinctive cognitive techniques that are producing the majority of the benefits, or are the other non-cognitive elements also a key part of the treatment?

The meta-analysis by Ekers and colleagues provides a tantalising piece of the puzzle. They conducted a comprehensive quantitative review of 17 clinical trials of behaviour therapy (BT) for depression, which did not include specific cognitive therapy techniques. The authors reasoned that if BT produces the same effects as the more comprehensive CBT package, then this would suggest that the cognitive elements are unnecessary for achieving therapeutic gains. As CBT is a more complicated treatment to train therapists to implement, BT would be an attractive alternative given its apparent parsimony. The results of the meta-analysis appeared to confirm this premise, with BT being equally as effective as CBT, and more effective than control interventions, including treatment as usual, brief psychodynamic therapy and supportive therapy. The dropout rates between BT and the other treatments were similar, and BT appeared to work well regardless of initial severity, length of treatment, or level of therapist training. The authors did not specifically test the comparability of BT and antidepressant medications.

Another meta-analysis reached similar conclusions about the equivalence of CBT and BT and a recent large-scale clinical trial of BT showed that it was as effective as antidepressant medication for depression.[3] While BT and CBT were equally as effective for lower-levels of depression severity, patients with severe depression specifically benefited more from BT than CBT.[3] Behavioural activation, which involves graded activity and goal scheduling, is the central component of BT for depression and may be related to a more general mechanism of change that is operating in many different treatments for depression, especially CBT. The UK's National Health Service recently allocated £173 million to train more CBT therapists for treating depression and anxiety.[4] Ekers and colleagues' meta-analysis suggests that the money may be best spent training better behaviour therapists.

Brandon A Gaudiano, PhD
Alpert Medical School of Brown University and Butler Hospital, Providence, Rhode Island, USA

Competing interests: None.

1. **Butler AC**, Chapman JE, Forman EM, *et al*. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clin Psychol Rev* 2006;**26**:17–31.
2. **Kirsch I**, Deacon BJ, Huedo-Medina TB, *et al*. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008;**5**:e45.
3. **Dimidjian S**, Hollon SD, Dobson KS, *et al*. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *J Consult Clin Psychol* 2006;**74**:658–70.
4. **Talk therapy plans under fire.** BBC News [Internet]. 6 July 2008. Available at http://news.bbc.co.uk/go/pr/fr/-/2/hi/health/7486132.stm (accessed 6 July 2008).

Copy of Ekers et al 2011 Publication BA delivered by the non specialist: phase II RCT

BJPsych | The British Journal of Psychiatry (2011) 198, 66-72. doi: 10.1192/bjp.bp.110.079111

Behavioural activation delivered by the non-specialist: phase II randomised controlled trial

David Ekers, David Richards, Dean McMillan, J. Martin Bland and Simon Gilbody

Background
Behavioural activation appears as effective as cognitive–behaviour therapy (CBT) in the treatment of depression. If equally effective, then behavioural activation may be the preferred treatment option because it may be suitable for delivery by therapists with less training. This is the first randomised controlled trial to look at this possibility.

Aims
To examine whether generic mental health workers can deliver effective behavioural activation as a step-three high-intensity intervention.

Method
A randomised controlled trial (ISRCTN27045243) comparing behavioural activation ($n = 24$) with treatment as usual ($n = 23$) in primary care.

Results
Intention-to-treat analyses indicated a difference in favour of behavioural activation of −15.79 (95% CI −24.55 to −7.02) on the Beck Depression Inventory-II and Work and Social Adjustment Scale (mean difference −11.12, 95% CI −17.53 to −4.70).

Conclusions
Effective behavioural activation appears suitable for delivery by generic mental health professionals without previous experience as therapists. Large-scale trial comparisons with an active comparator (CBT) are needed.

Declaration of interest
None.

Depression will be the second largest cause of disease burden by 2020.[1] It affects between 5 and 10% of the population and is the third most common reason for primary care consultation.[2] Psychological treatments, particularly cognitive-behavioural therapy (CBT), are recommended to treat depression;[3,4] however, less than 10% of those affected receive such treatment.[2] The standard CBT approach to depression is Beck's cognitive therapy,[5] which uses both behavioural and cognitive strategies to identify, question and modify maladaptive thought processes, life rules and core beliefs. In a landmark study in 1996, Jacobson *et al* compared the full version of CBT with a reduced version that included some but not all cognitive components, and a third intervention, termed behavioural activation, that relied entirely on behavioural strategies.[6] There was no evidence of any differences in effectiveness between treatments at post-treatment or follow-up.[7] These findings led Jacobson *et al*[6] and Jacobson & Gortner[8] to put forward a parsimony argument in favour of behavioural activation: if CBT and behavioural interventions are equally effective, then behavioural ones may be preferable because they are simpler to deliver and can therefore be delivered more economically by professionals with less training. Were this to be the case, this would have substantial implications for the organisation and delivery of treatments.[8] The provocative finding that much of what occurs in the leading psychological treatment for depression may be an unnecessary complication has led to a renewed interest in behavioural treatments. In a further study in 2006, Dimidjian and colleagues[9] replicated the finding that CBT and behavioural activation were comparably effective, and a series of recent meta-analyses[10–12] have also come to this conclusion. Although a main impetus for the renewed interest is the possibility of developing a more parsimonious treatment for depression, it is notable that no study reviewed in the meta-analyses has explored the parsimony argument in a controlled clinical trial. In earlier studies the treatments were delivered by clinicians who had previous experience of delivering therapy, with experience often amounting to several years. Although behavioural activation may indeed by simpler to deliver, it may be the experience of the therapist that counts. The aim of the current study was to examine whether generic mental health workers, without previous experience in therapy, could deliver effective behavioural interventions. We looked at the impact on depression symptom level, functioning and treatment satisfaction. This is, as far as we are aware, the first study to directly assess the parsimony argument offered in favour of behavioural activation.

Method

The Northumberland local research ethics committee and local NHS research governance departments approved this study. The trial is registered with the International Standard Randomised Controlled Trials Registry, number ISRCTN27045243 and complies with updated CONSORT recommendations.[13] It was a 'phase II' randomised controlled trial of behavioural activation facilitated by generic mental health workers compared with usual care for adults with depression.

Recruitment, participants and randomisation

We recruited potential participants aged 18 or over from either general practices directly or from primary care mental health services over a 9-month period. Practices were based in a mix of rural and urban settings. These practices opted to participate in the trial after receiving information about it. Participants identified with depression and either on a stable dose or no dose of antidepressant medication for 6 weeks preceding identification were supplied with study information and if requested were referred into the trial by their general practitioner (GP) or mental health worker. Following consent, eligibility was confirmed by the use of a standardised computer-based assessment tool, the Clinical Interview Schedule Revised (CSIR),[14] to confirm a ICD–10[15] diagnosis of depression. Exclusion criteria consisted of suicidal risk, psychotic symptoms, diagnosis of bipolar disorder, organic

brain disease or the use of alcohol/non-prescription drugs requiring clinical intervention. Baseline measures were then taken.

Following assessment, participants were randomised to two arms through an allocation concealment process independent of the study team using a block randomisation system in blocks of four. Taking into account increased risk of a moderating effect of baseline severity of depression[16] and because of the small sample sizes in this study, stratification based on baseline depression severity was conducted. Participants were allocated into two groups prior to randomisation based on Beck Depression Inventory (BDI–II)[17] scores ($\leqslant$25 and $\geqslant$26). General practitioners and participants were informed of allocation automatically by letter.

Sample size

We calculated the sample size required based on those studies incorporating a delayed start to psychological interventions with variable levels of concurrent usual care rather than active placebo identified in our previous systematic review.[11] A standardised between-group effect size (Cohen's *d*) of −0.84 (CI −1.27 to −0.41) was observed in a sample of nine studies (282 participants) indicating that with alpha set at 0.05 to obtain 80% power, a sample size of 23 was required in each group.

Measures

The primary clinical outcome used for depression symptoms was the Beck Depression Inventory (BDI–II),[17] with a score range of 0 to 63 (0–13 minimal, 14–19 mild, 20–28 moderate, 29–63 severe). Secondary outcome measures looked at functioning using the Work and Social Adjustment Scale (WSAS)[18] and satisfaction using the Client Satisfaction Questionnaire (CSQ–8).[19] Assessments were collected by a research worker masked to treatment allocation at baseline, 1-, 2- and 3-month follow-up. To reduce the risk of bias further we used self-completed assessments of depression symptom level, functioning and satisfaction.

Interventions

Behavioural activation

Participants received 12 1-hour face-to-face sessions of behavioural activation over a 3-month period as a step-three intervention.[3] Sessions followed a 12-session protocol based on two behavioural approaches developed in previous research.[20,21] Behavioural activation consisted of a structured programme increasing contact with potentially antidepressant environmental reinforcers through scheduling and reducing the frequency of negatively reinforced avoidant behaviours. A shared formulation was created based on a behavioural model in the early stages of treatment that was developed with the participant throughout the 12 sessions. Subsequent specific techniques incorporated in the 12-session protocol were self-monitoring, identifying 'depressed behaviours', developing alternative goal-orientated behaviours and scheduling. In addition, the role of avoidance and rumination was addressed through functional analysis and alternative responses were developed. The overall goal of behavioural activation was to re-engage participants with stable and diverse sources of positive reinforcement from their environment and to understand the behavioural activation rationale, thus developing depression management strategies for future use (the treatment manual is available from the author on request).

Behavioural activation therapists

Behavioural activation in this study was delivered by two qualified mental health nurses with no previous formal psychotherapeutic training or experience. Both had worked in a range of services in in-patient and community settings with 3 and 6 years experience since qualification. They received 5 days of training in behavioural activation and 1 hour of clinical supervision fortnightly from the principal investigator (D.E.). Training focused on the rationale and skills required to deliver a 12-session protocol of behavioural activation for depression. It included sections on behavioural learning theory and its application to depression, developing individualised behavioural activation formulations and specific techniques used in sessions. Training was a mix of presentation and role-play, with repeated practise and feedback. Competency assessment at the end of training was based on role-played treatment scenarios.

Usual care

Participants were followed up by their GP or primary care mental health worker and offered interventions deemed appropriate for their condition as per normal practice. At 3-month follow-up, control participants were offered behavioural activation therapy as delivered in the intervention arm.

Adherence assessment

All treatment sessions were audiotaped in the intervention arm. Recordings were stratified for study phase (early, mid, late) and therapists; 20% were then randomly selected by a research assistant masked to session content. Recordings of 38 h of therapy were then assessed by independent accredited cognitive–behavioural therapists with extensive experience in both CBT and behavioural activation. As no validated competency assessment tool is available for behavioural activation, we designed a brief checklist of treatment fidelity in this trial. Assessors specifically examined session and homework content against treatment protocols. They indicated if behavioural activation was the overall modality applied and if other therapeutic models were prominent in the therapy (such as cognitive therapy). We assigned scores of 1 where under each heading behavioural activation was dominant and scores of 0 if other therapy modes were prominent. After reviewing each tape, assessors decided if the session could be classed as behavioural activation with assigned values of 1 (yes) and (0) no.

Data analysis

Descriptive statistics are presented as means and standard deviations for psychometric scales (BDI–II, WSAS, CSQ–8) and counts and percentages for categorical variables (depressed/not depressed). The primary outcome variable, severity of depression at 3-month follow-up, was compared between groups using analysis of covariance on individual baseline depression (BDI–II) scores. Social functioning was compared at 3-month follow-up using analysis of covariance on individual baseline social adjustment scale scores. Satisfaction was measured at 3 months and compared between groups using an independent samples *t*-test. For continuous variables we present between-group mean end-point differences, both in terms of scores on the instrument and as standardised effect sizes (Cohen's *d*) and assigned values to effect size as per normal convention (small 0–0.32, medium 0.33–0.55, and large 0.56 and above).[22] The data analysis approach was decided *a priori* using analysis of covariance to counter potential baseline variance that may influence results because of the small sample sizes in this study.

Missing data

Missing data presents a common threat to the results of many trials, especially psychotherapy trials where sample sizes are small. Traditional approaches to dealing with missing data such as including completers only or last-observation-carried-forward (LOCF) can significantly bias results.[23,24] To deal with such problems, we used, where possible, multiple imputation for our intention-to-treat analysis.[25] We conducted an intention-to-treat analysis replacing missing data using multiple imputation by chained equations, as described by Royston using 100 imputations.[26] We incorporated baseline instrument scores (BDI–II, WSAS), age, gender, problem duration and allocation in modelling.

For the clinical significance analyses it was not possible to use the multiple imputations method as this approach does not supply individual participant-level data. We therefore report LOCF analyses for clinical significance as it is likely to be a conservative analysis.

Clinical significance

We used Jacobson & Truax[27] procedures for calculating reliable and clinically significance change to quantify clinical improvement in depressive symptoms on the BDI–II; this is recommended as a standard reporting strategy for all published research involving psychological interventions.[28] This requires a pre- to post-treatment improvement in scores that is unlikely to be because of the inherent unreliability of the measure (reliable change) accompanied by a movement from a clinical range to a non-clinical one (clinically significant change). In calculating reliable and clinically significant change criteria, we used data from the BDI–II manual[17] for clinical means, standard deviations and the reliability estimate (Cronbach's alpha), and data from Dozois *et al*[29] for the non-clinical mean and standard deviation. On the basis of these data, a participant had to improve by ten points or more from pre- to post-treatment to show reliable change and in addition had to score 17 or above pre-treatment and 16 or below post-treatment to meet criteria for clinically significant change (see Jacobson & Truax[27] for details of calculations).

As an additional measure of clinical improvement we used the response and remission criteria given in Dimidjian *et al*.[9] Response was defined as an improvement of at least 50% or more and remission as a score of ⩽10 on the BDI–II. Odds ratios with 95% confidence intervals were used to compare clinically significant change in the two groups.

Results

Baseline characteristics

In total, 68 participants were referred to the trial of whom 21 were excluded (17 did not meet diagnostic criteria, 2 refused randomisation, 2 had significant suicidal ideation (as measured by a score of ⩾2 on question 9 of the BDI–II)). Forty-seven participants met the inclusion criteria and proceeded to randomisation. Of these, 23 were allocated to behavioural activation and 24 to a control group. No differences were observed in scores at baseline between the two groups on BDI–II (mean for behavioural activation 35.57 (s.d. = 9.60) and for usual care 35.08 (s.d. = 9.60)), WSAS (mean for behavioural activation 26.39 (s.d. = 7.30) and for usual care 25.13 (s.d. = 7.30)), CSIR (mean for behavioural activation 31 (s.d. = 10.99) and for usual care 33.13 (s.d. = 8.22)) or problem duration (mean for behavioural activation 186.91 weeks (s.d. = 358.49) and for usual care 195.21 weeks (s.d. = 404.64)). Baseline participant characteristics are presented in Table 1 and show that the participants in the trial are representive of individuals with long-term severe depression, with substantial impairment of functioning. Data were collected from 38 participants at 3-month assessment, 16 in the behavioural activation arm and 22 in the control arm. Of those opting out of the study, three did so post-randomisation (one behavioural activation group, two usual care group), three at 1 month (all in the behavioural activation group), and three at 2 months (all in the behavioural activation group). There were no significant differences between completers and those dropping out of treatment on baseline BDI–II depression scores (mean for those who dropped out 36.55 (s.d. = 10.77) and for those who did not 35.21 (s.d. = 9.43)) or duration of problem (mean for those who

Table 1 Characteristics of participants at baseline

Baseline characteristic	Behavioural activation (*n* = 23)	Treatment as usual (*n* = 24)	All (*n* = 47)
Age, years: mean (range)	46.43 (24-63)	43.08 (28-63)	44.72 (24-63)
Gender, *n* (%)			
Male	8 (35)	10 (41.7)	18 (38)
Female	15 (65)	14 (58.3)	29 (62)
Employment, *n* (%)			
Full-time	13 (56.5)	8 (33.3)	21 (44.7)
Part-time	1 (4.3)	7 (29.2)	8 (17)
Housewife/husband	1 (4.3)	1 (4.2)	2 (4.3)
Carer	0	1 (4.2)	1 (2.1)
Retired	3 (13)	3 (12.5)	6 (12.8)
Unemployed	4 (17.4)	2 (8.3)	6 (12.8)
Incapacity benefit	1 (4.3)	2 (8.3)	3 (6.4)
Duration of problem in weeks, mean (s.d.)	186.91 (358.49)	195.21 (404.64)	191.15 (378.61)
Baseline Beck Depression Inventory-II score, mean (s.d.)	35.57 (9.60)	35.08 (9.60)	35.32 (9.50)
Baseline Work and Social Adjustment Scale score, mean (s.d.)	26.39 (7.30)	25.13 (7.70)	25.74 (7.46)
Baseline Clinical Interview Schedule Revised score, mean (s.d.)	31 (10.99)	33.12 (8.22)	32.09 (9.63)
Prescribed antidepressants, *n* (%)	15 (65)	17 (71)	32 (68)
Baseline Clinical Interview Schedule Revised (ICD–10) diagnosis, *n* (%)			
Mild depression	1 (4.3)	2 (8.3)	3 (6.4)
Moderate depression	13 (56.5)	9 (37.5)	22 (46.8)
Severe depression	8 (34.8)	13 (54.2)	21 (44.7)
Mixed anxiety and depression	1 (4.3)	0	1 (2.1)

dropped out 182 weeks (s.d. = 439) and for those who did not 193 weeks (s.d. = 369)). Of the 23 participants randomised to behavioural activation, 11 received all 12 sessions. Of those with missed sessions, three received one to three sessions, three received four to six sessions and five received seven to nine sessions. The study flow is presented in Fig. 1.

Treatment integrity

All reviewed sessions scored 1 (behavioural activation dominant) in relation to session and homework content and 0 in relation to other therapy modes being prominent; with all sessions scored 1 being classed as an example of behavioural activation.

Additional interventions

Antidepressant medication was prescribed at baseline to 17 (71%) participants in the usual care group and 15 (65%) participants in the behavioural activation group compared with 15/24 (62.5%) and 12/23 (52%) respectively during the intervention phase. Six participants in the usual care group had follow-up from a community psychiatric nurse. Two participants in the behavioural activation group had one session each with a psychiatrist.

Depression symptom level post-treatment on the BDI–II

A one-way between-groups analysis of covariance was conducted with participant's scores on the BDI–II pre-treatment used as the covariate (behavioural activation group $n = 16$, usual care group $n = 22$). After adjusting for baseline BDI–II scores there was a significant difference in favour of behavioural activation of −15.65 points on the BDI–II (95% CI −6.90 to −24.41, $P = 0.001$) representing a large effect size (Cohen's $d = -1.15$, 95% CI −0.45 to −1.85) (Table 2).

Multiple imputation analysis of missing data on the BDI–II

Intention-to-treat analyses with multiple imputation showed a mean difference on post-BDI–II scores of −15.78 in favour of behavioural activation (95% CI −24.55 to −7.02, $P = 0.001$), with all randomised participants (behavioural activation group $n = 23$, usual care group $n = 24$) included in the analysis.

Clinically significant improvement on the BDI–II

Using LOCF ($n = 47$), 65.2% of the behavioural activation group showed reliable improvement compared with 33.3% of the control group (odds ratio (OR) = 3.8, 95% CI 1.1–12.5). Although more of the treatment group (43.5%) met criteria for reliable and clinically significant change than the control group (20.8%), the confidence interval for the odds ratio included 1 (OR = 2.9, 95% CI 0.8–10.6). Response rates were higher in the treatment group (47.8% *v.* 16.7%, OR = 4.6, 95% CI 1.2–17.7) and were on the border of significance for remission (39.1% *v.* 12.5%, OR = 4.5, 95% CI 1.0–19.6). Four participants (16%) in the usual care group demonstrated deterioration at 3 months, which was not observed in the behavioural activation group. Figure 2 summarises pre- to post-treatment change against reliable and clinically significant criteria.

Functioning post-treatment on the WSAS

A one-way between-groups analysis of covariance was conducted with participant's scores on the WSAS pre-treatment used as the covariate (behavioural activation group $n = 16$, usual care group $n = 22$). After adjusting for baseline WSAS scores, there was a significant difference in favour of behavioural activation of −11.56 points (95% CI −4.79 to −18.33, $P = 0.001$).

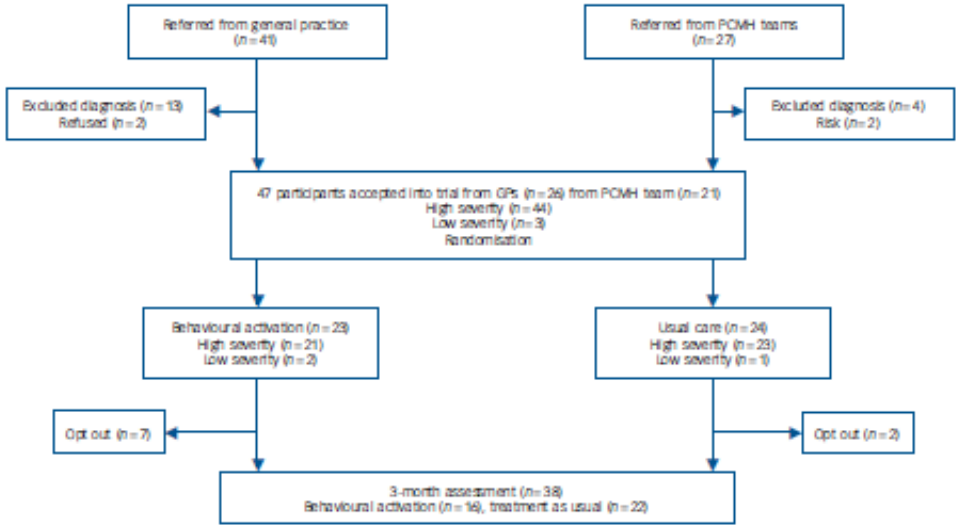


Fig. 1 Study flow chart. PCMH, primary care mental health; GPs, general practitioners.

Table 2 Analysis of outcome scores of behavioural activation v. usual care at 3-month assessment

	Pre-treatment						Post-treatment							
	Behavioural activation			Usual care			Behavioural activation			Usual care				
	Mean	(s.d.)	*n*	Mean	(s.d.)	*n*	Mean	(s.d.)	*n*	Mean	(s.d.)	*n*	Mean difference (95% CI)	*P*	Standardised mean difference (95%CI)
BDI-II															
Completers	35.57	(9.60)	23	35.08	(9.60)	24	11.93	(11.84)	16	27.40	(14.01)	22	−15.65 (−24.41 to −6.90)	0.001	−1.15 (−1.85 to −0.46)
Intention-to-threat[a]							n/a		23	n/a		24	−15.78 (−24.55 to −7.02)	0.001	n/a
WSAS															
Completers	26.39	(7.30)	23	25.13	(7.70)	24	11.12	(9.64)	16	22.68	(10.07)	22	−11.56 (−18.33 to −4.79)	0.001	−1.14 (−1.84 to −0.45)
Intention-to-threat[a]							n/a		23	n/a		24	−11.12 (−17.53 to −4.70)	0.001	n/a
CSQ															
Completers		n/a			n/a		29.13	(3.70)	16	24.32	(3.94)	22	4.81 (2.23 to 7.38)	0.001	n/a

BDI-II, Beck Depression Inventory–II; WSAS, Work and Social Adjustment Scale; CSQ-8, Client Satisfaction Questionnaire–8; n/a, not applicable.
a. Intention-to-treat using multiple imputation incorporates repeated imputed values of post scores for analysis hence individual post-treatment mean (s.d.) values are not produced.

Fig. 2 Participants meeting reliable and clinically significant change criteria by group.

Reliable improvement requires a score below the lowest diagonal. Scores below the lowest diagonal and below the horizontal reference line meet criteria for reliable and clinically significant change. BDI-II, Beck Depression Inventory–II.

Multiple imputation of missing data on the WSAS

Intention-to-treat analysis with multiple imputation showed a mean difference on post-WSAS scores of −11.12 in favour of behavioural activation (95% CI −17.53 to −4.70, $P=0.001$), with all randomised participants (behavioural activation group $n=23$, usual care group $n=24$) included in the analysis.

CSQ–8

Analyses were conducted for the 38 participants (16 in the behavioural activation group and 22 in the usual care group) who completed the CSQ–8 for post-treatment assessment. We observed a mean difference in post-treatment CSQ–8 scores in favour of behavioural activation of 4.81 (95% CI 2.23–7.38, $P=0.001$) showing a higher level of general satisfaction in the behavioural activation group.

Discussion

Main findings

We found that behavioural activation was an effective therapy for depression compared with usual care when delivered by generic mental health staff trained to follow a behavioural activation protocol. Behavioural activation as a step-three high-intensity intervention delivered by generic mental health staff had significantly greater benefits in terms of our primary outcome of depression symptom level and our secondary outcomes of functioning and satisfaction. Reliable and clinically significant change and response and recovery criteria also indicated greater change in the behavioural activation group. We obtained similar results when we used multiple imputation to account for missing data. This paper represents an important addition to the behavioural activation evidence base since this is the first trial to test directly the parsimony hypothesis first advanced by Jacobson *et al* over 10 years ago.[6] Our results support the tentative findings

of Boer *et al*[30] that paraprofessionals may be able to deliver psychological therapies as effectively as professionals. Large, adequately powered trials are required to examine this further.

The various clinical significance criteria used in this study also provide support for behavioural activation delivered by generic workers in comparison with usual care. Differences between the groups were significant or close to significant for nearly all comparisons using both Jacobson & Truax[27] reliable and clinically significant change criteria and the response and remission criteria. In addition, we found an effect size of −1.15 on depression symptom level post-treatment favouring behavioural activation. This compares favourably with an overall effect size of −0.70 of 12 studies (459 participants) comparing behavioural activation with controls, using experienced therapists in our previously reported meta analysis.[17]

We also observed similar findings in terms of functioning; although both groups improved, in the behavioural activation group improvement was substantial, whereas in the usual care group this was marginal. There was a significant difference in favour of behavioural activation at 3 months, suggesting less functional impairment in this group.

Satisfaction with behavioural activation appeared extremely good with a mean of 29 on a 32-point scale, significantly better than usual care, although this also received a reasonably positive evaluation. This finding would suggest that alongside the clinical gains achieved by behavioural activation delivered by generic mental health workers, those receiving the intervention found the experience very acceptable.

Participants in the study had high baseline depression of long duration representing a highly complex clinical group with longer-term and recurrent depression. Therapists were representative of the vast pool of generic mental health workers in that they had no previous therapy training, were relatively recently qualified and employed at the base level of registered psychiatric nurses. They delivered behavioural activation to a high standard in accordance with a 12-session protocol. Based on independent evaluation it would appear that with very little training it is possible to equip staff with the skills to deliver behavioural activation that is both acceptable and clinically effective.

Study limitations

This was an exploratory study and there are a number of limitations of note. First, the relatively small numbers of participants and therapists recruited limits the generalisability of results. We did however base numbers on power calculations from our previous meta-analysis suggesting a sample size of 23 in each arm was sufficient to detect previously observed effect sizes. Despite this, it would have been beneficial to have recruited more participants to allow for our completers analysis to reflect these numbers rather than our intention-to-treat analysis. The small sample size may account for the wide 95% confidence intervals found in this study post-treatment. These should be considered when reflecting on findings; future larger trials should provide more precise estimates of difference.

We found that more people dropped out in the behavioural activation group than in the usual care group. This in itself is not surprising as behavioural activation is an active intervention relying on the person receiving treatment to complete homework on a regular basis. Of those discontinuing treatment, three did so in the first month and three in the second. The drop-out rate in this study is similar to those seen in large data-sets of CBT provision in primary care.[31] Usual care in contrast does not involve as much investment from the participant and is nested within a person's overall healthcare, hence reducing the likelihood of drop out. Larger sample sizes however would have allowed us to have a more precise estimate of this finding as it is likely our limited sample was underpowered to accurately estimate likely drop-out from behavioural activation delivered by this workforce.

We used a self-report measure as our analysis of depression symptom level post-intervention. This is a source of potential information bias as in psychological therapy trials participants will be aware of their treatment allocation; this should be balanced in future studies by a repeated diagnostic interview. The lack of follow-up is also a limiting factor in the interpretation of the study. Previous studies of behavioural activation have demonstrated its durability to be equal to other therapies such as CBT; however, such studies have been delivered by experienced therapists. The aim of this study was to explore feasibility and parsimony of behavioural activation, although we found results supporting this we were unable to conduct follow-up assessments, which must be accounted for in future research. We were also unable to incorporate any validated measure of behavioural activation competence. This would have been beneficial to further asses the quality of behavioural activation administered and any possible contamination of the treatment modality.

Clinical implications and future research

This study is the first randomised controlled clinical trial to test feasibility of dissemination of behavioural activation to a wider mental health workforce and as such represents a major step forward in our understanding of this intervention. We have demonstrated that with limited training, generic mental health workers can be trained to deliver clinically effective behavioural activation to people with severe long-standing depression. If these findings can be replicated and translated into routine healthcare, then the clinical and cost implications of this finding are substantial for this prevalent and disabling condition. Now that such feasibility has been shown, future research with a larger sample and multiple therapists should investigate its longer-term durability and compare this behavioural activation delivery mode with an active psychological treatment such as CBT.

David Ekers, MSc, Dip CBT, RMN, Tees Esk and Wear Valleys NHS Trust and Mental Health Research Centre, Durham University, Co Durham; **David Richards**, PhD, RN, School of Psychology, University of Exeter, Exeter; **Dean McMillan**, DclinPsy, PhD, Department of Health Sciences, Hull York Medical School and Department of Health Sciences, University of York, York; **J. Martin Bland**, MSc PhD, Department of Health Sciences, University of York, York; **Simon Gilbody**, DPhil, FRCPsych, Psychological Medicine and Health Services Research, Hull York Medical School and Department of Health Sciences, University of York, York, UK

Correspondence: David Ekers, Mental Health Research Centre, Durham University Health Centre, Chester Le Street, Co Durham, DH3 3UR, UK. Email: david.ekers@tewv.nhs.uk

First received 23 Feb 2010, final revision 2 Aug 2010, accepted 11 Aug 2010

References

1 World Health Organization. *The World Health Report 2001: Mental Health: New Understanding, New Hope*. WHO, 2001.

2 Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer Y. *Office of National Statistics: Psychiatric Morbidity Among Adults Living in Private Households*. HMSO, 2001.

3 National Institute for Health and Clinical Excellence. *Management of Depression in Primary and Secondary Care*. NICE, 2009.

4 Hollon S, Thase M, Markowitz J. Treatment and prevention of depression. *Psychol Sci Public Interest* 2002; 3: 1–39.

5 Beck A. *Cognitive Therapy and the Emotional Disorders*. International Universities Press, 1976.

6 Jacobson N, Dobson K, Truax P, Addis M, Koerner K, Gollan J, et al. A component analysis of cognitive-behavioral treatment for depression. *J Consult Clin Psychol* 1996; 64: 295–304.

7 Gortner E, Gollan J, Dobson K, Jacobson N. Cognitive-behavioral treatment for depression: relapse prevention. *J Consult Clin Psychol* 1998; 66: 377–84.

8 Jacobson N, Gortner E. Can depression be de-medicalized in the 21st century: scientific revolutions, counter revolutions and the magnetic field of normal science. *Behav Res Ther* 2000; 38: 103–17.

9 Dimidjian S, Hollon S, Dobson K, Schmaling K, Kohlenberg R, Addis M, et al. Randomized trial of behavioural activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *J Consult Clin Psychol* 2006; 74: 658–70.

10 Cuijpers P, van Straten A, Warmerdam L. Behavioural activation treatments of depression: a meta analysis. *Clin Psychol Rev* 2007; 27: 318–26.

11 Ekers D, Richards D, Gilbody S. A meta analysis of behavioural therapy for depression. *Psychol Med* 2008; 38: 611–23.

12 Mazzucchelli T, Kane R, Rees C. Behavioral activation treatments for depression in adults: a meta-analysis and review. *Clin Psychol Sci Pract* 2009; 16: 383–411.

13 Schulz K, Altman D, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: 698–702.

14 Lewis G, Pelosi A, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardised assessment tool for use by lay interviewers. *Psychol Med* 1992; 22: 465–86.

15 World Health Organization. *The ICD–10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. WHO, 1992.

16 Elkin I, Gibbons R, Shea M, Sotsky S, Watkins J, Pilkonis P, et al. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol* 1995; 63: 841–7.

17 Beck A, Steer R, Brown G. *Manual for Beck Depression Inventory II (BDI–II)*. Psychology Corporation, 1996.

18 Mundt JC, Marks IM, Shear MK, Greist JM. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *Br J Psychiatry* 2002; 180: 461–4.

19 Attkisson C, Greenfield T. The UCSF Client Satisfaction Scales: I. The Client Satisfaction Questionnaire-8. In *The Use of Psychological Testing for Treatment Planning and Outcome Assessment, 3rd edn* (ed M Maruish): 799–812. Lawrence Erlbaum Associates, 2004.

20 Marsill C, Addis M, Jacobson N. *Depression in Context. Strategies for Guided Action*. Norton, 2001.

21 Hopko D, Lejuez C, Ruggiero K, Eifert G. Contemporary behavioural activation treatments for depression: procedures, principles and progress. *Clin Psychol Rev* 2003; 23: 699–717.

22 Lipsey M, Wilson D. The efficacy of psychological, educational and behavioural treatment. *Am Psychol* 1993; 48: 1181–209.

23 Nüesch E, Trelle S, Richenbach S, Reichenbach S, Rutjes A, Bürgi E, Scherer M, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 2009; 339: b3244.

24 Streiner D, Geddes J. Intention to treat analysis in clinical trials when there are missing data. *Evid Based Ment Health* 2001; 4: 70–1.

25 Sterne J, White I, Carlin J, Spratt M, Royston P, Kenward M, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; 339: b2393.

26 Royston P. Multiple imputation of missing values. *Stata J* 2004; 4: 227–41.

27 Jacobson N, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991; 59: 12–9.

28 Evans C, Margison F, Barkham M. The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evid Based Ment Health* 1998; 1: 70–2.

29 Dozois J, Dobson K, Ahnberg JA. A psychometric evaluation of the Beck Depression Inventory–II. *Psychol Assess* 1998; 10: 83–9.

30 Baer P, Wiersma D, Russo S, Bosch R. Paraprofessionals for anxiety and depressive disorders. *Cochrane Database Syst Rev* 2005; 2: CD004688.

31 Clark D, Layard R, Smithies R, Richards D, Suckling R, Wright B. Improving access to psychological therapy: initial evaluation of two UK demonstration sites. *Behav Res Ther* 2009; 47: 910–20.

100 words

Edvard Munch (1863–1944) *The Scream*

Alexandra Pitman

Edvard Munch is best known for *The Scream*, 1893, an image endlessly reproduced in the media to depict mental anguish. Explanations of the meaning behind the image abound, mainly focusing on an outpouring of emotion in response to suffering. Munch's own explanation is revealed in his diaries, which recall the melancholy of a walk along a bridge with friends. Trembling in fear at the fiery sunset, he sensed 'how an infinite scream was going through the whole of nature'. This dehumanised figure, into which viewers project their own neuroses, is not screaming but blocking out the scream of its existence.

The British Journal of Psychiatry (2011) 198, 72. doi: 10.1192/bjp.198.1.72

COBRA study protocol section based upon research reported in thesis

3.2.1 Preparatory work prior to the COBRA application.

In preparation for the COBRA trial, we conducted a systematic review, meta analysis and phase II pilot trial of BA.[1,2]

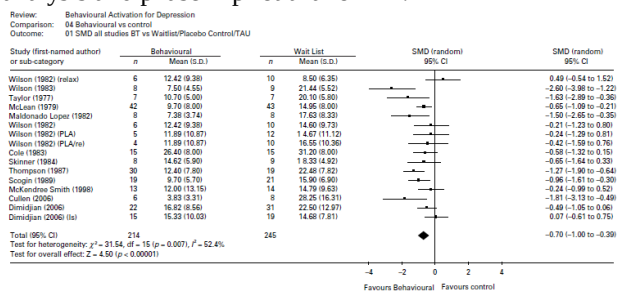
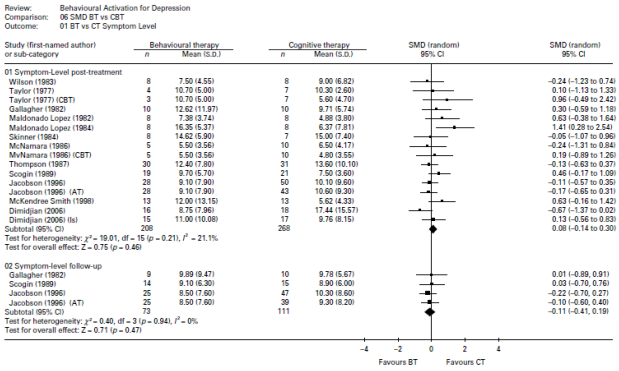
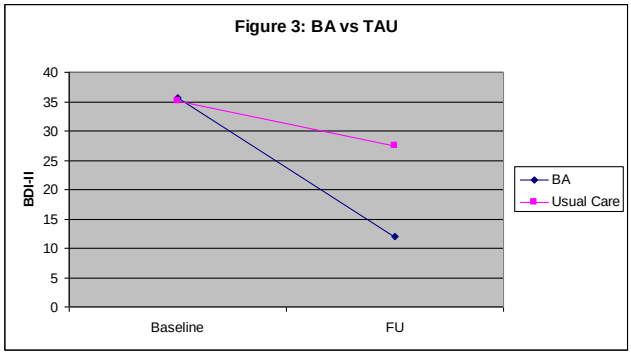


Fig. 1. Behavioural therapy (BT) *versus* wait list/control/placebo symptom-level post-treatment.



3.2.2 Systematic review and meta analysis: in our meta analysis of RCTs,[2] we first found a clinical effect size in terms of a reduced depression score of −0.70 SD units from twelve studies (n=459; 95% CI -1.00 to -0.39; p<0.001) comparing behavioural treatments to controls using experienced therapists (figure 1). We then found twelve studies comparing behavioural treatments with CBT (n=476) and showed that behavioural treatments had equivalent outcomes to CBT (pooled SMD 0.08; 95% CI -0.14 to 0.30, p=0.46), (additional figure). In a subsequent meta-regression analysis, these behavioural treatments demonstrated a greater level of effectiveness at more severe levels of depression (meta-regression b-coefficient -0.05; 95%CI -0.10 to -0.01; p=0.04).

However, many of the trials were of limited methodological quality, all were under-powered for comparing treatments and most did not utilise diagnostic interviews for trial inclusion. Treatments in many cases did not conform to modern clinical protocols for BA. Long term outcomes were rarely reported with average follow-up only to four months. Therefore, the existing trial data are insufficient to provide certainty that BA should be a first line treatment for depression and these limitations led to NICE regarding the evidence for BA as equivocal and of insufficient strength to recommend BA for first-line routine NHS depression treatment.[10] Consequently, NICE [p256] made a clear research recommendation "*to establish whether behavioural activation is an effective alternative to CBT*" using a study which is "*large enough to determine the presence or absence of clinically important effects using a non-inferiority design*".[10]



3.2.3 Pilot Phase II Trial: in order to test uncertainties around our main COBRA hypothesis – that BA will be equivalently effective to CBT and more cost effective – we piloted BA in a phase II RCT to examine the parsimony argument directly, i.e. whether generic mental health workers, without previous experience in therapy, can effectively treat depressed people using a full high-intensity BA therapeutic protocol.[1] We compared BA against usual care. BA was delivered by NHS

AfC grade 5 mental health workers with no previous formal training or psychotherapeutic experience, who received five days training in BA and subsequent one hour clinical supervision fortnightly from David Ekers (nurse consultant, educator and COBRA applicant). Intention to treat analyses (figure 3) indicated a difference in favour of BA of −15.79 (n=47; 95% CI −24.55 to −7.02) on depression (Beck Depression Inventory-II), an effect size of −1.15 SD units (95% CI − 0.45 to −1.85). This compares favourably to the overall effect size of −0.70 comparing BA to controls using experienced therapists in our meta analysis above.[2] The mental health workers demonstrated excellent fidelity to the protocol when audio recordings were assessed by independent accredited cognitive behavioural therapists with extensive experience in BA.

3.2.4 Implications: the implications for COBRA are thus: we have demonstrated that data from a number of small trials indicates that BA and CBT may have similar effectiveness when delivered by specialist therapists; we have also demonstrated that generic mental health workers with no prior training in psychological therapies can deliver a full BA protocol and achieve results from treating patients with BA which are at least as powerful as those achieved by specialist workers compared to usual care. We now need to combine these research strands and test BA against CBT using generic mental health workers to deliver BA in an adequately powered non-inferiority RCT using gold standard diagnostic and depression severity outcomes. We need to include a cost effectiveness analysis to see if BA should join CBT as a first line treatment for major depression with additional cost advantages to the NHS. If CBT and BA are equally effective, then BA may be preferable because its simplicity in delivery and mechanism of change allows for more cost effective delivery by less specialised NHS mental health workers. We also need to plan a trial with sufficient length of follow-up to measure impact on long-term outcomes and costs. Finally, we need to ensure that the training, supervision and treatment fidelity checks ensure that the trial is an adequate test of the two interventions, BA and CBT. As noted by NICE [p256] "*the results of this study will have important implications for the provision of psychological treatment in the NHS*."[10]

References

Altman, D. (2009) Missing outcomes in randomized trials: addressing the dilemma *Open Medicine*, **3**(2), ISSN 1911-2092

Altman, D. and Bland, M. (1996) Presentation of numerical data. *British Medical Journal*, **312**, 572.

Altman, D. and Bland, M. (1999a) How to randomise. *British Medical Journal*, **319**, 703-704.

Altman, D. and Bland, M. (1999b) Treatment allocation in controlled trials: why randomise? *British Medical Journal*, **318**, 1209.

Altman, D. and Bland, M. (2007) Missing Data. *British Medical Journal*, **334**, 424.

Amenson, C. and Lewinsohn, P. (1981) An investigation into the observed sex difference in prevalence of unipolar depression. *Journal of Abnormal Psychology*, **90**, 1-13.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders. Fourth Edition (DSM-IV)*, American Psychiatric Association.

Andrews, G. (2001) Should Depression be Managed as a Chronic Disease. *British Medical Journal*, **322**, 419-421.

Anello, C. and Fleiss, J. (1995) Exploratory or analytic meta-analysis: should we distinguish between them? *Journal of Clinical Epidemiology*, **48**, 109-116.

Attkisson, C. and Greenfield, T. (2004) The UCSF Client Satisfaction Scales: I. The Client Satisfaction Questionnaire-8. In *The use of psychological testing for treatment planning and outcome assessment (3rd. Ed.)* (Ed, Maruish, M.), Lawrence Erlbaum Associates, Mahwah, NJ.

Baigenta, C., Harrellb, F., Buysec, M., Embersona, J. and Altman, D. (2008) Ensuring trial validity by data quality assurance and diversification of monitoring methods. *Clinical Trials*, **5**(1), 49-55.

Barton, G., Hodgekins, J., Mugford, M., Jones, P., Croudace, T. and Fowler, D. (2009) Measuring the benefits of treatment for

psychosis: validity and responsiveness of the EQ–5D. *British Journal of Psychiatry,* **195**, 170-177.

Bebbington, P., Dunn, G., Jenkins, R., Lewis, G., Brugha, T., Farrell, M. and Meltzer, H. (2003) The influence of age and sex on the prevalence of depressive conditions: report from the National Survey of Psychiatric Morbidity. *International Review of Psychiatry* **15**, 74-83.

Beck, A. (1976) *Cognitive Therapy and the Emotional Disorders,* International Universities Press: New York.

Beck, A., Rush, A., Shaw, B. and Emery, G. (1979) *Cognitive Therapy of Depression,* Guilford, New York.

Beck, A., Steer, R. and Brown, G. (1996) *Manual for Beck Depression Inventory II (BDI-II),* Psychology Corporation, San Antonio, TX.

Beck, A., Ward, C., Mendelson, M., Mock, J. and Erdbaugh, J. (1961) An inventory for measuring depression. *Archives of General Psychiatry,* **4**, 561-571.

Beekman, A., Penninx, B., Deeg, D., Ormel, J. and van Tilburg, W. (2000) Depression and physical health in later life: results from the Longitudinal Aging Study Amsterdam *Journal Of Affective Disorders* **46**, 219-231.

Begg, C., Cooper, H. and Hedges, L. V. (1994) The Handbook of Research Synthesis (ed. H. Cooper and L. V. Hedges). *Russell Sage Foundation: New York, NY*, 399-409.

Berlin, J. (1997) Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. *Lancet,* **350**, 185-186.

Bhugra, D. and Mastrogianni, A. (2004) Globalisation and mental disorders. Overview with relation to depression. *British Journal Of Psychiatry,* **184**, 10-20.

Bird, A. (2006) *We need to talk. The case for psychological therapy on the NHS.* Mental Health Foundation, London.

Blackburn, I., James, I., Milne, D., Baker, C., Standart, S., Garland, G. and Reichelt, R. (2001) The revised cognitive therapy scale (CTS-

R): psychometric properties. *Behavioural and Cognitive Psychotherapy*, **29**(4), 431-446.

Bland, M. (2004) Clinstat. York University, York, UK.

Bland, M. and Altman, D. (1994) One and two sided tests of significance. *British Medical Journal,* **309**, 248.

Borenstein, M., Hedges, L. V. and Rothstein, H. (2007) Meta-Analysis Fixed effect vs. random effects. (Ed, Meta-analysis.com).

Bosmans, J., Hermens, M., Bruijne, M., van Hout, H., Terluin, B., Bouter, L., Stalman, W. and van Tulder, M. (2008) Cost-effectiveness of usual general practitioner care with or without antidepressant medication for patients with minor or mild-major depression. *Journal of Affective Disorders,* **111**, 106-112.

Bostwick, J. and Pankratz, V. (2000) Affective disorders and suicide risk: a re-examination. *American Journal of Psychiatry,* **157**, 1925-1932.

Boutron, I., Guittet, L., Estellat, C., Moher, D., Hrobjartsson, A. and Ravaud, P. (2007) Reporting Methods of Blinding in Randomized Trials Assessing Non-pharmacological Treatments. *PLoS Med,* **4**(2), e61.

Bower, P., Wallace, P., Ward, E., Graffy, J., Miller, J., Delaney, B. and Kinmonth, L. (2009) Improving recruitment to health research in primary care. *Family Practice,* **26**(5), 391-397.

Briggs, A. (1999) Handling uncertainty in economic evaluation. *British Medical Journal,* **319**, 120.

Briggs, A., Clarke, T., Wolstenholme, J. and Clarke, P. (2003) Missing presumed at random: cost-analysis of incomplete data. *Health Economics,* **12**, 377-392.

Briggs, A., O'Brien, B. and Blackhouse, G. (2002) Thinking outside the box: Recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Annual Review of Public Health,* **23**, 377-401.

British Medical Association & Royal Pharmaceutical Society of Great Britain (2009) *British National Formulary 57*, London

Brooks, R. (1996) EuroQol:the current state of play. *Health Policy,* **16**(3), 53-72.

Brown, R. and Lewinsohn, P. (1984) A psychoeducational approach to the treatment of depression: Comparison of group, individual, and minimal contact procedures. *Journal of Consulting and Clinical Psychology,* **52**(5), 774-783.

Brugha, T., Jenkins, R., Bebbington, P., Meltzer, H., Lewis, G. and Farrell, M. (2004) Risk factors and the prevalence of neurosis and psychosis in ethnic groups in Great Britain. *Social Psychiatry & Psychiatric Epidemiology* **39**, 939-946.

Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L. and Klassen, T. (2006) Single data extraction generated more errors than double data extraction in systematic reviews. *Journal of Clinical Epidemiology* **59**, 697-703.

Campbell, M., Fitzpatrick, R., Haines, A., Kinmouth, A., Sandercock, P., Spiegelhalter, D. and Tyrer, P. (2000) Framework for design and evaluation of complex interventions to improve health. *British Medical Journal,* **321**, 694-696.

Carter, A., Griffin, G. and Carter, T. (2006) A survey identified publication bias in the secondary literature. *Journal of Clinical Epidemiology,* **59**, 241-245.

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*.

Centre for Reviews and Dissemination (2009) *Systematic reviews: CRD's guidance for undertaking reviews in health care.* University of York, York.

Chalmers, I. (2001) Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in theraputic experiments. *International Journal of Epidemiology* **30**, 1156-1164.

Chew-Graham, C., Baldwin, R. and Burns, A. (2004) Treating depression in later life. *British Medical Journal,* **329**, 181-182.

Churchill, R., Hunot, V., Corney, R., Knapp, M., McGuire, H., Tylee, A. and Wessley, S. (2001) A systematic review of controlled trials of the effectiveness and cost effectiveness of brief psychological treatments of depression. *Health Technology Assessment,* **5**, 35.

Clark, D., Layard, R., Smithies, R., Richards, D., Suckling, R. and Wright, B. (2009) Improving access to psychological therapy: Initial evaluation of two UK demonstration sites. *Behaviour Research and Therapy,* **47**, 910-920.

Cochrance Collaboration (2003) Review Manager (RevMan) [computer program]. Version 4.2 for Windows. *The Nordic Cochrane Centre: Copenhagen*.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences (2nd ed.),* Lawrence Earlbaum Associates, Hillsdale NJ.

Cole, M. (1983) Behavioural treatment of chronic depression. *Dissertation Abstracts International, ,* **45**(3).

Cole, M., Bellavance, F. and Mansour, A. (1999) Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. *American Journal of Psychiatry* **156**, 1182-1189.

Counsell, C. (1997) Formulating questions and locating primary studies for inclusion in systematic reviews. *Annals of Internal Medicine,* **127**, 380-387.

Counsell, C. (1998) Formulating the questions and locating the studies for inclusion in systematic reviews In *Systematic reviews. Synthesis of best evidence for health care decisions* (Eds, Mulrow, C. and Cook, D.), American College of Physicians, Philadelphia, USA, pp. 67-79.

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

Cullen, J., Spates, R., Pagoto, S. and Doran, N. (2006) Behavioural activation treatment for major depressive disorder: a pilot investigation. *Behavior Analyst Today,* **7**, 151-166.

Curtis, L. (2009) *Unit costs of health and social care 2009.* Personal Social Services Research Unit, University of Kent, Canterbury.

Das-Munshi, J., Goldberg, D., Bebbington, P., Bhugra, D., Brugha, T., Dewey, M., Jenkins, R., Stewart, R. and Prince, M. (2008) Public health significance of mixed anxiety and depression: beyond current classification. *British Journal of Psychiatry*, **192**, 171-177.

Davis, J., Rawana, J. and Capponi, D. (1989) Acceptability of behavioural staff management techniques. *Behaviour Research and Therapy*, **4** 23-44.

Deeks, J., Higgins, J. and Altman, D. (2009) Chapter 9: Analysing data and undertaking meta-analyses. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 (updated 2009)* (Eds, Higgins, J. and Green, S.), The Cochrane Collaboration.

Department of Health (2006) *From values to action: The Chief Nursing Officer's review of mental health nursing.* DoH, London.

Dimidjian, S. and Davis, K. (2009) Newer variations of cognitive-behavioral therapy: behavioral activation and mindfulness-based cognitive therapy. *Current Psychiatry Reports*, **11**, 453-458.

Dimidjian, S., Hollon, S., Dobson, K., Schmaling, K., Kohlenberg, R., Addis, M., Gallop, R., McGlinchey, J., Markley, D., Gollan, J., Atkins, D., Dunner, D. and Jacobson, N. (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. (1989) A meta analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, **57**, 414-419.

Dobson, K., Dimidjian, S., Kohlenberg, R., Rizvi, S., Hollon, S., Schmaling, K., Gallop, R. and J, G. (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), 468-477.

Dolan P, Gudex C, Kind P and Williams A (1995) *A social tariff for EuroQoL: Results from a UK General Population Survey.* Economics, C. f. H., University of York, York.

Doll, R. (1998) Controlled trials: the 1948 watershed. *British Medical Journal,* **317**, 1217-1220.

Dozois, J., Dobson, K. and Ahnberg, J. A. (1998) A Psychometric Evaluation of the Beck Depression Inventory-II. *Psychological Assessment,* **10**, 83-89.

Drummond, C., Coulton, S., James, D., Godfrey, C., Parrott, S., Baxter, J., Ford, D., Lervy, B., Rollnick, S., Russell, I. and Peters, T. (2009) Effectiveness and cost-effectiveness of a stepped care intervention for alcohol use disorders in primary care: pilot study. *British Journal of Psychiatry,* **195**, 448-456.

Drummond, M., Sculpher, M., Torrance, G., O'Brien, B. and Stoddart, G. (2005) *Methods for the economic evaluation of health care programmes: third edition,* Oxford University Press, Oxford

Dupont, W. and Plummer, W. (2009) PS Power and Sample Size Calculator. Vanderbilt University.

Easterbrook, P., Berlin, J., Goalan, R. and Matthews, D. (1991) Publication bias in clinical research. *Lancet,* **337**(867-872).

Egger, M., Davey-Smith, G., Schneider, M. and Minder, C. (1997) Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal,* **315**, 629-634.

Ekers, D. (2010) Training depression case managers. In *Oxford Guide to Low Intensity CBT Interventions* (Eds, Bennet-levy, J., Richards, D., Christensen, C., Farrand, P., Griffiths, K., Kavanagh, D., Klein, B., Proudfoot, J., Ritterband, L., White, J. and Williams, C.), Oxford University Press, Oxford, UK.

Ekers, D., Richards, D. and Gilbody, S. (2008) A Meta Analysis of Behavioural Therapy for Depression. *Psychological Medicine,* **38**(5), 611-623.

Elkin, I., Gibbons, R., Shea, M., Sotsky, S., Watkins, J., Pilkonis, P. and Hedeker, D. (1995) Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of

Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, **63**, 841-847.

EuroQol Group (1990) EuroQol: A new Instrument for the measurement of Health Related Quality of Life. *Health Policy*, **16**, 199-208.

Evans, C., Margison, F. and Barkham, M. (1998) The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence-based Mental Health,* **1**, 70-72.

Everitt, B. and Wessley, S. (2008) *Clinical Trials in Psychiatry,* Wiley, Chichester.

Fava, M. and Kendler, K. (2000) Major depressive disorder. *Neuron,* **28**, 335-341.

Fenwick, E. and Byford, S. (2005) A guide to cost-effectiveness acceptability curves. *British Journal of Psychiatry,* **187**, 106-108.

Fenwick, E., Claxton, K. and Sculpher, M. (2001) Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics,* **10**, 779-787.

Ferster, C. (1973) A functional analysis of depression. *American Psychologist,* **28**, 857-870.

Fletcher, J., Bower, P., Gask, L., Richards, D. and Saunders, T. (2006) *Primary Care Services for Depression: A Guide to Best Practice.* Care Services Improvement Partnership, Hyde, Cheshire.

Furukawa, T., Barbui, C., Cipriani, A., Brambilla, P. and Watanabe, N. (2006) Inputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology,* **59**, 7-10.

Galandi, D., Schwarzer, G. and Antes, G. (2006) The demise of the randomised controlled trial: bibliometric study of the German-language health care literature 1948 to 2004. *BMC Medical Research Methodology,* **6**, 30.

Gallagher, D. and Thompson, L. (1982) Treatment of major depressive disorder in older adult outpatients with brief psychotherapies. *Psychotherapy,* **19**, 482-490.

Gardner, P. and Oei, T. (1981) Depression and self-esteem: An investigation that used behavioral and cognitive approaches to the

treatment of clinically depressed clients. *Journal of Clinical Psychology*.

Gellatly, J., Bower, P., Hennessy, S., Richards, D., Gilbody, S. and Lovell, K. (2007) What makes self help interventions effective in the management of depressive symptoms? Meta analysis and meta regression. *Psychological Medicine,* **37**(9), 1217-1228.

Gilbody, S., Bower, P., Fletcher, J., Richards, D. and Sutton, A. (2006) Collaborative care for depression. A cumulative meta-analysis and review of longer-term outcomes *Archives of Internal Medicine* **166**, 2314-2321.

Glanville, J. (2002) Phase 3: Identification of Research. In *Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. Report 4, 2nd edition* Centre for Reviews and Dissemination, University of York, York.

Glanville, J. and Sowden, A. (2002) Phase 0: Identification of the need for a review In *Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. Report 4, 2nd edition* Centre for Reviews and Dissemination, University of York, York.

Gloaguen, V., Cottraux, J., Chucherat, M. and Blackburn, I. (1998) A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders,* **49**, 59-72.

Glover, G., Webb, M. and Evison, F. (2010) *Improving Access to Psychological Therapies: A review of the progress made by sites in the first rollout year.* North East Public Health Observatory, Durham University.

Goldberg, D. and Huxley, P. (1992) *Common Mental Disorders: a Bio-Social Model,* Routledge, London.

Gortner, E., Gollan, J., Dobson, K. and Jacobson, N. (1998) Cognitive-behavioral treatment for depression: relapse prevention. *Journal of Consulting and Clinical Psychology,* **66**, 377-384.

Gøtzsche, P., Hróbjartsson, A., Maric, K. and Tendal, B. (2007) Data extraction errors in meta-analyses that use standardized mean

differences. *Journal of the American Medical Association,* **298**, 430-437.

Green, S. and Higgins, J. (2008) Chapter 2: Preparing a Cochrane review. In *Cochrane Handbook for Systematic Reviews of Interventions* The Cochrane Collaboration.

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

Hathaway, S. and McKinley, J. (1943) *The Minnesota Multiphasic Personality Inventory,* Psychological Corporation, New York.

Hayes, S., Strosahl, K. and Wilson, K. L. G. P. (1999) *Acceptance and Commitment Therapy: An experiential approach to behaviour change,* Guilford Press, London.

Healy, D. (1997) *The antidepressant era* Harvard University Press, Cambridge MA.

Hedges, L. and Vevea, J. (1998) Fixed- and random-effects models in meta-analysis. *Psychological Methods,* **3**(4), 486-504.

Henderson, S., Andrews, G. and Hall, W. (2001) Australia's mental health: an overview of the general population survey. *Australian & New Zealand Journal of Psychiatry,* **34**(2), 197-205.

Hernstein, R. (1970) On the law of effect. *Journal of the Experimental Analysis of Behavior,* **13**, 243-266.

Higgins, J. and Deeks, J. (2008) Chapter 7: Selecting studies and collecting data. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1* The Cochrane Collaboration.

Higgins, J., Deeks, J. and Altman, D. (2008) Chapter 16: Special topics in statistics. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1* The Cochrane Collaboration.

Higgins, J. and Thompson, S. (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 1539-1558.

Higgins, J. and Thompson, S. (2004) Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine,* **23**, 1663-1682.

Higgins, J., Thompson, S., Deeks, J. and Altman, D. (2003) Measuring inconsistency in meta-analyses. *British Medical Journal,* **327**, 557-560.

Hirschfeld, R., Keller, M., Panico, S., Arons, B., Barlow, D., Davidoff, F., Endicott, J., Froom, J., Goldstein, M., Gorman, J., Guthrie, D., Marek, R. and Maurer, T. (1997) The National Depressive and Manic-Depressive Association consensus statement on the under treatment of depression. *Journal of the American Medical Association,* **277**, 333-340.

Hollinghurst, S., Kessler, D., Peters, T. and Gunnell, D. (2005) Opportunity cost of antidepressant prescribing in England: analysis of routine data. *British Medical Journal,* **330**, 999-1000.

Hollinghurst, S., Peters, T., Kaur, S., Wiles, N., Lewis, G. and Kessler, D. (2010) Cost-effectiveness of therapist-delivered online cognitive–behavioural therapy for depression: randomised controlled trial. *British Journal of Psychiatry,* **197**, 297-304.

Hollon, S. and Garber, J.-.-. (1990) Cognitive therapy for depression: a social cognitive perspective. *Personality and Social Psychology Bulletin,* **16**, 58-73.

Hollon, S., Muñoz, R., Barlow, D., Beardslee, W., Bell, C., Bernal, G., Clarke, G., Franciosi, L., Kazdin, A., Kohn, L., Linehan, M., Markowitz, J., Miklowitz, D., Persons, J., Niederehe, G. and Sommers, D. (2002a) Psychosocial intervention development for the prevention and treatment of depression: Promoting innovation and increasing access. *Biological Psychiatry,* **52**, 610-630.

Hollon, S., Thase, M. and Markowitz, J. (2002b) Treatment and prevention of depression. *Psychological Science in the Public Interest,* **3**, 1-39.

Hopewell, S., Louden, K., Clarke, M., Oxman, A. and Dickersin, K. (2008) *Publication bias in clinical trials due to statistical significance or direction of trial results.* Cochrane Database of Systematic Reviews, The Cochrane Collaboration.

Hopko, D., Lejuez, C., Lepage, J., Hopko, S. and McNeil, D. (2003) A brief behavioral activation treatment for depression; a randomized pilot trial within an inpatient psychiatric hospital. *Behavior Modification*, **27**, 458-469.

Hopko, D., Lejuez, C., Ruggiaro, K. and Eifert, G. (2003) Contemporary behavioural activation treatments for depression: procedures, principles and progress. *Clinical Psychology Review*, **23**, 699-717.

Horwitz, R. (1995) Large-scale randomized evidence: large, simple trials and overviews of trials: discussion. A clinician's perspective on meta-analyses (comment). *Journal of Clinical Epidemiology*, **48**, 41-44.

Hunt, C., Shepherd, L. and Andrews, G. (2001) Do doctors know best? Comments on a failed trial. *The Medical Journal of Australia*, **174**, 144-146.

Isaacs, D. and Fitzgerald, D. (1999) Seven alternatives to evidence based medicine. *British Medical Journal*, **319**, 18-25.

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

Jacobson, N. and Gortner, E. (2000) Can depression be de-medicalised in the 21st century: scientific revolutions, counter revolutions and the magnetic field of normal science. *Behaviour Research and Therapy*, **38**, 103-117.

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

Jacobson, N. and Truax, P. (1991) Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, **59**, 12-19.

Jadad, A., Moore, R., Carroll, D., Jenkinson, C., Reynolds, D., Cavaghan, D. and McQuay, H. (1996) Assessing the quality of

reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials,* **17**, 1-12.

Judd, L., Akiskal, H., Maser, J., Zeller, P., Endicott, J., Coryell, W., Paulus, M., Kunovac, J., Leon, A., Mueller, T., Rice, J. and Keller, M. (1998) A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Archives of General Psychiatry,* **55**, 694-700.

Jüni, P., Holenstein, F., Sterne, J., Bartlett, C. and Egger, M. (2002) Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology,* **31**(115-123).

Jüni, P., Witschi, A., Bloch, R. and Egger, M. (1999) The Hazards of Scoring the Quality of Clinical Trials for Meta-analysis. *Journal of the American Medical Association,* **282**, 1054-1060.

Kanter, J., Busch, A. and Rusch, L. (2009) *Behavioural Activation,* Routledge, London and New York.

Kanter, J., Manos, R., Bowe, W., Baruch, D., Busch, A. and Rusch, L. (2010) What is behavioural activation? A review of the empirical literature. *Clinical Psychology Review,* **30**, 608-620.

Katon, W., Schoenbaum, M., Fan, M., Callahan, C., Williams, J., Hunkeler, E., Harpole, L., Zhou, X., Langston, C. and Unutzer, J. (2005) Cost-effectiveness of improving primary care treatment of late life depression. *Archives of General Psychiatry,* **62**, 1313-1320.

Kendrick, T., Simons, L., Mynors Wallice, L., Gray, A., Lathlean, J., Pickering, R., Harris, S., Rivero-Arias, K., Gerard, K. and Thompson, C. (2006) Cost-effectiveness of referral for generic care or problem-solving treatment from community mental health nurses, compared with usual general practitioner care for common mental disorders. *British Journal of Psychiatry,* **189**, 50-59.

Kennedy, N., Abbott, R. and Paykel, E. (2004) Longitudinal syndromal and sub-syndromal symptoms after severe depression. *British Journal of Psychiatry,* **184**, 330-336.

Kessler, R., Berglund, P., Demler, O., Jin, R., Merikangas, K. and Walters, E. (2005) Lifetime prevalence and age-of-onset

distributions of DSM-IV disorders in the National Co morbidity Survey Replication. *Archives of General Psychiatry*, **62**, 593-602.

Kessler, R., Stang, P., Wittchen, H., Ustun, T., Roy-Burne, P. and Walters, E. (1998) Lifetime panic-depression comorbidity in the National Co morbidity Survey. *Archives of General Psychiatry* **55**, 801-808.

Khan, K., Glanville, J. and Kleijnen, J. (2002) Phase 1: Preparation of a proposal for a systematic review In *Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews. Report 4, 2nd edition* Centre for Reviews and Dissemination, University of York.

Khan, K. and Kleijnen, J. (2002) Phase 4: Selection of Studies. In *Undertaking systematic reviews of research on effectiveness CRD's guidance for those carrying out or commissioning reviews: Report 4 (2nd edition)* Centre for reviews and dissemination, University of York

Klerman, G., Weissman, M., Rounsaville, B. and Chevron, E. (1984) Interpersonal Psychotherapy of Depression. *Basic Books: New York, NY*.

Kupfer, D. (1991) Long-term treatment of depression. *Journal of Clinical Psychiatry* **52**, 28-34.

Larsen, D., Attkisson, C., Hargreaves, W. and Nguyen, T. (1979) Assessment of client/patient satisfaction: development of a general scale. *Evaluation and Program Planning*, **2**, 197-207.

Layard, R. (2004) Mental health: Britain's biggest social problem? Strategy Unit Seminar Paper: RL414c. *http://www.strategy.gov.uk/downloads/files/mh_layard.pdf*.

Layard, R. (2005) Mental health: Britain's biggest social problem? Strategy Unit Seminar Paper: RL414c. *http://www.strategy.gov.uk/downloads/files/mh_layard.pdf*.

Layard, R. (2006) The case for psychological treatment centres. *British Medical Journal*, **332**, 1030-1032.

Lejuez, C., Hopko, D., Acierno, R., Daughters, S. and Pagoto, S. (2011) Ten Year Revision of the Brief Behavioral Activation Treatment for

Depression: Revised Treatment Manual *Behavior Modification,* **35**(2), 111-161.

Lejuez, C., Hopko, D. and Hopko, S. (2001) A brief behavioral activation treatment for depression: Treatment manual. *Behavior Modification,* **25**(2), 255-286.

Lewinsohn, P. and Graf, M. (1973) Pleasant activities and depression. *Journal of Consulting and Clinical Psychology,* **41**, 261-295.

Lewinsohn, P., Hoberman, H., Teri, L. and Hautzinger, M. (1985) An integrative theory of depression. In *Theoretical Issues in Behavior Therapy* (Eds, Reiss, R. and Bootzin, S.), Academic New York, pp. 331-59.

Lewinsohn, P., Sullivan, M. and Grosscup, S. (1980) Changing reinforcing events: an approach to the treatment of depression. *Psychotherapy Theory Research and Practice,* **17**, 322-334.

Lewinsohn, P. M. (1974) A behavioral approach to depression. In *Psychology of depression: Contemporary theory and research* (Eds, Friedman, R. J. and Katz, M. M.), John Wiley & Sons, Oxford, pp. 157−185.

Lewis, G., Pelosi, A., Araya, R. and Dunn, G. (1992) Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychological Medicine*, **22** 465-486.

Light, R. and Pillemer, D. (1984) *Summing up. The science of reviewing research,* Harvard University Press, Cambridge, Massachusetts, and London, England.

Lipsey, M. and Wilson, D. (1993) The efficacy of psychological, educational and behavioural treatment. *American Psychologist,* **48**, 1181-1209.

Lloyd, K., Jenkins, R. and Mann, A. (1996) Long-term outcome of patients with neurotic illness in general practice. *British Medical Journal,* **313**, 26-28.

Loannidis, J. (1998) Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *Jounral of the American Medical Association,* **279**, 281-286.

Lorant, V., Croux, C., Weich, S., Deliege, D., Mackenbach, J. and Ansseau, M. (2007) Depression and socio-economicrisk factors:7-year Longitudinal population study. *British Journal Of Psychiatry*, **190**, 293-298.

Lorant, V., Deliege, D., Eaton, W., Robert, A., Philippot, P. and Ansseau, M. (2003) Socioeconomic inequalities in depression: a meta-analysis. *American Journal of Epidemiology* **157**, 98-112.

Lovell, K., Cox, D., Haddock, G., Jones, C., Raines, D., Garvey, R., Roberts, C. and Hadley, S. (2006) Telephone administered cognitive behaviour therapy for treatment of obsessive compulsive disorder: randomised controlled non-inferiority trial. *British Medical Journal*, **333**, 883-887.

Lovell, K. and Richards, D. (2000) Multiple Access Points and Levels (MAPLE): Ensuring Choice Accessibility & Equity for CBT Services. *Behavioural and Cognitive Psychotherapy*, **28**(4), 379-391.

Luborsky, L., Mark, D. and Hole, A. (1995) Manual for supportive expressive dynamic therapy for depression: adaptation of the general SE manual. In *Psychodynamic Psychotherapy for Axis 1 Disorders* (Eds, Barber, J. and Crits-Christoph, P.), Basic Books, New York.

MacPhillamy, D. and Lewinsohn, P. (1982) The pleasant events schedule: Studies in reliability, validity and scale intercorrelation. *Journal of Consulting and Clinical Psychology*, **50**, 363-380.

Maldonado A (1982): Behavioral Therapy and Depression. *: Revista de Psicología General y Aplicada*.

Maldonado Lopez, A. (1982) Behavioural therapy and depression. *Revista dePsicologia General y Aplicada*, **37**, 31-56.

Maldonado Lopez, A. (1984) Behavioural therapy and depression: an experimental analysis of the interaction between cognitive and behavioural therapies and pharmacological therapy in depressed people. *Revista dePsicologia General y Aplicada*, **39**, 517-535.

Marks, I. (2000) Forty years of psychosocial treatments. *Behavioural and Cognitive Psychotherapy*, **28**(4), 323-334.

Martell, C., Addis, M. and Jacobson, N. (2001) Depression in Context. Strategies for Guided Action. *Norton: New York*.

Martell, C., Dimidjian, S. and Herman-Dunn, R. (2010) *Behavioral activation for depression: a clinicians guide*, Guilford, New York.

Mason, V., Shaw, A., Wiles, N., Mulligan, J., Peters, T., Sharp, D. and Lewis, G. (2007) GPs' experiences of primary care mental health research: a qualitative study of the barriers to recruitment. *Family Practice,* **24**(5), 518-525.

Mayo, E. (1993) *The human problems of an industrial civilization. Volume 3. 2nd edition.*, MacMillan, New York.

Mazzucchelli, T., Kane, R. and Rees, C. (2009) Behavioral Activation Treatments for Depression in Adults: A Meta-analysis and Review. *Clinical Psychology Science and Practice,* **16**(4), 383-411.

McCarney, R., Warner, J., Iliffe, S., Aan Haselen, R., Griffin, M. and Fisher, P. (2008) The Hawthorne Effect: a randomised, controlled trial. *BMC Medical Research Methodology,* **7**(30).

McCrone, P., Knapp, M., Proudfoot, J., Clash, R., Cavanagh, K., Shapiro, D., Ilson, S., Gray, J., Goldberg, D., Mann, A., Marks, I., Everitt, B. and Tylee, A. (2004) Cost-effectiveness of computerised cognitive behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *British Journal of Psychiatry,* **185**, 55-62.

McKendree-Smith, N. (1998a) Cognitive and behavioral bibliotherapy for depression: An examination of efficacy and mediators and moderators of change. *Dissertation Abstracts International.*

McKendree-Smith, N. (1998b) Cognitive and behavioural bibliotherapy for depression: an examination of efficacy and mediators and moderators of change. *Dissertation Abstracts International*.

McLean P and Hakstian A (1979): Clinical depression: comparative efficacy of outpatient treatments. *: Journal of consulting and clinical psychology*.

McLean, P. and Hakstain, R. (1979) Clinical depression: comparative efficacy of outpatient treatment. *Journal of Consulting and Clinical Psychology* **47**, 818-836.

McManus, S., Meltzer, H., Brugha, T., Bebbington, P. and Jenkins, R. (2009) *Adult psychiatric morbidity in England 2007: Results of a household survey.* The NHS Information Centre for Health and Social Care, Leeds.

McNamara, K. and Horan, J. (1986) Experimental construct validity in the evaluation of cognitive and behavioral treatments for depression. *Journal of Counselling in Psychology.*, **33**, 23-30.

Medical Research Council (1998) *MRC guidelines for good clinical practice in clinical trials.* Medical Research Council, London.

Medical Research Council (2000) *A framework for development and evaluation of RCTs for complex interventions to improve health* Medical Research Council, London.

Medical Research Council (2008) *Developing and evaluating complex interventions: new guidance.* Medical Research Council, London.

Milne, D. (2010) Can We Enhance the Training of Clinical Supervisors? A National Pilot Study of an Evidence-Based Approach. *Clinical Psychology and Psychotherapy,* **17**, 321-328.

Milne, D., Keegan, D., Westerman, C. and Dudley, M. (2000) Systematic process and outcome evaluation of brief staff training in psychosocial interventions for severe mental illness. *Journal of Behavior Therapy and Experimental Psychiatry,* **31**, 87-101.

Moher, D., Cook, D., Eastwood, S., Olkin, I., Drummond, R. and Sroup, D. (1999) Improving the quality of reporting the results of meta-analysis of randomised controlled trials: The QUOROM statement. *Lancet,* **354**(27), 1896-1900.

Moher, D., Jadad, A., Nichol, G., Penman, M., Tugwell, P. and Walsh, S. (1995) Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Controlled Clinical Trials,* **16**, 62-73.

Moher, D., Liberatl, A., Tetzlaff, J., Altman, D. and The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine,* **151**(4), 264-269.

Moher, D., Pham, B., Lawson, M. and Klassen, T. (2003) The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment,* **7**, 1-90.

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V. and Ustun, B. (2007) Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet,* **370**, 851-858.

Munt, J., Marks, I., Shear, K. and Greist, J. (2002) The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *British Journal of Psychiatry,* **180**, 461-464.

National Institute for Health and Clinical Excellence (2008) *Guide to the methods of technology appraisal.* NICE, London.

National Institute for Health and Clinical Excellence (2009) *NICE clinical guideline 91:Depression in adults with a chronic physical health problem.* NICE, London.

National Institute for Health and Clinical Excellence (2011) *Common mental health disorders: identification and pathways to care.* NICE, London.

National Institute of Clinical Excellence (2009) *Management of Depression in Primary and Secondary Care.* NICE, London.

National Research Ethics Service (2007) *Information sheets and consent forms. Guidance for researchers and reviewers. Version 3.2.* National Patient Safety Agency, NHS, London.

Nicholson, A., Kuper, H. and Hemingway, H. (2006) Depression as an aetiologic and prognostic factor in coronary heart disease: a meta-analysis of 6362 events among 146,538 participants in 54 observational studies. *European Heart Journal,* **27**, 2763-2774.

Nuesch, E., Trelle, S., Richenbach, S., Rutjes, A., Burgi, E., Scherer, M., Altman, D. and Juni, P. (2009) The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *British Medical Journal,* **339**, b3244, doi: 10.1136/bmj.b3244

Office for National Statistics (2000) *Key Health Statistics from General Practice 1998. Analyses of morbidity and treatment data, including time trends, England and Wales.* In *Series MB6 No. 2,* ONS, London.

Ostler, K., Thompson, C., Kinmonth, A., Peveler, R., Stevens, L. and Stevens, A. (2001) Influence of socio-economic deprivation on the prevalence and outcome of depression in primary care: the Hampshire Depression Project. *British Journal of Psychiatry,* **178**, 12-17.

Padfield, M. (1976) The comparative effects of two counseling approaches on the intensity of depression among rural women of low socioeconomic status *Journal of Counselling Psychology*, **23**, 209-214.

Pincus, H. and Pettit, A. (2001) The societal costs of chronic major depression. *Journal of Clinical Psychiatry,* **62**(6), 5-9.

Pollack, M. (2005) Comorbid anxiety and depression. *Journal of Clinical Psychiatry* **66**, 22-29.

Rehm, L. (1978) Mood, pleasant events, and unpleasant events: two pilot studies. *Journal of Consulting and Clinical Psychology,* **46**, 854-859.

Richards, D., Lankshear, A., Fletcher, J., Rogers, A., Barkham, M., Bower, P., Gask, L., Gilbody, S. and Lovell, K. (2006) Developing a U.K. protocol for collaborative care: a qualitative study. *General Hospital Psychiatry*, **28** 296– 305.

Richards, D., Lovell, K., Gilbody, S., Gask, L., Torgerson, D., Barkham, M., Bower, P., Bland, J., Lankshear, A., Simpson, A., Fletcher, J., Escott, D., Hennessy, S. and Richardson, R. (2008) Collaborative care for depression in UK primary care: A randomised controlled trial. *Psychological Medicine,* **38**(2), 279-288.

Richards, D. and Whyte, M. (2008) *Reach Out: National Programme Student Materials to Support the Delivery of Training for Practitioners Delivering Low Intensity Interventions,* Rethink, UK.

Richardson, W., Wilson, M., Nishikawa, J. and Hayward, R. (1995) The well-built clinical question: a key to evidence based decisions. *ACP Journal Club,* A12-A13.

Rogers, C. (1961) On Becoming a Person. *Houghton Mifflin: Boston, MA.*

Rosenthal, R. (1979) The 'file drawer problem' and tolerance for null results. *Psychological Bulletin,* **86**, 638-641.

Rothstein, H., Sutton, A. and Borenstein, M. (2005) *Publication bias in meta-analysis: prevention, assessment and adjustments,* Wiley, Chichester.

Royston, P. (2004) Multiple imputation of missing values. *The Stata Journal,* **4**(3), 227-241.

Saarni, S., Suvisaari, J., Sintonen, H., Pirkola, S., Koskinen, S., Aromaa, A. and Lonnqvist, J. (2007) Impact of psychiatric disorders on health related quality of life: general population survey. *British Journal of Psychiatry,* **190**, 326-332.

Sackett, D. and Wennberg, J. (1997) Choosing the best research design for each question. *British Medical Journal,* **315**, 1636.

Sapin C, Fantino B, Nowicki M and Kind P (2004) Usefulness of EQ-5D in Assessing Health Status in Primary Care Patients with Major Depressive Disorder

In *Health and quality of life outcomes,* Vol. 2 Biomed central open access.

Schulz, K., Altman, D. and Moher, D. (2010) CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal,* **340**, 698-702.

Schulz, K. and Grimes, D. (2002) Allocation concealment in randomised trials: defending against deciphering *Lancet* **359**, 614-618.

Scogin, F., Jamison, C. and Gochneaur, K. (1989) Comparative efficacy of cognitive and behavioral bibliotherapy for mildly and moderately depressed older adults. *Journal of Consulting and Clinical Psychology,* **57**, 403-407.

Shaw, B. F. (1977) Comparison of cognitive therapy and behavior therapy in the treatment of depression. *Journal of Consulting and Clinical Psychology,* **45**(4), 543-551.

Singleton, N., Bumpstead, R., O'Brien, M., Lee, A. and Meltzer, Y. (2001) *Office of National Statistics: Psychiatric Morbidity Among Adults Living in Private Households, 2000.* HMSO, London.

Skinner, B. (1953) *Science and human behavior,* MacMillan, Oxford, England.

Skinner, B. (1974) *About Behaviourism,* Alfred Knopf, New York.

Skinner, D. A. (1984) Self-control of depression: A comparison of behavior therapy and cognitive behavior therapy. *Dissertation Abstracts International,* **45**(3).

Song, F., Eastwood, A., Gilbody, S., Duley, L. and Sutton, A. (2000) Publication and related biases. *Health Technology Assessment,* **4**, 1-115.

Spielmans, G., Pasek, I. and McPhall, J. (1980) What are the active ingredients in cognitive and behavioral psychotherapy for anxoius and depressed children?A meta-analytic review. *Clinical Psychology Review,* **27**, 642-654.

SPSS for Windows (2008) SPSS version 17. Chicago.

Stata Corporation (2003) Version 10 [program]. College Station Texas.

Sterling, T., Rosenbaum, W. and Weinkam, J. (1995) Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *American Statistician,* **49**, 108-122.

Sterne, J. and Egger, M. (2001) Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology,* **54**, 1046-1055.

Sterne, J., White, I., Carlin, J., Spratt, M., Royston, P., Kenwad, M., Wood, A. and Carpenter, J. (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal,* **339**, doi: 10.1136/bmj.b2393.

Streiner, D. and Geddes, J. (2001) Intention to treat analysis in clinical trials when there are missing data. *Evidence Based Mental Health,* **4**, 70-71.

Sutton, A., Abrams, K. and Jones, D. (1998) Systematic reviews of trials and other studies. *Health Technology Assessment,* **2**, 19.

Tang, J. and Liu, J. (2000) Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology*, **53**, 477-484.

Taylor, F. and Marshall, W. (1977) Experimental analysis of a cognitive-behavioral therapy for depression. *Cognitive Therapy and Research*, **1**, 59-72.

Taylor, S., Repetti, R. and Seeman, T. (1997) Health psychology: what is an unhealthy environment and how does it get under the skin? *Annual Review of Psychology* **48**, 411-447.

Terrin, N., Schmid, C. and Lau, J. (2005) In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology*, **58**, 894-901.

The Sainsbury Centre for Mental Health (2003) *The economic and social costs of mental illness. Policy Paper 3.* The Sainsbury Centre for Mental Health, London.

Thomas, C. and Morris, S. (2003) Cost of depression among adults in England in 2000. *British Journal of Psychiatry*, **183**, 514-519.

Thompson (2005) Endnote

Thompson, L., Gallagher, D. and Breckenridge, J. (1987) Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology*, **55**, 385-390.

Thompson, S. and Higgins, J. (2002) How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, **21**, 1559-1573.

Tramèr, M., Reynolds, D., Moore, R. and McQuay, H. (1997) Impact of covert duplicate publication on meta-analysis: a case study. *British Medical Journal*, **315**, 635-640.

Tudor Hart, J. (1971) The inverse care law. *Lancet*, **297**(7696), 405-412.

Vickers, A. and Altman, D. (2001) Analysing controlled trials with baseline and follow up measurements. *British Medical Journal*, **323**, 1123-1124.

Von Korff, M. and Goldberg, D. (2001) Improving outcomes in depression. *British Medical Journal*, **323**, 948-949.

Wampold, B., Mondin, G., Moody, M., Stich, F., Benson, K. and Ahn, H., 203–215. (1997) A meta-analysis of outcome studies comparing

bona fide psychotherapies: Empirically, “all must have prizes”. *Psychological Bulletin,* **122**, 203-215.

Ward, E., King, M., Lloyd, M., Bower, P. and Friedli, K. (1999) Conducting randomized trials in general practice: methodological and practical issues. *British Journal of General Practice,* **49**, 919-922.

Watson, J. and Torgerson, D. (2006) Increasing recruitment to randomised trials: a review of randomised controlled trials. *BMC Medical Research Methodology,* **6**(34).

Weich, S. and Lewis, G. (1998) Poverty, unemployment, and common mental disorders: population based cohort study. *British Medical Journal,* **317**, 115-119.

Weich, S., Nazroo, J., Sproston, K., McManus, S., Blanchard, M., Erens, B., Karlsen, S., King, M., Lloyd, K., Stansfeld, S. and Tyrer, P. (2004) Common mental disorders and ethnicity in England: the EMPIRIC study. *Psychological Medicine,* **34**, 1543-1551.

WHO International Consortium in Psychiatric Epidemiology (2000) Cross-national comparisons of the prevalences and correlates of mental disorders. WHO International Consortium in Psychiatric Epidemiology. *Bulletin of the World Health Organisation* **78**, 413-426.

Wilhelm, K., Mitchell, P., Niven, H., Finch, A., Wedgewodd, L., Scimone, A., Blair, I., Parker, G. and Schofield, P. (2006) Life events, first depression onset and the serotonin transporter gene. *British Journal of Psychiatry,* **188**, 210-215.

Wilson, P. (1982) Combined pharmacological and behavioural treatment of depression. *Behavior Research and Therapy,* **20**, 173-184.

Wilson, P., Goldin, J. and Charbonneau Powis, M. (1983) Comparative efficacy of behavioral and cognitive treatments of depression. *Cognitive Therapy and Research.,* **7**, 111-124.

World Health Organisation (1992) *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines,* World Health Organisation, Geneva.

World Health Organisation (2001) *The World Health Report 2001: Mental Health: New Understanding, New Hope*, World Health Organisation, Geneva.

World Health Organisation (2002) *The World Health Report 2002: Reducing Risks, Promoting Healthy Life.* World Health Organisation, Geneva.

World Health Organization (2004) WHO Mental Health Survey consortium, Prevalence severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *Journal of the American Medical Association*, **291**(21), 2581-2590.

World Medical Association (2008) WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects: 59th WMA General Assembly, Seoul

Zeiss, A., Lewinsohn, P. and Munoz, R. (1979) Nonspecific improvement effects in depression using interpersonal skills training, pleasant activity schedules, or cognitive training. *Journal of Consulting and Clinical Psychology*, **47**, 427-439.